Standards for Cancer Registries Volume III

STANDARDS FOR COMPLETENESS, QUALITY, ANALYSIS, MANAGEMENT, SECURITY AND CONFIDENTIALITY OF DATA

Edited by Jim Hofferkamp, CTR   August 2008

SPONSORING ORGANIZATIONS:
Canadian Association of Provincial Cancer Agencies
Canadian Partnership Against Cancer
Centers for Disease Control and Prevention
College of American Pathologists
National Cancer Institute
National Cancer Registrars Association
Public Health Agency of Canada

SPONSORS WITH DISTINCTION:
American Cancer Society
American College of Surgeons
American Joint Committee on Cancer

NAACCR
North American Association of
Central Cancer Registries

North American Association of Central Cancer Registries

North American Association of Central Cancer Registries

Standards for Cancer Registries
Volume III

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

Sponsoring Organizations
Canadian Association of Provincial Cancer Agencies
Canadian Partnership Against Cancer
Centers for Disease Control and Prevention
College of American Pathologists
National Cancer Institute
National Cancer Registrars Association
Public Health Agency of Canada

Sponsors With Distinction
American Cancer Society
American College of Surgeons
American Joint Committee on Cancer

Edited by:
Jim Hofferkamp, CTR
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 and selected users of central cancer registry data, this Volume provides the record layouts and specifications for a number of standard NAACCR record formats, including: the standard record layouts for data exchange among central cancer registries; an update/correction record layout; and an analysis record layout that provides standard recodes for grouping selected variables such as race and primary site, as well as algorithms for converting data from one version of the International Classification of Diseases for Oncology to another.

❖ **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.

❖ **Volume IV,** *Standard Data Edits*
This standard document currently is only made available electronically as program code and a database. It documents standard computerized edits for data corresponding to the data standards Volume II.

❖ **Volume V,** *Pathology Laboratory Electronic Reporting*
Recommends message or format standards for electronic transmission of reports (pathology, cytology and hematology) from pathology laboratories to central cancer registries.

Copies of all standards documents can be viewed or downloaded from NAACCR's website at: http://www.naaccr.org. For additional paper copies, write to the NAACCR Executive Office at: 2121 W. White Oaks Drive, Suite B, Springfield, IL, 62704-6495.

TABLE OF CONTENTS

NAACCR BOARD OF DIRECTORS
2007-2008

President:
Susan T. Gershman, MS, MPH, PhD, CTR   2007-09
Massachusetts Cancer Registry
Telephone: (617) 624-5646
E-mail: susan.gershman@state.ma.us

Treasurer:
Maria J. Schymura, PhD   2007-09
New York State Cancer Registry
Telephone: (518) 474-2255
E-mail: mjs08@health.state.ny.us

Acting Executive Director:
Betsy A. Kohler, MPH, CTR
New Jersey State Cancer Registry
Telephone: (609) 588-3500
E-mail: betsy.kohler@doh.state.nj.us

Representative, Sponsoring Member Organization:
Elizabeth Ward, PhD   2006-08
American Cancer Society
Telephone: (404) 327-6552
E-mail: eward@cancer.org

Representatives at Large:
Susan Bolick-Aldrich, MSPH, CTR   2005-09
South Carolina Central Cancer Registry
Telephone: (803) 898-3626
E-mail: bolicks@dhec.sc.gov

Dennis Deapen, DrPH   Interim
Los Angeles Cancer Surveillance Program
Telephone: (323) 422-1574
E-mail: ddeapen@hsc.usc.edu

Mignon Dryden, CTR   2006-08
Cancer Registry of Northern California
Telephone: (530) 345-2483
E-mail: mdryden@healthcollaborative.org

Maureen MacIntyre, BSN, MHSA   2006-08
Cancer Care Nova Scotia
Telephone: (902) 473-6084
E-mail: maureen.macintyre@ccns.nshealth.ca

Howard J. Martin, PhD   2007-09
Virginia Cancer Registry
Telephone (804) 864-7865
E-mail: jim.martin@vdh.virginia.gov

Lilia C. O'Connor, MBA, RHIT, CTR   2003-08
California Cancer Registry
Telephone: (916) 779-0355
E-mail: lilia@ccr.ca.gov

Frances E. Ross, CTR   2005-09
Kentucky Cancer Registry
Telephone: (859) 219-0773
E-mail: fer@kcr.uky.edu

NAACCR REGISTRY OPERATIONS COMMITTEE
2007-2008

Chairs:

Robin Otto, RHIA, CTR
Pennsylvania Cancer Registry
Telephone: (717) 783-2548
E-mail: rootto@state.pa.us

Donald Shipley, MS
Oregon State Cancer Registry
Telephone: (971) 673-0986
E-mail: donald.k.shipley@state.or.us

NAACCR Staff:

Monica Thornton, BS
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: mthornton@naaccr.org

Jim Hofferkamp, BA, CTR
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: jhofferkamp@naaccr.org

Shannon Vann, BS, CTR
NAACCR, Inc.
Telephone: (217) 698-0800
E-mail: svann@naaccr.org

Committee Members:

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

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

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

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

Joanne Harris, CTR
Metropolitan Detroit Cancer Surveillance System
Telephone: (313) 578-4213
E-mail: harrisj@med.wayne.edu

Ali Johnson, CTR
Vermont Cancer Registry
Telephone: (802) 865-7749
E-mail: ajohnso@vdh.state.vt.us

Carol Hahn-Johnson, BS, CTR
NCI-SEER
Telephone: (301) 402-6226
E-mail: johnsoca@mail.nih.gov

Mary Jane King, MPH, CTR
Ontario Cancer Registry
Telephone: (416) 217-1242
E-mail: maryjane.king@cancercare.on.ca

Gary Levin, BA, CTR
Florida Cancer Data System
Telephone: (305) 243-4073
E-mail: glevin@med.miami.edu

Mary Lewis, CTR
Centers for Disease Control and Prevention
Telephone: (770) 488-4827
E-mail: bkf5@cdc.gov

Donna Morrell, CTR
Los Angeles Cancer Surveillance Program
Telephone: (323) 442-2324
dmorrell@hsc.usc.edu

Serban Negoita, MD, MPH
New York State Cancer Registry
Telephone: (518) 474-2255
E-mail: sxn04@health.state.ny.us

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

Joan Phillips, CTR
CDC-NPCR
Telephone: (770) 488-4739
E-mail: ggq8@cdc.gov

Cathy Rimmer, BA, CTR
NCRA Representative
Telephone: 336) 718-8462
E-mail: ccrimmer@novanthealth.org

Randi Rycroft, BA, BS, CTR
Colorado Central Cancer Registry
Telephone: (303) 692-2542
E-mail: randi.rycroft@state.co.us

Wendy Scharber, RHIT, CTR
Registry Widgets
Telephone: (763) 439-9408
E-mail: wendy@registrywidgets.com

Bette Smith, BTM, RHIT, CTR,
Ohio Cancer Incidence Surveillance System
Telephone: (614) 644-7058
E-mail: bette.smith@odh.ohio.gov

Brad Wohler
Florida Cancer Data System
Telephone: (305) 243-4600
E-mail: bwohler@med.miami.edu

NAACCR VOLUME III WORKGROUP
2007-2008

Chair-Deirdre Rogers MS, CTR
Mississippi Cancer Registry
Telephone: (601) 815-5479
E-mail: dbrogers@crr.umsmed.edu

Peggy Adamo, RHIT, CTR
National Cancer Institute/SEER Program
Telephone: (301) 435-4971
E-mail: adamom@mail.nih.gov

Mindy Burch, CTR
Community Cancer Care, Inc.
Telephone: 317-833-1268
E-mail: Mlbccc@aol.com

Lois Dickie, CTR
National Cancer Institute/SEER Program
Telephone: (301) 402-4285
E-mail: dickielo@mail.nih.gov

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

Jim Hofferkamp, BA, CTR
NAACCR, Inc.
Telephone: (217) 698-0800 x 5
E-mail: jhofferkamp@naaccr.org

Joan Phillips, CTR
Centers for Disease Control and Prevention
Telephone: (770) 488-4739
E-mail: ggq8@cdc.gov

Herman Menck, MBA
Marina Del Ray, CA 90292
Telephone: (310) 823-0070
E-mail: menckh@aol.com

Serban Negoita, MD, MPH
New York Cancer Registry
Telephone: (518) 474-2255
E-mail: Sxn04@health.state.ny.us

Lilia O'Connor, MBA, RHIT, CTR
California Cancer Registry
Telephone: (916) 779-0355
E-mail: lilia@ccr.ca.gov

Donald Shipley, MS
Oregon State Cancer Registry
Telephone: (971) 673-0986
E-mail: donald.k.shipley@state.or.us

Karen Starratt
Cancer Care Nova Scotia
Telephone: (902) 473-6058
E-mail: karen.starratt@ccns.nshealth.ca

PREFACE

One of the primary goals of the North American Association of Central Cancer Registries, Inc. (NAACCR) Registry Operations Committee (ROC) has been to review, update, and revise *Standards for Cancer Registries Volume III: Standards for Completeness, Quality, Analysis, Management, Confidentiality and Security 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 2008 Edition focus on:

❖ The addition of Chapter 6: Security and Confidentiality.

❖ Revisions to reflect updates to all chapters.

❖ Table 2: NPCR - CSS 2008 Data Submission Specifications.

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

❖ Table 4: Site/Histology Recode Based on International Classification of Childhood Cancer, Third Edition (ICCC-3) Based on ICD-O-3.

❖ Table 5: Standard Site Analysis Categories for Mortality Data (ICD-9 and ICD-10) SEER Cause of Death Recode 1969+ (3/25/2004).

❖ Table 6: SGC Codes for Canadian Provinces and Territories.

❖ Table 7: Standard Populations.

❖ The removal of Table 8: Standard Treatment Analysis Categories.

❖ Appendix H: Major-Minor Discrepancy Definitions for Colon has been changed to reflect only collaborative stage discrepancy definitions for colon primaries.

❖ The Inventory of Best Practices Assurance of Confidentiality and Security, previously found in Appendix J, was removed. Appendix J now includes examples of data use agreements.

❖ Updated reference information to include the most recent reference materials.

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

Robin Dietrich Otto, RHIA, CTR
Co-Chair, Registry Operations Committee

Donald K. Shipley, MS
Co-Chair, Registry Operations Committee

CHAPTER 1:

INTRODUCTION

The mission of the North American Association of Central Cancer Registries, Inc. (NAACCR) is to: (1) promote uniform data standards for cancer registration, (2) provide education and training, (3) certify population-based registries, (4) aggregate and publish data from central cancer registries, and (5) 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 NAACCR membership. The purpose of these standards is to increase the quality, comparability, and utility of cancer incidence data in North America.

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; 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: (1) the *Procedure Guidelines for Cancer Registries* (Series I-V), (2) training programs and educational CDs/DVDs, and (3) *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 interpretations and practices into account. These standards were formulated based on the following principles:

❖ *Model:* The central cancer registry model 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 consensus that any new registry should adopt these standards and existing registries should take them under serious consideration.
   – **SHOULD:** Experience has shown that other characteristics are strongly recommended, but not absolutely required as the **MUST** characteristics described above. These are designated by "**SHOULD"** in the standards. Some of the problems addressed as **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 "**MAY."**

Detailed discussions of methods have been omitted from this document when they are available elsewhere. References are provided for the sources 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/territorial 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: (1) reporting requirements, (2) patient record access, (3) enforceability, (4) data quality and data standards, (5) confidentiality and disclosure of data, (6) liability, and (7) specification of funding source. Section 2.1.1.1 provides a further explanation of these issues. 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 the federal agency Statistics Canada 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. The annual reporting on patterns of cancer occurrence is a joint activity of Statistics Canada, Public Health Agency of Canada, Canadian Cancer Society, National Cancer Institute of Canada, and the provincial/territorial cancer registries. The annual report, *Canadian Cancer Statistics*, can be found online at www.cancer.ca.

The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) 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, as well as 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 NCI also funds a cancer registry that, with technical assistance from SEER, collects information on cancer cases among Alaska Native populations. 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," 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 for the Cancer Registries Amendment Act; see Appendix D for the Benign Brain Tumor Cancer Registries Amendment Act). Prior to funding, the national legislation requires assurances from states that they will "provide for the authorization under State law of the statewide cancer registry, including publication of regulations." The national legislation mandates 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. The National Program of Cancer Registries (NPCR) is administered through the Centers for Disease Control and Prevention (CDC) and addresses three specific goals for its registries: (1) completeness, (2) timeliness, and (3) 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 in the United States 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. 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: (1) sharing of tumor records with other population-based registries, (2) facilitation of death clearance and other record linkages, and (3) preparation of reports for individual facilities that report all their tumors.

❖ For conciseness and the flexibility to make changes over time, laws and regulations **SHOULD** reference the more detailed documents containing reporting requirements, such as: (1) required reporting format; (2) registry data collection and coding manuals; and (3) 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: (1) hospitals or other facilities providing screening, diagnostic, palliative, or therapeutic services to patients who have reportable tumors; and (2) physicians, surgeons, and all other health care providers who diagnose or provide treatment for patients with reportable tumors.

> *Exception:* Patients previously reported by another hospital or facility that provides screening, diagnostic, or therapeutic services.

❖ All cancers **SHOULD** be reported to the state or provincial/territorial 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 (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

The 180-day standard is consistent with the requirements of the NPCR and the American College of Surgeons (ACoS) Commission on Cancer (CoC) Approvals Program for hospital cancer programs. Standard practices for reporting timelines in Canada are established by individual provincial/territorial cancer registries. There is no uniform national standard at this time. For more information on timeliness of reporting, see Section 2.3.4..

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 is restricted. The guidelines **SHOULD NOT** be so strict that approved researchers are denied access to the confidential data (see NAACCR's *Data Use and Confidentiality Task Force Report*). Regulations could make provisions for the following:

❖ Central registries **SHOULD** make reported data available 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 provincial/territorial cancer experience.

❖ Central registries **MAY** exchange patient-specific data with the reporting facility, any other cancer-control agency or clinical facility involved in the patient's care 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's 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 NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*)

❖ Multiple primary rules (see Section 2.1.2.2.)

❖ 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 ICD-O. At a minimum, all neoplasms with a behavior code of 2 or 3 in ICD-O-3 **SHOULD** be designated as reportable. Effective January 1, 2004, and later, benign and borderline intracranial and central nervous system tumors **SHOULD** be designated as 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, interest, and sufficient resources are available to collect all histologically confirmed high-grade squamous intraepithelial lesions and equivalent terms (see NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*).

❖ Any benign neoplasm 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: (1) all reporting facilities or practitioners required to report, (2) all cancer registrars in the coverage area, (3) all medical records or cancer registrar training programs or schools in the area, and (4) 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 to determine the number of primaries for each 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. For cases diagnosed January 1, 2007, and after, NAACCR endorses the 2007 Multiple Primary and Histology Coding Rules as the *de facto* standard for both central and hospital-based registries. For cases diagnosed prior to January 1, 2007, NAACCR endorses the SEER Program rules as the *de facto* standard in the United States.

The 2007 Multiple Primary and Histology Coding 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 primaries than those that are reported using the 2007 Multiple Primary and Histology Coding Rules.

The CCR follows the 2007 Multiple Primary Rules for cases diagnosed January 1, 2007, and after. For cases diagnosed prior to January 1, 2007, the CCR rules are followed by most of the provincial/territorial cancer registries, but some registries follow the IARC rules, some follow the SEER Multiple Primary and Histology Coding Rules, and some have developed their own rules. The CCR rules are different from the SEER Multiple Primary and Histology Coding Rules. Therefore, when data are published in Canada for cases diagnosed prior to January 1, 2007, the IARC rules are used to count multiple primaries, because this is the lowest common denominator. The CCR rules for cases diagnosed prior to January 1, 2007, do not assess the time of diagnosis, nor the behavior. Further details can be found in CCR's *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 and 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 usually excluded from routine treatment and survival statistics, most prominently those first diagnosed or treated somewhere other than the reporting hospital (see CoC's *Facility Oncology Registry Data Standards* [*FORDS*, Revised 2007]).

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. These criteria include: (1) both inpatient and outpatient cases, (2) patients seen only in the emergency room (including patients who are dead on arrival), (3) tumors diagnosed at autopsy, (4) patients seen for consultation only, (5) surgery centers, (6) physicians, (7) stand-alone centers, and (8) pathology laboratories (including 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 CCR. 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 important estimate to obtain is the number of case reports (not incidence cases) that central registry staff will be responsible for identifying and abstracting. This includes hospitals that are not submitting their own reports and also may include the following: (1) federal facilities not subject to state and provincial/territorial reporting requirements; (2) non-hospital sources such as clinics, physician offices, pathology laboratories, nursing homes, and coroner's offices; or (3) facilities outside of 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: (1) increase or decrease in the defined population, (2) new treatment facilities or the closures and mergers of existing facilities, (3) increase or decrease in physicians treating cancer patients, and (4) national standards for estimating completeness.

❖ *New or Expanding Central Cancer Registries:* New or 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

Currently, no firm standard has been developed for estimating the number of data collection staff within a central registry. However, the following criteria **SHOULD** be assessed when determining the number of staff necessary to perform data collection at reporting facilities:

❖ *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 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 types 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 laptop 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 be required, depending on the educational background and experience of the data collection staff. Standards for training are addressed in Section 2.2.9..

2.1.3.5. Standards for Data Collection 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 (CTRs) or those who are CTR-eligible **SHOULD** be used for performing 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: (1) tumor diagnosis, staging, and treatment; (2) data management; (3) epidemiology and statistics; (4) hardware and software applications; and (5) security and confidentiality.

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 in U.S. registries **SHOULD** be informed about the Cancer Information Service at 1-800-4-CANCER. The central registry **MAY** provide access to online forums and online resources such as the National Library of Medicine's MEDLARS® databases. These include Physician Data Query (PDQ®), CANCERLIT®, and MEDLINE®. Other resources for continuing education include the ACoS, NAACCR, NCRA, NPCR, and 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 CCR Technical Workshop, 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. In addition, 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/territorial, 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 the 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, 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 is covered separately in Section 2.2.8..

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 NAACCR's *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: (1) independent pathology laboratories (histopathology and hematology laboratories), (2) ambulatory surgery centers, (3) radiation therapy centers, and (4) 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 are followed back to the treating physician or facility (for additional information, see Section 2.2.3., Physician Reporting).

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 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 probably 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 which the registry works on a day-to-day basis. Broad support from the general public, voluntary agencies, and community special interest groups also can 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 long-term relationships 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 American 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 geographic locations 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 to obtain 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 the 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 for 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 in Section 2.2.5.2. 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 (see 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:* Written agreements between registries covering the uses and confidentiality of exchanged data. These agreements **MAY** be informal, simply requesting data and affirming the confidential nature of the data, or the agreements may be more formal legal documents, depending upon the laws governing release of data. For an example, see Appendix F: Sample Case Sharing Agreement.

❖ *Exchange Media:* Data **MAY** be exchanged between central registries across a variety of media. These exchange media include: (1) electronic files of data on diskette, CD ROM, DVD, or tape; (2) electronic files of data transferred by e-mail or website; (3) copies of paper abstracts; or (4) printed reports generated from computer systems (see Section 6.2.6. for information on data security).

❖ *Exchange Format:* The North American 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. When submitting to NAACCR, the Canadian provincial/territorial registries use the NAACCR format; when Canadian provincial/territorial registries exchange data with other Canadian provincial/territorial registries, the CCR record layout is used.

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 in an electronic format 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.

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 privacy regulations such as HIPPA in the United States or the Canadian Personal Information Protection and Electronic Documents Act 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 or Canadian 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: (1) direct mailings to individuals; (2) meetings with groups, such as the staff of large clinics or specialty laboratories; (3) presentations at scheduled meetings, such as hospital staff meetings or medical society meetings; and (4) regional presentations and orientation workshops organized by the central registry.

❖ These communications **SHOULD** be directed to: (1) all relevant physicians (e.g., pathologists, medical oncologists, dermatologists, general surgeons and surgical specialists, and radiation oncologists); and (2) all related facility personnel (e.g., hospital administrators, health information service administrators, and cancer registry managers).

In situations where 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 as well as areas where patients are being diagnosed and treated. Facility openings, closings, and mergers as well as the establishment of new screening programs all can affect the 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.

2.2.8. Death Clearance Follow-Back

Death clearance is defined as the process of matching registered deaths in a population against reportable conditions in the central cancer registry database for two purposes: (1) ascertainment of death information for persons in the central cancer registry (referred to as death clearance match), and (2) identification of all deaths with a reportable condition mentioned as a cause of death which are not found in the central cancer registry (referred to as death clearance follow-back).

This section addresses death clearance follow-back. Death clearance match for the purpose of obtaining information on tumor records already registered is described in Section 2.2.12..

Death clearance follow-back is an essential step in achieving complete population-based reporting. It serves as a check on the completeness of reporting from all sources and often identifies tumor records that should be reported from those sources but were not.

Death certificates included in the death clearance follow-back process are those that contain a diagnosis of *in situ* or malignant cancer, or benign or borderline intracranial or central nervous system (CNS) tumor, but are not found in the central registry database. They are patient non-matches, meaning the patient is in the vital records mortality file with a cause of death corresponding to a reportable condition but is not in the central registry database, or they are tumor non-matches, in which the patient is in the central registry database but the reportable condition in the mortality file is not found in the central registry database.

2.2.8.1. Standards

The official mortality file from the state/province vital records office used to conduct death clearance follow-back **MUST** include all causes of death (underlying cause and all other contributing causes) so that all potentially missed cases can be identified.

