AMERICAN ASSOCIATION OF CENTRAL CANCER REGISTRIES AACCR

Standards for Cancer Registries
Volume II

• • •

Data Standards and Data Dictionary

Edited By
Herman R. Menck and Jennifer E. Seiffert

February 14, 1994

Sponsoring Organizations

American Association of Cancer Institutes
American Cancer Society
American College of Surgeons
Association of Community Cancer Centers
Centers for Disease Control and Prevention
National Cancer Institute
National Cancer Registrars Association
Statistics Canada

AMERICAN ASSOCIATION OF CENTRAL CANCER REGISTRIES AACCR

Standards for Cancer Registries
Volume II

• • •

Data Standards and Data Dictionary

Edited By
Herman R. Menck and Jennifer E. Seiffert

February 14, 1994

Sponsoring Organizations

American Association of Cancer Institutes
American Cancer Society
American College of Surgeons
Association of Community Cancer Centers
Centers for Disease Control and Prevention
National Cancer Institute
National Cancer Registrars Association
Statistics Canada

Edited By

Herman R. Menck, MBA, CDP
American College of Surgeons
55 East Erie St.
Chicago, IL 60611

Jennifer E. Seiffert, MLIS, CTR
Standards and Technical Assistance
AACCR Cancer Surveillance and Control Program
601 North 7th St. / MS 592
P.O. Box 942732
Sacramento, CA 94234-7320

Comments and suggestion on this and other AACCR standards documents are welcome. Please send your comments to the editors or any member of the AACCR Executive Board.

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

Volume I, *Data Exchange Standards and Record Description*. Intended for programmers, this provides the record layout and specifications for the standard for data exchange.

Volume III, *Standards for Completeness, Quality, Analysis, and Management of Data*. Intended for central registries, this provides detailed standards for many aspects of the operation of a population-based cancer registry.

For additional copies, write to the AACCR Cancer Surveillance and Control Program at the above address.

Support for production and distribution of these standards was provided through cooperative agreement U50/CCU906484 between the Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, and the American Association of Central Cancer Registries, administered by the California Public Health Foundation.

These standards were adopted by the AACCR Executive Board on February 9, 1994.

CONTENTS

AACCR EXECUTIVE BOARD
1993 - 1994

President

John L. Young, DrPH, CTR
Chief, Cancer Surveillance Section
California Cancer Registry
California Dept. of Health Services
601 North 7th St. / MS592
P.O. Box 942732
Sacramento, CA 94234-7320

Executive Secretary

Brenda K. Edwards, PhD
Associate Director
Surveillance Program, DCPC
National Cancer Institute
EPN 343J
9000 Rockville Pike
Bethesda, MD 20892

Treasurer

Sue A. Watkins, ART, CTR
Director
Tri-Counties Regional Cancer Registry
Santa Barbara County Health Care Svcs.
300 North San Antonio Rd.
Santa Barbara, CA 93110

Immediate Past President

William McWhorter, MD
Utah Cancer Registry
Salt Lake City, Utah

Representatives-at-Large

Rosemary Dibble, CTR
Administrator
Utah Cancer Registry
420 Chipeta Way, Suite 190
Salt Lake City, UT 84108

Gilbert Friedell, MD
Director
Lucille Parker Markey Cancer Center
306 MRISC Building
800 Rose St.
Lexington, KY 40536-0098

Leslie Gaudette, BSc
Senior Analyst
Canadian Cancer Registry
Health Status Section
Canadian Centre for Health Information
Statistics Canada
18-A, R.H. Coats Building
Ottawa, Ontario K1A OT6
Canada

Kathleen M. McKeen
Director
State Health Registry of Iowa
100 Westlawn Bldg. South
Iowa City, IA 52242

Dee W. West, PhD
Executive Director
Northern California Cancer Center
32960 Alvarado Niles Rd., Suite 600
Union City, CA 94587-3106

MEMBERS OF DATA STANDARDS DOCUMENT SUBCOMMITTEE

Chairs

Herman R. Menck, MBA, CDP
Chair, Data Standards Document Subcommittee
American College of Surgeons

Thomas Tucker, MPH
Chair, AACCR Uniform Data Standards Committee
Kentucky Cancer Registry

Subcommittee Members

Catherine C. Boring, MS, MPH
American Cancer Society

Mary Crist, CTR
American College of Surgeons

Larry Derrick
Utah Cancer Registry

Rosemary L. Dibble, CTR
Utah Cancer Registry

Barry Gordon, PhD
California Cancer Registry and C/NET

Suzanna S. Hoyler, CTR
Washington Hospital Center

Carol Johnson, CTR
American College of Surgeons

William G. Kraybill, MD
Washington University School of Medicine

Mary L. Lerchen, DrPH
Centers for Disease Control and Prevention

Jerri Linn Phillips, MA, CTR
Wisconsin Cancer Reporting System

Lynn A. Gloeckler Ries, MS
SEER Program, National Cancer Institute

Jennifer E. Seiffert, MLIS, CTR
AACCR Cancer Surveillance and Control Program

J. Mark Shields, MD, MPH, MS, CTR
Centers for Disease Control and Prevention

Jan Snodgrass, CTR
Illinois State Cancer Registry

John Young, DrPH
California Cancer Registry

Kathleen E. Zuber-Ocwieja, ART, CTR
American College of Surgeons

ADDRESSES OF MAJOR STANDARD-SETTING ORGANIZATIONS

American Association of Central Cancer Registries
Cancer Surveillance and Control Program
Standards and Technical Assistance
601 N. 7th St. / MS592
P.O. Box 942732
Sacramento, CA 94234-7320
(916) 327-4663
FAX (916) 327-4657

American Cancer Society
Department of Epidemiology and Statistics
1599 Clifton Road, N.E.
Atlanta,GA 30329
(404) 329-7658
FAX (404) 325-1467

American College of Surgeons
Commission on Cancer
55 East Erie Street
Chicago, IL 60611
(312) 664-4050
FAX (312) 440-7144

American Joint Committee on Cancer
Executive Office
55 East Erie Street
Chicago, IL 60611

Centers for Disease Control and Prevention
Division of Cancer Prevention and Control
MS K55
NCCDPHP
4770 Buford Highway NE
Atlanta, GA 30341-3724
(404) 488-4795
FAX (404) 488-4727

National Cancer Institute
SEER Program
Executive Plaza North, 343J
9000 Rockville Pike
Bethesda, MD 20892
(301) 496-8510
FAX (301) 402-0816

National Cancer Registrars Association
505 East Hawley Street
Mundelein, IL 60060
(708) 566-0833
FAX (708) 566-7282

PREFACE

Publication of AACCR's data standards is a landmark for the organization and for efforts at standardization of cancer registry data in North America. This volume documents the achievements of a process of collaboration, consensus-building, and compromise among all of the major organizations involved in setting standards for cancer registries, including AACCR, the American College of Surgeons, the Centers for Disease Control and Prevention, the National Cancer Institute, and the National Cancer Registrars Association. It represents the status of agreed-upon data items and codes as of January 1994, and highlights areas where more work is needed. This volume was drafted by a subcommittee of AACCR's Uniform Data Standards Committee during 1992 and 1993 and was adopted by the AACCR Executive Board in February 1994.

It is our hope that making these consensus standards available to a wide audience will make it easier for individual hospital and central registries and software providers to adopt them and will raise awareness of the value of standardization and the costs involved in lack of standardization. Ultimately, our goal is to improve the quality, comparability, and usefulness of cancer registry data. AACCR is involved in a variety of standard-setting activities, and this volume is the second in a series of three that document standards in several areas (see chapter I).

On behalf of the AACCR Executive Board, I want to extend sincere thanks to the individuals (listed on page v) who worked for many hours over two years to pull this material together. Herman Menck, of the National Cancer Data Base (NCDB), deserves AACCR's special thanks for having the vision to propose this document and then agreeing to chair the subcommittee to prepare it. NCDB provided financial support for several subcommittee meetings, without which the document would not have been completed. I would also like to recognize Jennifer Seiffert, who chaired the Uniform Data Standards Committee during most of the time this document was being prepared. Her leadership, organization, and attention to detail were crucial to the document's final form.

John L. Young, DrPH
President, AACCR

ABBREVIATIONS USED

AACCR	American Association of Central Cancer Registries
ACoS	American College of Surgeons
ACS	American Cancer Society
AJCC	American Joint Committee on Cancer
CDC	Centers for Disease Control and Prevention
COC	Commission on Cancer (of the American College of Surgeons)
CTR	Certified Tumor Registrar
DAM	*Data Acquisition Manual* (of the American College of Surgeons)
EDITS	Exchangeable-edits, Data-dictionary, and Information Translation Standard
EOD	Extent of Disease
FTRO	*Fundamental Tumor Registry Operations Program* (of the American College of Surgeons)
IACR	International Association of Cancer Registries
IARC	International Agency for Research on Cancer
ICD	International Classification of Diseases
ICD-O, ICD-O-1, ICD-O-2	*International Classification of Diseases for Oncology,* and the 1st and 2nd editions, respectively
NCDB	National Cancer Data Base
NCI	National Cancer Institute
NCRA	National Cancer Registrars Association
N.d.	No date (bibliographic term: no ascertainable date of publication)
NTRA	National Tumor Registrars Association, former name of NCRA
NOS	Not Otherwise Specified
N.p.	No place (bibliographic term: no ascertainable place of publication)
SEER	Surveillance, Epidemiology, and End Results (Program of the National Cancer Institute)
TNM	Tumor, Nodes, and Metastasis: staging system of AJCC and UICC (see reference 5)
UICC	Union Internationale Contre le Cancer (in English, International Union Against Cancer)
WHO	World Health Organization

• • •

Chapter I

PROBLEM STATEMENT, GOALS, AND SCOPE OF THIS DOCUMENT

The Problem

In the last decade, increased efforts to pool data collected by different cancer registries for different purposes have drawn attention to problems encountered as a result of insufficient data standardization. Three examples follow:

- Electronic submission of hospital registry data to state or other central registries. Central registries recognized that data quality and efficiency of collection could be improved with reporting of data via diskette or modem. Many set up systems for receiving data from multiple sources. Often these data were collected using different software, different data sets, different codes, and different coding rules. Central registries experienced a new frustration of mapping the submissions into their own data systems, while software providers were frustrated at having to prepare submissions for multiple state registries which were yet again different from each other and followed different models of collecting data electronically.

- American Association of Central Cancer Registries (AACCR) Data Publication and Evaluation Committee activities. AACCR undertook to receive statistical analysis files from its member registries in the standard AACCR data exchange record layout (reference 1) in order to prepare descriptive epidemiological data about the participating areas. However, data sets of the participants differed; the original codes used, format of the data, edits applied, and coding rules differed; and a significant amount of work has been required to produce comparable summary statistics.

- National Cancer Data Base (NCDB). This is a joint project of the Commission on Cancer (COC) of the American College of Surgeons (ACoS) and the American Cancer Society (ACS) which pools data submitted voluntarily by participating hospitals to address questions of clinical interest. Data are submitted in the AACCR data exchange record layout. Once again, problems have been discovered in codes, format, and data sets requiring effort and interpretation before the data can be successfully pooled.

Examples of the simplest problems encountered in pooling data include the lack of standardization regarding the use of blanks in fields, and the inconsistent use of blanks,

dashes, and 9s to code "unknown" data. More substantial discrepancies no doubt are present as well, but they are less easy to detect and correct.

Data items used by different individual registries or software systems vary in their definition and codes, even when they have the same name and attempt to represent the same information. Differences in some of the data elements have made comparison or exchange of data difficult if not impossible. Hospitals are sometimes caught between conflicting standards when they are both reporting to a central registry using one set of standards and attempting to maintain a database consistent with a different set of ACoS standards.

It has become clear throughout the cancer registry community that lack of standardization has a substantial cost, and it limits more widespread use of valuable data.

The Proposed Solution

Many of AACCR's sponsoring organizations, including the National Cancer Institute (NCI), the Centers for Disease Control and Prevention (CDC), and the American College of Surgeons, have recognized that increasing standardization of the way registries collect their data is an essential step in decreasing costs of data collection; more efficiently using of increasingly-limited human resources needed for data collection, management and analysis; and obtaining more useful data which can be compared across registries and geographic areas.

Preparation of a statement of consensus on data standards for cancer registries was proposed by NCDB and the AACCR Data Exchange Committee, resulting in AACCR appointing a subcommittee of its Uniform Data Standards Committee to prepare one. At the same time, CDC entered into a cooperative agreement with AACCR, and one of the projects to be accomplished under that agreement was preparation of broader standards for population-based cancer registries. The two efforts thus became complementary, with the intent of producing separate but related documents that together would specify AACCR-agreed-upon standards. The result of these efforts is the present series of three standards documents.

Goals for the Data Standards Document

The goal of this document is to define the data standards for cancer registration as used by central registries, hospital based registries, and other groups.

Objectives of the standardization effort include:

- To provide a resource to help ensure uniform data collection
- To reduce the need for redundant coding and data recoding between agencies
- To facilitate collection of comparable data among groups

- To provide a resource document to help registries which are establishing or revising their databases
- To serve as the single reference for data standards for cancer registration
- To encourage the adoption of these standards by all parties

This document will be used by new and existing hospital and central cancer registries wishing to assure that the definitions and codes used within their programs are standard and consistent with those used by regional and national data bases. Other potential users include registry software providers and those using registry data, especially if they are combining data from multiple sources or exchanging data. National standard-setting groups, such as AACCR, COC, CDC, ACS, and NCI will also benefit.