If the underlying cause of death is the only cause of death coded, the central registry **MUST** review all death certificates to ensure that the reportable condition is not listed as a contributing cause of death. In some states, a flag is set by vital records whenever a reportable condition is mentioned on the death certificate; this can be very helpful in eliminating manual review of hardcopy death certificates.

Death clearance follow-back **MUST** be completed at least once each year for a specified year of deaths. For the death clearance follow-back process to be complete for a specified year, all potential incidence cases identified during the matching process **MUST** be resolved as a death certificate only (DCO), incidence case, or excluded as non-reportable as defined below:

❖ *DCO Case:* A case for which the only information the central registry has is a death certificate stating that the patient has the reportable condition.

❖ *Incidence Case:* A case first identified as a non-matched cancer death for which confirmation of the diagnosis and other information are obtained through follow-back to clinical sources.

❖ *Non-Reportable Case:* A case first identified as a non-matched cancer death but after further investigation does not meet reporting criteria.

All DCO and incidence cases **MUST** be abstracted and included in the central registry files submitted for required calls for data.

Timing must be planned carefully. The goals are to link every reportable condition from the time period against every death from that period, avoiding unnecessary follow-back and distributing the follow-back workload across a reasonable time. The timing for performing final death clearance follow-back is based on when the final mortality file is complete and when the central registry database is complete for the diagnosis year corresponding to the year of deaths. Careful timing of the process to ensure that all central registry and mortality records are available before the death clearance follow-back process begins will eliminate an excess of non-matched cases requiring follow-back.

Central registries may find it beneficial to conduct death clearance follow-back more than once per year. The mortality 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 records 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 mortality or central registry files. Additional linkage or linkages then **MUST** be performed when the central registry considers its case file to be complete and the vital records office considers the mortality file complete for the year. Canada also has a national death clearance process that is conducted annually through the CCR. This involves a national-level, electronic linkage to the Canadian Mortality Data Base, maintained by Statistics Canada.

For new central registries, the first years conducting death certificate follow-back are the most difficult and time consuming. The mortality file will contain a greater number of potential incidence cases that are not in the central registry database than in succeeding years (e.g., the diagnosis date was prior to the central registry reference date, the reporting source was not required to report at the time the patient was diagnosed, or the facility that should have reported the case to the central registry failed to report it). As the central registry matures, more of these cases will have been diagnosed after the central registry reference date or after non-hospital facilities were required to report, or will have been identified as a "missed case" during a casefinding audit.

The central registry also **SHOULD:**

❖ Include a tumor linkage comparison in the death clearance follow-back process (i.e., verify that for patients in both the registry file and the mortality file, the reportable conditions are the same primary). If there are discrepancies, follow-back is necessary to determine if the patients had additional reportable tumors that should be registered.

❖ Apply the ICD-O rules (manually and/or with an automated mortality classification system) to classify and select the underlying and other contributing causes of death. Across jurisdictions, the number of codes kept in the vital records database may differ.

❖ Employ standard coding for DCO cases as specified in the most recent version of NAACR's *Standards for Cancer Registries, Volume II: Data Standards and Data Dictionary*.

❖ Analyze the results of the death clearance follow-back process, 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.

2.2.8.2. Required Components

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

❖ The central registry **SHOULD** establish a formal agreement with the state, provincial/territorial, or national vital records 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 mortality files and tumor records and identify matches, non-matches, and potential matches with a reportable condition as a cause of death.

❖ The central registry **MUST** have an adequate number of staff, trained in casefinding and abstracting, to perform follow-back. Factors that influence the number of staff needed for casefinding and follow-back include length of time the central registry has performed death clearance follow-back, software used by the central registry, level of automation of death clearance and follow-back processes, and size of the central registry database. Whether a central registry performs casefinding on an annual basis or more frequently also may influence the number of trained staff needed to perform these functions.

❖ The central registry **MUST** track the progress and results of follow-back. Tracking preferably **SHOULD** be automated, but **MAY** be manual.

2.2.9.   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. The 2007 Multiple Primary and Histology Rules are used to determine multiple primaries for all cases diagnosed on or after January 1, 2007. This is 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 (see http://training.seer.cancer.gov/). Canada has adopted the new 2007 Multiple Primary and Histology Rules for all cases diagnosed January 1, 2007, and after (see Section 2.1.2.2.).

2.2.9.1. 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: (1) criteria for case reportability, and (2) rules for multiple primary determinations.

Training **SHOULD** be provided to central registry staff and to staff in all reporting facilities in which they 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's reporting requirements.

Training **MAY** be offered as formal education, online training, through professional association meetings, or at workshops scheduled by the central registry. Professional publications and central registry newsletter articles **MAY** be used as resources to help address reporting problems.

See Section 3.2.4. for other training standards.

2.2.10. Monitoring Completeness of Reporting and Ensuring Compliance

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.

2.2.10.1. Standards

The central registry **SHOULD** monitor 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.

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

❖ Completeness of reporting for each facility, each county, and the entire coverage area.

❖ Status of screening of the casefinding sources, such as: (1) type of pathology report (i.e., surgical specimens, cytologies, autopsies, bone marrows, etc.); (2) disease and operations indices; and (3) 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 during which some casefinding sources were not reviewed.

When the number of reported tumors deviates widely from the expected number, 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.

2.2.11. 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.

2.2.11.1. 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 also **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.

2.2.12. Patient Follow-Up

Registries intending to evaluate survival and/or quality of life **MUST** follow all registered patients for life (carcinomas *in situ* of the cervix uteri and basal and squamous skin cancers, if registered, often are not followed). Methods of obtaining follow-up will vary due to local considerations, such as the number of tumor records 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.

2.2.12.1. Standards

The choice of methods or sources for obtaining patient follow-up **SHOULD** be driven by: (1) the availability of the method or source to the central registry, and (2) the effectiveness of the method or source.

Patients who are categorized 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.

❖ 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.

❖ 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: (1) hospitals; (2) local/family physicians; (3) specialist physicians; (4) nursing homes; (5) telephone books; and (6) the Internet (e.g., online telephone books, reverse directories, genealogy, social security number search, and 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 matching 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 eliminating non-reportable cases, such as: (1) deaths not caused by cancer but coded as cancer deaths, (2) out-of-jurisdiction residents, and (3) cancers diagnosed before the central cancer registry reference date.

Step 3 is resolving 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 calculating the DCO rate as follows:

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

The percentage of DCO cases, or percent DCO, 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 otherwise would 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 central 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 activities have not been performed. DCO percentage rates of more than 3 probably are a result of underreporting from other sources, 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 lead to incorrect conclusions about the cancer burden in the population. There are a number of ways that central registry staff can determine the level of data completeness in the cancer registry: (1) calculating the percentage of cases identified by DCO; (2) analyzing collected data to be sure they follow known patterns (e.g., incidence is greater than mortality); and, most importantly, (3) 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 the 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 been refined. The use of this method makes the following assumptions: (1) cancer death rates are complete, and (2) the ratio of SEER incidence to U.S. mortality rates is 80 percent and is similar within race-sex site groups (with a 20 percent allowance for variation in case fatality).

For a complete list of assumptions and the calculation method, see Appendix G: Method To Measure Completeness.

All calculations and analyses addressed in this section and in Section 2.3.3. assume that duplicate records for persons and tumors are 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 their judgment to determine whether 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 appears to be rural whites living 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 more than 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 of these cases are clinically diagnosed and about 1 to 1.5 percent are DCO cases. In Canada, the percentage of cases microscopically confirmed is between 80 to 95 percent. The 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 the 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 are collected, processed, analyzed, and reported depends on many factors, some of which are within the registry's control, and some of which are not. Historically, abstracting began 6 months after the cancer diagnosis because treatment usually was complete within those 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 tumor records from facilities, within limits. Many facilities are capable of concurrent reporting and can complete abstracts in real time. Electronic pathology reporting has expedited case identification and the abstracting process for some reporting facilities and central cancer registries.

Transmission of cases from a reporting facility to the central cancer registry also affects 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; 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 Chapter 3: Data Quality; Chapter 4: Data Analysis and Reporting; and Chapter 5: Data Management. 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

❖ *ACoS/CoC: Cancer Program Standards*, 2004 Revised Edition, Standard 3.3: For each year between surveys, 90 percent of cases are abstracted within 6 months of the date of first contact.

❖ *SEER 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.

❖ *CDC/NPCR:* 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.

❖ *NAACCR:* 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.

❖ *CCR:* As of March 1, 2003, the annual CCR Call for Data (new and updated cases) deadline is 14 months from the calendar year end (i.e., all 2001 data due no later than March 1, 2003).

2.3.5. Casefinding Audit Results

Casefinding audits are studies involving independent re-ascertainment 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: (1) physician offices; (2) clinics and outpatient facilities, including radiation therapy centers and surgery treatment centers; (3) freestanding and out-of-state pathology laboratories; (4) hospices; and (5) facilities that are outside of the coverage area of a central registry but that treat 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. These will allow repetition of the study at a later time or in another area or group of facilities. 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 study 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. Another reason prostate was chosen is because it is often diagnosed and treated in non-hospital settings and, therefore, is at a greater risk for being missed. Pathology and cytology reports were the source documents for the audit. Limiting this audit to the most common sites and source documents achieved the goal of ensuring that we review at least 5 percent of the eligible cases while 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. Review of source records was limited to pathology and bone marrow reports. The objectives of the 2004 casefinding study were to estimate casefinding completeness for all SEER central registries for cases diagnosed in 2001, estimate timeliness of reporting eligible cases diagnosed in 2001 for all primary sites, and compare those estimates to the SEER standards for casefinding.

Further, the 2004 casefinding study sought to audit: (1) a representative proportion of cases in each registry; (2) the newly reportable Hematopoietic diseases; (3) the newly reportable benign brain and CNS diseases; and (4) hospital and "pathology only," or non-hospital cases.

The field portion of the audit took place between July and October 2004. The auditors reviewed a total of 21,825 cases and identified 554 missed and late cases. The SEER standard for casefinding is two percent or fewer missed plus late cases. The individual registry missed plus late case rates ranged from a high of 6.4 percent to a low of 0.11 percent. Seven of the 14 registries met the SEER standard for casefinding based on the results of this audit.

The 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 workplan 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. The NPCR percent of case completeness formula is determined by the following formula:

100 - (number of missed cases X 100/total number of reportable cases found by auditors)

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 Practices 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 Practices Workshop 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 for Y +1. A is the number of follow-up dates on or after January 1, (Y+1) (includes alive and dead). 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/CoC:* 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, those aged 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 ASSURANCE

3.1. STRUCTURAL REQUIREMENTS

3.1.1. General Requirements

Data quality assurance encompasses the personnel and activities that focus on the assessment and improvement of data quality. Every feature of the central registry's operations, including the laws and regulations under which the registry operates, impacts data quality. Examples include:

- ❖ 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.

3.1.1.1. Standards

The registry **MUST** have a quality assurance program with specified activities integrated into basic central registry operations. The central registry's budget **MUST** specify—and adequately fund—quality control staff and activities.

Definitions for the quality assurance program **SHOULD** include: (1) assignment of a qualified individual to perform quality control activities, (2) schedule for routine edits and reports, and (3) 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.

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 staff **MUST** consist of personnel who have adequate qualifications to conduct registry business in a timely, competent manner. Registry staff should be competent in the following areas: (1) knowledge of cancer registration, (2) data evaluation and analysis (including statistics and epidemiology), (3) skills related to cancer surveillance software, (4) training and professional development, and (5) organizational and communication skills.

3.1.2.1. Standards

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

❖ *CTRs:* 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.

❖ *Quality Control Manager/Supervisor/Coordinator:* The central registry **MUST** 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.

❖ *Data Analyst(s):* An individual who trained in both data analysis and 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 most often is 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: (1) data items definitions, codes, and formats; (2) coding rule interpretations and procedures; and (3) 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. CCCR, 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. 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 judgment, analysis, and/or interaction with reporting facilities. In most cases for which technology use has increased, data quality has improved.

Central registry computer software systems **MUST** provide a repository for data and the tools to generate case listings, research datasets, or other registry end products. It also is a major focal point for quality assessment processes. One basic function of central registry software is maintenance of data integrity. Careful and effective data management as well as 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 each data item. 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 5.9.).

❖ *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 operations 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 Quality Assessment 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: (1) the utility of local data is compromised when data categorization is not comparable, (2) a standard edit contributes to comparable data, and (3) errors in primary editing steps cannot be fixed by subsequent edits.

Standards for edits are discussed in Section 5.9.1. 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 5.9.):
– 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.

❖ *Sampling/Case Listing:* The central registry system **SHOULD** allow drawing of samples for quality assessment studies by any desired characteristic.

❖ *Staff:* The central registry **MUST** have staff trained 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: (1) prompts; (2) coding choice lists; (3) online help; (4) edits: single-field, multi-field, multi-record, or multi-database; and (5) error messages.

Auto-coding is convenient but can be risky, especially for histology variables or surgical procedure names for which modifiers to a root word can change a histology or treatment 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's *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 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: (1) patient care evaluation, (2) cancer surveillance research, (3) etiologic research, (4) cancer control program evaluation, and (5) outcomes research.

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: (1) data and data translations codes, (2) item sequence and record layout, and (3) electronic media specifications.

NAACCR's recommended exchange format is presented in *Standards for Cancer Registries Volume I: Data Exchange Standards and Record Description*, and in *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 clarification, minimizes permanent information loss, and increases the usefulness of the data.

Item, internal consistency, and inter-record edits **SHOULD** be applied routinely before new records are added to the database. 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 5.9.).

Information on EDITS is located on the Registration Standards page of the NAACCR website (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). For more information on record consolidation, see Section 5.10..

3.2.   PROCESS STANDARDS

3.2.1.   Standards for Coded Data Items

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 coding systems 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 format and using standard codes.

3.2.2.   Standards for Text Data Items

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 items **SHOULD** be edited using the appropriate edit standards for that item. The EDITS metafile contains standard edits for each of the standard-setting organizations. Some states include state-specific edits as well. 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 that data collection is accurate, consistent, and complete (see Section 2.2.9.).

3.2.4.1. Required Components

Training **MUST** be provided to the central registry staff involved in data collection and quality control as well as 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:* Instructions on reportable neoplasms, casefinding procedures, abstracting requirements, application of multiple primary rules, ICD-O coding, staging, and 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.

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/DVDs on core central registry analysis.

❖ Training and educational media.

❖ Web-based training modules.

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

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

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

3.2.5. Quality Assessment Activities

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

❖ *Process Control*: Statistical process control involves the prospective monitoring of rationally aggregated results (e.g., percent unknown primaries) that should exhibit stability over time or across geographic areas, reporting facilities, and subpopulations (e.g., males versus females). Process controls do not necessarily focus on errors. Process control design requires statistical expertise, including the specification of an appropriate probability model, the selection of a sampling plan, the selection of appropriate control-charting procedures, and the specification of control limits.

❖ *Special Assessments:* Central registries **MAY** 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 (Dryden MM, Brogan K. Quality Control. In: Central Cancer Registries Design, Management and Use. Editors: Menck HR, Deapen D, Phillips JL, Tucker TC. Second Edition. Dubuque, IA: Kendall/Hunt; 2007: 211-223).

❖ *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 results can be statistically represented by comparing the results to proficiency 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 resulting actions to be taken, **SHOULD** be fully documented. Measures of central registry quality that should benefit from the formal development of process controls include, but are not limited to, the following:

❖ Visual review rejection rates.

❖ Duplicate entry/recoding/rejection rates.

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

❖ Missing data and the use of unknown or ill-defined codes for data items considered critical for 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, a computer can assist in the acquisition, management, and charting of process control data. 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/recoding audits to assess overall data reliability (see Sections 2.2.10. and 2.3.5. for discussions of casefinding audits). 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: (1) introduction and rationale; (2) statement of purpose; (3) sampling plan, including sample size considerations, stratifications, and randomization; (4) eligibility criteria and study population; (5) procedures to be followed for study execution; and (6) 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 the 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 any 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 at the central registry, information about the corrections **SHOULD** be returned to the abstractor for review. Discrepancy reports or error reports from edits also **MAY** be returned.

The central registry **SHOULD** provide the results of recoding audits, casefinding audits, and reabstracting audits with analysis of discrepancies and recommendations for improvement to abstractors. Feedback on the 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 assessment activities as feedback to other aspects of registry functioning. For example, the central registry **SHOULD:** (1) interpret the results of quality monitoring, and incorporate the conclusions when revising training materials, documentation, or item definitions as needed; and (2) 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 often are used to retrospectively assess the accuracy (agreement with source medical records) and reproducibility (agreement among data collectors) of registry data. Audits are studies on a sample of cases and **MUST** be conducted in accordance with a study protocol that states the study objectives, describes the sampling scheme, and outlines plans for the analysis. 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 that is actually selected. A random number table, found as a reference in statistics handbooks, may be used to randomly assign numbers to each facility.

❖ *Stratified Random 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 may be grouped by geographic location. The desired number of groups is selected using a random number generator, and from the groups selected, the desired number of hospitals is selected at random.

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 a specific study (see Appendix H for sample major-minor collaborative stage discrepancy definitions for colon primaries developed by SEER). Such studies require arbitration or reconciliation mechanisms 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 information 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. (Dryden MM, Brogan K. Quality Control. In: Central Cancer Registries Design, Management and Use. Editors: Menck HR, Deapen D, Phillips JL, Tucker TC. Second Edition. Dubuque, IA: Kendall/Hunt; 2007: 211-223).

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 previously 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 abstractor and coder 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. Results from this study method help identify ambiguity or inadequacy of existing data definitions and rules as well as 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

The kappa statistic measures 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 item. 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 complete disagreement. For most targets, values greater than 0.75 represent excellent agreement. Values below 0.40 represent poor agreement. Values between 0.40 and 0.75 represent fair-to-good agreement (Fleiss, J.L. Statistical Methods for Rates and Proportions, Second Edition. New York: John Wiley & Sons; 1981). Other useful statistics of reliability are percent agreement and percent positive agreement (Szklo M, Nieto FJ. Epidemiology Beyond the Basics. Gaithersburg, MD: Aspen Publications; 2000: 371-75).

3.3.3. Unknown Values

The proportion of tumors with unknown values for various data items **SHOULD** be an indicator of data quality. Unknown values can result from problems with: (1) data collection system or access to necessary source documents, (2) equivocal definitions of data items and/or code values, and (3) 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 tumor characteristics, disease stage, class of case, and type of reporting source.

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 the analysis will be affected. Will incidence and survival rates be affected? 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	$\geq 95\%$	-1.0	$\geq 90\%$	-1.0
2. Passing Edits	100%	0	$\geq 97\%$	0
3. DCOs	$\leq 3\%$	0.4	$\leq 5\%$	0.4
4. Timeliness	Within 23 months		Within 23 months	
5. Duplicate Records	$\leq 1/1,000$	0.4	$\leq 2/1,000$	0.4
6. Missing Data Fields – Sex, Age, State/Province and County	$\leq 2\%$	0.4	$\leq 3\%$	0.4
7. Race	$\leq 3\%$	0.4	$\leq 5\%$	0.4

Data submission specifications required for publication in national cancer statistics and public data use files have been established by the CDC-NPCR program and are summarized in Table 2.

Table 2. NPCR-CSS 2008 Data Submission Specifications.

Criteria	NPCR 12-Month Standard	NPCR 24-Month Standard	USCS[*] Publication Criteria	U.S. County Public-Use File Criteria[**]	Measurement Error
Percentage Completeness of Case Ascertainment[†]	>=90%	>=95%	>=90%	>=90%	-1.0%
Percentage Missing or Unknown					
Age	N/A	<=2%	<=3%	<=3%	-0.4%
Sex	N/A	<=2%	<=3%	<=3%	-0.4%
Race	N/A	<=3%	<=5%	<=5%	-0.4%
County	N/A	<=2%	N/A	<=3%	-0.4%
Percentage Death Certificate Only (DCO) [‡]	N/A	<=3%	<=5%	<=5%	-0.4%
Unresolved Duplicates (per 1,000)[§]	N/A	<=1	N/A	N/A	-0.4
Percentage Passing Core Edits[+]	>=97%	>=99%	>=97%	>=97%	N/A

[*] U.S. cancer statistics.
[**] See NPCR-CSS *Data Release Policy*, December, 2007.
[†] Case completeness estimates will be calculated using the NAACCR method and adjusted for duplicates if the duplicate rate was derived from a sample of the incidence file. Adjustment will not occur if duplicates were identified and corrected on the entire database.
[‡] The registry must perform death clearance.
[§] Based on the results of NAACCR duplicate protocol.
[+] Only "core" single-field, inter-field, and inter-record edits will be used to evaluate data.
N/A = Not applicable.

CHAPTER 4:

DATA ANALYSIS AND DISSEMINATION

4.1. STRUCTURAL REQUIREMENTS

4.1.1. 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 rather than estimates from the national census or to use officially approved race or ethnic categories, especially if the estimates and categories are used by all other government programs in that geographic area.

4.1.1.1. General Requirements

The level 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.

4.1.1.2. 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. and 4.2.2.).

State/provincial, territorial, and local governmental agencies often are good sources for additional information about the size and characteristics of a population in the central registry's area. Some agencies or jurisdictions employ demographers who can serve as 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.

4.1.1.3. 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. Another result of this effort was the NAACCR Asian/Pacific Islander Identification Algorithm (NAPIIA), to enhance the identification of specific Asian sub-populations for patients coded as Asian, NOS. For ease of use, NAACCR combined the two algorithms into the NAACCR Hispanic and Asian/Pacific Islander Identification Algorithm (NHAPIIA). Employing standardized approaches 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.

Another example of race information enhancement 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 most often is interested in the population that resides within a defined geographic area. SEER and NPCR registries also now submit files to the Indian Health Service for linkage to enhance coding of American Natives and decrease inaccuracy of miscoding for this racial group.

SEER has developed new guidelines to reduce the lack of consistency in interpolating races from 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.

4.1.1.4. 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. 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.

4.1.2. 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.

4.1.2.1. Standards for Number and Type of Staff

The central registry **MUST** have access to the appropriate 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 that 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.

4.1.2.2. 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.

4.1.2.2.1. 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.

4.1.2.2.2. 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 forums so that staff have rapid access to the most current information.

4.1.2.2.3. 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. ANALYSIS OF CANCER REGISTRATION DATA

4.2.1. Analysis Categories and Recoded Groups

Analysis of cancer registry data **SHOULD** include standardized data categories, analysis methods, and outcome classifications. The selection of standard categories for analysis and presentation **MUST** be compatible with NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary*. In addition, categorization of cancer data **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; 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 computer-based tool to view individual cancer records and 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) converting text data to the specific binary format required by SEER*Stat, and (2) creating 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; 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. Registries **SHOULD** use the SEER recoding scheme of cancer site categories for routine analyses.

Another important standard is the grouping used by the World Health Organization (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 could be different from adults and cell type is more important than the organ site; thus, there is a different set of cancer categories. The standard is the *International Classification of Childhood Cancer*, 3rd edition (ICCC-3), shown in Table 4. The ICCC-3 is based on ICD-O-3, 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 5 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-3 Codes.

SEER Site Recode ICD-O 3 (1/27/03)			
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

SEER Site Recode ICD-O 3 (1/27/03)			
Site Group	**ICD-O-3 Site**	**ICD-O-3 Histology (Type)**	**Recode**
Liver and Intrahepatic Bile Duct			
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

SEER Site Recode ICD-O 3 (1/27/03)			
Site Group	**ICD-O-3 Site**	**ICD-O-3 Histology (Type)**	**Recode**
Male Genital System			
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

SEER Site Recode ICD-O 3 (1/27/03)			
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 4. Site/Histology Recode Based on *International Classification of Childhood Cancer*, 3rd Edition (ICCC-3) Based on ICD-O-3*.

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
I Leukemias, myeloproliferative diseases, and myelodysplastic diseases			
(a) Lymphoid leukemias			
(a.1) Precursor cell leukemias	9835, 9836, 9837	C000-C809	001
(a.2) Mature B-cell leukemias	9823, 9826, 9832, 9833, 9940	C000-C809	002
(a.3) Mature T-cell and NK cell leukemias	9827, 9831, 9834, 9948	C000-C809	003
(a.4) Lymphoid leukemia, NOS	9820	C000-C809	004
(b) Acute myeloid leukemias	9840, 9861, 9866, 9867, 9870-9874, 9891, 9895-9897, 9910, 9920, 9931	C000-C809	005
(c) Chronic myeloproliferative diseases	9863, 9875, 9876, 9950, 9960-9964	C000-C809	006
(d) Myelodysplastic syndrome and other myeloproliferative diseases	9945, 9946, 9975, 9980, 9982-9987, 9989	C000-C809	007
(e) Unspecified and other specified leukemias	9800, 9801, 9805, 9860, 9930	C000-C809	008
II Lymphomas and reticuloendothelial neoplasms			
(a) Hodgkin lymphomas	9650-9655, 9659, 9661-9665, 9667	C000-C809	009
(b) Non-Hodgkin lymphomas (except Burkitt lymphoma)			
(b.1) Precursor cell lymphomas	9727-9729	C000-C809	010
(b.2) Mature B-cell lymphomas (except Burkitt lymphoma)	9670, 9671, 9673, 9675, 9678-9680, 9684, 9689-9691, 9695, 9698, 9699, 9731-9734, 9761, 9762, 9764-9766, 9769, 9970	C000-C809	011
(b.3) Mature T-cell and NK-cell lymphomas	9700-9702, 9705, 9708, 9709, 9714, 9716-9719, 9767, 9768	C000-C809	012
(b.4) Non-Hodgkin lymphomas, NOS	9591, 9760	C000-C809	013
(c) Burkitt lymphoma	9687	C000-C809	014
(d) Miscellaneous lymphoreticular neoplasms	9740-9742, 9750, 9754-9758	C000-C809	015
(e) Unspecified lymphomas	9590, 9596	C000-C809	016
III CNS and miscellaneous intracranial and intraspinal neoplasms			
(a) Ependymomas and choroid plexus tumor			
(a.1) Ependymomas	9383, 9391-9394	C000-C809	017
(a.2) Choroid plexus tumor	9390	C000-C809	018

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(b) Astrocytomas	9380	C723	019
	9384, 9400-9411, 9420, 9421-9424, 9440-9442	C000-C809	019
(c) Intracranial and intraspinal embryonal tumors			
(c.1) Medulloblastomas	9470-9472, 9474, 9480	C000-C809	020
(c.2) PNET	9473	C000-C809	021
(c.3) Medulloepithelioma	9501-9504	C700-C729	022
(c.4) Atypical teratoid/rhabdoid tumor	9508	C000-C809	023
(d) Other gliomas			
(d.1) Oligodendrogliomas	9450, 9451, 9460	C000-C809	024
(d.2) Mixed and unspecified gliomas	9380	C700-C722, C724-C729, C751, C753	025
	9382	C000-C809	025
(d.3) Neuroepithelial glial tumors of uncertain origin	9381, 9430, 9444	C000-C809	026
(e) Other specified intracranial and intraspinal neoplasms			
(e.1) Pituitary adenomas and carcinomas	8270-8281, 8300	C000-C809	027
(e.2) Tumors of the sellar region (craniopharyngiomas)	9350-9352, 9582	C000-C809	028
(e.3) Pineal parenchymal tumors	9360-9362	C000-C809	029
(e.4) Neuronal and mixed neuronal-glial tumors	9412, 9413, 9492, 9493, 9505-9507	C000-C809	030
(e.5) Meningiomas	9530-9539	C000-C809	031
(f) Unspecified intracranial and intraspinal neoplasms	8000-8005	C700-C729, C751-C753	032
IV Neuroblastoma and other peripheral nervous cell tumors			
(a) Neuroblastoma and ganglioneuroblastoma	9490, 9500	C000-C809	033
(b) Other peripheral nervous cell tumors	8680-8683, 8690-8693, 8700, 9520-9523	C000-C809	034
	9501-9504	C000-C699, C739-C768, C809	034
V Retinoblastoma	9510-9514	C000-C809	035
VI Renal tumors			
(a) Nephroblastoma and other nonepithelial renal tumors			
(a.1) Nephroblastoma	8959, 8960	C000-C809	036
(a.2) Rhabdoid renal tumor	8963	C649	037
(a.3) Kidney sarcomas	8964-8967	C000-C809	038
(a.4) pPNET of kidney	9364	C649	039

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(b) Renal carcinomas	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, 8560-8576	C649	040
	8311, 8312, 8316-8319, 8361	C000-C809	040
(c) Unspecified malignant renal tumors	8000-8005	C649	041
VII Hepatic tumors			
(a) Hepatoblastoma	8970	C000-C809	042
(b) Hepatic carcinomas	8010-8041, 8050-8075, 8082, 8120-8122, 8140, 8141, 8143, 8155, 8190-8201, 8210, 8211, 8230, 8231, 8240, 8241, 8244-8246, 8260-8264, 8310, 8320, 8323, 8401, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8576	C220, C221	043
	8160-8180	C000-C809	043
(c) Unspecified malignant hepatic tumors	8000-8005	C220-C221	044
VIII Malignant bone tumors			
(a) Osteosarcomas	9180-9187, 9191-9195, 9200	C400-C419, C760-C768, C809	045
(b) Chondrosarcomas	9210, 9220, 9240	C400-C419, C760-C768, C809	046
	9221, 9230, 9241-9243	C000-C809	046
(c) Ewing tumor and related sarcomas of bone			
(c.1) Ewing tumor and Askin tumor of bone	9260	C400-C419, C760-C768, C809	047
	9365	C400-C419	047
(c.2) pPNET of bone	9363, 9364	C400-C419	048
(d) Other specified malignant bone tumors			
(d.1) Malignant fibrous neoplasms of bone	8810, 8811, 8823, 8830	C400-C419	049
	8812, 9262	C000-C809	049
(d.2) Malignant chordomas	9370-9372	C000-C809	050
(d.3) Odontogenic malignant tumors	9270-9275, 9280-9282, 9290, 9300-9302, 9310-9312, 9320-9322, 9330, 9340-9342	C000-C809	051
(d.4) Miscellaneous malignant bone tumors	9250, 9261	C000-C809	052
(e) Unspecified malignant bone tumors	8000-8005, 8800, 8801, 8803-8805	C400-C419	053

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
IX Soft tissue and other extraosseous sarcomas			
(a) Rhabdomyosarcomas	8900-8905, 8910, 8912, 8920, 8991	C000-C809	054
(b) Fibrosarcomas, peripheral nerve sheath tumors, and other fibrous neoplasms			
(b.1) Fibroblastic and myofibroblastic tumors	8810, 8811, 8813-8815, 8821, 8823, 8834-8835	C000-C399, C440-C768, C809	055
	8820, 8822, 8824-8827, 9150, 9160	C000-C809	055
(b.2) Nerve sheath tumors	9540-9571	C000-C809	056
(b.3) Other fibromatous neoplasms	9491, 9580	C000-C809	057
(c) Kaposi sarcoma	9140	C000-C809	058
(d) Other specified soft tissue sarcomas			
(d.1) Ewing tumor and Askin tumor of soft tissue	9260	C000-C399, C470-C759	059
	9365	C000-C399, C470-C639, C659-C768, C809	059
(d.2) pPNET of soft tissue	9364	C000-C399, C470-C639, C659-C699, C739-C768, C809	060
(d.3) Extrarenal rhabdoid tumor	8963	C000-C639, C659-C699, C739-C768, C809	061
(d.4) Liposarcomas	8850-8858, 8860-8862, 8870, 8880, 8881	C000-C809	062
(d.5) Fibrohistiocytic tumors	8830	C000-C399, C440-C768, C809	063
	8831-8833, 8836, 9251, 9252	C000-C809	063
(d.6) Leiomyosarcomas	8890-8898	C000-C809	064
(d.7) Synovial sarcomas	9040-9044	C000-C809	065
(d.8) Blood vessel tumors	9120-9125, 9130-9133, 9135, 9136, 9141, 9142, 9161, 9170-9175	C000-C809	066
(d.9) Osseous and chondromatous neoplasms of soft tissue	9180, 9210, 9220, 9240	C490-C499	067
	9231	C000-C809	067
(d.10) Alveolar soft parts sarcoma	9581	C000-C809	068
(d.11) Miscellaneous soft tissue sarcomas	8587, 8710-8713, 8806, 8840-8842, 8921, 8982, 8990, 9373	C000-C809	069
(e) Unspecified soft tissue sarcomas	8800-8805	C000-C399, C440-C768, C809	070
X Germ cell tumors, trophoblastic tumors, and neoplasms of gonads			
(a) Intracranial and intraspinal germ cell tumors			

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(a.1) Intracranial and intraspinal germinomas	9060-9065	C700-C729, C751-C753	071
(a.2) Intracranial and intraspinal teratomas	9080-9084	C700-C729, C751-C753	072
(a.3) Intracranial and intraspinal embryonal carcinomas	9070, 9072	C700-C729, C751-C753	073
(a.4) Intracranial and intraspinal yolk sac tumor	9071	C700-C729, C751-C753	074
(a.5) Intracranial and intraspinal choriocarcinoma	9100	C700-C729, C751-C753	075
(a.6) Intracranial and intraspinal tumors of mixed forms	9085, 9101	C700-C729, C751-C753	076
(b) Malignant extracranial and extragonadal germ cell tumors			
(b.1) Malignant germinomas of extracranial and extragonadal sites	9060-9065	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	077
(b.2) Malignant teratomas of extracranial and extragonadal sites	9080-9084	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	078
(b.3) Embryonal carcinomas of extracranial and extragonadal sites	9070, 9072	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	079
(b.4) Yolk sac tumor of extracranial and extragonadal sites	9071	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	080
(b.5) Choriocarcinomas of extracranial and extragonadal sites	9100, 9103, 9104	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	081
(b.6) Other and unspecified malignant mixed germ cell tumors of extracranial and extragonadal sites	9085, 9101, 9102, 9105	C000-C559, C570-C619, C630-C699, C739-C750, C754-C768, C809	082
(c) Malignant gonadal germ cell tumors			
(c.1) Malignant gonadal germinomas	9060-9065	C569, C620-C629	083
(c.2) Malignant gonadal teratomas	9080-9084, 9090, 9091	C569, C620-C629	084
(c.3) Gonadal embryonal carcinomas	9070, 9072	C569, C620-C629	085
(c.4) Gonadal yolk sac tumor	9071	C569, C620-C629	086
(c.5) Gonadal choriocarcinoma	9100	C569, C620-C629	087
(c.6) Malignant gonadal tumors of mixed forms	9085, 9101	C569, C620-C629	088
(c.7) Malignant gonadal gonadoblastoma	9073	C569, C620-C629	089

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(d) Gonadal carcinomas	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8143, 8190-8201, 8210, 8211, 8221-8241, 8244-8246, 8260-8263, 8290, 8310, 8313, 8320, 8323, 8380-8384, 8430, 8440, 8480-8490, 8504, 8510, 8550, 8560-8573, 9000, 9014, 9015	C569, C620-C629	090
	8441-8444, 8450, 8451, 8460-8473	C000-C809	090
(e) Other and unspecified malignant gonadal tumors	8590-8671	C000-C809	091
	8000-8005	C569, C620-C629	091
XI Other malignant epithelial neoplasms and malignant melanomas			
(a) Adrenocortical carcinomas	8370-8375	C000-C809	092
(b) Thyroid carcinomas	8010-8041, 8050-8075, 8082, 8120-8122, 8130-8141, 8190, 8200, 8201, 8211, 8230, 8231, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8430, 8440, 8480, 8481, 8510, 8560-8573	C739	093
	8330-8337, 8340-8347, 8350	C000-C809	093
(c) Nasopharyngeal carcinomas	8010-8041, 8050-8075, 8082, 8083, 8120-8122, 8130-8141, 8190, 8200, 8201, 8211, 8230, 8231, 8244-8246, 8260-8263, 8290, 8310, 8320, 8323, 8430, 8440, 8480, 8481, 8500-8576	C110-C119	094
(d) Malignant melanomas	8720-8780, 8790	C000-C809	095
(e) Skin carcinomas	8010-8041, 8050-8075, 8078, 8082, 8090-8110, 8140, 8143, 8147, 8190, 8200, 8240, 8246, 8247, 8260, 8310, 8320, 8323, 8390-8420, 8430, 8480, 8542, 8560, 8570-8573, 8940, 8941	C440-C449	096
(f) Other and unspecified carcinomas			
(f.1) Carcinomas of salivary glands	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C079-C089	097
(f.2) Carcinomas of colon and rectum	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C180, C182-C189, C199, C209, C210-C218	098

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(f.3) Carcinomas of appendix	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C181	099
(f.4) Carcinomas of lung	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C340-C349	100
(f.5) Carcinomas of thymus	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C379	101
(f.6) Carcinomas of breast	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C500-C509	102
(f.7) Carcinomas of cervix uteri	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C530-C539	103
(f.8) Carcinomas of bladder	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C670-C679	104
(f.9) Carcinomas of eye	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C690-C699	105

Site Group	ICD-O-3 Histology (Type)	ICD-O-2/3 Site	Recode for Extended Classification
(f.10) Carcinomas of other specified sites	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C000-069, C090-C109, C129-C179, C239-C339, C380-C399, C480-C488, C510-C529, C540-C549, C559, C570-C619, C630-C639, C659-C669, C680-C689, C700-C729, C750-C759	106
(f.11) Carcinomas of unspecified site	8010-8084, 8120-8157, 8190-8264, 8290, 8310, 8313-8315, 8320-8325, 8360, 8380-8384, 8430-8440, 8452-8454, 8480-8586, 8588-8589, 8940, 8941, 8983, 9000, 9010-9016, 9020, 9030	C760-C768, C809	107
XII Other and unspecified malignant neoplasms			
(a) Other specified malignant tumors			
(a.1) Gastrointestinal stromal tumor	8936	C000-C809	108
(a.2) Pancreatoblastoma	8971	C000-C809	109
(a.3) Pulmonary blastoma and pleuropulmonary blastoma	8972, 8973	C000-C809	110
(a.4) Other complex mixed and stromal neoplasms	8930-8935, 8950, 8951, 8974-8981	C000-C809	111
(a.5) Mesothelioma	9050-9055	C000-C809	112
(a.6) Other specified malignant tumors	9110	C000-C809	113
	9363	C000-C399, C470-C759	113
(b) Other unspecified malignant tumors	8000-8005	C000-C218, C239-C399, C420-C559, C570-C619, C630-C639, C659-C699, C739-C750, C754-809	114

* Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. *International Classification of Childhood Cancer*, Third Edition. *Cancer* 2005;103:1457-67.

Table 5. 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
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

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9**	**ICD-10**
	(1979-1998)*	**(1999+)***
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
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

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9**	**ICD-10**
	(1979-1998)*	**(1999+)***
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
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

Cancer Causes of Death		
Cancer Causes of Death	**ICD-9**	**ICD-10**
	(1979-1998)*	**(1999+)***
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

Table 5. Standard Site Analysis Categories for Mortality Data (ICD-9 and ICD-10) SEER Cause of Death Recode 1969+ (3/25/2004). (Continued)

Non-Cancer Causes of Death			
Non-Cancer Causes of Death	**ICD-8**	**ICD-9**	**ICD-10**
	(1968-1978) *	**(1979-1998) ***	**(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 Conditions	490-493, 519.3	490-496	J40-J47

Non-Cancer Causes of Death			
Non-Cancer Causes of Death	**ICD-8** (1968-1978) *	**ICD-9** (1979-1998) *	**ICD-10** (1999+)*
Stomach and Duodenal Ulcers	531-533	531-533	K25-K28
Chronic Liver Disease and Cirrhosis	571	571	K70, K73-K74
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 Cause of Death (COD). Those deemed valid but not meeting the definition of any above grouping are classified as Other COD.

** 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.

† 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.

‡ External causes of injury and poisoning.

Source: SEER 2003.

4.2.1.2. Standards for Age Categories

The age distribution of cancer patients most often is 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. Malignancies occurring between the age of 15 and 19 years usually are referred to as adolescent cancers. Most often childhood cancers include pediatric and adolescent cancers, and cover age categories 0, 1-4, 5-9, 10-14, and 15-19. For some pediatric cancers, single-year age groups are desired when incidence rates change dramatically within the 5-year interval. Most registries use 85+ 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 and Trend Analysis

Time periods used to present cancer statistics should overlap periods used by governmental agencies to estimate population counts. In North America, federal agencies provide population estimates per calendar year; and, most commonly, surveillance periods cover 1, 3, or 5 calendar years. Consequently, cancer statistics should be presented by calendar year or multiple consecutive calendar years. The choice of time period(s) should ensure the stability and comparability of cancer statistics over time. Stability and comparability depend on both the size of the population covered and the length of time the registry has been in existence. Central cancer registries that have covered large populations may have sufficient data to evaluate cancer statistics on a year-by-year basis. In contrast, registries with a small population base may have insufficient data to present stable reliable statistics in such detail. Three- or 5-year averages may be used to reduce random variation in statistics created from small numbers. It is not recommended to calculate temporal trends for short time periods; most often, 10 years is an acceptable period for trend analysis.

Cancer registries data may be used to calculate trends over time based on cancer rates, although the trend may reflect changes in cancer proportions as well. Central cancer registries may present the percent change of a rate and/or the annual percent change (APC) of a rate. The percent change should be calculated by taking the difference between the average rate of the first n years and the average rate of the last n years. The difference then is divided by the average rate of first n year and the result is converted to a percentage. The APC should be calculated by fitting a regression model to the natural logarithm of the cancer rate, using the calendar year as an independent variable. Percent change and annual percent change may be calculated using software provided by the NCI (www.seer.cancer.gov/seerstat). The software allows for calculation of confidence intervals around APC and statistical testing using APC. More sophisticated trend analysis includes jointpoint models (www.srab.cancer.gov/joinpoint) and spatio-temporal trends (www.srab.cancer.gov/satscan).

Registry staff **SHOULD** consult with staff experienced in cancer epidemiology 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 code includes the province/territory (2 digits), census division (2 digits), and census subdivision (3 digits). Census divisions are a level of geographic classification between the province or territory and the municipality. Municipalities are roughly comparable to census subdivisions. 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 6.

Table 6. 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 ACoS supports efforts to standardize the collection of first course of treatment information such as surgery, radiation therapy, systemic therapy, palliative care, and other treatment procedures. The current treatment codes and coding rules were developed by the ACoS with input from medical specialty organizations, SEER, NPCR, NCRA, NAACCR, and software providers. Major national programs such as the NCDB, SEER, and NPCR use these codes and coding structures for the data items that they require. Details on the codes and rules can be found in publications from each organization such as COC's FORDS (revised for January 1, 2007). Although NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary* includes these treatment data items in the NAACCR record layout, detailed codes for the site-specific items such as surgery to the primary site can only be found in manuals such as *FORDS 2007* or *SEER Coding Manual 2007*. There is no widely accepted categorization system for analysis and dissemination of cancer treatment information. For research reports and data quality assessment activities, treatment information **SHOULD** be categorized according to procedures that constitute the standard of care for a given tumor site and cancer stage group. The standard of care **MAY** be defined per recommendation of national organizations (e.g., National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology™ or NCI's PDQ®).