The AACCR has previously published the *AACCR National Standard for Cancer Data Exchange: Record Description* which contains its data exchange record layout. Version 3.0 of this standard has been incorporated into AACCR's standards series as volume I, and retitled *Data Exchange Standards and Record Description* (reference 1). The present document uses the same structure and philosophy as the data exchange standards. Where there is an existing standard for an item or type of data, the standard is incorporated by reference. Where a variety of standards are in use, there is provision for alternate coding schemes, but the different items are kept separate or the data are tagged internally to indicate the coding standard used.

The AACCR data exchange standard incorporates three record types which are combinations of standard components, such as demographic information, patient confidential information, and text. Thus the different purposes and constraints of data exchange can be accommodated without the requirement for separate program formats.

Scope of This Document: What We Mean by Standards

A variety of types of standards for cancer registries can be specified. There are standards that apply to the data collected and other standards that can be established for other aspects of registry process, such as standards for death clearance procedures, follow-up methods, or quality control. Yet another standard would address the completeness of coverage of a population-based central registry, and still another the qualifications and adequacy of staffing.

The scope of the present document is limited to standards regarding data rather than procedures, and is in fact further limited to a subset of possible data standards that AACCR considers important to establish.

Data standards can be categorized in the following types:

- Reportability. This type of standard specifies the rules for which cases are to be included in the registry. See chapter III.

- Data Items or Elements to be Included. This type of standard consists of a list of required or recommended data items which a registry should collect and include in its data base. Chapter VIII contains standards for data items.

  Example: "Sex" is a standard data element on the list in chapter VIII.

- Record Layout/Data Exchange. This type of standard identifies the position of the data item in a standard record, which in this document is the AACCR data exchange record layout (reference 1). These positions are indicated in chapter VIII.

  Example: "Sex" is in character position 66 in the AACCR data exchange record layout.

- Codes. This type of standard identifies allowable values, their meanings, and data-entry formats for data items. Chapter IX lists these standards for each data item.

  Example for the item "Sex":

  Allowable Values and Format:
  1-4, 9

  Codes:

1	*Male*
2	*Female*
3	*Other (Hermaphrodite)*
4	*Transsexual*
9	*Not Stated*

When it is necessary to collect more specific information than that represented by the standard codes, every effort should be made to ensure that the more specific codes will accurately collapse into the categories represented by the standard codes. This approach permits diversity without compromising inter-registry comparability or metanalyses.

- Coding Rules. This type of standard provides the rules and rule interpretations for deciding which code is correct for a given case. This document does not contain coding rules, but points to them as they exist in the documentation of other standard-setting organizations. For each data item in chapter IX, this document lists a "Source of Standard", and the documentation of this source should be consulted for coding rule standards.

Hypothetical Example: a coding rule might state what code to assign when the medical record states the patient is female and the death certificate states male.

- Data Edits. Standards for data edits specify allowable values and format and also the necessary relationships among data items. This document does not provide detailed item-specific data edit standards, but does provide a discussion of their importance and work in process toward standardization. See chapters IV and V.

  Example: When "Sex" is code 1 (male), "Primary Site" cannot be C56.9 (ovary).

- Standardized Groupings for Analysis. Many registry data items are assigned more numerous codes than are required for routine analyses. For example, most analyses using "Age" will group the data into 5- or 10-year age groups. Use of standardized analysis groupings facilitates comparisons of data. Items for which standardized groupings are crucial include "Primary Site", "Race", "Age", and treatment items.

  AACCR's standards for analysis groupings are included in volume III of this series, *Standards for Completeness, Quality, Analysis, and Management of Data* (reference 20).

Volume III of AACCR's standards (reference 20) addresses the following categories of standards:

- Access to Source Data and Completeness of Reporting. These address reportable laws and regulations, confidentiality policies, non-hospital reporting, physician reporting, death clearance procedures, and the like.

- Data Quality. These include, for example, standards for number and qualifications of staff, training, instruction manuals, adherence to established data standards, and quality control activities and results.

- Data Analysis and Reporting. Examples of analysis standards included are specification of standard analysis categories and standards for timeliness of data publication.

- Data Management. These address the capabilities required of a registry data management system to perform necessary functions efficiently.

• • •

Chapter II

HISTORICAL BACKGROUND AND STATUS OF U.S. STANDARDS

Standard-Setting Organizations and Other Standards Documents

Several organizations have played a major role in the development of data standards.

- American College of Surgeons. Since the 1950s the American College of Surgeons has had the leading role in establishing standards for hospital-based cancer programs and the cancer registries that are a part of such programs. Through the Commission on Cancer's (COC) Approvals Program, they have recommended data sets and codes for items to be collected, and have published the standard references used by hospital registries. Their publications include:

  - *Cancer Program Manual* (reference 25) which stipulates which cases are to be included in the registry and which data items are to be collected. It defines a core data set of required items, a fuller recommended data set, and an expanded data set.

  - *Data Acquisition Manual* (DAM) (reference 3) which presents standard codes and coding rules for the items in the data sets.

  - *Fundamental Tumor Registry Operations Program* (FTRO) (reference 28), which is a complete basic educational program for cancer registrars, contained in 15 separate modules.

  Up to this point, the College has recommended but has not required cancer programs to use the codes listed in the *Data Acquisition Manual*.

- SEER Program. The NCI's SEER Program has collected standardized data from a number of population-based registries since 1973, covering about 10% of the U.S. population. They require that their registries submit data in a standard format using standardized definitions and codes. They apply standardized edits to data submissions and supply the participating registries with complete sets of edits and documentation. The SEER staff have performed frequent audits to verify that the standards have been followed. They publish their own code manual (reference 2) and extent of disease manuals (references 6, 7, and 8), and these are widely used outside the SEER Program as well. However, the individual

SEER registries have not used standardized data collection methods or data management methods locally, and they differ in the extent to which they impose data requirements on the reporting facilities in their areas.

The SEER Program has also produced and distributed standards documents for the registry community as a whole, and these are in widespread use throughout hospital-based and central registries. For example, SEER has produced:

- *Summary Staging Guide for the Cancer Surveillance, Epidemiology and End Results Reporting (SEER) Program* (reference 9), the existing standard for localized-regional-distant staging.

- A series of eight self-instructional manuals for cancer registrars (reference 23) covering many aspects of abstracting, coding, and analyzing registry data. Book 8 in the series is a comprehensive lists of drugs used in treating cancer and is the standard reference for drug-treatment coding rules.

- Field trial editions of the *International Classification of Diseases for Oncology* (references 13, 14, and 15).

- Conversion schemes that have facilitated conversion of topography and morphology codes in a consistent way for 20 years.

- World Health Organization. WHO publishes the *International Classification of Diseases* (ICD) (references 10, 17, and 18) and the *International Classification of Diseases for Oncology* (ICD-O) (references 12 and 16), both world-standard diagnosis coding systems.

- American Joint Committee on Cancer. AJCC produces the *Manual for Staging of Cancer* (reference 5), the U.S. standard for TNM staging. (See chapter VI for a discussion of coding stage.)

- National Cancer Registrars Association (NCRA). (Formerly the National Tumor Registrars Association.) NCRA has been instrumental since its founding in 1974 in the training and certification of staff for cancer registries. The first Certification Exam for Tumor Registrars was offered in 1983. Nearly 1,500 Certified Tumor Registrars (CTRs) have successfully completed the exam which evaluates technical knowledge of methods of cancer data collection, management, and quality control, as well as ICD-O topography and morphology coding and AJCC and SEER staging systems.

  NCRA has produced guidelines for a college curriculum in cancer registry management, an extensive survey of hospital registry staffing, and several short course workshops to encourage high standards of data collection. In addition, NCRA's annual conferences are designed for continuing education of both individuals new to registry work and those with many years experience.

- American Association of Central Cancer Registries. AACCR is an organization of organizations established in 1989, the members of which are cancer registries and related organizations, including ACoS, ACS, CDC, NCI, and NCRA. AACCR provides a forum for registries in the U.S. and Canada to discuss common problems and share data. It has several active committees which set registry standards. The Data Exchange Committee has produced and maintains the standard for registry data exchange (reference 1). The Uniform Data Standards Committee recommends standard data items, codes, and formats.

- Centers for Disease Control and Prevention. CDC has contributed substantial resources to the further development of data standards, both through their cooperative agreement with AACCR and their funding and development of the EDITS project, a software system that facilitates coordination of data standards. (See chapter V.)

- American Cancer Society (ACS). The ACS has historically supported the development of standardized cancer classification systems and published the first code manual for the morphology of neoplasms in 1951. The ACS supports programs of the American College of Surgeons including the Fundamental Tumor Registry Operations education program. The ACS participates as a sponsoring organization in the American Joint Committee on Cancer, and as such is committed to the ongoing development, publication and promotion of uniform staging classifications.

Full bibliographic citations for the above standards documents are given in chapter VII.

Historical Background of Standards Coordination

In the 1980s the Commission on Cancer and the SEER Program of NCI began working together to make the codes and definitions in their manuals consistent. At that time, hospital-based cancer registries often used the Commission's codes and coding rules, and SEER central registries used SEER codes and coding rules. The two systems were not always in agreement. For hospitals with a COC-approved program which also participated in a SEER registry, two data sets were sometimes required.

COC and SEER attempted to define one common set of data item definitions, field lengths, and codes for use by both SEER registries and hospital-based registries. The collaboration resulted in the publication in 1988 of both the COC's *Data Acquisition Manual* and the *SEER Program Code Manual* with data items and codes in substantial agreement. Having more congruent data sets allowed easier data sharing and data comparisons. The achievement was very useful, and maintaining congruence whenever possible has continued to be a top priority for these two groups.

During the same time period, the California Cancer Registry was developing a statewide automated system which would allow facilities to report electronically to the state registry system. One region of California was a SEER registry at that time, and a large number of hospitals maintained COC-approved programs. To facilitate implementation of standards within its program, the California Cancer Registry requested that SEER and COC

establish a formal committee to pursue data standardization and requested membership on the committee.

When the American Association of Central Cancer Registries was established, the committee function was transferred to the AACCR's Uniform Data Standards Committee. Membership on the committee was expanded to include other registries. Over a relatively few years the committee has continued to make progress toward standardization. A major success occurred when all of the participating groups agreed to implement the second edition of ICD-O simultaneously, for cancer cases diagnosed 1992 and later. In 1993 they convened a multidisciplinary conference to address the issue of collecting data on pre-invasive cervical cancers which resulted in specific recommendations for member registries.

The committee provides a national forum for discussion of data issues and a mechanism for reaching consensus. AACCR member registries and sponsoring organizations are all notified of decisions reached by the committee, and they are provided with all of the pertinent standards documents.

The Centers for Disease Control and Prevention added another strong voice for standardization. They have participated in committee activities of AACCR and entered into a cooperative agreement with AACCR with a major focus of setting standards and identifying innovative ways to facilitate their adoption. The EDITS project described in chapter V is an example of the innovative approach CDC is supporting.

At the time of this publication, the major organizations agree in principle that their data standards will be consistent wherever possible. There are, however, areas where agreement has not been reached. These are discussed in detail in chapter VI, "Unresolved Issues".

Despite the progress made toward standardization and the near-universal agreement that standardization is desirable, much remains to be done. Implementation of existing hospital registry standards is not uniform. Central registries that are not SEER participants have adopted a wide variety of idiosyncratic codes and definitions. SEER and COC will continue to publish separate coding manuals on different update schedules. Standardized data edits are just beginning to be adopted. Coding rules and rule interpretations are sometimes determined informally and marginally documented.

AACCR hopes that by documenting existing standards, recommending standards where they do not yet exist, and publishing the results in a concise and authoritative form that it will be easier for registries and software providers to move forward to the next steps in achieving comparable data that can be more widely used.

In Canada, cancer registries at the provincial and territorial level joined together with Statistics Canada, a federal agency, to form the Canadian Council of Cancer Registries. This process, begun in 1986, has led to development of common national standards for the Canadian Cancer Registry which is being implemented with a reference date of January 1, 1992. A Data Quality Committee, which reports to the Council, is responsible for making recommendations to set national standards, and will review and monitor data quality and resolve any inconsistencies in procedures, coding, or other activities affecting data comparability.

• • •

Chapter III

STANDARDS FOR CASE INCLUSION AND REPORTABILITY

Hospital-based and population-based central registries follow different standards in these areas, albeit with many similarities. A more thorough discussion of the differences can be found in chapter 4 of *Central Cancer Registries: Design, Management and Use* (reference 19). For hospital-based registries, the Commission on Cancer stipulates which cases are to be included in approved registries, while most population-based registries follow the standards set by the SEER Program. The *Cancer Program Manual* (reference 25), the DAM (reference 3), and SEER code manuals (references 2 and 6) should be consulted for more details.

Standards covering reportability and which cases to include are defined in terms of the following:

- Reference Date. The reference date is the effective date cancer registration starts in a specified population at risk or in a specific facility. It is not the date the registry is organized or actually performs the work. Cases diagnosed on or after the reference date must be included. The reference date should usually be January 1 of a calendar year, but is sometimes another date.

- Residency. For a population-based registry, it is essential to include all cases occurring in the population at risk, and rules must be in place for determining who in fact is a member of that population. The goal is to use rules for the cancer cases that correspond to those used by the Census Bureau in enumerating the population. The registry must have rules for determining residency of, for example, part-year residents, institutionalized persons, homeless persons, military personnel, and students. See the SEER code manual for specific instructions.

  AACCR recommends that population-based registries include in their database case reports of non-residents from facilities in their catchment area to allow for sharing of cases which may otherwise go unreported with other population-based registries; to facilitate death clearance and other record linkages; and to allow preparation of reports to individual facilities which include all of their cases.