4.2.1.6. Standards for Grouping by Stage of Disease

Collaborative stage was implemented on 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 in this section within the context of three key elements of epidemiologic inquiry: (1) person, (2) place, and (3) time. Analyses usually are based on tumor records (i.e., independent primary tumors). Each person may be diagnosed with more than one primary tumor. Some analyses focus on persons rather than tumors.

❖ *Person:* Reports of cancer data **SHOULD** document the demographic characteristics of the case represented in the report. At a minimum, these characteristics should include sex, age, and race/ethnicity (United States). A person may be associated with 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 cases represented in the report. Typically, the area of coverage follows political boundaries such as provinces, states, counties, cities, or census entities (see Section 6.3.2.1. 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 quantitative measure used with cancer registry data is the simple enumeration of tumors. Knowledge of the number of tumors can be of great use for health planning purposes, in which it is important to measure the burden of cancer on existing health care resources and 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:
(1) percent distribution of tumors by stage of disease at diagnosis, (2) proportion of tumors with a histologically confirmed diagnosis, and (3) 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 which tumors are more common. Usually, tumors of the breast, lung, colon and rectum, and prostate together will account for more than one-half of all cancers, with each site 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 the true distribution of tumors in the general population.

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 incidence of a specific cancer in relation to all cancers between 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

Standardization is a set of techniques used to remove the effects of difference in the distribution of age or other confounding variables between 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 a 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 standardization was carried out.

❖ *Indirect Method:* This is used to compare study populations for which the specific rates either are 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 proportions. 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 have a higher risk than younger persons) and is useful for comparing 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 in calculation of malignant neoplasm incidence rates. As an exception, count *in situ* urinary bladder tumors toward the urinary bladder incidence rate.

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 4.1.1. 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 7.

❖ *Canadian Standard:* Canada's 1991 and 1996 populations are used to standardize rates for routine comparisons within Canada. The 1991 and 1996 populations are shown in Table 7. 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 Publication No. 15, 1976) is the world standard used by the IARC, also shown in Table 7. This is useful for international comparisons. There also is a WHO 2000-2005 standard that is not used for cancer registration data.

Table 7. 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. 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 a small number 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 vital statistics bureaus based on information reported through death registration. However, because of their expertise and focus on cancer and the need for confidentiality associated 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; 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 of care following a cancer diagnosis. When preparing survival rates: (1) select cases based on the purpose of the study, (2) all inclusions and exclusions **MUST** be accounted for, and (3) 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 represent 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 in the paragraphs that follow.

❖ *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. This method measures the excess mortality that a cancer-patient 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. This method works well when there are no major differences between the cancer cohort and general populations with regard to other risk factors, except for the cancer itself. This method does not work well in lung cancer, because patients have a higher risk of death from heart disease compared to the general population due to past smoking behavior.

❖ *Kaplan-Meier:* The Kaplan-Meier 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 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 (e.g., 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 prognostic factors, including the patient's age, stage of disease at diagnosis, histologic type of cancer, treatment, and comorbidities. 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 who has 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).

4.3.   DISSEMINATION OF CANCER REGISTRATION DATA

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 **MAY** obtain copies of reports and newsletters from other registries to use as models when developing their own publications. Most cancer registries **MAY** 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.; for a discussion of use and release of confidential data, see Section 6.3..

4.3.1. Standards for Type and Frequency of Reports

4.3.1.1. Summary of Central Registry Data

Central cancer registries **SHOULD** assemble a comprehensive summary of the cancer burden (i.e., incidence of primary tumors, cancer deaths) 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, mortality, 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.

4.3.1.2. 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, 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, and primary site. It is helpful to provide patient lists sorted alphabetically, by the facility's accession number, and by cancer type. In addition to reports as described above, the registry **MAY** consider providing patient follow-up information to hospitals, such as the results of death clearance and other follow-up activities. This can be of great value to hospital cancer registries in reducing follow-up workload. 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.

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

4.3.1.3. Reports to Physicians

Reports to individual physicians **MAY** include descriptive statistics for their specialty (i.e., melanoma for dermatologists). Physicians may make special data requests or request follow-up information; these reports **SHOULD** be generated upon request (see Section 4.3.1.6.).

4.3.1.4. 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.

4.3.1.5. 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.

4.3.1.6. 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.

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.

4.3.1.7. Occasional or Special Topic Reports

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

4.3.2.   Standards for Narrative Text

An important component of any report is the narrative text that accompanies the presentation of the data. As outlined in Section 4.3.2.2., the narrative guides the reader by documenting methods used to produce the report, highlighting important findings, and interpreting the results.

4.3.2.1. 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.

4.3.2.2. 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.

4.3.2.3. 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: (1) completeness of case ascertainment; (2) proportion of error-free records based on

standardized edits; (3) proportion of death certificate only cases; (4) timeliness of data; (5) rate of duplicate case records; and (6) proportion of cases with unknown or missing race, sex, county or age information.

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.

4.3.3.   Standards for Displaying the Data

4.3.3.1. 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.

4.3.3.2. Graphs and Charts

The graphical presentation of data often is more intuitively appealing than a table full of numbers (i.e., 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 may 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.

4.3.3.3. 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 requires 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*).

4.3.3.4. 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.

4.3.4. 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 (see Section 6.3.2.4.). In addition, participating facilities/organizations **SHOULD** be provided with a courtesy review of the publication prior to release.

All questions regarding the quality of the data **MUST** be brought to the attention of the quality control staff and **SHOULD** be resolved before any 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, participating facilities, and 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.

4.3.5. Electronic Publication and Distribution of Registry Data

In addition to publishing summary data, registries **MAY** disseminate data in an electronic form that allows users flexibility in querying the data. Epidemiologists, biostatisticians, public health officers, and students could benefit from the ability to formulate and run their own queries of cancer registry data. The registry may have the additional burden of training the users of their data products to prevent incorrect analyses and inappropriate conclusions.

4.3.5.1. Types of Electronic Publication

Two types of electronic publication frequently are used to disseminate cancer registry data: (1) de-identified data files, and (2) query systems.

4.3.5.1.1. De-Identified Data Files

A de-identified data file is an electronic file with individual record-level data concerning cancer patients for which direct and in direct identifiers have been removed. De-identified data files may be restricted access or public use files. Section 6.3.2.3.3., Standards for Protecting Confidentiality in De-Identified Data Files, provides information on confidentiality protection for de-identified files.

A restricted access file is a de-identified file for which the user **MUST** apply for access through a formal mechanism and attest to abide by the confidentiality provisions in a data use agreement. As examples, Appendix J presents the Data Confidentiality Agreement for NAACCR Researchers and the SEER Public Use File Agreement. Restricted access files **MAY** be provided to universities, medical schools, health departments, physicians, epidemiologists, voluntary cancer societies, and science journalists. SEER and NPCR have restricted access files available for researcher use.

A public use file is a de-identified file for which there is not a specific application process and that is available to all users who request it. The level of disclosure risk is higher with public use files compared to restricted access files, and consequently, a higher degree of confidentiality protection may be needed. For example, county of residence at diagnosis may be provided on a restricted access file but a registry may chose to redact county of residence from a public use file.

4.3.5.1.2. Query Systems

Another approach for electronic dissemination of data is through a query system or software. This can be a web-based query system such as CINA+ Online and CDC Wonder, or PC-based system such as SEER*Stat. Some features that make query systems useful are built-in recodes to appropriate analysis groups, suppression of statistically insignificant or meaningless results, and suppression of cells with small numbers. The potential user base is broadened when the analysis software is provided and is user friendly.

Query systems may be based on record-level data or on pre-tabulated data. SEER*Stat is an example of a query system based on record-level data. If the underlying data are accessible to the user, such a query system would need to include similar confidentiality protection as de-identified data files. Query systems based on pre-tabulated data, such as CINA+ Online, need to include similar confidentiality protection as summary statistics (see Section 6.3.2.3.2., Standards for Protecting Confidentiality in Summary Statistics).

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.

4.3.5.2. Distribution Methods

Public use data should be available on CDs/DVDs or via an Internet-accessible 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. The SEER Program and some state central registries (e.g., New York State Cancer Registry) also provide public use data files.

4.3.5.3. Standards

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 J 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, advances in computer software and hardware have increased the efficiency of data collection and improved data quality, standardization, and accessibility. These advances also have facilitated the collaborative pooling of data. Computers have enhanced the ability to more fully use the rich resource of cancer registry data.

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. However, 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.

This chapter 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 chapter 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; these differences are not addressed.

This chapter 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 here.

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. Non-standard systems are more costly to maintain, often include hidden costs, and sometimes contribute to problems associated with 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 described 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 5.1.4.2.; and EDITS in Section 5.9.).

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, CD/DVD, or other electronic media. The central registry **SHOULD** encourage the use of NAACCR's data exchange standard for such transmissions. (See NAACCR's *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary* and *NAACCR Standards for Cancer Registries Volume V: Pathology Laboratory Electronic Reporting).*

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 5.10. and 5.11.).

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 tools. 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 5.12.).

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 2.2.8.). 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 for 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 Sections 5.1.5.2.3. and 6.4.).

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 NCDB. The registry **MUST** use NAACCR's data exchange standard whenever possible (see Appendix J for a sample data sharing agreement).

5.1.3.17. Communications

The system **SHOULD** provide telecommunications capabilities for importing and exporting 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.). The standard for data transmission from pathology laboratories to the central registry can be found in NAACCR's *Standards for Cancer Registries Volume V: Pathology Laboratory Electronic Reporting.*

5.1.4.2. Standard Edits

The central registry **SHOULD** use standard data edits (see Section 5.9.).

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: (1) use of standard analysis categories, (2) application of standard statistical methods, (3) provision for the use of multiple population standards, and (4) 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, case reporting, and the required reportable data items are predictable. It is possible to anticipate the amount of disk space that 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 increases in size 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 Section 6.3. for detailed discussions of handling confidential data).

The *Annotated Bibliography on Confidentiality Protection in Data Release* developed by the NAACCR Confidentiality Subcommittee of the Data Use and Research Committee may be used to develop and maintain policies and procedures concerning data confidentiality. This document may be found at www.naaccr.org/confidentiality/index.asp.

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 **SHOULD** be conducted.

❖ 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.

For additional information on security-related issues, see Section 6.2..

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: (1) maintenance of equipment; (2) upgrades of equipment and software for improved performance; (3) implementation of new standards as they become available; and (4) implementation of new technologies and software that will enhance the efficiency, effectiveness, and security of the registry and its data.

5.2. HARDWARE REQUIREMENTS

Although the task of making specific hardware recommendations is not part of NAACCR's mission, and any recommendations in this area would be outdated as soon as they were made, it is helpful to present guidelines and considerations to assist with hardware purchases. Prior to considering hardware options, the central registry's planners **SHOULD** consider all of the points listed in Sections 5.1.1. and 6.4.13., as well as the following questions and suggestions:

❖ 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?

❖ Should separate hardware platforms be used based on the class of software installed (i.e., registry database software versus statistical packages versus office automation applications)?

❖ 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, or may allow for attractive pricing due to volume buying power.

❖ Some hardware systems require annual maintenance support contracts to insure reliable uptime performance. Is the registry funded to allow for this recurring cost? Care should be taken to obtain a commitment from the vendor for projected annual fees for these maintenance support agreements over the life of the hardware.

Infrastructure security is a critical consideration when specifying hardware solutions. The guidelines for registry security procedures detailed in Chapter 6 **MUST** be observed.

5.2.1. Computer Systems

Lower cost, commodity hardware platforms built using standards-based components and offered from multiple vendors **SHOULD** be considered. The lower cost affords a shorter and more realistic lifecycle for the hardware, allowing for replacement sooner and reducing higher-priced maintenance costs that often are incurred on aging hardware systems. The registry **SHOULD** build in hardware replacement costs every 3 to 5 years. A policy **MUST** be in place to ensure that no sensitive data remain on an obsolete computer system's data storage prior to discarding or reallocating the system. Various mechanisms often are employed to clean the data off of devices—a more practical and cost effective method may be to use certified data destruction services to physically shred the media.

Computer systems should be researched and built to the full projected requirements during the life cycle of the devices. Purchasing a system at a lower cost, with the vendor-suggested option of a later upgrade of hardware components at additional costs, **SHOULD** be avoided.

The selection and purchase of computer hardware **SHOULD** be one of the last decisions made when building a central registry data system. The selection of operating systems, database management systems, and other commercial software products can dictate some of the hardware options required. If hardware is selected first, the central registry may find limitations in software selections available for that hardware.

5.2.2. Servers

The current trend of using rack-mounted, small footprint servers can conserve space and centralize data servers in climate- and access-controlled areas of the registry. Low power consumption chip sets offer adequate system performance while reducing power and cooling requirements.

5.2.3. Workstations and Laptops

Whether a registry uses an in-house resource to procure and configure user workstations and laptops or if it outsources these tasks, standardization **SHOULD** be considered. Standardizing on one hardware platform reduces the different potential repair and maintenance issues that a registry is likely to encounter during the life of the hardware.

To maximize potential cost savings by securing bids on a larger group of systems, and to reduce employee "computer envy" when only some are provided with a new system, a registry **SHOULD** consider batch purchases of systems. With the large number of workstations and laptops typically used in a central registry, standardizing on fewer models of hardware (and associated operating systems) will help with IT support issues.

Registries **SHOULD** consider installing firmware-based, "whole disk" types of encryption tools that offer the highest level of data protection for information stored on hard disk drives. Any computer can use this type of encryption, and laptop devices are prime candidates for added assurance that stolen laptops cannot have data retrieved, even when the hard drive is removed and installed in another computer.

5.2.4. Peripheral Hardware

Besides the computer itself, careful attention **SHOULD** be given when considering the various hardware components that are attached or networked to a computer.

5.2.5. 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. When projecting costs for high-quality color printers, registries should include the total cost for operation expenses, including the toner cartridge prices, print fuser costs and the lifespan for these expendable printer components.

High-end, large-format printers that often are used to make posters and large displays also are available. Often purchased with year end, excess funds, it is cautioned that the difficulty of operation, expensive media and ink costs, office space required to house them, and infrequent need to produce the posters all should be considered before committing to buying these types of printers.

Service providers can provide high-quality printing results for infrequent presentation needs and often are a cost-effective alternative to using registry staff time to produce such materials. A good example is the many online printers specializing in full-color, tri-fold brochure printing. Competition between the providers usually will afford a much lower cost of production than can be obtained by printing them in-house.

5.2.6. Communications and Data Exchange

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 the implementation of internal e-mail, and may be able to communicate with a registry's website, online forums, or external e-mail with other organizations around the globe. Basic, low-end data transfer can be accomplished by sending a diskette, CD, DVD, or other electronic media 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.

All data exchange activities **MUST** involve using industry-recognized standards for data encryption, such that interception of the data transfer by unintended recipients will not jeopardize confidentiality. This concept applies to all modes of data exchange

5.2.7. Data Backup

In a central registry operation, the quantity of information that requires backup is substantially greater than that in a hospital registry, pathology laboratory, or a 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/DVD burner, or communications hardware to allow transfer of data to another machine for backup.

Whether achieved by hardware or software solutions, backup media containing registry data files **MUST** be encrypted, such that unauthorized retrieval from the backup media is impossible. Central registries **MUST** have procedures in place to recover information from a backup.

The central registry also **MUST** carefully evaluate physical storage needs to provide a secure backup solution. A fireproof safe for storing backup files and offsite storage **MUST** be required. If the backup solution incorporates a high-speed connection to a secure, remote data center, this would help with the offsite storage requirement necessary to provide the foundation for disaster recovery (see Section 6.5.).

5.2.8.   Basic Hardware Requirements

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

Careful computer software selection is important for cancer registries and can require significant resources of both staff time and financial investment. Site licenses, educational discounting, and consideration of open-source software packages can be taken into account when researching software purchases. Although standards may not exist in the cancer registry field, seeking information from and the experiences of other cancer registries using the software prior to purchasing can be useful.

Infrastructure security is a critical consideration when specifying software solutions. Guidelines for registry security procedures detailed in Sections 5.1.3.2.3. and 6.2. **MUST** be observed.

5.3.1.   Database Management Software

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. When designed properly, a DBMS eliminates or drastically reduces data duplication. Elimination of duplication saves storage space and frequently can reduce processing requirements. It also helps to avoid data update anomalies and improve data integrity. Since all computer processes communicate and act upon the registry data as controlled by the DBMS, a consistent set of rules and uniform definition of data items helps to ensure high-quality data.

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. The number of concurrent users, the number of independent activities that occur concurrently, and the complexity of the database design also factor into this overhead. Today, standards-based hardware that will provide for the performance needs of central cancer registries is relatively inexpensive and readily available today.

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: (1) define and store specific information about the database structure, (2) provide a wide variety of methods for accessing data, (3) provide security features to protect access to the data, (4) enable control over concurrent operations, and (5) 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.

5.3.1.1. Relational Database Management Systems

The vast majority of current DBMSs employ the relational model, which is based on a mathematically derived view of structuring data, and follow principles of set theory. It was first described by Edgar Codd in 1969, but was not implemented in software systems until much later. Database software applications that deliver this model came to be called Relational Database Software Management Systems (RDBMSs). Most (if not all) large scale, multi-user cancer registry database systems today are built using an RDBMS.

As described in this section, the basis for RDBMSs involves breaking down data into smaller, yet interrelated units. This is fundamentally different from the older-style systems often used in legacy computing environments. Referred to sometimes as "flat file" or "hierarchical–type" systems, information was stored in what could be thought of as one very wide record. Limitations of this type of system can easily be seen using registry data as an example. Using a flat file approach, one needs to define the maximum number of different tumors a patient can have during the initial design work of the system. After predicting this maximum number, designers then define redundant groups of variables (i.e., tum1, tum2, tum3, tum1grade, tum2grade, tum3grade) in the wide record to accommodate the maximum number of tumors. So, for the majority of patients who only have one tumor, most of these redundant groups of variables were left empty in this flat file, wide record example. What would happen if the maximum number of different tumors turns out to be wrong and one patient develops one more tumor than was allotted for? The file structure would need to be modified in this case to allow for the new maximum.

The relational database model views the database as a set of 2-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. Relations between tables are accomplished by placing indexed values in both tables, therefore ensuring a way to always associate the proper data from one table to the other. In RDBMS terminology, these are the keys that are stored in each table. Consider a simplified example, using three tables often used in cancer registries. 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 developed properly, clearly shows all of the entities, their attributes, and the relation to all of the other entities. By testing this self-documenting model with sample data, it is possible to determine whether 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—in instances for which they can be determined—future central registry needs.

Additional features found in an RDBMS, and which **SHOULD** be considered when evaluating the software choices are:

❖ Some form of file, record, and/or column-locking system to insure data integrity in a multi-user environment, whereby updates to shared data can be synchronized so that only one user can be updating the shared data at a time.

❖ Transaction processing that maintains 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.

❖ Transaction logging and recovery that provides data integrity protection in the event of a failure occurring while transaction changes are actually being written to a database.

❖ The ability to design referential integrity into the database schema that describes the relations between the tables. This ensures integrity between coupled tables, such that one cannot delete a record in one table if a critical referring link in another table would be left orphaned.

❖ Data normalization steps to ensure that no duplicate data are stored within the database, by breaking down information into a higher and higher number of tables. Complex mathematical theory dictates the many forms of data normalization. Care should be taken not to "over do" this step, as the practical performance liabilities incurred when a program re-joins many tables back together can sometimes negate the theoretical benefits offered.

❖ User-specific access limitations. Using database configuration methods called roles and views, the user is only allowed to see and act upon data that are required by their user credentials.

❖ Common interface language tools. RDBMSs all share a common access language, called Structured Query Language (SQL). American National Standards Institute standards exist to help with consistency between vendor versions. However, many of the major vendors have their own, enhanced versions of SQL, which should be investigated when selecting an RDBMS vendor.

5.3.2.   Data Exchange and Communications Software

Software purchases are a sizeable consideration for central cancer registries. Categories of commonly used software are provided in this section.

5.3.2.1. Data Extraction and Manipulation Tools

Data formatting, sorting, and data import and export are just a few of the functions served by this class of software. Many utility programs will be included with the operating system; others will be part of a good DBMS system. SQL utilities offered by third-party vendors often offer cleaner and more feature-rich manipulation options, such as interoperability between RDBMSs, and the ability to move data quickly between formats. This can be very useful when extracting data for use in graphical or statistical packages, or when sending data to outside agencies or researchers.

5.3.2.2. Datafile Transfer Tools

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/DVDs 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 as and faster than postal or courier services. File transfer protocol (FTP) transfers through the Internet can transfer larger files very quickly.

5.3.3.   Statistical Analysis 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. Statistical researchers and staff typically bring with them the needed skills for using these more specific analysis tools.

5.3.4.   Geographic Information Systems

The central registry **SHOULD** have a geographic information system (GIS) available to map, store, link, manipulate, and analyze geographic data. The majority of North American central cancer registries are using GIS for a range of different purposes. Specially trained staff will be necessary to properly use a GIS.

Two excellent reference documents are *Central Cancer Registries: Design, Management and Use,* Second edition, Chapter 17-Geographic Information System, which is available for purchase, and *Using Geographic Information Systems Technology in the Collection, Analysis and Presentation of Cancer Registry Data: A Handbook of Basic Practices* (October 2002), which is available on the NAACCR website.

5.3.5.   Office Automation Software

The central registry **SHOULD** provide users with a common suite of inter-office applications to help carry out day-to-day activities. Word processing, spreadsheet, and data presentation tools are commonly used for these tasks. Using a consistent version of the software across the enterprise will protect against version incompatibility and feature mismatch situations. The costs of office automation software often can be reduced by specifying the software as a bundled component of new computer bid costs, with vendors often reducing costs when larger numbers of systems are purchased together.

5.3.6.   Basic Software Requirements

The registry **MUST** have computer software 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 cancer reporting.

The central registry's software **MUST** adequately protect the accuracy of registry data.

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 **MUST** be involved in planning and overall system design.

5.4.1. Standards for Number and Type of Staff

The central registry **SHOULD** 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 the requisite technical knowledge.

Central registry personnel **SHOULD** 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 adversely affect 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, the NCRA, and other local, state, provincial, and national organizations can provide excellent training opportunities (see Appendix E for education and training resources 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 answer questions that arise. Current pertinent reference books, journals, and other periodicals **SHOULD** be available immediately. The central registry also **SHOULD** provide access to online services and online forums 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 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 reportable neoplasm records usually is the process of abstracting directly onto a computer. Computerized data collection combines all abstracting tasks—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 might employ a variety of data collection methods for new 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, keyed entry from data collection forms, the use of imaging software to scan abstracts, and other methods. Regardless of the method(s) 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 5.9.; 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; physicians' offices; coroners' offices; state vital statistics departments; other local, state, and federal governmental agencies; other central registries; and outside vendors. Data received by the central registry **MAY** include:

❖ 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).

❖ Reportable neoplasm 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 received by the central registry:

❖ As specified in NAACCR's *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 CDs/DVDs. 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.

❖ Reporting facility specifics.

❖ Data from confidential resources (i.e., linking with DMV, Medicare, etc.).

❖ Cancer identification.

❖ Stage.

❖ Prognostic factors.

❖ Treatment.

❖ Follow-up for each tumor.

❖ Supporting text for all coded fields and diagnostic processes.

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 data input 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 5.9.)

❖ *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 the New York State Identification and Intelligence System (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 records with existing records in the database to identify duplicate or subsequent records or previously unreported neoplasms and to assign unique accession numbers to the new 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 for lack of adequate information within the submitted record).
  – Reports indicating the status of data submissions, with items such as numbers of tumor records sent in each 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 Section 3.1.4.).

5.5.2.3. Standards for Follow-Up Data Input Files

5.5.2.3.1. Definition

If 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 by the reporting facility 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 data 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 5.11. for updating guidelines).

❖ *Error Reports:* Produce error reports for incoming follow-up records that fail edits.

❖ *Management Information:* Provide the means to identify database records with follow-up information that has been updated 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 records. These files contain the changes made to required data items after the 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 data input files:

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

❖ *Editing and Updating:* Provide the ability to either manually or automatically apply an incoming correction record to the corresponding database record after editing for intrafield and interfield consistency (see Section 5.11. for updating guidelines).

❖ *Error Reports:* Produce error reports for incoming correction records that fail edits.

❖ *Management Information:* Provide the means to identify database records with information that 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 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 data input files:

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

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

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

5.5.2.6. Standards for Death Clearance Input Files

5.5.2.6.1. Definition

Death clearance processing involves the use of data about individuals for whom death certificates were filed. The purpose is to provide new information about previously reported records (follow-up) and obtain new record information for previously unreported patients or reportable neoplasms (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 records, provide the ability to: (1) automatically apply the incoming death information to the database record, when appropriate, after editing for compatibility and consistency; and (2) update other items coded in the death record, such as race and birthplace, when the database record contains unknown or non-specific values and the death record is more specific (see the discussion of consolidation in Section 5.10.).

❖ *Error Reports:* For death certificate records that link to database records, provide error reports on records that fail edits.

❖ *Suspense:* For death certificate records that do not link to database records but are records that should have been reported, provide the ability to suspend the death records in the database for further follow-back investigation (see Sections 2.2.8. and 5.5.2.8.).

❖ *Management Information:* Provide the means to identify records for which death information has been applied to the existing 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 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, sometimes using a commercial vendor (see Section 5.3.4. and the NAACCR downloadable document *Using Geographic Information Systems Technology in the Collection, Analysis, and Presentation of Cancer Registry Data-A Handbook of Basic Practices, October 2002* and for additional information on GIS see *Central Cancer Registries: Design, Management and Use* 2nd ed, Chapter 17-Geographic Information System).

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 record.

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

❖ *Error Reports:* Produce error reports for incoming geocoded records that consistently fail edits.

❖ *Management Information:* Provide the means to identify records for which geocoded information has been applied 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 areas (such as dissemination areas, census tracts, and census subdivisions). Special procedures are used, including generation of reports, for a small percentage of records that cannot be directly converted, so that consistent and valid codes may be applied.

5.5.2.8. Standards for Limited Information Input Files

5.5.2.8.1. Definition

These files contain limited information about reportable neoplasm records. The information may not have been reported because it is not yet complete (e.g., a record identified through rapid case ascertainment); the patient record may have been ascertained from a source with limited information, necessitating follow-back to other sources (e.g., a reportable neoplasm identified through a pathology laboratory); or the patient may have been overlooked by the facility responsible for reporting it (e.g., a record 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 information input files:

❖ *Editing:* Edit the incoming data for very basic content.

❖ *Suspense:* Provide the ability to suspend the records in the database for further investigation.

❖ *Reports:* Provide reports of the suspected reportable neoplasms 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 records so that the limited information records can be deleted if records have been added to the database from another source.

❖ *Deletion:* Provide the ability to delete a limited information record if the neoplasm is found to be non-reportable.

❖ *Management Information:* Provide the means to identify disposition of limited information 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 as well as other system functions, including table variables and population denominator files (see Section 4.1.1.). 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, including: (1) management reports that allow for monitoring of the database and central registry operations; (2) standard reports to give feedback to or request information from reporting sources; and (3) 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 tables presenting the:

❖ Number of records reported for each reporting facility and for other sources of records (such as DCO cases or physician-only cases) by month and year of admission (or, for DCO cases, month and year of death.)

❖ Difference between the number of 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.

❖ Records from all reporting sources by month and year of diagnosis.

❖ Distribution of incidence by year of diagnosis by site for comparison with other registries.

❖ Number of records by process completed (e.g., number inspected or visually reviewed, number in suspense, etc.), by date received in the central registry to monitor workflow.

❖ Interval between diagnosis date and date abstracted, and between diagnosis date and the date the record was entered in the central registry system, by facility, to show timeliness of abstracting.

❖ 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 records to the registry. These reports can be transmitted to the facilities electronically or in hardcopy form (see Section 4.3.1. 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 acknowledgment of the central registry's receipt of a record submission (e.g., date and number of records received) so that the facility can verify that its records were received and were readable.

❖ A table presenting the number of 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 (see Section 4.3.1.) 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 patients identified through the screening of pathology laboratory reports of patients who 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** provide listings of records in the database that meet specified criteria and can be sorted 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-Primary Site-Admission Displays

The system **MUST** be able to display all the data values that are stored for a specific patient, primary site, or admission.

5.6.4.3. Frequencies

The system **SHOULD** allow for the 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, incidence, 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 database, whether adding data to an existing record or adding new records, an accurate record linkage mechanism is needed to ensure that the additional data are correctly associated with the existing record(s). Efficient record linkage procedures on the same individual and the same tumor are essential, as central registries often receive multiple reports for the same person and tumor due to multiple reporting sources. If a record is added to the database without adequate checking for redundancy, this will generate duplicate patients or tumors in the registry's database, and overcounts will occur. However, if distinct records are linked together mistakenly, this will result in undercounting patients or tumors.

These overcounts and/or undercounts diminish the quality of the registry and result in inaccurate counts and rates being generated. 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

Regardless of the type of linkage being run, the key fields used for the record linkage should be assessed for quality before use in the linkage, 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 to determine match status for questionable matches that need to be reviewed manually. Checks for consistency in coding conventions between the files for the fields being compared should be performed (e.g., M, F, and U versus 1, 2, and 9 for the coding of sex). Data formatting in the files being compared also should be standardized, so that variants in punctuation, spacing, and case will not cause matches to be missed. For example, "Van Houten" and "Vanhouten," or "O'Hara" and "Ohara" can be standardized to "VANHOUTEN" and "OHARA" to eliminate non-matches due to irregularities in data entry.

Although probabilistic linkage often is recommended over deterministic methods for all linkages—and in particular when coding errors, reporting variations, missing data or duplicate records exist in the data being linked—deterministic methods sometimes are more feasible to implement. Central registries can achieve

accurate, efficient patient linkage using record linkage that is performed deterministically, probabilistically, or using a combination of both approaches.

5.7.1.1. Deterministic Linkage

A deterministic record linkage involves the comparison of two records on several key fields (e.g., social security number, last name, first name, etc.). Two fields are considered a match only if their values are identical.

However, slight variations often exist in the data between the two files for the same fields, or fields are missing from one of the files. These variations would prevent a match from being identified via deterministic linkage. The deterministic linkage algorithm is pre-defined, and results in either a match or non-match. It is good for use in production environments, and can be easily incorporated into a data system. However, it may miss a significant number of true matches due to variations within the data being linked.

Deterministic record linkage is best suited for the linkage of files that contain no errors and no missing data. However, there are some approaches that will increase the effectiveness of the deterministic matching process. Matching on partial fields (e.g., the last 4 digits of the social security number or the year component of the date of birth) may help find matches that would be overlooked due to variations in the remaining components of the fields.

Performing multiple passes of linkages also can improve deterministic linkage results. Multiple passes involve specifying a specific set of matching variables, running the linkage, isolating all identified non-matches, running a subsequent linkage with different matching variables on the remaining non-matches, and repeating this process until the number of non-matches is pared down to only true non-matches.

In addition, algorithms can be developed that run against deterministic linkage results from the matching of multiple individual fields to determine the overall match status for a pair of records. These algorithms commonly take into account the match/non-match status on different groupings of specific fields to determine whether the record pair is a true match. For example, an algorithm might specify that a record pair is a true match if the last and first name, and social security number match, or if the first name, last name and sex matches and date of birth is +/- 4 years, etc. Or, an algorithm might generate a score for the potential record pairs based on the results across multiple matching fields, and automatically assign match status based on the score a record pair receives. Many production environments, such as those used for central registry patient linkage, commonly employ a deterministic approach in combination with such algorithms.

5.7.1.2. Probabilistic Linkage

A probabilistic record linkage also involves the comparison of two records on several key fields; however, in this case a probability is estimated that two records represent the same person (match versus non-match). The probabilistic linkage process involves finding the records in one file that appears to match records in the other file and calculating a linkage score that indicates, for any pair of records, how likely it is that they both refer to the same person.

The total score for a linkage between any two records is the sum of the scores generated from matching individual fields. The score assigned to the matching of an individual field is determined by the probability that the fields agree given that a comparison pair is a match (M probability; similar to sensitivity), and reduced by the probability that the fields agree given that a comparison pair is not a match (U probability; similar to specificity). Agreement between the fields argues for linkage (and results in a higher score);

disagreement argues against linkage (and results in a lower score). Full agreement between the fields results in a higher score than partial agreement. Some types of partial agreements are stronger than others.

Deterministic linkages only compare fields of the two records being matched; the remainder of the information in the files being linked is disregarded. Probabilistic linkages, on the other hand, can take advantage of information available from the data files being linked when comparing the two fields being matched. For example, the probabilistic algorithm can take into account the distribution of the last names in the data files being linked when generating the linkage score, enabling the linkage score to be value specific; a match on the more frequent name of "Jane" would result in a lower score than a match on the less frequent name of "Janiqua." Probabilistic scores also are field specific; a match on a field such as birth date would result in a higher linkage score than a match on the field of sex, because date of birth is more specific than sex.

Probabilistic record linkage involves such a large number of comparisons between the files being linked that probabilistic linkage of large files can require a considerable amount of time and computing resources. Blocking is an initial probabilistic linkage step that reduces the number of record comparisons between files. Blocking involves sorting and deterministically matching two files by one or more "blocking" variables, and subsequently making the probabilistic comparisons only between the records that matched on one or more of the blocking variables. Last name, first name, social security number, and date of birth are common blocking variables. The main objective of blocking is to exclude very unlikely record pairings from the resource-intensive probabilistic comparisons to save both time and computing resources.

5.7.2. Phonetic Systems

A phonetic system also may be applied to the name-matching fields to reduce the number of matches missed due to variations in spelling, typographical errors, and misspellings of the name fields between two records being linked. Phonetic coding involves coding a string based on how it is pronounced. There are two main phonetic coding systems that are commonly used, Soundex and NYSIIS.

Soundex is the oldest, most widely accepted phonetic coding system; it is more than 120 years old and was first applied to 1880 census data. The Soundex phonetic code for a name consists of a letter followed by three numbers—the letter is the first letter of the name, and the numbers encode the remaining consonants. Additional letters are disregarded. Zeroes are added to the end of the code, if necessary, to produce a 4-character code. For example, the name Washington is coded W-252 (W, 2 for the S, 5 for the N, 2 for the G, remaining letters disregarded). Generating a Soundex code is simple and fast.

The NYSIIS was developed in New York State in 1970 and has become very widely used. NYSIIS maps similar phonemes to the same letter and unlike Soundex, maintains relative vowel positioning. One advantage of NYSIIS is that it generates a string that can be pronounced by the reader without decoding. For example, the name Deborah Walker is translated to DABARA WALCAR. NYSIIS is an improvement over the Soundex algorithm in that it is more distinctive. Names are more likely to have the same Soundex than the same NYSIIS. NYSIIS has a reported accuracy increase of 2.7 percent over Soundex. Studies suggest that NYSIIS performs better than Soundex when Spanish names are used. However, because Soundex is less distinctive, it may bring more pairs for comparison together.

Using phonetic coding when deterministically linking name fields softens the rule that the name fields must match exactly and will help to link records for the same person that would not have been linked by deterministically matching on the original name field.

5.8.   LINKING PATIENTS VERSUS LINKING TUMORS

Multiple records submitted for the same patient need to be linked. Records that describe the same tumor must be identified so that they can be consolidated, and records describing different neoplasms for the same patient need to be stored as separate tumors. The task of tumor linkage and consolidation is hard to fully automate; it involves comparisons of many fields, such as primary site, histology, and the date of diagnosis. Nuances associated with assigning morphology and the ambiguous terminology rules used in determining the date of initial diagnosis can make this procedure so complex that it may require manual intervention (see Section 5.9.).

5.8.1.   Software

Many commercial probabilistic record linkage software packages are available for purchase and use by central cancer registries. An alternative to commercial software is Link Plus, a stand-alone, probabilistic record linkage program available free of charge from the CDC/NPCR website: (www.cdc.gov/cancer/npcr/). Link Plus was designed specifically for cancer registry linkages and accepts NAACCR-formatted files, fixed-width, and delimited files.

Regardless of the program used for linkage, the efficiency of the program is a key feature. Ease of use also is important, especially if the central registry does not have a record linkage expert on staff. Faster, more efficient linkage processes allow for more linkages to be conducted at less cost and reduced staff time, resulting in more accurate counts in the data and more time for increased utilization of cancer registry data.

5.8.2.   Standards

The central registry **MUST** have an effective record linkage system for linking patients and neoplasms. Record linkages can be carried out manually, by computer, or by a combination of both. 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. With advances in computing technology and resources, computerized tumor (neoplasm) linkage is likely to become more practical for all central registries, regardless of size or available resources.

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 Standards Section of the NAACCR website at www.naaccr.org).

5.9   DATA EDITING

Data editing is an essential aspect of the central registry's overall data management. Data quality edits are discussed in Section 3.1.4.4. The electronic NAACCR EDITS Metafile includes the standard single and inter-field edits of all major cancer registration standard-setting organizations and can be downloaded from the NAACCR website (www.naaccr.org/standards/volumeiv).

5.9.1.   Standards

The central registry system **SHOULD** employ a complete set of standard edits to evaluate a registry database on file. One way to conduct the evaluation is to use the NAACCR EDITS Metafile. 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 (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.

Data editing **MAY** be performed interactively, as a batch process, or both, and **MUST** be applied at several points in the data flow to: (1) tumor records, before submission to the central registry; (2) newly submitted tumor records before they are linked against the central registry database; (3) database tumor records after linkage; and (4) database tumor records after consolidations, corrections, or any other changes have been made.

The central registry's edits **MUST** allow for over-ride flags for situations in which the edit identifies a rare condition that needs review but may be correct. The over-ride 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 that might require over-ride flags as ''inconsistencies'' rather than ''errors,'' because the data are not necessarily incorrect.

The NPCR has developed the EDITS software tools to help with the writing, implementation, and publishing of standard edits. EDITS is a freeware utility suite supported by the NPCR and includes an edit-writer application (EditWriter3), an application program interface (API), and a generic driver program that incorporates the EDITS API (GenEdits Plus). EditWriter3 is used to write and distribute edits and edit sets via the EDITS Metafile. GenEDITS Plus is used to apply edits to data files using the metafile produced by EditWriter3. The API can be used with various database management systems for real-time validation of cancer data using the edits provided in the metafile. NPCR EDITS software can be downloaded from the NPCR utilities website (www.cdc.gov/cancer/npcr/tools/edits/).

5.10. 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 **MAY** include multiple abstracts from hospitals, radiation therapy centers, or other providers, or they can include information from death certificates or from other registries. Values for the same data item, patient, and tumor may be identical from each source, but they also may be contradictory or complementary. A large task of the central registry **SHOULD** be 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. In addition, it should be emphasized that record consolidation is distinct from and subsequent to tumor linkage, although both operations **MAY** be performed nearly at the same time during an automated or manual process. More resources on record consolidation are available from the Registry Resources Section on the NAACCR website (www.naaccr.org).

Examples:

Hospital A:

Name:	SMITH JOHN FITZGERALD
Date of Birth:	2/10/27
Social Security Number:	999-99-9999
Date of Diagnosis:	3/99/07
Text:	Carcinoma of the colon diagnosed March of 2007 by biopsy elsewhere, treated at Hospital A by sigmoid colectomy on 4/15/07.

Hospital B:

Name:	SMITH JACK
Date of Birth:	10/2/27
Social Security Number:	123-45-6789
Date of Diagnosis:	2/6/07
Text:	Biopsy of the sigmoid colon on 2/6/07 showing adenocarcinoma.

Death Certificate:

Name:	SMITH F. JOHN
Date of Birth:	2/10/27
Social Security Number:	123-45-6789
Date of Death:	9/12/07
Cause of Death	153.9 (Malignancy of the colon)

Record From Neighboring State/Province/Territory Registry:

Name:	SMITH JOHN F
Date of Birth:	2/10/27
Social Security Number:	999-99-9999
Text:	Patient presented to our facility on 5/1/07 with a diagnoses of T3 N2 M0 colon cancer. He had been diagnosed 2 months prior at another facility. Patient was seen at Major University Medical Center for chemotherapy. Start date was 5/5/07.

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:

Name:	SMITH JOHN FITZGERALD
Social Security Number:	123-45-6789
Date of Birth:	2/10/27
Date of Diagnosis:	2/6/07
Primary Site:	C18.7 (Sigmoid colon)
Histologic Type:	8140/3 (Adenocarcinoma)
Date First Course Treatment:	4/15/07
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 data items, 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 histology) or those used for subsequent calculations (e.g., date of birth, diagnosis for calculation), it is important to establish one value to be used in analysis. Some central registries have found it helpful to store separately all values that have ever been submitted, so that the system can reproduce a record as it was originally submitted by a facility.

5.10.1. Standards

Standards for item-specific consolidation rules, either for computer application or for manual application, have not been developed but many existing systems can be used as models (e.g., CRS Plus™ - TLC function). Some general principles include:

❖ 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 consolidation algorithm cannot determine the correct value.

❖ Known values **SHOULD** be preferred over unknown values and more specific values **SHOULD** be preferred over less specific values. However, this rule should be applied with caution, because existing but invalid 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. Moreover, the incorrect number can create problems in linkage projects.

❖ Consolidation **SHOULD** take into account the sequence of events, the class of case, and the type of reporting source.

❖ Consolidated values for tumor characteristics (e.g., primary site) **SHOULD NOT** contradict reliable text information concerning staging and the treatment of cancer (e.g., it is improper to consolidate primary site to C049 - Floor of the Mouth, NOS when the surgery text indicates tonsillectomy).

5.11. 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 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's *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.

5.11.1. Potential Problems With Automated Updates

5.11.1.1   *Keeping the Different Datasets Synchronized*

If review of hospital data is performed centrally and some data items are accepted while others are 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 becomes 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.

5.11.1.2. *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 for which different software packages are being used, the central registry will have to develop procedures to reduce the volume of review workload.

5.11.1.3. *Data Ownership*

Proprietary ownership of data between the hospital and central registry is complex (see Section 6.3.2. for a discussion on release of registry data).

5.11.1.4. *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.

5.11.1.5. *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 inconsistencies and the transactions can be reviewed and resolved by quality control staff.

5.11.1.6. *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 **SHOULD** be reviewed 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.

5.11.1.7. *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 made automatically, there is a risk of having 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.

5.11.1.8. *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 or 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.

5.11.1.9. *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), some of the cancer registry software vendors consider this to be the same treatment. Other vendors consider 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 manually determine 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.

5.11.1.10. *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.

5.12.   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.

5.12.1. Standards

The central registry **MUST** develop the technical, procedural, and administrative capacity to perform linkages with external files.

5.12.1.2. *Linkage With Death (Mortality) Files*

Linkage with death files is a particular instance related to the general linkage problem, one that **SHOULD** 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 **SHOULD** 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 2.2.8.. Because both the death file and the registry file are dynamic, timing of the linkages is important.

5.12.1.3. *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 2.2.12.. Linkage generally will involve 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 include linking a cohort against the central registry to determine the occurrence of cancer among the cohort, or linking another disease registry (e.g., an AIDS registry) against the central registry to ascertain the occurrence of cancer among individuals with AIDS. Confidentiality precautions **MUST** be followed stringently in all such investigations (see Section 6.3.).

5.13. 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.

5.13.1. 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.