  Hospital-based registries are less concerned with residency of the patient than the reason for the admission, and hospital registries may exclude certain categories of patients that the central registry must include, for example, patients admitted to a designated hospice unit or transient patients who receive care to avoid

interrupting a course of therapy. Also, COC does not require complete abstracting of cases that are "nonanalytic" for the facility. Therefore, for the central registry, clear rules which are well documented, widely distributed, and accepted are essential to prevent missed cases.

- Reportable List. This is another area where COC and SEER requirements differ. Both standards require inclusion of all neoplasms in the *International Classification of Diseases for Oncology*, second edition (reference 12), with a behavior code of 2 or 3 (in situ or malignant). However, the two programs have different lists of exceptions, in the areas of skin cancer and carcinoma in situ of the cervix. For a detailed discussion of the differences, see chapter 4 of *Central Cancer Registries: Design, Management and Use* (reference 19).

- Multiple Primary Rules. In order to compare cancer rates for two registries, it is important that identical rules have been used for counting multiple tumors in the patient, whether in the same organ, opposite sides of paired organs, different subsites, or different sites, and whether at the same or different times. The SEER Program rules are the de facto standard in the U.S. for both central and hospital-based registries. See the SEER Program code manual for details.

  SEER rules are not identical to the international standard recommended by IARC and IACR (reference 24, page 78). The IARC rules have the effect of defining fewer cases than are defined using SEER rules.

CIN and the Bethesda System

Diagnostic terminology for pre-invasive cervical neoplasia has changed significantly, from the 4-tiered system of dysplasia and carcinoma in situ, to the 3-tiered system of cervical intraepithelial neoplasia (CIN), to the 2-tiered Bethesda System, with high- and low-grade squamous intraepithelial lesions (SIL). Registries have differed in which of these terms they considered synonymous with carcinoma in situ and hence reportable. Consequently, data as presently collected are not comparable over time or across registries.

AACCR convened a multidisciplinary working group in April 1993 to review the problem and make recommendations for its membership. The recommendation was that "population-based registries discontinue routine collection of data on pre-invasive cervical neoplasia unless there is strong local need and interest and sufficient resources are available to collect all [high-grade squamous intraepithelial lesions] and its equivalent terms." (reference 22, page 5). AACCR has adopted this recommendation as its standard.

• • •

Chapter IV

RECOMMENDED DATA EDITS

The term "data edits" is used here to denote algorithms that check the content of data fields against an encoded set of acceptable possible contents and provide information on the quality or acceptability of the coded data. Data edits verify that only acceptable values are used for codes, and more importantly, enforce relationships among the values in related data items. Data edits can apply pass/fail criteria to data, so that a particular entry or set of entries is determined to be either acceptable or unacceptable, and if unacceptable, the entries must be changed until they pass the edits. Other edits may allow for looser criteria, where a grey zone of possibly-acceptable, or acceptable-upon-review is allowed, through, for example, the provision of a distinction between warnings and error conditions, or the provision of override flags in the database itself to indicate that a particular code or codes are to be accepted.

Four types of edits are usually discussed:

- Item edits (single field edits), where a single field or item is edited alone. For example, an edit of the item "Sex" would verify that only valid values are used in the field.

- Interfield edits (multi-field edits), where the contents of more than one item are checked against each other. For example, a common interfield edit checks the code for "Sex" against the code for "Primary Site", and identifies as an error prostate cancer in a female.

- Interrecord edits (multi-record edits), where multiple items in different records for multiple tumors for the same patient are checked for internal consistency. For example, an interrecord edit might verify that sequence numbers have been assigned in chronological order to the patient's cancers.

- Interdatabase edits (multi-database edits) where multiple items in different records in different related databases are checked for internal consistency. In registries with relational data structures, fields that might have been on the same record in a non-relational "flat" file may be on separate records in separate databases

A critical moment for supporting and enforcing data standards occurs at the moment the <ENTER> key is hit, after a value has been typed into an individual field. That moment is the best time to find any problems and bring them to the attention of the user. Most

software does this immediate editing to some degree, though rarely with the completeness of the edits that are performed at a later time, or in a distant place. More of a problem, however, is that each registry software provider writes his or her own edit algorithms, even when the intent is to conform to a single external standard.

At least six different obstacles to standardization of edits across registries, and to their optimal use to achieve standardized data, can be identified.

- Registry systems which may encode an edit from standard specifications are written in different computer languages, with possible differences in translation detail.
- Each implementation of an agreed-upon standard specification may be programmed differently, despite an intent to encode a standard meaning.
- Complete edits are not always performed at the time of data entry.
- Documentation of the edit algorithms used is often difficult for both data analysts and data collectors to obtain and use.
- Merged data collected via different data-entry tools may encourage "apples" and "oranges" to be equated, without the users' awareness.
- Because standards change often, synchronized implementation is unlikely due to the release schedules of software providers and their ability to respond to changes at a given time.

Only when the identical edits are applied to data can one reasonably expect identical results.

The SEER Program has for many years maintained a library of standardized edits written in IBM COBOL (reference 21), which it both applies to data submission from the participating SEER registries and makes available to its participating registries and other interested registries. In theory the SEER registries are able to implement SEER edits identical to those of the central program in their own locale, thus eliminating all edit-detectible errors before data submissions. In practice, all six problems listed above intrude at one time or another to prevent data congruence. Most registries are run on computers other than IBM mainframes, and some do not have COBOL compilers. Their systems differ in the timing with which edits are applied to the data. Synchronization of edit changes is a continuing practical problem.

Nonetheless, the SEER edits constitute the "gold standard" in terms of the logical algorithms they represent, and they are the only set of standardized edits available to AACCR-member registries at this time. The logic they specify reflects years of experience and expertise. For the data items in chapter IX for which SEER is a source of the standard, the SEER edits are the AACCR-designated standard. However, it must be kept in mind that the SEER data set does not contain all of the data items in chapter IX, and the SEER edits do not specify relationships among data items not included in its data set. The SEER edits are thus not complete, even for its own participant registries. To give an example, the SEER data set as submitted to the central registry does not include zip code

(it does include county and census tract), and therefore no edit is specified in the SEER set between zip code and county, although the individual participants have probably implemented such an edit.

The SEER edit package consists of the complete set of COBOL routines (IBM-standard COBOL) on magnetic tape or 3.5" HD floppy discs, and accompanying written documentation, containing:

- 54 item edits
- 64 interfield edits
- 15 interrecord edits
- report-generating routines and the driver programs

Many of the routines are quite large, with thousands of lines of code and very large internal matrices or tables. Others are much simpler. The Utah Cancer Registry has successfully compiled most of the edits under DOS for a PC application, although at this time this is not supported centrally by the SEER Program.

SEER's edits are revised as needed, with new edits added and modifications made continually. Updates are distributed two or more times per year. For more information, contact the SEER Program.

Further work is needed to achieve increased standardization of data edits. One promising approach is the EDITS project, which provides a mechanism for standardized transportable and updatable edits to be provided through a "public library". (See chapter V for a more thorough discussion of the EDITS project.) AACCR will pursue increased edit standardization in the future.

• • •

Chapter V

SOFTWARE COORDINATION OF STANDARDS

Introduction

One perspective on cancer registry work is that there is an absence of data standards. A more empirical view might be that there is an excess of standards – and an absence of systematic standards coordination.

The problem of standards proliferation is not one that can be solved by edict. Nor does it seem likely that owners of established positions can be moved to a single standard system by persuasion; the benefits to society of a single standard are often outweighed by local opportunities and constraints. Yet, the less comparable the data from different sources, the more limited is the usefulness of each collection effort.

The goal of software coordination of standards is to help limit standards proliferation when there is no compelling need to be different, and to provide comprehensive public documentation in a current and readily accessible form in those instances where standards must differ.

In the cancer registry field, standard setters have increased communication and coordination over the last several years, but numerous problems remain. Some examples:

- There remain areas of essential registry data for which standards have not been determined, the most obvious being stage, where the TNM or AJCC staging system, SEER Extent of Disease, and versions of a localized-regional-distant system are all in wide use. (See chapter VI.)

- Even where standards have been established, synchronization of implementation has not been achieved. Organizations which supply software and coding manuals (for example, SEER and COC/ACoS) are on different publication and updating schedules. Those implementing the standards have differing levels of resources to devote to the frequent updating that is necessary.

- There is no adequate mechanism to proactively distribute standards to all potential users. Each organization maintains its own distribution method and schedule. AACCR has filled the gap to a significant extent, but AACCR has not yet adopted a systematic distribution mechanism.

A group of co-developers in the cancer registry field realized that by taking a meta-level approach, software could be engineered to address the latter two problems. If the edit algorithms could be separated from the program code for the application, fewer resources would be needed to implement a standard locally. Moreover, a system of distributable edits would permit the individual software providers to focus their creative energies on the aspects of their cancer registry software that benefit from variety.

The EDITS Project

The EDITS project (Exchangeable-edits, Data-dictionary, and Information Translation Standard) was conceived as a way to optimize the software contribution to coordination (documentation and exchange) of standards by providing a set of public domain tools for creating and exchanging data standards. The purpose of this software is improved accuracy, efficiency, synchronization, timeliness, and standardization in data collection efforts involving multiple cancer registry systems.

EDITS will make it easier for standard setters to create and implement fully documented and distributable standards, and it will make it easier for providers to pick a predefined and authoritative standard from a computer-resident public library of standards.

The EDITS project has been sponsored and developed by the Centers for Disease Control and Prevention in consultation and cooperation with the primary persons and institutions responsible for cancer registry data standards.

The main parts of EDITS are:

- The Metafile–A Database of Data Standards. The Metafile is a comprehensive database of cancer-registry standards. It consists of a collection of tables that can contain all the information needed to test data fields for validity and acceptability, specifically:
    - Look-up tables
    - Translation tables
    - Choice lists
    - Data dictionary of standard fields
    - Local field name table
    - Error messages
    - Executable single- and multi-field validation logic
    - Text descriptions of edits
    - Sets of fields defining standard records
    - Standard-setter list
    - Description of local data storage
    - Data-entry help
    - Standards-documentation text
    - EDITS system help
    - EDITS language reference
- The Edit Engine–A Sub-Program. The Edit Engine provides the link between the Metafile and a data-entry software program and allows the program to make use

of standard or custom edits that have been defined in the Metafile. The Edit Engine may be accessed in several ways, including:

- Modification of existing software to include calls to Edit Engine C library functions for field and/or record level validation
- Inclusion of a batch edit driver in existing software written in a variety of languages
- Use of stand-alone batch programs

In each instance, the software provider controls what parts of the metafile are used, and how and when they are used. The software is designed to facilitate compliance to standards by making it easier to use shared standard edits than to write one's own, but it is not designed to force compliance if non-compliance is intended.

- The Metafile Editor–A Standards Database Tool. The Metafile Editor is a specially designed database editor, which standard setters and software providers will use to select, organize, and maintain data standards. The Editor and associated tools are used to change or add tables, algorithms, help screens, messages, and documentation. EDITS edit logic and supporting documentation from various sources may be extracted from the metafile for use outside EDITS, or they may be combined, edited, and reused in new EDITS entries.

  The software provider must also use the Metafile Editor to modify a Metafile for distribution with the application. A Master Metafile provides a starting point for the provider's own Metafile. The Editor compiles the run-time Metafile for use with the provider's data-entry or data-validation software. Finally, the Editor is also used to generate reports about Metafile contents.

- Public Availability. The EDITS software and the Metafile "public library of standards" are designed to be maintained for public access on an electronic bulletin board. This will permit broad and timely access to the most current data standards.

The EDITS Language

Algorithms for determining the validity of data items are specified in the EDITS programming language. Statements in this language, stored in the Metafile, are processed by an interpreting mechanism in the Edit Engine at appropriate times in the processing of a cancer registry record.

The EDITS language is very similar in syntax, structure, and philosophy to the programming language C. The EDITS language is essentially a subset of this general-purpose programming language, simplified for the special task of writing field-validation edits.

Because the Metafile Editor is itself an MS-DOS program, the Metafile must be modified under MS-DOS. However, the resulting run-time Metafile is usable on UNIX and other platforms.

For further information about the EDITS project, contact the Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, at (404) 488-4682.

Chapter VI

UNRESOLVED ISSUES

Despite the progress made toward data standardization as reflected in this document, some issues remain unresolved. These are explored in detail below.

Stage, TNM, and EOD

Currently, four major staging schemes are being widely used in cancer registries throughout the United States. The lack of comparability among these systems causes major problems for those collecting the data and for users of the data. The four schemes are:

- The American Joint Committee on Cancer's TNM system. Published as *The Manual for Staging of Cancer*, and now in its fourth edition (reference 5), this is a clinically oriented site-specific system which assigns a separate category for the T (tumor), N (nodes), and M (metastases). The TNM categories are then grouped into stages 0 through IV. The COC requires that approved hospital registries use TNM.

- SEER Extent of Disease (EOD). This site-specific 10-digit coding scheme (reference 6) is required for use by SEER registries and is used by some other state and central registries as well. EOD was designed to allow collapsing into several different stage groupings, including AJCC stage group.

- *Summary Staging Guide for the Cancer Surveillance, Epidemiology, and End Results Reporting (SEER) Program*. This manual (reference 9) is used by most registries for assigning localized-regional-distant stage, either instead of or in addition to the above schemes.

- SEER Historic Stage. When SEER data by stage have been published, the stage categories used were derived from categories used by an earlier program, the End Results Group. The categories are not identical to those in the *Summary Staging Guide*.