CHAPTER 6:

SECURITY AND CONFIDENTIALITY

6.1. STRUCTURAL REQUIREMENTS

It is the responsibility of every registry to protect its data from unauthorized access and release. The central registry's Director **MUST** be responsible for data security. There **SHOULD** be a Chief Technology Officer who works directly with the registry Director to ensure data security. 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. Each central registry **MUST** comply with all applicable security procedures and practices of its parent organization. Confidentiality policies and procedures are required in all phases of central registry operations and **MUST:** (1) protect the privacy of the individual patient, (2) protect the privacy of the reporting sources, (3) provide public assurance that the data will not be abused, and (4) abide by any confidentiality-protecting legislation or administrative rules that may apply.

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

Aspects of confidentiality policies and procedures that relate to authorized use of and release of data are addressed in Section 6.3.2., Issues in Research, Reporting, and Release of Registry Data.

Just as a written security plan is specified by this document, a written data confidentiality policy is necessary. As an alternative, the plan could be jointly titled "Security and Data Confidentially Policy/Plan." This would appropriately denote the security's primary goal of protecting confidentiality.

6.1.1. Objectives

The objectives of a security system **SHOULD** include: (1) establishment of security standards, (2) training of registry personnel in those standards, (3) adoption of policies and procedures to meet those standards, and (4) periodic update of policies and procedures as appropriate.

6.1.2. Need To Protect Confidentiality

Registry personnel work with confidential information, including patient identification and medical histories. Unauthorized disclosure of this private information is illegal and threatens the environment in which the central registry operates. The continuous transmission and processing of patient data inherent to registry operations makes the registry vulnerable to security lapses.

Patient confidentiality **MUST** be protected. A risk assessment of the vulnerability of the central registry **SHOULD** be conducted and included in the central cancer registry's security manual, as defined in Section 6.2.2.. Such an assessment **SHOULD** identify potential threats from natural, human, and environmental sources as well as vulnerabilities due to weaknesses in security configuration, policy standards, procedures, and degree of compliance with both technical and non-technical requirements.

6.1.3. Legislation

This section describes security legislation. Privacy legislation is covered in Section 6.3.2..

6.1.3.1. Definitions

❖ *Access*: Individuals who have permission to see the electronic health information are able to access it when needed.

❖ *Confidentiality*: The concept that data and/or information are protected from unauthorized persons or processes. 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.

❖ *Covered Entities* (*under HIPAA):* Include health plans, health care clearinghouses, and health care providers that transmit health information electronically in connection with certain defined HIPAA transactions, such as claims or eligibility inquiries. Central cancer registries are not considered covered entities.

❖ *HIPAA:* A law requiring, among other things, uniform federal privacy protections for individually identifiable health information.

❖ *Integrity:* The electronic health information is what it is supposed to be, has not been changed by anyone who does not have permission to change it, and is an accurate representation of facts and events.

❖ *Public Health Authority* (*Under HIPAA):* "An agency or authority of the United States, a State or territory, a political subdivision of a State or territory, or an Indian tribe, or a person or entity acting under a grant of authority from or contract with such public agency, including the employees or agents of such public agency or its contractors or persons or entities to whom it has granted authority, that is responsible for public health matters as part of its official mandate...Such agencies are authorized by law to collect or receive such information for the purposes of preventing or controlling disease, injury, vital events such as birth or death, and the conduct of public health surveillance, public health investigations, and public health interventions." Central cancer registries are considered public health authorities because state laws mandate their duties.

❖ *Security or Security Measures*: Encompass all of the administrative, physical, and technical policies and procedures that protect the confidentiality of the central cancer registry.

6.1.3.2. Registry Operations Affected by Security Legislation

The registry security operations affected by legislation include data collection, data storage, quality assurance, data access, and cancer related reports. Security legislation related to data release activities is described in Section 6.3..

6.1.3.3. Legal Tools Relevant to Securing Data

6.1.3.3.1. HIPAA

The HIPAA Security Rule (45 CFR part 160 and Part 164, subparts A and C) requires that covered entities in the United States protect and control access to electronic health information.

The HIPAA Privacy Rule addresses how all health information—whether the health information is in written, spoken, or in an electronic form—is used and disclosed. It creates minimum nationwide standards for making sure that an individual's health information is kept private. This section focuses on the HIPAA Security Rule.

Compliance with the Security Rule was required by no later than April 20, 2005, for all HIPAA-covered entities, except for small health plans, which had an extra year to comply.

The Privacy Rule applies to the disclosure of protected health information by covered entities as required by law; state cancer registries are not covered entities. Public health reporting under the authority of state law is specifically exempted from HIPAA rules. The provision of the Privacy Rule authorizing disclosure of protected health information as required by law is an exception to the requirement for written authorization.

Specific safety measures used to help protect and control access to electronic health information are described in Section 6.2.3.. Section 6.3. describes standards for confidentiality protection in data release.

6.1.3.3.2. Interjurisdiction Data Exchange Agreements

A legal agreement **MUST** be used when information concerning cancer patients, who are residents of one state, province, county, or country is exchanged between central cancer registries. For example, an interstate data exchange agreement **MUST** be used when information concerning cancer patients is exchanged between state central cancer registries. The following security provisions **SHOULD** be used:

❖ 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 to anyone not employed in the direct operation of the recipient registry. Employees and contractors may include those involved in the processing, administration, quality control review, and the statistical surveillance of cancer incidence data.

❖ Notify the exchange registry if, in the conduct of approved research or other activities, there is an unintended release of a cancer patient's identifying information. Should such a release take place, the receiving registry **MUST** be notified in writing within 48 hours of the release of the data.

6.1.3.3.3. Interagency Data Exchange Agreements

Central registries may sign legal documents to facilitate receiving data from another organization. The legal documents often specify conditions for acceptable use of the data and/or restrictions on the data's re-release. The registry **MUST** build both administrative and technical controls to ensure the registry's future ability to comply with agreed-upon terms. For example, this could involve flagging data items in the registry's data management system so that the registry knows how to handle the item when re-release situations arise. The included metadata also could be used to indicate that efforts should be made to collect the data item from a non-restricted source.

6.1.3.3.4. Veterans Health Administration Directive

The Veterans Health Administration (VHA) has special restrictions for data security, use, and release. The specifications typically come in the form of a VHA Directive—which requires a data transfer agreement between the VHA and the central cancer registry—and a standing letter of request for data from the central cancer registry.

6.1.3.3.5. Provincial/Territorial Health Information Protection Acts - Canada

In recent years, many of the Canadian provinces/territories have established specific legislation to address the protection of health information. For example, in 2003, the Province of Saskatchewan enacted this type of legislation that must be interpreted and applied by the central registry. Registries must investigate their specific jurisdictions to ensure that they are aware of any applicable legislation or regulations that will impact operations.

6.2.   REGISTRY POLICIES AND PROCEDURES

This section describes the development and use of security functions within operations, excluding information technology.

6.2.1.   Definitions

- ❖ *Appropriate Data Users:* Persons not employed or contracted by the central registry who require access to confidential data.

- ❖ *Central Cancer Registry Security Manual:* Hardcopy and/or electronic means of documentation for all security policies and procedures for the central registry.

- ❖ *External Consultant:* Person(s) contracted to do work for the central registry who work at a different physical location than the central registry.

- ❖ *Incident Reporting Form:* Means of recording security-related incidents.

- ❖ *Internal Consultant:* Person(s) contracted to do work for the central registry who work at the same physical central registry location.

- ❖ *Personal Security Compliance Form:* Documents participation in the Security Awareness Program.

- ❖ *Security Awareness Program:* Education and training of all persons who have access to confidential data.

- ❖ *Security File:* Place where a copy of security documentation is kept.

- ❖ *Staff Persons:* Personnel who are directly employed by the central registry.

6.2.2.   Confidential Data Flow

The flow of confidential cancer information occurs in many directions through various cancer registry operations. Data may be transmitted to many persons both inside and outside of the cancer registry. Strong

internal data (casefinding) and external data flow (analytic files) controls **SHOULD** be adopted to ensure confidentiality.

These controls **SHOULD** ensure that confidential data only are transmitted or downloaded by a limited number of authorized persons. Internal data flow, including unencrypted e-mail transmissions or unencrypted files on a server, also **SHOULD** be included in these precautions.

	Submitters	Data	Central Registry	Data	Appropriate Data Users
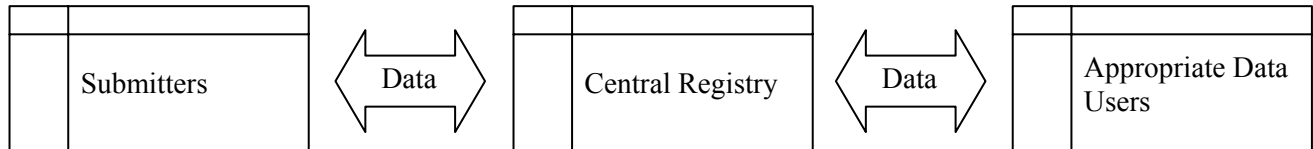
6.2.3. Access Control

The goal of access control is to allow access by authorized individuals and devices to central registry data and to disallow access to all others. Such controls **SHOULD** include physical barriers/safeguards, security locks, an alarm system, and visitor control. A central registry **MUST** establish specific guidelines in managing the controls for data access.

Registry data **MUST** be collected and stored in a physically secure place with strict access controls. Appropriate security measures **MUST** be in place at the physical location where data are stored. For example, a security guard on duty to control building access, card-key passes for access to designated offices, and card-key passes at night and on weekends to prevent unauthorized intrusion to the facility at all entrances. All registry personnel **SHOULD** be required to wear a photo ID all of the time.

A well thought-out fire plan **SHOULD** be enforced for rooms where data servers are located.

6.2.4. Registry Staff Security Policies and Training

The intent of staff security policy and training, hereinafter referred to as the Security Awareness Program, **SHOULD** be to reduce the risk of human error and misuse of cancer information to an acceptable level.

Specific security roles and responsibilities **MUST** be documented. These roles and responsibilities include general responsibilities for all employees and contractors, as well as specific responsibilities for protecting cancer information and performing tasks related to security procedures or processes.

A Security Awareness Program **MUST** be developed, implemented, and maintained that addresses the security education needs of all employees, contractors, and users as appropriate.

All individuals who have access to cancer information **MUST** receive security awareness training to ensure that they are knowledgeable of security procedures, their role and responsibilities regarding the protection of cancer information, and the proper use of information-processing facilities to minimize security risks.

At least annually, employees and users **SHOULD** sign a Personal Security Compliance Form documenting participation in the Security Awareness Program, which **SHOULD** be maintained as a hardcopy in a secure file. This declaration **SHOULD** remain in effect after cessation of employment. The central registry Director **SHOULD** maintain a list of staff members indicating the nature and extent of their access to registry data.

6.2.5. Consultant Security Policies and Training

Internal users, including contractors, often have direct access to confidential data. Job definitions **SHOULD** include whether access to confidential information is required. Non-disclosure agreements **SHOULD** be used whenever necessary. The internal user is responsible for any disclosure of information in their care, and **MUST** ensure the necessary precautions are taken during and after access to confidential information.

All registry staff **SHOULD** sign a Non-Disclosure and Data Confidentially Agreement annually. This document **SHOULD** include the need for annual security awareness training. It is important to be sure the annual review and confidentiality pledge include a thorough review of both the registry's security and data confidentiality plans. As risks and threats to data continually change, the security and data confidentiality plans will routinely change. At times, it could be necessary for staff to review these materials more often than annually.

6.2.6. E-mail, Fax, and Internet Use

Electronic mail (e-mail) is the electronic equivalent of an open postcard. Information transmitted over the Internet passes as readable bytes available to anyone who can read them. As the traffic passes from one network to another, the probability of a user's e-mail message being read increases. Additionally, if the message ends up in the wrong mailbox, one can unintentionally reveal information that should not be released.

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** use two envelopes, putting the confidential information in a separate tear-free envelope marked "Confidential," with a contact telephone number enclosed in the mailing envelope.

Central registries **MUST** prohibit sending patient information or passwords by e-mail and **MUST NOT** allow confidential or proprietary patient information or passwords to be included in email without encryption.

Central registries **SHOULD** discourage sending confidential information by fax unless a secure fax line is available. Fax server software can be configured to create a facsimile transaction stored on a secure server such that the fax can be routed to the recipient's confidential inbox, as opposed to receiving a fax over the main fax machine. In the event that a secure fax line is not available, the fax machine used for receiving confidential information **SHOULD** be located in a secure, limited-access area.

Central registries **SHOULD** consider the use of pre-coded numbers to eliminate dialing errors, cover sheets so that data are not physically exposed, testing fax machines to ensure correct number and function, and clearing fax memory buffers after use to prevent recovery of confidential information. When sending/receiving a fax to/from reporting facilities in which confidential data are involved, registry personnel **SHOULD** use an official cover sheet, confirm the fax number before sending the fax, and obtain acknowledgment that the fax was received.

Central registries **SHOULD** develop policies defining staff access to the Internet and authorized uses. Undesirable uses **SHOULD** be listed and discouraged. Workstations and laptops that have Internet connections **SHOULD** be continuously inventoried, locked, and controlled.

If the central registry provides Internet access, then security systems such as an intrusion detection system (IDS) and other measures **SHOULD** be maintained (see Section 6.4.).

6.2.7. Media Disposal

Central registries **MUST** establish appropriate disposal procedures for both electronic and paper-based media. Designated personnel **MUST** be charged with media-disposal responsibilities to ensure accountability and promote compliance with disposal policies. Such policies **MUST** prohibit employees from discarding media containing sensitive information along with regular garbage to avoid accidental disclosure. Registries **MUST** use paper/media shredders onsite or use collection and disposal services to ensure that the media are rendered unreadable and unlikely to be reconstructed. Registries that contract with third parties **SHOULD** use care in selecting vendors to ensure adequate employee background checks, controls, and experience. Contracts with third-party disposal firms **SHOULD** address acceptable disposal procedures. The disposal of registry information in the United States **SHOULD** meet the requirements of the HIPAA 501(b) guidelines.

Computer-based media present unique disposal problems. Policies and procedures **SHOULD** comprehensively address all of the various types of electronic media in use. Residual data frequently remain on media after erasure. Because these data can be recovered, additional disposal techniques **SHOULD** be applied to sensitive cancer records. Physical destruction of the media, for instance by subjecting a compact disk to microwaves, can make the data unrecoverable. Additionally, data sometimes can be recovered after overwriting. Overwriting may be preferred when the media will be re-used. Central registries **SHOULD** base their disposal policies on the sensitivity of the information contained on the media and, through policies, procedures, and training, ensure that the actions taken to securely dispose of computer-based media adequately protect the data from the risks of reconstruction. Management **SHOULD** log the disposal of sensitive media, especially computer-based media. Logs **SHOULD** record the party responsible for performing disposal as well as the date, media type, hardware serial number, and method of disposal.

6.2.8. Internal Auditing

Central registries **MUST** establish policies and procedures to conduct internal IT security audits on a periodic basis. The primary role of the internal IT audit staff is to assess independently and objectively the controls, reliability, and integrity of the IT environment within the registry. These assessments can help maintain or improve the efficiency and effectiveness of the registry IT risk management, internal controls, and security. The parent organization may fulfill this function.

Internal auditors **SHOULD** evaluate IT plans, strategies, policies, and procedures to ensure adequate management oversight. Additionally, they **SHOULD** assess the day-to-day IT controls to ensure that transactions are recorded and processed in compliance with acceptable accounting methods and standards and are in compliance with policies set forth by management. Auditors also **SHOULD** perform operational audits, including system development audits, to ensure that internal controls are in place, policies and procedures are effective, and registry personnel operate in compliance with approved policies. Auditors **SHOULD** identify weaknesses, review management's plans for addressing those weaknesses, monitor their resolution, and report material weaknesses.

6.2.9. Disclosure Accounting

Failure to observe the confidentiality policies **MUST** result in firm disciplinary action, including potential 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.

Formal incident or malfunction reporting and response procedures **MUST** be established to define the actions to be taken when an incident occurs. Compliance with the disclosure accounting policy is mandatory. The following **SHOULD** be included: (1) symptoms of the problem, (2) any messages displayed, and (3) the risk level of the violation.

The source **SHOULD** be isolated, if appropriate, until the problem has been identified and resolved; the incident **MUST** be reported immediately to the appropriate manager.

Feedback mechanisms **MUST** be implemented to ensure that individuals reporting incidents are notified of the results after the incident has been resolved and closed.

An incident management process **MUST** be established to track the types and volumes of security incidents and malfunctions. This information will be used by the appropriate manager to identify recurring or high-impact incidents and record lessons learned. This may indicate the need for additional controls to limit the frequency, damage and cost of future incidents. This information also **SHOULD** be taken into account in the policy review process.

All users **MUST** know the procedure for reporting security incidents, threats, or malfunctions that may have an impact on the security of cancer registry information. All staff and contractors **MUST** report any observed or suspected incidents to the appropriate manager. Approaches to incident management **MUST** be documented and procedures **MUST** be clearly identified to ensure responsibilities are defined, resulting in a prompt and organized response to security incidents.

Incident response procedures **MUST** be clearly identified to promote effective response to security incidents. They also **MUST** include procedures for information system failure, denial of service, disclosure of confidential information, and compromised software systems.

Practicing and drilling on incident response procedures is important. "Security work" in general involves three major areas: (1) prevention, (2) detection, and (3) response. A sound plan needs to attend to all three areas and not concentrate too heavily on prevention and the use of information systems technologies.

6.2.10. Corrective/Disciplinary Action

Each employee/contractor **MUST** understand his/her role and responsibilities regarding information security issues and protecting confidential data. Any compromise or suspected compromise of this policy **MUST** be reported to the appropriate management, as required by this policy.

Access authorization for user accounts involved in a compromise may be suspended during the time when a suspected violation is under investigation.

Areas where compliance with the policy requirements is not met **SHOULD** be documented and a plan **SHOULD** be developed to address the deficiencies.

Managers and supervisors **MUST** ensure that all security processes and procedures within their areas of responsibility are followed.

6.2.11. Registry Security Policy Manual

A registry security policy manual **MUST** be developed containing all security policies and procedures for the registry. This manual **SHOULD** be the basis for security training for users/contractors. The manual and supporting policies and standards **SHOULD** be reviewed on an annual basis, at a minimum.

6.3. DATA USE AND RELEASE

6.3.1. Definitions

❖ *Data User Agreement:* An agreement signed by data users that specifies the appropriate and inappropriate use of the data, usually specifying that the data user will not re-release the data to other individuals and not attempt to re-identify the records.

❖ *De-Identified:* A record for which direct identifiers such as name, address, and social security number, as well as indirect identifiers such as street address, have been removed.

❖ *Disclosure Risk:* The risk that an individual on a data file or in summary statistics will be identified and have confidential information disclosed.

❖ *Individual Level Data:* Data in which each record represents one patient or tumor.

❖ *Public Use File:* Data file of individual–level, de-identified data that are accessible by a member of the general public without an application process or data user agreement.

❖ *Restricted Access File:* Data file of individual level de-identified data that is accessible only after an application and review process and/or with a data user agreement.

❖ *Summary Statistics:* Data in tabular form, in which the cells of the table are counts of cancer cases/deaths, rates, or other summary measures.

6.3.2. Issues in Research, Reporting, and Release of Registry Data

Confidentiality is the cancer registry's responsibility to the patients whose data are 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. regarding standards for confidentiality and disclosure of data. See Appendix I for a copy of the NAACCR Policy Statement 99-01: Confidentiality.

Maintaining patient confidentiality while collecting and using high-quality data presents significant challenges. The *Annotated Bibliography on Confidentiality Protection in Data Release* developed by the NAACCR Confidentiality Subcommittee of the Data Use and Research Committee can be found at www.naaccr.org/confidentiality/index.asp. This document may be used to develop and maintain policies and procedures concerning data confidentiality.

The HIPAA Privacy Rule governs the use and disclosure of some health-related information. This federal law clearly defines the covered entities to which the Privacy Rule applies (i.e., a health plan, a health care clearinghouse, or a health care provider).

Because U.S. central cancer registries do not perform any of these functions, they generally are not covered entities and the HIPAA Privacy Rule governs neither their activities nor the information that they hold, including the release of registry data. A HIPAA frequently asked questions (FAQ) document has been developed for registries and can be found on the NAACCR website: www.naaccr.org. This FAQ document is updated as necessary to reflect ongoing interpretations and revision to these rules.

6.3.2.1. Definition of Confidential Data

In terms of patient-specific data, confidential data are any information that would allow public identification of the cancer patient or would disclose information about the cancer patient that would otherwise be considered private or protected from disclosure by law. 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 inadvertently may 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.

6.3.2.2. 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 the 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.

6.3.2.3. 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: (1) hospital cancer registrar requests a list of all prostate cancer patients who have been treated at his or her facility, or (2) physician requests a list of patients that he or she has treated for breast cancer.

❖ Central cancer registries **MUST** abide by their specific laws or regulations that may have specific procedures for the 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 6.3.2.3.2.).

❖ Central registries **MUST** document in their security manual the procedures and criteria for the release of registry data to researchers who request access to data.

6.3.2.3.1. Inappropriate Uses of Confidential Information

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.

6.3.2.3.2.   Standards for Protecting Confidentiality in Summary Statistics

Reports of summary statistics generally do not raise confidentiality concerns. However, confidential information can be conveyed inadvertently through summary statistics. Registries **SHOULD** institute a policy to evaluate the disclosure risk of summary statistics and apply an established technique to minimize the risk prior to publication. The central cancer registry **SHOULD** institute a policy to protect against disclosure in 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 6 (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. These types of cells inadvertently could 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 do not need to be suppressed. The Confidentiality and Data Access Committee of the Federal Committee on Statistical Methodology has prepared a checklist tool that

statistical agencies such as cancer registries can use to review the disclosure risk of data products, which can be found at http://www.fcsm.gov/committees/cdac/.

6.3.2.3.3. *Standards for Protecting Confidentiality in De-Identified Data Files*

Data files that contain individual-level, de-identified records can be released as public use files or as restricted access files. Confidentiality concerns for these record level files include the capability of identifying a patient from the data file and 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. Re-identification is a particular concern when the file includes personal characteristics such as age and race or when it includes information on residence coded for small geographic areas such as Zip codes or census tracts. NAACCR has developed the Record Uniqueness Program to aid registries in identifying the percent of records that are outliers on de-identified data files. The Record Uniqueness Program is available for download on the NAACCR website (www.naaccr.org). It contains complete instructions for use and interpretation. The Confidentiality and Data Access Committee's checklist mentioned in the previous section also is applicable to de-identified files.

6.3.2.4. **Standards for Use of Registry Data for Research**

6.3.2.4.1. *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 4.3.5.). When non-confidential data are not sufficient to answer the question, the central registry **MUST** determine whether the researcher is qualified to use cancer registry data for research purposes. The central cancer registry may be adversely affected 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 scientific investigators. Registry data **SHOULD NOT** be made available for scientific research until 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 **MUST** ensure 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 (United States) 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.

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).

6.3.2.4.2. *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 **MUST** retain the right to review any reports prior to their dissemination to ensure that confidentiality has been respected.

6.3.2.4.3. *Patient Contact for Participation in Epidemiologic Studies*

Central cancer registries can identify cancer patients as potential subjects for epidemiologic studies that involve contacting and interviewing patients. In these instances, the study **MUST** meet all of 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 make a decision regarding study participation and that his/her physician does not have the right to make that choice on the patient's behalf.

Consequently, in many current epidemiologic studies, passive physician contact is used. 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 active physician permission before contacting the patient. This requires that the physician give permission before the patient can be contacted.

Additionally, practice varies among registries as to whether or not informed patient consent needs to be obtained prior to the release of identifying information to researchers. Some registries perform patient consent prior to releasing identifying information to researchers, while other registries release the identifying information and informed consent is obtained by the researcher. Registries **MUST** establish policies on release of identifying information in accordance with the laws governing the registry.

6.4. INFORMATION TECHNOLOGY POLICIES AND PROCEDURES

6.4.1. Security Policies

This section describes the development and use of security functions within information technology.

There **MUST** be written security policies that are shared with staff and included in training. These policies **SHOULD** at a minimum cover the items below along with the appropriate use of registry equipment and resources as well as disciplinary actions for non-compliance.

6.4.2. Definitions

❖ *Authentication:* Process used to identify a user.

❖ *Authorization:* Process of granting permissions to a user.

❖ *Encryption:* Process of converting readable information into undistinguishable text that only the owner of the encryption key can convert back to readable information.

❖ *Malware:* Generic term for unwanted software such as viruses, spyware, trackware, and root kits.

❖ *PDA:* Personal digital assistant, palmtop computer.

❖ *Security Boundary:* The limit at which the registry controls the equipment, configuration, and security of the computers and network.

❖ *UserID:* Unique set of letters and/or numbers that identify an individual.

6.4.3.   User Authentication and Authorization

Most systems and applications control and track access by a UserID (authentication). Because this UserID controls access (authorization), policies and procedures **MUST** be in place to grant a unique ID to each user and users **MUST** be trained to protect this ID and not share it. Policies **MUST** also be in place to promptly disable UserIDs when an individual is terminated or the access is no longer needed.

6.4.4.   Passwords

To prove that the person typing the UserID is the correct person, a password **MUST** be required to complete the authentication process. There are additional methods such as biometrics (i.e., fingerprint) or hardware tokens (card with a changing number on it), but these are rare and used only in high-security settings.

Because the password is key to proving that the user is authentic, good password policy and training are paramount. The challenge is setting the requirements for the password to a difficult enough level such that it cannot be easily guessed or hacked and at the same time not overly burdensome so that the owner does not have to write the password down and leave it in proximity to their computer for reference. The minimum policy **SHOULD** be 8 characters or more and complex (meaning it requires 3 of the 4 character sets—upper-case letters, lower-case letters, numbers, symbols), or 11 characters or more without the complexity associated with the 8-character approach. All passwords **SHOULD** expire in 60-90 days and the system **SHOULD** not allow a password to be reused. Training also **SHOULD** be included to assist registry staff with selecting effective passwords and to instruct staff not to share their password (not even with technical support staff or supervisors).

6.4.5.   Least Privileges

The concept of least privilege involves granting the smallest amount of privileges or permissions to each user required to complete their job.

Although it is easier to grant all users full access to all data and administrative rights to the computers, it greatly increases risks. A well thought-out and documented set of permissions by work duty **SHOULD** be developed and implemented. This will ease administration of users and increase security. Granting users user rights and not administrative rights to their computers greatly reduces the chance of having inappropriate or malicious software installed on computers. This may increase IT support for installing applications, but also should reduce problems, performance issues, and down time.

Malicious software usually is installed inadvertently by regular users because they were unaware that the software was inappropriate or because it appeared to be harmless on a web page or e-mail attachment. IT support **SHOULD** maintain an awareness of current malware and take appropriate actions to minimize the risk of installation on the registry's information system.

6.4.6. Network Access

Foreign systems connecting to a registry's network pose a significant threat to the network and data because the status of the network or machines is unknown. Policies **SHOULD** be in place to block or at least restrict any access to the registry's network and to data from machines that are not maintained by the registry. A technical solution **SHOULD** be provided if possible.

Registry external consultants **SHOULD NOT** be allowed to connect unless they can prove compliance with registry policies and that their system is free of malware. If remote access is available, then additional controls **SHOULD** be in place to heavily restrict and monitor this access.

6.4.7. Activity Logging

An activity log **SHOULD** record users who are accessing systems. Most operating systems, server applications, and network appliances have the ability to record these activities. This logging capability **SHOULD** be enabled to detect inappropriate access, be monitored by an appropriately trained staff person, and archived for a minimum of 2 years.

6.4.8. Encryption

Most systems by default store and transmit data in a readable, clear-text format. Because communications can be listened to, and laptops, storage devices, PDAs, and servers can be accessed or stolen, encryption is the best protection. Encryption is a method of mathematically scrambling the data so that only the appropriate user or system can understand it. The encryption method used **SHOULD** be certified under FIPS 140 (http://csrc.nist.gov/publications/PubsFIPS.html).

Encryption **MUST** be used for all digital communications (Web, FTP, e-mail, etc.) or storage media (laptops, USB drives, CD-ROMs, DVDs, PDAs, tape backups, etc.) containing confidential data that leave the security boundary of the registry facility. Communications and confidential data stored within the registry security boundary **SHOULD** be encrypted if possible. Confidential data that are backed up **MUST** follow the same encryption requirements and be encrypted if they are not stored within the protection of the registry's offices.

6.4.9. Firewall

A firewall is a device or software that blocks (selectively if so configured) access from an outside source to the local machine or network. Most current networks have a firewall between their network and the Internet. Windows XP® Service Pack 1 and above has a firewall built in that can be enabled to protect the local machine. All registry networks **MUST** have a firewall protecting the network servers and workstations from unrestricted access from the Internet. If a laptop will connect to a non-registry network, it **MUST** have a firewall enabled on the local machine to protect it from outside access.

6.4.10. Intrusion Detection System

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.

An intrusion detection system, or intrusion prevention system (IPS, if it automatically blocks after detection), listens to communications to detect if malicious behavior is occurring, such as a worm or other exploitation, and will either notify an administrator or block the activity. An IDS/IPS **SHOULD** be installed either outside

the firewall (to monitor what attempts are being made to compromise the network), inside the firewall (to monitor what compromise attempts made it through the firewall), or both.

6.4.11. Data, Equipment, and Media Retention and Disposal

Policies and procedures **SHOULD** be in place to document how long to keep a particular type of data, equipment, and media. When the retention period is over, the data, equipment, and media need to be securely destroyed as described Section 6.2.7..

6.4.12. Application Specific

6.4.12.1. Intranet (Internal Network Applications)

Registries **SHOULD** publish and follow a list of approved commercial grade or select supportable open-source applications and limit the user's ability to install their own applications to reduce the risk of malware.

6.4.12.2. Internet (Internet Accessible Applications)

If Internet-accessible applications and data will be provided by a registry, special care needs to be implemented to protect the servers and data. This can be very challenging and somewhat risky, because a human error, programming mistake, or bug in a program potentially could expose data to anyone on the Internet. Extra care needs to be taken to implement proper authentication, authorization, logging, secure programming, and server and firewall hardening and maintenance. See above.

6.4.12.3. Application Maintenance

Most applications and operating systems are updated for enhancements and patched for security issues on a regular basis. Not keeping systems updated increases the potential for systems to be exploited. A process **SHOULD** be developed to evaluate and install any operating system or application update or patch that the manufacturer ranks as critical, within 30 days of its release. All machines **SHOULD** be properly updated every 90 days. Periodic audits to verify that all machines have been properly updated also **SHOULD** be conducted.

6.4.12.4. Malware Protection

Malware is a catch-all term for a group of malicious software including viruses, spyware, addware, and trackware that are detrimental to the security and confidentiality of systems. Most current commercial-grade antivirus programs include some spyware, addware, and trackware protection. All systems **MUST** have some form of antivirus software running that continuously monitors for infections, performs a full system scan once per week, and updates its signatures (list of viruses it checks for) daily.

6.4.13. Hardware Specific

6.4.13.1. Laptops

Due to their portability, special precautions for laptops need to be followed over and above what would be done for a desktop machine or server. Laptops that may contain confidential data **MUST** have the data encrypted. It is much safer to encrypt the entire hard drive as opposed to folders or files. Laptop physical security also is an issue—the safest location is in the physical possession of the user. Laptops **SHOULD**

remain in the possession of the user at all times and when not in use, secured in a locked cabinet or cabled to an immobile object. Laptops left in cars, even in the trunk, are prime targets for loss.

6.4.13.2. PDAs

PDAs that synchronize e-mail or are used for other applications or for file storage may contain confidential data, so encryption **SHOULD** be set up and the device **SHOULD** remain in the user's possession, or be secured in some other way. Any access to the PDA **MUST** require at least a password.

6.4.13.3. Portable Storage

Portable storage such as CD-ROMs, DVDs, USB keys, USB hard drives, etc., have become very popular and easy to use. The disadvantage of their ease and portability is that they can easily be lost or stolen. Policies and procedures **MUST** be written to restrict use of portable storage to where appropriate and include encryption requirements.

6.4.13.4. Wireless Network

Wireless networks are becoming popular for home Internet access and many businesses are installing them. Wireless networks are inherently insecure because they broadcast the signal and anyone can listen to the communication. Therefore, the use of wireless networks **SHOULD** be minimized and any communication over wireless networks **MUST** be encrypted. Wireless networks can be set up using their own encryption and **SHOULD** be configured with a minimum of wi-fi protected access encryption. Even if the wireless network provides encryption, the use of other security measures such as a virtual private network or secure website access is recommended.

Wireless networks **SHOULD** be further secured by restricting access only to authorized wireless clients using the client hardware's media access control (MAC) address. Only wireless network access points that support MAC address restriction **SHOULD** be used on a registry network. The wireless access authorization list **SHOULD** be monitored and updated as wireless clients are retired or no longer require wireless access.

6.5.   DISASTER RECOVERY

This section describes emergency preparedness, disaster recovery, and continuity of registry functions.

6.5.1. Guidelines

Every registry **SHOULD** develop a plan for responding to events that could cause an interruption of registry activities and endanger the lives and property of staff and their families. Further, the plan **SHOULD** consider possible disaster recovery options and how the registry might continue operations offsite.

In any plan, each registry **SHOULD** address, at a minimum, each of the operational and IT issues outlined in this section.

6.5.2. Registry Operations Role

Registry Directors/managers **SHOULD** begin the registry preparedness plan by identifying and incorporating the preparedness plan of the hosting institution into that of the registry.

Whether housed in a state, federal, or private/university building, registry managers **SHOULD** identify the key decision makers within building management and obtain a copy of the institutional and/or building emergency response plan. Registry Directors/managers **MUST** be aware of the preparedness plan(s) of their hosting institution(s) and be able to communicate the details to all registry staff.

6.5.3. Points of Contact

Registry Directors/managers **SHOULD** maintain a list of points of contact for emergency situations. The point of contact **SHOULD:**

❖ Decide a procedure for building closure/reopening.

❖ Incorporate the registry's hosting institution plan as appropriate.

❖ Identify the key individuals within the property management group from which the registry rents space (if the registry rents space in a privately owned building).

❖ Identify any other key property management personnel.

Before completing the emergency plan, each registry **SHOULD:**

❖ Identify all emergency exits and the locations of fire extinguishers and any automated external defibrillators (AEDs).

❖ Verify that registry staff know how to use a fire extinguisher and portable AEDs.

❖ Identify if anyone in the registry or the building has emergency medicine or cardiopulmonary resuscitation training.

Once the plan is established, each registry **SHOULD:** (1) post all of the procedures for new staff or visitors who may not be aware of these emergency procedures, (2) make an evacuation plan for registry staff to follow, and (3) verify that registry staff have reliable transportation or means of evacuation.

Also, the following issues **SHOULD** be considered depending on the registry's location: (1) if the registry is located in an earthquake- or tornado-prone area, establish procedures for registry staff to take shelter inside the building; and (2) if the registry is located along the Gulf or Atlantic Coasts, verify that registry staff have personal evacuation plans.

6.5.4. Communication

Registry Director/managers **SHOULD** decide how to manage, maintain, and use communications systems to keep in touch with institutional decision makers and registry staff.

Communication is one of the most critical yet vulnerable aspects of an emergency. Policies and procedures **SHOULD** describe how information and instructions about the emergency will be disseminated. To account for registry staff, regroup, and resume operations, registry Directors/managers **SHOULD** employ multiple methods of communication.

Every employee of the registry also is an employee of a larger institution. Registry Directors/managers **MUST** understand how the registry's institution expects them to stay in touch in the event of an emergency. If registry Directors/managers are unaware of requirements to contact their human resources office in an emergency, they could be subject to disciplinary action and/or loss of pay.

6.5.4.1. Phone Trees/Call Lists

As part of the policy, registry Directors/managers **MUST** establish a master contact list that every staff member can access. The following **SHOULD** be considered: (1) decide who should maintain the list; (2) determine how often the list should be updated; and (3) determine how much contact information is to be collected (name, address, phone number, alternate phone number, next-of-kin name, next-of-kin phone number).

6.5.4.2. Text Messaging Systems

To better maintain contact with staff, registries should consider text messaging as part of their emergency plans. Registry Directors/managers should discuss the following:

❖ If it is an option, decide which system will work best for the registry.

❖ Determine whether that system has been tested.

❖ Decide who is to be responsible for this option (design, training, and testing).

❖ Verify that staff are able to participate.

6.5.4.3.   Websites/E-mail

For the communication lines to remain open among staff, the registry's website **MUST** be updated on a regular basis. Each registry **SHOULD:**

❖ Verify that all staff members have an e-mail address.

❖ Request that staff maintain an alternate e-mail address that is not dependent on the hosting institution (Yahoo, MSN, gmail, etc.).

❖ Determine who is responsible for maintaining the list of e-mail addresses.

❖ Establish guidelines to follow regarding the list.

❖ Consider other contact possibilities and purposes.

6.5.4.4. Alternative Methods of Communication

The registry **SHOULD** consider the following alternatives to create a more efficient method in communicating with staff: (1) satellite phone; (2) cell phone with a phone number outside of the immediate area code(s); (3) websites such as Wiki sites, Yahoo groups, etc.; and (4) a 1-800 number for emergency use.

Registries **SHOULD NOT** assume that they can rely on land-line, cellular or Internet-based communications. An earthquake, flood, hurricane, etc. can cause extensive damage to cellular towers, telephone lines, electrical

lines, power stations, and data centers. There is no guaranteed method of communication, but the more options presented to registry staff, the better the chances of staying in contact.

6.5.4.5. Point Person

The registry **SHOULD:** (1) establish a point person (e.g., at the NPCR/SEER project office, NAACCR office, regional registry office, state office, etc.); and (2) instruct registry staff to report to this person as soon as possible.

6.5.5.   Designate Staff Role in Emergency Preparedness Plan

Registry Directors/managers **SHOULD** decide who has what role in emergency preparedness, disaster recovery, and continuity of registry activities. Every staff member can play an important role in the preparedness plan.

❖ Each staff member **SHOULD** have a readily available checklist of their duties in the event of an emergency.

❖ Identify the responsibilities of each staff member. Generally, the IT/IS staff will be the most involved, as protecting and securing registry data is of paramount importance.

❖ Assign a backup person to each task. A clear hierarchy of responsibilities makes preparation more efficient, but emergency situations are unpredictable. For every task assigned in the planning process, consider assigning a backup person.

❖ Develop a policy specifying what can/cannot be taken out of the building in the event that the premises must be evacuated. In addition, a simple inventory list of any equipment removed **SHOULD** be maintained to account for all registry equipment.

❖ Develop a policy for securing and protecting registry assets left behind in the event that the office must be evacuated.

❖ Consider how emergencies that arise over weekends or off hours should be handled. There may be civil service or other restrictions on what staff can be asked to do during non-work hours.

6.5.6.   Plan To Regroup and Reestablish

Registry Directors/managers **SHOULD** have a plan for how to regroup and reestablish registry operations after the emergency. Because the registry has a communication plan, all staff can be quickly located, their safety verified, and they can resume some level of work.

❖ Identify what equipment is immediately available for use, how quickly the database can be restored, and what additional resources are required.

❖ Contact NAACCR, NPCR, SEER, etc., to update them on the registry's status.

❖ Depending on the extent and level of damage, the facilities that report to the registry may not consider re-establishing the cancer registry a priority. Further, they may have suffered more damage than the central office and need central registry assistance identifying cancer patients and the treatment that they may

require. Establishing a new rapport can build good will and possibly even speed up the reestablishment of the hospital registry.

6.5.7. IT/IS Role

Technical staff **SHOULD** work closely with registry managers to develop the registry's preparedness plan. Registry data are irreplaceable. The data **MUST** be protected in the event of an emergency. Equipment can be replaced, but the data contained thereon cannot.

6.5.8. IT/IS Data Back-Up Plan

Technical staff **MUST** have a data backup plan for the central registry database and **SHOULD** have a plan for systematic, centralized backup of staff computers.

❖ If the data are backed up to tapes, ensure that they are stored in a secure, fireproof area.

❖ If the data are backed up by the institution and not by registry staff, verify that an institutional policy is established for tape storage, tape recall, and data recovery.

❖ If widespread damage occurs to the institution, evaluate the probable promptness of the institution's recovery.

❖ Develop a plan for how the registry plans to restore the database after an emergency.

❖ In the event of an evacuation, establish a plan to move critical backup equipment offsite so that data recovery can proceed. Storing backup tapes offsite is most effective if the machine that reads the tapes also is available offsite to restore the data.

❖ Establish a plan for offsite storage of data. If the registry is located in an earthquake- or hurricane-vulnerable area, consider an arrangement for emergency backup to a site hosted by an organization willing and able to abide by the registry's data security requirements.

6.5.9. Evacuation of Equipment

Technical staff **SHOULD:**

❖ Identify and label all equipment with a notice as to whether it should be taken out of the building or stored in a secure area in the event of an evacuation.

❖ Identify a "safe room" for equipment storage.

❖ Prioritize equipment for possible evacuation and clearly label each item.

❖ Instruct registry staff members on how to read the equipment labels, dissemble their equipment, and move it to safety or prepare it for evacuation.

Users with laptops **SHOULD** be required to remove them from the office in the event of an emergency. If an evacuation is in effect, users with laptops **MUST** leave with the laptop and all laptop accessories. Laptops not routinely used in the office may not have the latest service packs, virus definitions, etc., installed. Consider a policy that requires all laptops to be brought in regularly for routine updates. This also provides an

opportunity to back up any important files and/or discuss with the user how to backup their laptop on a more regular basis.

6.5.10. Equipment Recovery

Technical staff **SHOULD** develop a plan for the possible recovery and removal of equipment if, after an emergency, the office can no longer be used by registry staff.

❖ If an equipment-recovery mission can be attempted, remove priority equipment.

❖ Evaluate the status of the equipment left behind.

❖ Construct a clear inventory of what was damaged, what was left behind, and what was removed and to where: this can be invaluable evidence when requesting insurance, Federal Emergency Management Agency, or other reimbursements.

6.5.11. Re-Establishing the Registry Database

In the event the registry cannot return to their office, technical staff **SHOULD** develop a plan for reestablishing the registry database and providing connectivity to staff from remote sites.

❖ Regrouping and resuming registry operations **MUST** be a priority. This will be difficult if nearly all equipment is destroyed or off limits due to building structural failure, but it is not impossible.

❖ Consider replacing desktops with laptops so that users are more mobile. However, there needs to be sufficient warning before the emergency, so that users are able to leave with their laptops.

❖ Recruit an organization to host the registry database until the registry has a new location.

❖ Explore methods of providing remote connectivity so that users can connect to the database.

❖ Conduct routine drills on the execution of this disaster recovery plan. The skills and knowledge gained by actually bringing up the registry's database and limited operations at a different location is very valuable. Almost certainly the first time this is done, it will cause the plan to be substantially modified, and thus improved.

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, or 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:

LOUISIANA TUMOR REGISTRY LAW, 2001 REVISION

PART XXVII. CANCER AND CARDIO-PULMONARY
DISEASES PROGRAMS

§1299.80. Definitions

As used in this Part:

(1) "President" shall mean the president of the Louisiana State University System or his designee.