These schemes were designed for different purposes and at different times, and they are not easily compared. There have been several editions of the TNM manual, and

implementation has not been synchronized. The SEER Program has recently published the *Comparative Staging Guide for Cancer* (reference 4) as an attempt to present comprehensive site-specific comparisons of the schemes to aid in data collection and interpretation. According to the guide:

> Changes over time in methods of cancer screening, diagnosis, staging, and treatment have affected the distribution of stage of disease, but there have also been changes over time in the classification schemes themselves that can complicate data analysis and obscure the meaning of time trends. For all these reasons, comparing cancer registry data by stage over time or across registries, or using pooled data collected by different registries, is problematic. (reference 4, page I.3)

Various other staging schema are also in use. Several oncologic sub-specialties have developed staging systems applying to a limited number of cancer sites.

Population-based registries need to monitor changes in stage distribution over long periods of time to assess the effects of cancer-control efforts. AACCR, through its Uniform Data Standards Committee and its cooperative agreement with CDC, intends to convene a multidisciplinary advisory group to recommend a new standard staging scheme which will be:

- simple
- anatomically-based
- stable over time
- compatible at the stage-group level with AJCC stage categories

Hispanic Ethnicity

There is agreement on the standard data item "Spanish Surname/Origin" and its codes. However, there is substantial variation in how the code is determined among registries. Some registries record an ethnicity listed in the medical record, and some have the abstractor code ethnicity based on all information available, including the surname, birthplace, or stated ethnicity. Other registries rely on a manual or computer matching of surname against a list of Spanish surnames from (usually) the 1980 census, sometimes linking maiden name also, but sometimes not. Another method used by some registries is the application of a computer algorithm to surnames to determine ethnicity.

Population-based registries must attempt to categorize their cases using a method that best approximates the method used by the Census Bureau to determine ethnicity of the population denominators, but a standard method has not been determined. The AACCR Uniform Data Standards Committee will continue to discuss this problem.

Canadian and Other Non-U.S. Data

The AACCR data standards thus far adopted do not do a good job of handling non-U.S. data. Changes will be needed to accommodate postal codes, standard abbreviations for provinces, and other fields, especially with regard to Canadian data. Future versions of

this document will review and incorporate standards established for the Canadian Cancer Registry by the Canadian Council of Cancer Registries.

Time Period for First Course of Treatment

The ACoS DAM defines first course of therapy as "limited to procedures that begin (or are planned) within four months of the date of initial diagnosis." The SEER Program code manual defines first course of treatment as "all cancer directed therapy administered to the patient within four months after the initiation of therapy." Thus patients whose treatment begins more than four months after diagnosis are handled differently by the two systems.

Each system allows coding of treatment given beyond its four-month period, if the treatment was part of the planned first course. However, determining this from the medical record is often problematic, and coding is probably not consistent across registries.

Occupation and Industry

Most population-based registries have found collection of occupation and industry data to be difficult and of limited utility. Some reasons include:

- Lack of good employment or occupational histories in patient records, especially for those over age 65
- Lack of standardization in patient records regarding documentation of usual occupation versus longest-held occupation or most recent occupation

A notable exception has been the research-oriented Cancer Surveillance Program of Los Angeles. Their successful collection and analysis of occupational data has required:

- Supplementing patient record information by obtaining paper copies of death certificates on all expired cases to obtain the listed occupation and industry
- Collecting the name of the employer from the patient record on all cases, and having registry staff research the industry of the employers and maintain a roster of employers and industries
- Having an active occupational cancer research program

The Los Angeles data have been extensively used and resulted in many publications. However, an effort such as theirs is expensive and not timely, since it relies on death certificates. Most registries do not have the resources for a comparable data collection and analysis effort.

The Cancer Registries Amendment Act (reference 27) specifies collection of occupation and industry data. However, since information from patient records alone is insufficient to generate useful, timely data, AACCR does not include occupation or industry in its recommended data sets. AACCR recommends that such data be collected only when special resources are available.

Other

Additional standards need to be set for some procedural methods that affect data exchange. These include:

- Reusability of accession number from a deleted case
- Assigning identification numbers to facilities and central registries
- Standardizing TNM field widths and justification rules

• • •

Chapter VII

REFERENCES

Code Manuals and Record Layouts

1. AACCR (American Association of Central Cancer Registries). 1994. *Data exchange standards and record description*, Version 3.0. Standards for Cancer Registries, vol. 1, edited by Barry Gordon. N.p.: American Association of Central Cancer Registries. February 9, 1994.

2. SEER (Surveillance, Epidemiology, and End Results Program). 1992. *The SEER program code manual.* rev. ed. [Bethesda]: National Institutes of Health, National Cancer Institute. NIH Pub. No. 92-1999.

3. ACoS (American College of Surgeons Commission on Cancer). 1990. *Data acquisition manual.* Chicago: American College of Surgeons Commission on Cancer.

Stage and Extent of Disease Manuals

4. SEER (Surveillance, Epidemiology, and End Results Program). 1993. *Comparative staging guide for cancer,* Major Cancer Sites, Version 1.1. [Bethesda]: National Institutes of Health, National Cancer Institute. June 1993.

5. AJCC (American Joint Committee on Cancer). 1992. *Manual for Staging of Cancer,* edited by O. H. Beahrs, D. E. Henson, R. V. P. Hutter, and B. J. Kennedy. 4th ed. Philadelphia: J. B. Lippincott Company. See also editions 1, 2, and 3.

6. SEER (Surveillance, Epidemiology, and End Results Program). 1992. *SEER extent of disease–1988: Codes and coding instructions.* 2d ed. [Bethesda]: National Institutes of Health, National Cancer Institute. NIH Pub. No. 92-2313. June 1992.

7. Shambaugh, E. M., Ries, L. G., and Young, J. L. 1984. *Extent of Disease: New 4-Digit Schemes: Codes and Coding Instructions.* N.p.: Demographic Analysis Section, Biometry Branch, National Cancer Institute. Revised March 31, 1984.

8. SEER (Surveillance, Epidemiology, and End Results Program). 1977. *Extent of Disease: Codes and Coding Instructions*. [Bethesda]: SEER Program. April 1977.

9. SEER (Surveillance, Epidemiology, and End Results Program). 1977. *Summary staging guide for the cancer surveillance, epidemiology and end results reporting (SEER) Program*. [Bethesda]: National Institutes of Health. April 1977, reprinted July 1986)

Disease Classifications

10. WHO (World Health Organization). 1992. *ICD-10: International Statistical Classification of Diseases and Related Health Problems*, 10th rev., vol. 1 [of 3]. Geneva: World Health Organization.

11. HCFA (Health Care Financing Administration). 1991. *The International Classification of Diseases: 9th Revision: Clinical Modification: ICD-9-CM*, 4th ed. N.p.: Health Care Financing Administration, Public Health Service, U.S. Department of Health and Human Services.

12. Percy, C., VanHolten, V., and Muir, C., eds. 1990. *International classification of diseases for oncology*. 2d ed. Geneva: World Health Organization.

13. Percy, C., VanHolten, V., eds. 1988. *International classification of diseases for oncology*, field trial edition. (March 1988)

14. Percy, C., and VanHolten, V., eds. 1987. *International classification of diseases for oncology, morphology*, field trial edition.

15. Percy, C., and VanHolten, V., eds. 1986. *International classification of diseases for oncology, morphology*, field trial edition.

16. WHO (World Health Organization). *International Classification of Diseases for Oncology*, 1st ed. 1976. Geneva: World Health Organization.

17. WHO (World Health Organization). 1977. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*, Based on the Recommendations of the Ninth Revision Conference, 1975 . . . Geneva: World Health Organization.

18. *Eighth Revision International Classification of Diseases, Adapted for Use in the United States*, 2 vol. 1968. [Washington, D.C.]: U.S. Department of Health, Education, and Welfare. (vol. 1 undated, vol. 2 December 1968)

Other References

19. Menck, H. R., and Smart, C. R., eds. In press. *Central cancer registries: Design, management, and use*. New York: Gordon & Breach.

20. AACCR (American Association of Central Cancer Registries). 1994. *Standards for completeness, quality, analysis, and management of data.* Standards for Cancer Registries, vol. 3, edited by J. E. Seiffert. N.p.: American Association of Central Cancer Registries.

21. SEER (Surveillance, Epidemiology, and End Results Program). 1993. *SEER Edit Documentation.* Bethesda: SEER Program, National Cancer Institute. May 1993.

22. AACCR (American Association of Central Cancer Registries). 1993. *Working group on pre-invasive cervical neoplasia and population-based cancer registries: Final subcommittee report.* AACCR Conference held April 5-6, 1993, Rockville, Maryland. Subsequently adopted by the AACCR Executive Board May 1993 and amended November 1993.

23. Shambaugh, E., ed.-in-chief. Various years. *SEER program: Self-instructional manual for cancer registrars.* [Bethesda]: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health. Various years.

Book One - Objectives and Functions of a Tumor Registry, 2d ed., 1980. [New edition is in preparation.]

Book Two - Cancer Characteristics and Selection of Cases, 3d ed., 1992.

Book Three - Tumor Registrar Vocabulary: The Composition of Medical Terms, 2d ed., 1993.

Book Four - Human Anatomy as Related to Tumor Formation, 2d ed., 1993.

Book Five - Abstracting a Medical Record: Patient Identification, History, and Examinations, 2d ed., 1993.

Book Six - Classification for Extent of Disease, 1977. [New edition is in preparation.]

Book Seven - Statistics and Epidemiology for Tumor Registrars, scheduled for publication in 1994.

Book Eight - Antineoplastic Drugs, 2d ed., 1986. [New edition is in preparation.]

24. Jensen, O. M., Parkin, D. M., MacLennan, R., Muir, C. S., and Skeet, R. G., eds. 1991. *Cancer registration: Principles and methods.* Lyon: International Agency for Research on Cancer. IARC Scientific Pub. No. 95.

25. ACoS (American College of Surgeons Commission on Cancer). 1991. *Cancer program manual.* Chicago: American College of Surgeons Commission on Cancer.

26. NTRA (National Tumor Registrars Association). 1989. *Registry staffing manual.* A joint study and report from the National Tumor Registrars Association, Inc. and the American College of Surgeons Commission on Cancer with financial support of the American Cancer Society. Mundelein, Ill.: National Tumor Registrars Association.

27. *Cancer Registries Amendment Act,* Public Law 102-515, Oct 24, 1992.

28. ACoS (American College of Surgeons Commission on Cancer). 1992. *Fundamental Tumor Registry Operations Program.* Chicago: American College of Surgeons Commission on Cancer.

• • •

Chapter VIII

RECOMMENDED DATA ITEMS AND RECORD LAYOUT

In this chapter and the next, the data elements and their codes which comprise the AACCR standards have been organized in the order they appear in the AACCR data exchange record layout (reference 1). The column positions of the data elements and their relative location within the layout are summarized in this chapter. This summary can be used as an index to the data elements in chapter IX.

The following table also summarizes in five columns various categories of data item standards.

The first two columns summarize the data items recommended by the COC and those currently used by the SEER Program. The next two columns include AACCR recommended standard data items for two kinds of central registries:

- Incidence-Only. A population-based surveillance registry that does not collect treatment or follow-up information.

- Full (Multipurpose). A population-based registry that collects treatment, follow-up, and detailed staging information in addition to all of the data items included in an incidence registry.

Finally, the fifth column lists those data items presently included in the AACCR data exchange standard, i.e., those items considered essential for exchange as reflected in the data exchange record layout.

The sections of the data exchange record and the table below are:

A. Record ID Section
B. Demographic Section
C. Cancer Identification Section
D. Hospital-Specific Section
E. Stage/Other Prognostic Factors Section
F. Treatment - 1st Course Section
G. Treatment - Subsequent Section
H. Follow-up/Recurrence Section
I. Death Information Section

J. Edit Overrides/Conversion History Section
K. System Administration Section
L. Special Use Section
M. Patient - Confidential Section
N. Hospital - Confidential Section
O. Physician - Confidential Section
P. Text - Diagnosis Section
Q. Text - Treatment Section
R. Text - Miscellaneous Section

Expanded definitions of the six columns in the table are as follows:

- COC HOSP (Hospital Registry). Core data requirements and additional recommendations for a registry seeking approval from ACoS COC. These data elements are designed to meet the needs of facilities collecting detailed clinical information (reference 25).

- SEER CENTRAL. Required data elements for a central registry affiliated with the NCI's SEER Program. These comprise the subset of the registries' data elements which are actually submitted to the NCI.

- AACCR STANDARDS: INCIDENCE CENTRAL. Data requirements for a population-based, surveillance registry that does not collect treatment or follow-up information. These recommendations have been reviewed and accepted by experts in the field, represented by the AACCR Uniform Data Standards Committee. These data items include those needed to produce age-, race-, sex-, and area-specific cancer incidence rates by cancer site and histologic type. General Summary Stage is included for cancer control studies. Administrative items are included that facilitate death clearance, patient matching, interfield edit reviews, and hospital admission tracking. Rationales for some of the individual items are included in chapter IX.

- AACCR STANDARDS: FULL CENTRAL. Data requirements for a population-based registry that collects treatment, follow-up, and detailed staging information in addition to all of the data items included in an incidence registry. AACCR recommends these data items for a full-service central registry which makes use of data collected by hospital-based registries.

  Many of the recommended items are also required by NCI's SEER Program. Also included are items that help support and analyze hospital-specific activities, such as Accession Number (Hosp), Dates of Admission and Discharge, Class of Case, and hospital-specific treatment codes and treatment dates. AJCC Stage Group and TNM elements are included because this detailed staging is required of most hospital registries. Patient follow-up is supported with patient-contact and following-physician information, and tumor status and recurrence fields are included. Administrative items include pre-converted ICD-O-1 codes, and the name of the vendor submitting data.

- AACCR STANDARDS: EXCHANGE. Minimum items required for a data exchange record.

- AACCR STANDARDS: COLUMN. This gives the placement in the record layout. It may be used as a reference to find the data item within the detailed list in chapter IX.

Recommended Data Items

			AACCR STANDARDS			
ITEM	COC HOSP	SEER CENTRAL	INCIDENCE CENTRAL	FULL CENTRAL	EX-CHANGE	COLUMN
A. Record ID Section						
Record Type	.	.	.	.	Y	1
Patient ID Number	.	Y	Y	Y	Y	2
Registry Type	.	.	.	.	Y	10
Registry ID	.	Y	.	.	Y	11
B. Demographic Section						
Addr at Dx–City	Y	.	Y	Y	.	19
Addr at Dx–State	Y	.	Y	Y	Y	39
Co. of Residence at Dx	Y	Y	Y	Y	Y	41
Addr at Dx–Zip	Y	.	Y	Y	.	44
Census Tract	.	Y	Y	Y	.	53
Census Tract Coding Sys	.	Y	Y	Y	.	59
Marital Status	Y	Y	Y	Y	.	60
Race	Y	Y	Y	Y	Y	61
Race Coding Sys - Current	.	.	.	.	Y	63
Race Coding Sys - Original	.	.	.	.	Y	64
Spanish/Hispanic Origin	Y	Y	Y	Y	Y	65
Sex	Y	Y	Y	Y	Y	66
Age at Diagnosis	Y	Y	Y	Y	.	67
Birth Date	Y	Y	Y	Y	Y	70
Birthplace	Y	Y	Y	Y	.	78
Religion	.	.	.	.	.	81
Occupation (Census)	.	.	.	.	.	83
Industry (Census)	.	.	.	.	.	86
Occupation (SOC)	.	.	.	.	.	89
Industry (SIC)	.	.	.	.	.	93
Occup/Ind Coding System	.	.	.	.	.	97
Smoking History	.	.	.	.	.	98

Recommended Data Items

			AACCR STANDARDS			
ITEM	**COC HOSP**	**SEER CENTRAL**	**INCIDENCE CENTRAL**	**FULL CENTRAL**	**EX-CHANGE**	**COLUMN**
B. Demographic Section (continued)						
Name-Derived Ethnicity	.	.	.	.	.	99
Reserved for expansion	.	.	.	.	.	101
C. Cancer Identification Section						
Tumor Record Number	.	Y	Y	Y	Y	107
Sequence Number	Y	Y	Y	Y	Y	109
Date of Diagnosis	Y	Y	Y	Y	Y	111
Primary Site	Y	Y	Y	Y	Y	119
Laterality	Y	Y	Y	Y	Y	123
Morphology -Type&Behavior	Y	Y	Y	Y	Y	124
Grade	Y	Y	Y	Y	Y	129
Site Coding Sys. - Current	.	.	.	.	Y	130
Site Coding Sys. - Original	.	.	.	.	Y	131
Morph Coding Sys. -Current	.	.	.	.	Y	132
Morph Coding Sys. -Original	.	.	.	.	Y	133
Diagnostic Confirmation	Y	Y	Y	Y	Y	134
Type of Reporting Source	.	Y	Y	Y	.	135
Accession Year	Y	.	.	Y	.	136
Reserved for expansion	.	.	.	.	.	138
D. Hospital-Specific Section						
Reporting Hospital (Dx)	.	.	Y	Y	.	142
Accession Number (Hosp)	Y	.	Y	Y	.	150
Abstractor	.	.	.	.	.	156
Date of Admission	.	.	Y	Y	.	159
Date of Discharge	.	.	.	Y	.	167
Class of Case	Y	.	.	Y	.	175
RX Hospital–Surgery Type	.	.	.	Y	.	176
RX Hospital–Radiation	.	.	.	Y	.	178

Recommended Data Items

			AACCR STANDARDS			
ITEM	**COC HOSP**	**SEER CENTRAL**	**INCIDENCE CENTRAL**	**FULL CENTRAL**	**EX-CHANGE**	**COLUMN**
D. Hospital-Specific Section (continued)						
RX Hospital–Chemo	.	.	.	Y	.	179
RX Hospital–Hormone	.	.	.	Y	.	180
RX Hospital–BRM	.	.	.	Y	.	181
RX Hospital–Other	.	.	.	Y	.	182
E. Stage/Other Prognostic Factors Section						
General Summary Stage	Y	.	Y	Y	Y	183
Loc/Reg/Distant Stage	.	.	.	.	.	184
EOD–Tumor Size	Y	Y	.	Y	.	185
EOD–Extension	Y	Y	.	.	.	188
EOD–Lymph Node Involv	Y	Y	.	.	.	190
Regional Nodes Positive	Y	Y	.	Y	.	191
Regional Nodes Examined	Y	Y	.	Y	.	193
EOD - old 13 digit	.	Y	.	.	.	195
EOD - old 2 digit	.	Y	.	.	.	208
EOD - old 4 digit	.	Y	.	.	.	210
T Code (Path-based)	Y	.	.	Y	.	214
N Code (Path-based)	Y	.	.	Y	.	216
M Code (Path-based)	Y	.	.	Y	.	218
AJCC Stage Group (Path)	Y	.	.	Y	.	219
T Code (Clinical)	Y	.	.	Y	.	221
N Code (Clinical)	Y	.	.	Y	.	223
M Code (Clinical)	Y	.	.	Y	.	225
TNM Edition Number	.	.	.	Y	.	226
AJCC Stage Group (Clin)	Y	.	.	Y	.	227
Distant Metastasis	Y	.	.	.	.	229
Residual Tumor	Y	.	.	.	.	232

Recommended Data Items

			AACCR STANDARDS			
ITEM	**COC HOSP**	**SEER CENTRAL**	**INCIDENCE CENTRAL**	**FULL CENTRAL**	**EX-CHANGE**	**COLUMN**
E. Stage/Other Prognostic Factors Section (continued)						
Alternate Staging Scheme	.	.	.	.	.	233
Tumor Marker 1	.	Y	.	Y	.	237
Tumor Marker 2	.	Y	.	Y	.	238
Reserved for expansion	.	.	.	.	.	239
F. Treatment - 1st Course Section						
RX Coding System - Current	.	.	.	.	.	251
RX Date–Surgery	.	.	.	Y	.	252
RX Date–Radiation	.	.	.	Y	.	260
RX Date–Chemotherapy	.	.	.	Y	.	268
RX Date–Hormone Therapy	.	.	.	Y	.	276
RX Date–BRM	.	.	.	Y	.	284
RX Date–Other	.	.	.	Y	.	292
RX Date–Started	Y	Y	.	Y	.	300
RX Summary–Surgery Type	Y	Y	.	Y	.	308
Reason for No Surgery	Y	Y	.	Y	.	310
RX Summary–Radiation	Y	Y	.	Y	.	311
RX Summary–Rad to CNS	Y	Y	.	Y	.	312
RX Summary–Surg/Rad Seq	Y	Y	.	Y	.	313
RX Summary–Chemo	Y	Y	.	Y	.	314
RX Summary–Hormone	Y	Y	.	Y	.	315
RX Summary–BRM	Y	Y	.	Y	.	316
RX Summary–Other	Y	Y	.	Y	.	317
First Course Calc. Method	.	.	.	.	.	318
Reserved for expansion	.	.	.	.	.	319
G. Treatment - Subsequent Section						
Subseq RX–Date (#1)	Y	.	.	.	.	338

Recommended Data Items

			AACCR STANDARDS			
ITEM	COC HOSP	SEER CENTRAL	INCIDENCE CENTRAL	FULL CENTRAL	EX-CHANGE	COLUMN
G. Treatment - Subsequent Section (continued)						
Subseq RX–Codes (#1)	Y	.	.	.	.	346
Subseq RX #2	Y	.	.	.	.	353
Subseq RX #3	Y	.	.	.	.	368
Subseq RX #4	Y	.	.	.	.	383
Reserved for expansion	.	.	.	.	.	398
H. Follow-up/Recurrence Section						
Date of Last Contact	Y	Y	Y	Y	.	418
Vital Status	Y	Y	Y	Y	.	426
Tumor Status	Y	.	.	Y	.	427
Quality of Survival	Y	.	.	.	.	428
Last Type of Follow-up	Y	.	.	Y	.	429
Follow-up Contact City	.	.	.	Y	.	430
Follow-up Contact State	.	.	.	Y	.	450
Follow-up Contact Zip	.	.	.	Y	.	452
Recurrence–Date (First)	Y	.	.	Y	.	461
Recurrence–Type (First)	Y	.	.	Y	.	469
Recurrence–Distant Sites	Y	.	.	Y	.	470
I. Death Information Section						
Cause of Death	Y	Y	Y	Y	.	473
ICD Revision Number	Y	Y	Y	Y	.	477
Autopsy	.	.	.	.	.	478
Place of Death	.	.	.	.	.	479
J. Edit Overrides/Conversion History Section						
Over-ride Age/Site	.	Y	Y	Y	.	482
Over-ride SeqNo/DxConf	.	Y	Y	Y	.	483
Over-ride Site/Lat/SeqNo	.	Y	Y	Y	.	484

Recommended Data Items

ITEM	COC HOSP	SEER CENTRAL	AACCR STANDARDS INCIDENCE CENTRAL	AACCR STANDARDS FULL CENTRAL	AACCR STANDARDS EX-CHANGE	COLUMN
J. Edit Overrides/Conversion History Section (continued)						
Over-ride Surg/DxConf	.	Y	Y	Y	.	485
Over-ride–Site/Type	.	Y	Y	Y	.	486
Over-ride–Histology	.	Y	Y	Y	.	487
Over-ride Report Source	.	Y	Y	Y	.	488
Over-ride Ill-define site	.	Y	Y	Y	.	489
Over-ride Leuk, Lymphoma	.	Y	Y	Y	.	490
Site (1973-91)	.	Y	.	Y	.	491
Morph (1973-91)	.	Y	.	Y	.	495
ICDO-2 Conversion Flag	.	Y	.	Y	.	501
Reserved for expansion	.	.	.	.	.	502
K. System Administration Section						
Date Case Completed	.	.	.	.	.	503
Date Case Last Changed	.	.	.	.	.	511
General Coding Procedure	.	.	.	.	.	519
SEER Coding Sys. - Current	.	.	.	.	Y	521
SEER Coding Sys. - Original	.	.	.	.	Y	522
ACoS Coding Sys. - Current	.	.	.	.	Y	523
ACoS Coding Sys. - Original	.	.	.	.	Y	524
Subseq Report For Primary	.	.	.	.	.	525
Vendor Name	.	.	.	Y	Y	526
AACCR Record Version	.	.	.	.	Y	536
Reserved for expansion	.	.	.	.	.	537
L. Special Use Section						
Site-Specific Studies	.	.	.	.	.	551
State-Specific Items	.	.	.	.	.	751

Recommended Data Items

ITEM	COC HOSP	SEER CENTRAL	AACCR STANDARDS INCIDENCE CENTRAL	AACCR STANDARDS FULL CENTRAL	AACCR STANDARDS EX-CHANGE	COLUMN
M. Patient - Confidential Section						
Patient's Last Name	Y	.	Y	Y	.	851
Patient's First Name	Y	.	Y	Y	.	866
Patient's Middle Initial	Y	.	Y	Y	.	880
Alias	Y	.	Y	Y	.	881
Spouse/Parent Name	.	.	Y	Y	.	896
Medical Record Number	Y	.	.	Y	.	946
Social Security Number	Y	.	Y	Y	.	957
Addr at Dx–No & Street	Y	.	Y	Y	.	966
Follow-up Contact Name	.	.	.	Y	.	991
Follow-up Contact Address	.	.	.	Y	.	1021
Patient Phone Number	Y	.	.	Y	.	1046
DC State File Number	.	.	Y	Y	.	1056
Maiden Name	Y	.	Y	Y	.	1062
Reserved for expansion	.	.	.	.	.	1077
N. Hospital - Confidential Section						
Hospital Referred From	Y	.	.	Y	.	1082
Hospital Referred To	Y	.	.	Y	.	1090
Last Follow-up Hospital	.	.	.	.	.	1098
Next Follow-up Hospital	.	.	.	.	.	1106
O. Physician - Confidential Section						
Physician (Attending)	Y	.	.	.	.	1114
Follow-up Physician	.	.	.	Y	.	1122
Reserved for expansion	.	.	.	.	.	1130
P. Text - Diagnosis Section						
Text–Dx Proc–PE	.	.	.	.	.	1151
Text–Dx Proc–X-ray/scan	.	.	.	.	.	1351

Recommended Data Items

			AACCR STANDARDS			
ITEM	**COC HOSP**	**SEER CENTRAL**	**INCIDENCE CENTRAL**	**FULL CENTRAL**	**EX-CHANGE**	**COLUMN**
P. Text - Diagnosis Section (continued)						
Text–Dx Proc–Scopes	.	.	.	.	.	1601
Text–Dx Proc–Lab Tests	.	.	.	.	.	1851
Text–Dx Proc–Op	.	.	.	.	.	2101
Text–Dx Proc–Path	.	.	.	.	.	2351
Text–Primary Site Title	.	.	.	.	.	2601
Text–Histology Title	.	.	.	.	.	2641
Text–Staging	.	.	.	.	.	2681
Q. Text - Treatment Section						
RX Text–Surgery	.	.	.	.	.	2981
RX Text–Radiation (Beam)	.	.	.	.	.	3131
RX Text–Radiation (Other)	.	.	.	.	.	3281
RX Text–Chemo	.	.	.	.	.	3431
RX Text–Hormone	.	.	.	.	.	3631
RX Text–BRM	.	.	.	.	.	3831
RX Text–Other	.	.	.	.	.	3931
R. Text - Miscellaneous Section						
Text–Remarks	.	.	.	.	.	4031
Text–Occup/Indus	.	.	.	.	.	4381
Place of Diagnosis	.	.	.	.	.	4481
Reserved for expansion	.	.	.	.	.	4531

• • •

Chapter IX

RECOMMENDED DATA ITEM CODES

In this chapter, data items are presented in the order of their location in the AACCR data exchange record layout, separated into the following sections:

A. Record ID Section
B. Demographic Section
C. Cancer Identification Section
D. Hospital Specific Section
E. Stage/Other Prognostic Factors Section
F. Treatment - 1st Course Section
G. Treatment - Subsequent Section
H. Follow-up/Recurrence Section
I. Death Information Section
J. Edit Overrides/Conversion History Section
K. System Administration Section
L. Special Use Section
M. Patient - Confidential Section
N. Hospital - Confidential Section
O. Physician - Confidential Section
P. Text - Diagnosis Section
Q. Text - Treatment Section
R. Text - Miscellaneous Section

A general description of each item is given, followed by the allowable values and standard format for data entry. Specific codes used and their meaning are listed. For many items, the document provides a brief rationale for collecting the data item or for using the codes listed.