(2) "Participating hospital" shall mean every hospital operating as such in the state of Louisiana.

(3) "Pathology laboratory" shall mean every pathology laboratory located or doing business in the state of Louisiana.

(4) "Office" shall mean the office of the president.

(5) "Board" shall mean the Louisiana Cancer and Lung Trust Fund Board.

(6) "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.

(7) "Radiation center" shall mean every freestanding radiation diagnostic and treatment facility in the state of Louisiana.

Added by Acts 1978, No. 653, §1. Amended by Acts 1982, No. 812, §1; Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1.

§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.

Added by Acts 1978, No. 653, §1. Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1.

§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, the Centers for Disease Control, and other national and international cancer surveillance programs designated by the Louisiana Tumor Registry in providing cancer 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 the 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.

Added by Acts 1978, No. 653, §1. Amended by Acts 1982, No. 812, §1; Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1; Acts 2003, No. 225, §1.

§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 this 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 carrying out the purposes of this Part.

(3) Repealed by Acts 1985, No. 345, §1, eff. July 9, 1985.

Added by Acts 1978, No. 653, §1. Amended by Acts 1982, No. 812, §1; Acts 1985, No. 345, §1, eff. July 9, 1985; Acts 2001, No. 197, §1.

§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 when 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.

Added by Acts 1978, No. 653, §1; Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1

§1299.85. 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 confidential and privileged. The president 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 president 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 president from publishing compilations relating to morbidity and mortality studies which do not identify case specific data or sources of information.

Added by Acts 1978, No. 653, §1; Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1.

§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.

Added by Acts 1978, No. 653, §1. Amended by Acts 1982, No. 812, §1; Acts 2001, No. 197, §1.

§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 nonaggregate, 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 they will keep nonaggregate, case-specific information confidential and privileged and 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.

Added by Acts 1978, No. 660, §2. Amended by Acts 1982, No. 812, §1; Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1.

§1299.88. Louisiana Cancer and Lung Trust Fund Board

A.(1) There is hereby created the Louisiana Cancer and Lung Trust Fund Board, which shall consist of the following members appointed and reappointed by the governor, to serve at his pleasure, upon recommendation of each institution and organization represented:

(a) A representative from Tulane University School of Medicine.

(b) A representative from the Louisiana State University School of Medicine, New Orleans.

(c) A representative from the Louisiana State University School of Medicine, Shreveport.

(d) A representative from the Alton Ochsner Medical Foundation.

(e) A representative of the American Cancer Society, Louisiana Division, Inc.

(f) A representative of the American Lung Association of Louisiana, Inc.

(g) A representative of the Leukemia Society of America, Inc., Louisiana Chapter.

(h) A representative of the Mary Bird Perkins Cancer Radiation and Research Foundation, Inc.

(i) A representative of the Xavier University School of Pharmacy.

(j) A representative of the Louisiana State Medical Society.

(k) A representative of the Acadiana Medical Research Foundation.

(l) A representative of the American Heart Association, Louisiana Affiliation.

(2) Each appointment by the governor shall be subject to Senate confirmation.

B. The board shall determine the eligibility of medical research programs and clinical investigation and training projects to receive funds; however, sufficient funds shall be allocated annually to the statewide registry program for reporting cancer cases under the provisions of R.S. 40:1299.80 et seq. Administration of funds shall be exercised by the office of the president.

C.(1) The board shall establish rules and regulations for its own procedures, establish policies for the operation of the statewide registry program for reporting cancer cases established under the provisions of R.S. 40:1299.80 et seq., establish criteria for review panels, and establish guidelines and deadlines for grant applications to be submitted. The appointment of review panels for the purpose of evaluating grant applications and making recommendations to the board on a priority basis shall be made before monies are allocated. Any member of the board or review panels with a direct conflict of interest shall excuse himself or herself from voting on any grant proposal.

(2) The board shall elect from among its members a chairman, a vice chairman, a secretary, and a treasurer. Any member may hold two of these positions. In the absence of the chairman, the vice chairman shall preside and in the absence of the chairman and vice chairman, the secretary shall preside.

(3) The members shall not receive compensation for their services but shall be entitled to reimbursement for expenses, including travel expenses, incurred in the discharge of their duties.

(4) Six members shall constitute a quorum for the transaction of business; however, no board action shall be taken by a vote of less than a majority of the full board.

(5) The secretary shall keep complete and accurate records of all meetings and actions taken by the board.

(6) The treasurer shall keep full and accurate financial records, make periodic reports to the board, and submit a complete annual report, in written form, to the secretary.

(7) Meetings of the board shall be held at regular intervals as provided in the bylaws. Emergency meetings may be held upon twenty-four hours actual notice and business transacted, provided that not less than a majority of the full board concurs in the proposed action.

D. For purposes of this Section, the following definitions shall apply:

(1) "Medical research" shall mean a program to determine the cause and prevention of disease.

(2) "Clinical investigation" shall mean the application of the results of medical research to treat patients.

(3) "Training" shall mean the educational preparation for a subspecialist career in cancer or lung disease.

E. A current report on the programs funded shall be made to the House Committee on Ways and Means and to the Senate Committee on Revenue and Fiscal Affairs, meeting jointly, prior to each regular session of the legislature.

F. Any member of the board or of a review panel, whether or not such member is compensated by the institution or organization he represents, shall recuse himself from participating in any discussion or voting regarding any matter relating to awarding a grant or contracting with the institution or organization he is appointed to represent. No member of the board or of a review panel who complies with the recusal provisions contained in this Subsection shall be deemed to have violated the Code of Governmental Ethics. The appointment of a compensated employee as a representative of a designated institution or organization shall not constitute a prohibited relationship under the provisions of the Code of Governmental Ethics.

Added by Acts 1980, No. 825, §1, eff. Aug. 1, 1980. Amended by Acts 1982, No. 812, §1; Acts 1984, 1st Ex. Sess., No. 14, §1, eff. July 1, 1984; Acts 1985, No. 929, §1; Acts 1989, No. 355, §1; Acts 1992, No. 227, §1; Acts 1992, No. 984, §12; Acts 1993, No. 1004, §1; Acts 2001, No. 197, §1.

§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.

Added by Acts 1983, No. 711, §1. Amended by Acts 1995, No. 1197, §1, eff. June 29, 1995; Acts 2001, No. 197, §1.

§1299.90. Annual lung cancer report

A. The Louisiana Cancer and Lung Trust Fund Board shall annually publish a comprehensive report on the incidence of lung cancer in Louisiana and the progress made in reducing or eliminating the high lung cancer rates in Louisiana. The report shall place special emphasis on the lung cancer rate in the southern portion of the state.

B. The report shall be submitted 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.

Acts 1984, No. 468, §1.

APPENDIX C:

CANCER REGISTRIES AMENDMENT ACT

The United States 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

Entitled the "Cancer Registries Amendment Act."

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

106 STAT. 3372   **PUBLIC LAW 102–515—OCT. 24, 1992**

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

106 STAT. 3372   **PUBLIC LAW 102–515—OCT. 24, 1992**

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;

106 STAT. 3372   **PUBLIC LAW 102–515—OCT. 24, 1992**

"(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

106 STAT. 3372   **PUBLIC LAW 102–515—OCT. 24, 1992**

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.

106 STAT. 3372   **PUBLIC LAW 102–515—OCT. 24, 1992**

"(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 United States 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.

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 "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.".

(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.

Grants.
42 USC 280e note.

Approved October 29, 2002.

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). The contact information for standard-setting organizations is listed below:

American College of Surgeons (ACoS)
633 N. Saint Clair Street
Chicago, IL 60611-3211
Telephone: (312) 202-5000
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: www.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
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
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
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
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, Suite B
Springfield, IL 62704-6495
Telephone: (217) 698-0800
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"):

$$C_{sk} = \frac{Observed\ T_{sk}}{Expected\ T_{sk}}$$

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:

$$Expected\ I_{skij} = (M_{skij})\frac{I_{SEERkij}}{M_{U.S.kij}}$$

$$Expected\ T_{sk} = \sum_{i=1}^{2}\sum_{j=1}^{N} Expected\ I_{skij}$$

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. NAACCR Method to Estimate Completeness. 2006-2007 NAACCR Webinar Series; March 15 2006. Available: www.naaccr.org/index.asp?Col_SectionKey=10&Col_ContentID=42.

APPENDIX H:

MAJOR-MINOR DISCREPANCY DEFINITIONS FOR COLON

Major-Minor Collaborative Stage Discrepancies for Colon

Data Item	Major	Minor	Unk to Known	Known to Unk
CS Tumor Size	000 [T0] to (001-995) and vice versa 998 to (001-995) and vice versa 999 to 000 and vice versa	Any within (001-995)	999 to 000-998	000-998 to 999
CS Extension	(00-11 [Tis]) to (12-80 [T1-T4]) and vice versa 10-16 to 20 or 40-46 or 50-80 and vice versa (12-16 or 30 [T1]) to 20 [T2] and vice versa (12-16 or 30 [T1]) to (40-46 [T3]) and vice versa (12-16 or 30 [T1]) to (50-80 [T4]) and vice versa 20 to (40-46) and vice versa 20 to (50-80) and vice versa (40-46) to (50-80) and vice versa 30 to 20 or 40-46 or 50-80 and vice versa 95 to 00-80 and vice versa 99 to (00-95) and vice versa	Any within 00-11 Any within 12-16 Any within 40-46 Any within 50-80	99 to 00-95	00-95 to 99
CS TS/Ext Eval	((0, 1, 5) to (2, 3, 6) and vice versa (2, 3) to 6 and vice versa 0-6 to 8 [a] and vice versa 9 to (2, 3, 6, 8) and vice versa	0 to (1, 5) and vice versa 1 to 5 and vice versa 2 to 3 and vice versa	9 to (0-8)	(0-8) to 9
CS Lymph Nodes	00 [N0] to (10-80 [N1-2]) and vice versa 03 [invalid] 05 [invalid] 90 [invalid] 98 [invalid] 99 to (00-80) and vice versa	Any within (10, 20, 30) (10, 20, 30) to 80 and vice versa	99 to 00-80	00-80 to 99
CS Nodes Eval	(0, 1, 5) to (2, 3, 6) and vice versa (2, 3) to 6 and vice versa 0-6 to 8 [a] and vice versa 9 to (2, 3, 6, 8) and vice versa	0 to (1, 5) and vice versa 1 to 5 and vice versa 2 to 3 and vice versa	9 to 0-8	0-8 to 9
Reg Nodes Pos	00 [neg] to 01-97 [pos] and vice versa 98 to 00-97 and vice versa (01-03 [N1]) to (04-90 [N2]) and vice versa 91-94 [invalid] 99 to (00-98) and vice versa	Any within 01-03 Any within 04-90 Any within 95-97 (95 or 97) to any within (01-90) and vice versa	99 to 00-98	00-98 to 99

Data Item	Major	Minor	Unk to Known	Known to Unk
Reg Nodes Examined	00 to any of (01-90, 95-98) and vice versa (96, 97, 98) to any of (01-90 or 95) and vice versa (01-90) to 95 and vice versa 91-94 [invalid] 99 to (00-95) and vice versa	Any within 01-90 Any within (96-98)	99 to 00-98	00-98 to 99
CS Mets at Dx	00 [M0] to (08, 10, 40, 50) [M1] and vice versa 04 [invalid] 51 [invalid] 99 to (00-50) and vice versa	Any within (08, 10, 40, 50)	99 to any of (00-50)	Any of (00-50) to 99
CS Mets Eval	(0, 1, 5) to (2, 3, 6) and vice versa (2, 3) to 6 and vice versa 0-6 to 8 [a] and vice versa 9 to (2, 3, 6, 8) and vice versa	0 to (1, 5) and vice versa 1 to 5 and vice versa 2 to 3 and vice versa	9 to any of (0-8)	(0-8) to 9
SSF1: CEA	000 [not done] to (010, 020, 030, 080 [done]) and vice versa 010 [abn] to 020 [normal] and vice versa 888 [invalid] 999 to (010, 020, 030) and vice versa	010 to 030 and vice versa 020 to 030 and vice versa	999 to (00-080) 080 to (010-030)	(000-080) to 999 (000-030) to 080
SSF2: none				
SSF3: none				
SSF4: none				
SSF5: none				
SSF6: none				

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 reanalysis 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:

DATA USE AGREEMENTS

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"), a California corporation. Researcher is engaged in research into the causes, control, or prevention of cancer, specifically described as follows:

_____

NAACCR 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 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 that has been accepted, in writing, by NAACCR. 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.

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 agrees that linkage to another database is not permitted for the purpose of identifying an individual on the file, but may be permitted if appropriate linkage is described in the proposal and this linkage is approved by the NAACCR IRB.

7. Researcher further agrees that all data provided under the provisions of this Confidentiality Agreement shall remain the sole property of NAACCR and may not be copied or reproduced in any form or manner without NAACCR's prior written consent.

8. Researcher shall indemnify NAACCR 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 in connection with any such failure, Researcher agrees that NAACCR may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR, at the expense of Researcher. NAACCR, 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.

9. Researcher will not take any action that will provide any Data furnished by NAACCR to any unauthorized individual or agency without the prior written consent of NAACCR.

10. 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. Also, Researcher will not provide any computer password or file access codes that protect the Data to any unauthorized person.

11. 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.

12. In the event that any attempt is made to obtain from Researcher any or all of the Data provided to Researcher by NAACCR by subpoena or other legal means, Researcher will notify NAACCR immediately. Researcher agrees that NAACCR may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR. NAACCR, 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.

13. Researcher's obligations hereunder shall remain in full force and effect and survive the completion of Researcher's research project described hereinabove.

14. The terms of this Confidentiality Agreement shall be binding upon Researcher, his/her agents, assistants and employees.

15. Notwithstanding any contrary language in this Confidentiality Agreement, Researcher acknowledges and agrees that Researcher's access to the Data maintained by NAACCR shall at all times be in the sole discretion of NAACCR.

16. NAACCR 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.

17.   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.

18.   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.

19.   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: _____

Received and accepted this _____day of _____, 200_____.

North American Association of Central Cancer Registries, Inc.
By: _____
Its: _____

<u>**SEER DATA USE AGREEMENT**</u>

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. Chicago: American College of Surgeons Commission on Cancer; 2004.

American College of Surgeons Commission on Cancer. Facility Oncology Registry Data Standards: Revised for 2007 (FORDS). Chicago: American College of Surgeons Commission on Cancer; 2002 (updated 2007).

American Joint Committee on Cancer. AJCC Cancer Staging Manual. Green F, et al., editors. Sixth Edition. New York: Springer-Verlag; 2002. (See also: Editions 1, 2, 3, and 4, which were published by Lippincott-Raven under the title Manual for Staging of Cancer; and Edition 5, published by Lippincott-Raven under the title of AJCC Cancer Staging Manual.)

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.

Benign Brain Tumor Cancer Registries Amendment Act, Pub. L. No. 107-260, 116 Stat. 1743; Oct. 29, 2002.

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.

HIPAA Administrative Simplification. 45 CFR Part 164, Subpart C – Security Standards for the Protection of Electronic Protected Health Information. www.hhs.gov/ocr/AdminSimpRegText.pdf. § 164.306 Security standards describes general rules and general requirements.

HIPAA Security Rule. www.cms.hhs.gov/SecurityStandard/Downloads/securityfinalrule.pdf. (HIPAA Administrative Simplification, www.hhs.gov/ocr/AdminSimpRegText.pdf.)

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.

Johnson, CH, Adamo, M, editors. SEER Program Coding and Staging Manual 2007. Bethesda, MD: National Institutes of Health, National Cancer Institute. NIH Pub. No.07-5581;2007.

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, Deapen D, Phillips JL, Tucker TC, editors. Central Cancer Registries: Design, Management and Use. Second Edition. Dubuque, IA: Kendall/Hunt Publishing Company; 2007.

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. Pathology Laboratory Electronic Reporting, Version 2.1. Havener L, editor. Standards for Cancer Registries Volume V. Springfield, IL: North American Association of Central Cancer Registries; September 2007.

NAACCR. Data Exchange Standards and Record Description. Version 11.2. In: Havener L, editor. Standards for Cancer Registries. Volume I. Springfield, IL: North American Association of Central Cancer Registries; June 2007.

NAACCR. Data Standards and Data Dictionary. Standards for Cancer Registries. Volume II. Version 11.2. In: Hofferkamp J, Havener L, editors. Standards for Cancer Registries Volume II. Springfield, IL; North American Association of Central Cancer Registries; April 2007.

NAACCR. NAACCR Method to Estimate Completeness. 2006-2007 NAACCR Webinar Series; March 15, 2006. Holly L. Howe. Available: www.naaccr.org/index.asp?Col_SectionKey=10&Col_ContentID=42.

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. Standard Data Edits. In: Seiffert JE, Capron S, and 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 (accessed 3/20/07).)

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.

Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule www.privacyruleandresearch.nih.gov/pr_02.asp.

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.

Title 45 of the Code of Federal Regulations, Part 46 at www.ohrp.osophs.dhhs.gov/humansubjects/guidance/45cfr46.htm.

Title 21 of the Code of Federal Regulations, Part 50 at www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/showCFR.cfm?CFRPart=50&showFR=1, Part 56 at www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/showCFR.cfm?CFRPart=56&showFR=1. Additional requirements are found in Title 21 of the Code of Federal Regulations, Part 312 at www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=312&showFR=1, and Part 812 at www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=812&showFR=1.

Title 42 of the Code of Federal Regulations, Part 2 at www.access.gpo.gov/nara/cfr/waisidx_02/42cfr2_02.html.

WHO. ICD-10: International Statistical Classification of Disease and Related Health problems. Tenth Revision, Second Edition Volume 1 [of 3]. Geneva: World Health Organization; 2004.

INDEX

Abstractors, 84
Accuracy, 88, 177
ACoS. *See* American College of Surgeons
AJCC. *See* American Joint Committee on Cancer
Ambiguous terminology, 7
American Cancer Society, i, 14, 16, 170
American College of Surgeons, i, 5, 11, 26, 29, 32, 76, 159, 175
American Joint Committee on Cancer, i, 69, 159, 175
Audits, casefinding, 21, 27
Audits, recoding, 38, 40
Autonomy, 89
Autopsy, 112
Canadian Association of Provincial Cancer Agencies, i
Canadian Cancer Registry, 3, 8, 9, 15
Canadian Cancer Society, 3
Canadian Council of Cancer Registries, 3, 32, 159, 175
Canadian Health Information Management Association, 12
Canadian Partnership Against Cancer, i
Cancer Registries Amendment Act, 4, 147, 148, 155, 156, 175
CANCERLIT, 11
Casefinding, 19, 21, 27
Caseload, estimating, 10
Cause of death, 110
CCCR. See Canadian Council of Cancer Registries
CCR. *See* Canadian Cancer Registry
CDC. *See* Centers for Disease Control and Prevention
Centers for Disease Control and Prevention, 4, 159
Central cancer registries, 1, 21, 41, 68, 76, 127, 129, 176
Certified Tumor Registrar, 11, 20, 32, 37
Class of case, 7
Clinically diagnosed, 9
CoC. *See* Commission on Cancer
Codes, xv, 36, 69
College of American Pathologists, i
Commission on Cancer, 5, 9, 29, 37, 69, 159
Community Advisory Boards, 14
Completeness, iii, 10, 20, 43, 78, 104, 166
Compliance, 20
Computer system, 36
Computerized data collection, 96
Confidential data, 122
Confidential data, definition, 118, 126
Confidentiality, 6, 80, 81, 88, 114, 117, 125, 126, 128, 162
Consolidation, 35, 36, 85, 108, 109, 176
Continuing education, 11, 47, 96
Correction record, 111

Counts, 20, 24, 70
County, 4, 43
CTR. *See* Certified Tumor Registrar
Data collection, 10, 11, 31, 42
Data edits, 33
Data entry, 96
Data Exchange Standards and Record Description, iii
Data quality, 41, 78
Data standards, 176
Database management system, 86, 92, 93, 94
Datasets, 34, 35, 112
DBMS. *See* Database management system
DCO. *See* Death certificate only
Death certificate only, 19, 23, 24, 25, 26, 86, 103, 105
Death clearance, 17, 18, 100, 176
Death rates, 74
Deletion, 98, 100, 102
Denominators, 72
Diagnostic confirmation, 7
Documentation, 31, 32, 34, 37, 78, 114, 115
EDITS, 35, 36, 84, 108
Edits, computerized, 34
Expected, 24, 71, 164
Federal facilities, 12
Follow-up, 21, 22, 29, 74, 97, 98, 99, 112, 113, 114
Funding, 5, 7, 89
Geocoding, 101, 102
Graphs, 79
Health Insurance Portability and Accountability Act, 6, 118, 119, 123, 126, 175, 177
Health Maintenance Organization, 10, 22
HIPAA. *See* Health Insurance Portability and Accountability Act
Histology, 50
HMO. *See* Health Maintenance Organization
IACR. *See* International Association of Cancer Registries
IARC. *See* International Agency for Research on Cancer
ICD, 4, 5, 7, 8, 37, 48, 49, 50, 54, 62, 65, 66, 67, 78, 157, 177
ICD-O, 4, 5, 7, 8, 37, 48, 49, 50, 54, 78, 157
ICD-O-2, 48, 49, 54
ICD-O-3, 8, 48, 50, 78
In situ, 4, 8, 17, 21, 29, 66, 72, 105
Incidence rates, 25, 71
Institutional Review Board, 7, 128, 129, 172
International Agency for Research on Cancer, 8, 72, 175, 176
International Association of Cancer Registries, 8, 175
IRB. *See* Institutional Review Board