The header for each item contains the standard item name, usually the name used in AACCR's *Data Exchange Standards and Record Description* (reference 1). Listed beneath the standard name are other names by which the same item is called, including the name used by the standard setter for the item. The name of the standard-setting organization for the item is then listed. The column numbers given correspond to positions in the AACCR data exchange record layout, as do the lengths in characters that are stated.

NOTE 1 The format for all dates is numeric (MMDDCCYY), with 99 for unknown day, month, year, or century (i.e., 1899 = year 1899, 9999 = year unknown).

NOTE 2 References for more detailed coding instructions for many items are implied by the "Source of Standard" listed, as follows:

AACCR. Usually AACCR's *Data Exchange Standards and Record Description* (reference 1).

COC. Usually the *Data Acquisition Manual (reference 3).*

SEER. Either the *SEER Program Code Manual* (reference 2) or the *SEER Extent of Disease–1988: Codes and Coding Instructions* (reference 6).

Reporting Registry. The documentation of the registry which created the record.

Other references are listed in the text as needed.

IX.A Record ID Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Record Type**	1	AACCR	1 - 1

Description:
Generated field which identifies which of the 3 AACCR data exchange record types is being used in a file of data exchange records. A file should have records of only one type.

Allowable Values and Format:
I, C, A

Codes:
- I Incidence-only record type: length = 850
- C Confidential record type (includes patient name, etc.): length = 1150
- A Full case abstract record type (includes text summaries): length = 5300

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Patient ID Number (Case Number)**	8	Reporting Registry	2 - 9

Description:
Unique number assigned by the central registry to an individual patient. All tumors for the same patient will carry this same number. May contain a check digit. AACCR recommends that this number not be reused when a patient is deleted from the files.

Rationale:
Central registries need a unique, unchangeable identification number which links records for the same patient across different tumors and different hospital reports.

Allowable Values and Format:
Numeric, right justified, zero-filled

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Registry Type**	1	AACCR	10 - 10

Description:
Generated code which best describes the type of the submitting registry for use when data from multiple registries are being pooled.

Allowable Values and Format:
1, 2, 3

Codes:
- 1 Central registry (population based)
- 2 Central registry or hospital consortium (not population-based)
- 3 Single hospital/free standing center

IX.A Record ID Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Registry ID**	8	AACCR	11 - 18

Description:
A numeric code assigned to each data source identifying who is sending the record and what population it is based on, for use when data is exchanged or pooled.

Allowable Values and Format:
Numeric, right justified

Codes:
For SEER areas, it may be their SEER Registry ID number. Other central registries should have unique numbers assigned to them. The AACCR will maintain the list of codes to prevent duplication.

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

Note: The field "Addr at DX-- No & Street" is located in section IX.M, the Patient-Confidential Section.

✓ **Addr at DX--City** | 20 | COC | 19 - 38
(City or Town)
(City)
(Town)

Description:

Name of city in which the patient resides at the time the reportable tumor is diagnosed. If patient resides in rural area, record the name of the city used in their mailing address. If the patient has multiple tumors, the city of residence may be different.

Allowable Values and Format:

Alpha only, no special characters, mixed case left justified, blank filled. If unknown, enter "unknown".

✓ **Addr at DX--State** | 2 | COC | 39 - 40
(State)

Description:

U.S. Postal Service abbreviation for the state, commonwealth, or country in which the patient resides at the time the reportable tumor is diagnosed. If a commonly accepted two-letter abbreviation does not exist for the country, leave blank. If the patient has multiple tumors, the state of residence may be different. No standard has yet been set for Canadian Provinces.

Allowable Values and Format:

Alpha only, upper case, or all blank

✓ **County** | 3 | FIPS | 41 - 43
(County of Residence at Dx)

Description:

Code for the county of the patient's residence at the time the reportable tumor is diagnosed. For more discussion of the problem of determining residency at diagnosis, consult the SEER code manual. If the patient has more than one tumor, the county codes may be different. Codes used are those of the Federal Information Processing Standards (FIPS) Publication, "Counties and Equivalent Entities of the United States, Its Possessions, and Associated Areas" No standard has yet been set for Canadian Provinces.

Allowable Values and Format:

Numeric

Special Codes:

998 Known place of residence, county code not available
999 County unknown

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Addr at DX--Zip (Zip Code)**	9	COC	44 - 52

Description:
United States Postal Service zip code for the state and city in which the patient resides at the time the reportable tumor is diagnosed. May use either the 5 digit or 9 digit extended zip code. Blanks follow the 5 digit code. If the patient has multiple tumors, the zip code may be different. No standard has yet been set for entry of Canadian postal codes.

Allowable Values and Format:
Numeric, left justified, blank filled

Special Codes:

888888888	Resident of foreign country
999999999	U.S. resident, zip code unknown

✓	**Census Tract**	6	SEER	53 - 58

Description:
Code for the census tract of the patient's residence at the time of diagnosis for the tumor being reported. If the patient has more than one tumor, the codes may differ. Codes used are those of the Census Bureau.

Rationale:
Central registries can calculate incidence rates for geographical areas having population estimates. The smallest units for which the US Census Bureau provides population data are census tracts. These rates can be used for general surveillance or special geographical and socioeconomic analyses.

Allowable Values and Format:
Numeric. Assume that the decimal point is between the 4th and 5th digits of the field, and zero-fill any unused digits, e.g., census tract '409.6' would be entered '040960'.

Special codes:

000000	Area not census-tracted
999999	Area census-tracted, but census tract is not available.

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Census Tract Coding Sys**	1	**SEER**	**59 - 59**

Description:
Identifies the set of Census Bureau's census tract definitions (boundaries) that were used to code a specific record. Codes correspond to the definitions used in a specific census. The years in parentheses below indicate the diagnosis years to which the set of definitions should usually be applied.

Rationale:
Census tracts for particular cases can change between 1980 and 1990 definitions. Either all cases must be recoded using the latest census definitions, or the coding system used on cases must be coded, so that time trends can be correctly analyzed.

Allowable Values and Format:
0-3

Codes:

0	Not tracted
1	1970 Census Tract Definitions (1973-77)
2	1980 Census Tract Definitions (1978-87)
3	1990 Census Tract Definitions (1988+)

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Marital Status** **(Marital Status at Diagnosis)** **(Marital Status at Initial Diagnosis)**	1	**SEER/COC**	**60 - 60**

Description:
Code for the patient's marital status at the time of diagnosis for the tumor being reported. If the patient has multiple tumors, marital status may be different for each tumor.

Rationale:
Marital Status is linked to sexual activity and to hormonal status as a surrogate for parity, and it is an indicator of lifestyles that may relate to certain cancers. It also helps in patient identification.

Allowable Values and Format:
1-5, 9

Codes:

1	Single (never married)
2	Married (including common law)
3	Separated
4	Divorced
5	Widowed
9	Unknown

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Race**	2	SEER/COC	61 - 62

Description:

Code for the race of the patient, separate from Spanish/Hispanic Origin (see below). All tumors for the same patient should carry the same race code.

Rationale:

Since racial origin has a significant association with cancer rates and outcomes, a comparison between areas with different racial distributions may require an analysis of race to interpret the findings. The Race codes listed here largely correspond to race categories used by the US Census bureau, so that race-specific incidence rates can be calculated. Even if one's own state population is missing many of the race categories, the full coding system should be used to allow accurate national comparison and collaboration.

Allowable Values and Format:

01-14, 20-22, 25-28, 30-32, 96-99

Codes:

01	White
02	Black
03	American Indian, Aleutian, or Eskimo
04	Chinese
05	Japanese
06	Filipino
07	Hawaiian
08	Korean
09	Asian Indian, Pakistani
10	Vietnamese
11	Laotian
12	Hmong
13	Kampuchean
14	Thai
20	Micronesian, NOS
21	Chamorran
22	Guamanian, NOS
25	Polynesian, NOS
26	Tahitian
27	Samoan
28	Tongan
30	Melanesian, NOS
31	Fiji Islander
32	New Guinean
96	Other Asian, incl. Asian, NOS and Oriental, NOS
97	Pacific Islander, NOS
98	Other
99	Unknown

NOTE: Codes 20 - 97 were adopted for use effective with 1991 diagnoses and code 14 for 1994 and later cases.

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Race Coding Sys - Current**	1	AACCR	63 - 63

Description:
Code that best describes how race is currently coded. If converted, this field shows the system it is converted to.

Allowable Values and Format:
1-5, 9

Codes:

1	4-value coding: 1=White, 2=Black, 3=Other, 9=Unknown
2	SEER < 1988 (1-digit)
3	1988+ SEER & COC (2-digit)
4	1991+ SEER & COC (added codes 20-97, additional Asian and Pacific Islander codes)
5	1994+ SEER & COC (added code 14, Thai)
9	Other

✓	**Race Coding Sys - Original**	1	AACCR	64 - 64

Description:
Code that best describes how race was originally coded. If later converted, this field shows the original codes used.

Allowable Values and Format:
1-5, 9

Codes:

1	4-value coding: 1=White, 2=Black, 3=Other, 9=Unknown
2	SEER < 1988 (1-digit)
3	1988+ SEER & COC (2-digit)
4	1991+ SEER & COC (added codes 20-97, additional Asian and Pacific Islander codes)
5	1994+ SEER & COC (added code 14, Thai)
9	Other

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Spanish/Hispanic Origin**	1	SEER/COC	65 - 65

Description:

Code identifying persons of Spanish or Hispanic origin. Some registries code the information from the medical record; others code ethnicity based on Spanish names; others use a mixture of methods. Persons of Spanish or Hispanic origin may be of any race, but these categories are generally not used for Native Americans, Filipinos, etc., who may have Spanish surnames. All records for a patient should contain the same code. See chapter VI, Unresolved Issues, for a discussion of problems associated with this field. See also the field Name-Derived Ethnicity.

Rationale:

See the rationale for the item, "Race". Ethnic origin has a significant association with cancer rates and outcomes. Hispanic populations have a different pattern of occurrence of cancer than other populations that may be included in the "white" category of the item "Race".

Allowable Values and Format:

0-7, 9

Codes:

0	Non-Spanish; non-Hispanic
1	Mexican (includes Chicano)
2	Puerto Rican
3	Cuban
4	South or Central American (except Brazil)
5	Other Spanish (includes European)
6	Spanish, NOS; Hispanic, NOS (there is evidence other than surname or maiden name that the person is Hispanic, but he/she cannot be assigned to any of the categories 1-5)
7	Spanish surname only (only evidence of person's Hispanic origin is surname or maiden name)
9	Unknown whether Spanish or not

NOTE: Code 7 was adopted for use effective with 1994 diagnoses. It does not include computer assignment of ethnicity (but see Name Derived Ethnicity).

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Sex**	1	SEER/COC	66 - 66

Description:
Code for sex of the patient.

Rationale:
The codes presented below are the de facto standard, since they have been used historically by both SEER and the COC.

Allowable Values and Format:
1-4, 9

Codes:

1	Male
2	Female
3	Other (Hermaphrodite)
4	Transsexual
9	Not Stated

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Age at Diagnosis**	3	SEER/COC	67 - 69

Description:
Age of the patient at diagnosis, in completed years. Different tumors may have different values.

Allowable Values and Format:
Numeric, right-justified, zero filled

Codes:

000	Less than one year old
001	One year old, but less than two years
002	Two years old
...	
...	(Show actual age)
...	
101	One hundred one years old
...	
...	
...	
120	One hundred twenty years old
...	
...	
...	
999	Unknown Age

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Birth Date (Date of Birth)**	8	SEER/COC	70 - 77

Description:
Date of birth of the patient.

Allowable Values and Format:
SEE NOTE 1, page 42.

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Birthplace (Place of Birth)**	3	SEER/COC	78 - 80

Description:
Code for place of birth of the patient. All records for a patient should contain the same code.

Rationale:
Place of birth is helpful identifying information during patient matching, and is also useful when reviewing race and ethnicity codes. It can also be associated with certain cancer rates and outcomes.

Allowable Values and Format:
Numeric

Codes:
See Appendix B of the *SEER Program Code Manual* or COC DAM Part Four for numeric and alphabetic lists of places and codes.

✓	**Religion**	2		81 - 82

AACCR has not adopted standards for this item.

✓	**Occupation (Census)**	3	Census	83 - 85

Description:
Code for the patient's occupation, using Census Bureau codes. AACCR has not adopted standards for this item.

✓	**Industry (Census)**	3	Census	86 - 88

Description:
Code for the patient's industry, using Census Bureau codes. AACCR has not adopted standards for this item.

✓	**Occupation (SOC)**	4	Dept. of Commrc.	89 - 92

Description:
Code for the patient's occupation, using Department of Commerce codes. AACCR has not adopted standards for this item.

✓	**Industry (SIC)**	4	Dept. of Commrc.	93 - 96

Description:
Code for the patient's industry, using Department of Commerce codes. AACCR has not adopted standards for this item.

✓	**Occup/Ind Coding System**	1		97 - 97

AACCR has not adopted standards for this item.

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Smoking History**	1		98 - 98
	AACCR has not adopted standards for this item.			
✓	**Name-Derived Ethnicity**	2		99 - 100
	AACCR has not adopted standards for this item.			

IX.B Demographic Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	6		101 - 106

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Tumor Record Number**	2	AACCR	107 - 108

Description:
This is a system-generated number assigned to each tumor. It should never change, even when tumor sequence is changed or a tumor is deleted.

Rationale:
Incoming information on a specific tumor cannot be easily linked unless there is an unvarying record number associated with each tumor. Since Sequence Number can sometimes change, e.g. when a late report comes in on an earlier tumor, or a tumor is deleted, another number is needed which will not change.

Allowable Values and Format:
Numeric, 01-99 only

✓	**Sequence Number**	2	SEER/COC	109 - 110

Description:
Code indicating the sequencing of reportable neoplasms in the patient's lifetime. Each tumor is assigned a different number. Reportable neoplasms not included in the registry are also allotted a sequence number, so the registry may, for example, contain one record for a patient with a sequence number of 2. (See COC DAM for discussion of alphabetic codes used to sequence benign and borderline tumors.)

Allowable Values and Format:
Alphanumeric

Codes:

00	One primary only
01	First of two or more primaries
02	Second of two or more primaries
..	
..	(Actual number of this primary)
..	
10	Tenth of ten or more primaries
11	Eleventh of eleven or more primaries
..	
..	
99	Unspecified sequence number

NOTE: See COC DAM for discussion of alphabetic codes used to sequence benign and borderline tumors.

✓	**Date of Diagnosis (Date of Initial Diagnosis)**	8	SEER/COC	111 - 118

Description:
Date of initial diagnosis by a recognized medical practitioner for the tumor being reported. For more discussion on determining date of diagnosis, consult the SEER code manual or COC DAM.

Allowable Values and Format:
SEE NOTE 1, page 42.

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Primary Site**	4	SEER/COC	119 - 122

Description:
Code for the primary site of the tumor being reported. AACCR adopted ICD-O-2 as the standard coding system for cases diagnosed in 1992 and later.

Allowable Values and Format:
Alphanumeric, 'C' followed by 3 digits.

Codes:
See the *International Classification of Diseases for Oncology*, Second Edition (ICD-O-2), Topography Section for the codes for primary site.

✓	**Laterality** **(Laterality at Diagnosis)** **(Paired Organ)**	1	SEER/COC	123 - 123

Description:
Code for the laterality (side of paired organs, or of the body) on which the primary tumor being reported originated.

Allowable Values and Format:
0-4, 9

Codes:

0	Not a paired site
1	Right: origin of primary
2	Left: origin of primary
3	Only one side involved, right or left origin unspecified
4	Bilateral involvement, lateral origin unknown: stated to be single primary; including both ovaries involved simultaneously, single histology; bilateral retinoblastomas; bilateral Wilms's tumors
9	Paired site, but no information concerning laterality; midline tumor

✓	**Morphology -Type&Behavior** **(Morphology: Histologic Type and Behavior Code)** **(Histology)**	5	SEER/COC	124 - 128
	Subfields:			
	Histologic Type	4	SEER/COC	124 - 127
	Behavior Code	1	SEER/COC	128 - 128

Description:
Codes for the histologic type and behavior of the tumor being reported.

Allowable Values and Format:
Numeric

Codes:
See the *International Classification of Diseases for Oncology*, Second Edition (ICD-O-2), Morphology Section for histologic type and behavior codes.

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Grade** (Grade, Differentiation, or Cell Indicator)	1	SEER/COC	129 - 129

Description:
Code for the grade or degree of differentiation of the tumor being reported. For lymphomas and leukemias, field is used instead to indicate T- or B-Cell origin.

Allowable Values and Format:
1-7, 9

Codes:
See the *International Classification of Diseases for Oncology*, Second Edition (ICD-O-2), Morphology Section for codes.

✓	**Site Coding Sys. - Current**	1	AACCR	130 - 130

Description:
Code that best describes how primary site currently is coded. If converted, this field shows the system it is converted to.

Allowable Values and Format:
1-4, 9

Codes:

1	ICD-8
2	ICD-9 (First Edition)
3	ICD-O-1 (1976)
4	ICD-O-2 (1990)
9	Other

✓	**Site Coding Sys. - Original**	1	AACCR	131 - 131

Description:
Code that best describes how primary site was originally coded. If later converted, this field shows the original codes used.

Allowable Values and Format:
1-4, 9

Codes:

1	ICD8
2	ICD9 (First Edition)
3	ICDO (1976)
4	ICDO Second Edition (1990)
9	Other

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Morph Coding Sys. -Current**	1	AACCR	132 - 132

Description:
Code that best describes how morphology is currently coded. If converted, this field shows the system it is converted to.

Allowable Values and Format:
1-4, 9

Codes:

1	ICDO (First Edition) (1976)
2	ICDO 1986 Field Trial
3	ICDO 1988 Field Trial
4	ICDO Second Edition (1990)
9	Other

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Morph Coding Sys. -Original**	1	AACCR	133 - 133

Description:
Code that best describes how morphology was originally coded. If later converted, this field shows the original codes used.

Allowable Values and Format:
1-4, 9

Codes:

1	ICDO (First Edition) (1976)
2	ICDO 1986 Field Trial
3	ICDO 1988 Field Trial
4	ICDO Second Edition (1990)
9	Other

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Diagnostic Confirmation**	1	SEER/COC	134 - 134

Description:
Code for the best method of diagnostic confirmation of the cancer being reported at any time in the patient's history.

Rationale:
It is often useful to calculate rates based on microscopically-confirmed cancers. On the other hand, full incidence calculations must also include cases that are only confirmed clinically. The percent of cases that are clinically diagnosed only is an indication of whether casefinding is including sources outside of pathology reports

Allowable Values and Format:
1-2, 4-9

Codes:

1	Positive histology
2	Positive exfoliative cytology, no positive histology
4	Positive microscopic confirmation, method not specified
5	Positive laboratory test/marker study
6	Direct visualization without microscopic confirmation
7	Radiography and other imaging techniques without microscopic confirmation
8	Clinical diagnosis only (other than 5, 6, or 7)
9	Unknown whether or not microscopically confirmed

✓	**Type of Reporting Source**	1	SEER	135 - 135

Description:
Code reflecting source documents used to abstract the cancer being reported.

Rationale:
The code in this field can be used to explain why information may be incomplete on a case. The field is also used to monitor the success of non-hospital case reporting and follow-back mechanisms. All population-based registries should have some DC-only, autopsy-only, and MD-only case reports each year, for cases where no hospital admission was involved, but too high a percentage can imply that follow-back to uncover missed hospital reports was not complete.

Allowable Values and Format:
1, 3-7

Codes:

1	Hospital Inpatient/Outpatient or Clinic
3	Laboratory Only (Hospital or Private)
4	Physician's Office/Private Medical Practitioner (LMD)
5	Nursing/Convalescent Home/Hospice
6	Autopsy Only
7	Death Certificate Only

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Accession Year**	2	COC	136 - 137

Description:

The last two digits of the year in which the reportable tumor was first registered. If patient has multiple tumors, the accession year may differ.

Rationale:

This is used by hospital registries to organize their case reporting into individual years. It differs from the first two-digits of the Accession Number, since the year is case-specific rather than patient-specific, and from the diagnosis year since it relates to the specific facility and not the tumor. Central registries that wish to compare their data with hospital case lists can make use of this field to create equivalent reports.

Allowable Values and Format:

Numeric

IX.C Cancer Identification Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	4		138 - 141

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Reporting Hospital (DX) (Hospital ID Number)**	8	COC	142 - 149

Description:
Code for the facility reporting the case.

Rationale:
This number is used to identify the reporting hospital in the central registry database.

Allowable Values and Format:
6-digit number, right-justified, blank-filled.

Codes:
Codes assigned by COC.

✓	**Accession Number (Hosp)**	6	COC	150 - 155
	Subfields:			
	Year	2	COC	150 - 151
	Number	4	COC	152 - 155

Description:
Unique number assigned by the hospital registry to identify the patient. The first two digits identify the first year the patient was seen at that institution for diagnosis or treatment of cancer. The following four digits identify the numerical order in which the first cancer was entered into the registry. In multiple primaries, the accession number remains the same for each primary.

Rationale:
Hospitals use this to identify cases. If the central registry preserves this number, they can then refer to it when communicating with the hospital. It also provides a way to link computerized follow-up reports from hospitals into the central database.

Allowable Values and Format:
Year: Valid year. Number: Numeric.

✓	**Abstractor**	3		156 - 158

Description:
An alphanumeric field that identifies the individual abstracting the case.

AACCR has not adopted standards for this item.

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Date of Admission**	8	AACCR	159 - 166

Description:
Date of first admission to the facility for diagnosis and/or treatment of reportable tumor. This may be the date of an outpatient visit. If autopsy-only or death-certificate-only case, use date of death.

Allowable Values and Format:
SEE NOTE 1, page 42.

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Date of Discharge**	8	AACCR	167 - 174

Description:
Date the patient was discharged from the facility after diagnosis and/or treatment of reportable tumor. If OPD or 1-day admission, date of discharge and date of admission are the same. If autopsy-only or death-certificate-only case, use date of death.

Allowable Values and Format:
SEE NOTE 1, page 42.

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Class of Case**	1	COC	175 - 175

Description:

For a hospital registry, divides cases into those included in reports on patient outcome (analytic) and those that are not included (nonanalytic). Class of Case codes 0-2 are analytic, i.e., were diagnosed and/or received all or part of their first course of treatment or had treatment planning at the reporting hospital. Class of Case codes 3-5 are nonanalytic, i.e., received all of their first course of therapy at another institution or were diagnosed at autopsy.

Rationale:

This field is important to hospitals, since nonanalytic cases are handled differently from analytic cases. It can help central registries understand what parts of the cancer diagnosis and treatment were carried out at the reporting hospital, and what should have been reported elsewhere.

Allowable Values and Format:

0-5, 8, 9

Codes:

0	First diagnosed at the reporting institution since the reference date of the registry and all of the first course of therapy elsewhere
1	First diagnosed and all or part of the first course of therapy at the reporting institution
2	First diagnosed elsewhere and all or part of the first course of therapy planned or given at the reporting institution after the reference date of the registry
3	First diagnosed and all of the first course of therapy elsewhere
4	First diagnosed and first course of therapy at the reporting institution before the reference date of the registry
5	First diagnosed at autopsy
8	By death certificate only
9	Unknown

Note: Codes 8 and 9 are used in central registries only.

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

Note: The following rationale applies to all of the RX Hospital fields.

Rationale:

If central registries wish to study the treatment given at particular hospitals, these hospital-level treatment fields must be used. The summary treatment fields, conversely, combine information across all hospitals that provided first course treatment on the case. Hospital-specific fields allow studies of detailed referral patterns and treatment by type of hospital. Knowing what part of the treatment was given at a particular hospital also helps resolve coding issues. The codes are identical to those used in the summary treatment fields, except codes 9 and 09 do not apply here.

✓ **RX Hospital--Surgery Type** 2 AACCR 176 - 177

Description:

Describes surgical procedures used to diagnose and/or treat reportable tumor or to alleviate symptoms or pain caused by tumor. This item reflects that portion of the first course of treatment given at the reporting facility.

Allowable Values and Format:

Numeric only, site-specific

Codes:

No Cancer Directed Surgery (used for all sites):

00 No surgical procedure
01 Incisional, needle, or aspiration biopsy of other than primary site
02 Incisional, needle or aspiration biopsy of primary site.
03 Exploratory ONLY (no biopsy)
04 Bypass surgery, -ostomy ONLY (no biopsy)
05 Exploratory and incisional or needle biopsy of primary site or other sites
06 Bypass surgery, -ostomy and incisional or needle biopsy of primary site or other sites
07 Non-cancer-directed surgery, NOS
08 Reconstructive surgery (for subsequent therapy only)
09 Unknown if surgery done

Types of cancer-directed surgery:
Site-specific codes in the range 10-90

See Appendix C of the *SEER Program Code Manual* (Revised Edition, June 1992) or COC DAM Part 3 for detailed site-specific 2-digit codes.

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Hospital--Radiation**	1	AACCR	178 - 178

Description:
Defines the type of radiation therapy the patient received as a part of the initial treatment for the reportable tumor at the reporting facility.

Allowable Values and Format:
0-5, 7-9

Codes:

0	None
1	Beam radiation
2	Radioactive implants
3	Radioisotopes
4	Combination of 1 with 2 or 3
5	Radiation, NOS - method or source not specified
7	Patient or patient's guardian refused
8	Radiation recommended, unknown if administered
9	Unknown if radiation therapy administered

✓	**RX Hospital--Chemo**	1	AACCR	179 - 179

Description:
Defines the type of chemotherapy the patient received as a part of the initial treatment for the reportable tumor at the reporting facility.

Allowable Values and Format:
0-3, 7-9

Codes:

0	None
1	Chemotherapy, NOS
2	Chemotherapy, single agent
3	Chemotherapy, multiple agents (combination regimen)
7	Patient or patient's guardian refused
8	Chemotherapy recommended, unknown if administered
9	Unknown if chemotherapy administered

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Hospital--Hormone**	1	AACCR	180 - 180

Description:
Defines the type of hormone therapy the patient received as a part of the initial treatment for the reportable tumor at the reporting facility.

Allowable Values and Format:
0-3, 7-9

Codes:

0	None
1	Hormones (including NOS and antihormones)
2	Endocrine surgery and/or endocrine radiation (if cancer is of another site)
3	Combination of 1 and 2
7	Patient or patient's guardian refused
8	Hormonal therapy recommended, unknown if administered
9	Unknown if hormone therapy administered

✓	**RX Hospital--BRM**	1	AACCR	181 - 181

Description:
Defines the type of biological response modifier therapy the patient received as a part of their initial treatment for the reportable tumor at the reporting facility.

Allowable Values and Format:
0, 1, 7-9

Codes:

0	None
1	Biological response modifier
7	Patient or patient's guardian refused
8	Biological response modifier recommended, unknown if administered
9	Unknown if BRM therapy administered

IX.D Hospital-Specific Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Hospital--Other**	1	AACCR	182 - 182

Description:
Defines any and all cancer-directed treatment not appropriately assigned to other specific codes and also experimental cancer-directed therapy given at the reporting facility as part of the first course of treatment.

Allowable Values and Format:
0-3, 6-9

Codes:

0	No other cancer-directed therapy except as coded elsewhere
1	Other cancer-directed therapy
2	Other experimental cancer-directed therapy (not included elsewhere)
3	Double-blind study, code not yet broken
6	Unproven therapy (including laetrile, krebiozen, etc.)
7	Patient or patient's guardian refused therapy which would have been coded 1-3 above
8	Other cancer-directed therapy recommended, unknown if administered
9	Unknown if other cancer-directed therapy administered

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

See chapter VI, Unresolved Issues, for a discussion of standardization of stage information.

✓ **General Summary Stage** — 1 — **AACCR** — **183 - 183**

Description:

Code for general summary stage at the initial diagnosis or treatment of the reportable tumor. This has traditionally been used by central registries to monitor time trends. For hospital registries, COC requires its use in the absence of a defined TNM scheme. For site-specific definitions of categories, see SEER *Summary Staging Guide*.

Rationale:

Stage information is important when evaluating the effects of cancer control programs. It is crucial in understanding whether changes over time in incidence rates or outcomes are due to earlier detection of the cancers. In addition, cancer treatment cannot be studied without knowing the stage at diagnosis.

To study historical trends in stage, the coding system must be relatively unchanged (stable) over time. The AJCC'S TNM system is updated periodically to maintain clinical relevance and maintain parity with changes in diagnosis and treatment. The surveillance registries often rely on the General Summary Stage, which they consider to be more "stable". General Summary Stage has been in widespread use, either as the primary staging scheme or a secondary scheme, in most central and hospital-based registries since 1976.

Allowable Values and Format:

0-5, 7, 9

Codes:

0	In situ
1	Localized
2	Regional by direct extension
3	Regional to lymph nodes
4	Regional (both 2 and 3)
5	Regional NOS
7	Distant metastases/systemic disease
9	Unstaged, unknown, or unspecified

✓ **Loc/Reg/Distant Stage** — 1 — **AACCR** — **184 - 184**

Description:

For use if no other staging is available.

Allowable Values and Format:

0-3, 9

Codes:

0	In situ
1	Local
2	Regional
3	Distant
9	Unstaged

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Extent of Disease** **(EOD)** **(10-Digit Extent of Disease -- 1988)** **(SEER Extent of Disease 1988)** **(Extent of Disease at Initial Diagnosis)**	10	SEER	185 - 194
	Subfields:			
	EOD--Tumor Size **(Size of Primary Tumor)** **(Size of Tumor)**	3	SEER/COC	185 - 187
	EOD--Extension **(Extension)**	2	SEER	188 - 189
	EOD--Lymph Node Involv. **(Lymph Nodes)**	1	SEER	190 - 190
	Regional Nodes Positive **(Number of Regional Nodes Positive)** **(Number of Positive Regional Lymph Nodes)** **(Pathology Review of Regional Lymph Nodes)**	2	SEER/COC	191 - 192
	Regional Nodes Examined **(Number of Regional Nodes Examined)** **(Number of Regional Lymph Nodes Examined)** **(Pathology Review of Regional Lymph Nodes)**	2	SEER/COC	193 - 194

Description:

Detailed site-specific codes for anatomic extent of disease used by SEER for cases diagnosed from 1988 forward. The subfields Regional Nodes Positive, Regional Nodes Examined, and Tumor Size are also included in the COC data set, separate from extent of disease.

Rationale:

Site-specific EOD codes provide extensive detail describing disease extent. The EOD codes can be grouped into different stage categories for analysis, e.g., summary stage categories consistent with those used in published SEER data since 1973, and more recently, AJCC stage groupings. The codes are updated as needed, but updates are usually backward-compatible with old categories. See *Comparative Staging Guide for Cancer.*

Allowable Values and Format:

Numeric

Codes:

See *SEER Extent of Disease 1988: Codes and Coding Instructions* (Second Edition, June 1992) or COC DAM Part Three for site-specific codes.

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**EOD - Old 13 Digit (13-Digit (Expanded) Site-Specific Extent of Disease (1973-82))**	13	SEER	195 - 207

Description:
Detailed site-specific codes for anatomic extent of disease used by SEER for selected sites of cancer for cases diagnosed 1973 - 1982.

Allowable Values and Format:
Numeric plus special characters

Codes:
See former EOD code manual for codes.

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**EOD - Old 2 Digit (2-Digit Nonspecific and 2-Digit Site-Specific Extent of Disease (1973-82))**	2	SEER	208 - 209

Description:
Site-specific codes for anatomic extent of disease used by SEER for cases diagnosed from 1973 - 1982, for cancer sites which did not have a 13-digit scheme.

Allowable Values and Format:
Numeric, plus special characters

Codes:
See former EOD code manual for codes.

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**EOD - Old 4 Digit (4-Digit Extent of Disease (1983-87))**	4	SEER	210 - 213

Description:
Codes for anatomic extent of disease used by SEER for cases diagnosed from 1983 - 1987, all sites.

Allowable Values and Format:
Numeric

Codes:
See former EOD code manual for codes.

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**TNM**	14	AJCC	214 - 227
	Subfields:			
	pAJCC:			
	T code (Path-based) (T)	2	AJCC	214 - 215
	N Code (Path-based) (N)	2	AJCC	216 - 217
	M Code (Path-based) (M)	1	AJCC	218 - 218
	AJCC Summary Stage (Path) (AJCC Stage Group)	2	AJCC	219 - 220
	cAJCC:			
	T Code (Clinical) (T)	2	AJCC	221 - 222
	N Code (Clinical) (N)	2	AJCC	223 - 224
	M Code (Clinical) (M)	1	AJCC	225 - 225
	AJCC Summary Stage (Clin) (AJCC Stage Group)	2	AJCC	226 - 227

Description:

Detailed site-specific codes for Tumor (T), Nodes (N), and Metastases (M) as defined by the American Joint Committee on Cancer. Clinical and pathological stage data are given separate fields in the AACCR Data Exchange Record Layout.

Rationale:

The COC requires that AJCC TNM staging be used in its approved cancer programs. The AJCC developed its staging system for evaluating trends in the treatment and control of cancer. This staging is used by physicians to estimate prognosis, to plan treatment, to evaluate new types of therapy, to analyze outcome, to design follow-up strategies, and to assess early detection results.

Allowable Values and Format:

Alphanumeric, plus blanks. Convert AJCC Roman numerals to Arabic numerals. Use upper case for stage group and mixed case for letters in the TNM elements. For TNM elements, truncate least significant subdivision of the category from the right as needed. If an entry is fewer characters than allowed, left-justify and blank fill.

Codes:

See AJCC's *Manual for Staging of Cancer*.

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**TNM Edition Number**	1	AACCR	228 - 228

Description:
Code indicating the edition of the AJCC manual used to stage the case.

Rationale:
TNM codes have changed over time and conversion is not always simple. Therefore a case-specific indicator is needed to allow grouping of cases for comparison.

Allowable Values and Format:
2-4, 9

Codes:

2	Second Edition (published 1983)
3	Third Edition (published 1988)
4	Fourth Edition (published 1992), recommended for use for cases diagnosed 1993+
9	Edition unknown

✓	**Distant Metastasis (Site or Sites of Distant Metastasis)**	3	COC	229 - 231

Description:
Codes for up to three sites of distant metastasis. Treated as 3 individual subfields, each with a one-digit code for a site of metastasis.

Allowable Values and Format:
0-9 in each of three positions. If fewer than 3 sites, left-justify and zero-fill.

Codes:

0	None
1	Peritoneum
2	Lung
3	Pleura
4	Liver
5	Bone
6	Central nervous system
7	Skin
8	Lymph nodes (distant)
9	Other, generalized, carcinomatosis, not specified, unknown
000	No distant metastasis

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Residual Tumor**	1	COC	232 - 232

Description:
Reflects status of primary tumor after surgical resection.

Allowable Values and Format:
0-2, 8, 9

Codes:
0 No residual tumor
1 Microscopic residual tumor
2 Macroscopic residual tumor
8 Not applicable
9 Unknown

✓	**Alternate Staging Scheme**	4		233 - 236

Description:
Field for entry of stage of disease in using another scheme, such as Dukes for colon. Left justified. AACCR has not adopted standards for this item.

✓	**Tumor Marker 1** **(Tumor Marker One (Estrogen Receptor Assay))** **(Estrogen Receptor Status)**	1	SEER/COC	237 - 237

Description:
For breast cancer only, code for the estrogen receptor status of the tumor.

Allowable Values and Format:
0-3, 8, 9

Codes:
For Breast Cases Only:
0 None done
1 Positive
2 Negative
3 Borderline; undetermined whether positive or negative
8 Ordered, but results not in chart
9 Unknown or no information

For All Other Cases:
9 Not applicable

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Tumor Marker 2** **(Tumor Marker Two (Progesterone Receptor Assay))** **(Progesterone Receptor Status)**	1	SEER/COC	238 - 238

Description:
For breast cancer only, code for the progesterone receptor status of the tumor.

Allowable Values and Format:
0-3, 8, 9

Codes:
For Breast Cases Only:
0 None done
1 Positive
2 Negative
3 Borderline; undetermined whether positive or negative
8 Ordered, but results not in chart
9 Unknown or no information

For All Other Cases:
9 Not applicable

IX.E Stage/Other Prognostic Factors Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	12		239 - 250

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Coding System - Current**	1	AACCR	251 - 251

Description:
Code describing how treatment type for this case is now coded.

Allowable Values and Format:
1, 2, 9

Codes:
1 1-digit surgery codes
2 2-digit surgery codes (1989+ SEER and COC manuals)
9 Other

✓	**RX Dates**	48	AACCR	252 - 299
	Subfields:			
	RX Date--Surgery	8	AACCR	252 - 259
	RX Date--Radiation	8	AACCR	260 - 267
	RX Date--Chemotherapy	8	AACCR	268 - 275
	RX Date--Hormone Therapy	8	AACCR	276 - 283
	RX Date--BRM	8	AACCR	284 - 291
	RX Date--Other	8	AACCR	292 - 299

Description:
Date of initiation for each separate modality of treatment that is part of the first course of treatment. AACCR has not adopted standards for these items, but SEE NOTE 1, page 42.

Rationale:
It is useful to record separately the dates on which different treatment modalities were started. It helps when evaluating whether a treatment was part of the first course of therapy and when checking the correctness of the Radiation Sequence with Surgery coding.

✓	**RX Date--Started** **(Date Therapy Initiated)** **(Date Started)**	8	SEER/COC	300 - 307

Description:
Date of initiation of the first cancer-directed therapy for the cancer being reported.

Allowable Values and Format:
SEE NOTE 1, page 42.

Special Codes:
00000000 No cancer-directed therapy
99999999 Unknown if any cancer-directed therapy was administered

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Summary--Surgery Type (Site-Specific Surgery) (Surgery)**	2	SEER/COC	308 - 309

Description:
Site-specific codes for the type of surgery performed as part of the first course of treatment. This includes treatment given at all facilities as part of the first course of treatment.

Allowable Values and Format:
Numeric, site-specific

Codes:
No Cancer Directed Surgery (used for all sites):
00 No surgical procedure
01 Incisional, needle, or aspiration biopsy of other than primary site
02 Incisional, needle or aspiration biopsy of primary site.
03 Exploratory ONLY (no biopsy)
04 Bypass surgery, -ostomy ONLY (no biopsy)
05 Exploratory and incisional or needle biopsy of primary site or other sites
06 Bypass surgery, -ostomy and incisional or needle biopsy of primary site or other sites
07 Non-cancer-directed surgery, NOS
08 Reconstructive surgery (for subsequent therapy only)
09 Unknown if surgery done

Types of cancer-directed surgery:
Site-specific codes in the range 10-90

See Appendix C of the *SEER Program Code Manual* (Revised Edition, June 1992) or COC DAM Part Three for detailed site-specific 2-digit codes.

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Reason for No Surgery (Reason for No Cancer-Directed Surgery)**	1	SEER/COC	310 - 310

Description:
Code for the reason no cancer-directed surgery was performed.

Allowable Values and Format:
0-2, 6-9

Codes:
0 Cancer-directed surgery performed
1 Cancer-directed surgery not recommended
2 Contraindicated due to other conditions; Autopsy Only case
6 Unknown reason for no cancer-directed surgery
7 Patient or patient's guardian refused
8 Recommended, unknown if done
9 Unknown if cancer-directed surgery performed; Death Certificate Only case

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Summary--Radiation** **(Radiation)** **(Radiation Therapy)**	1	SEER/COC	311 - 311

Description:
Codes for the type of radiation therapy performed as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0-5, 7-9

Codes:

0	None
1	Beam radiation
2	Radioactive implants
3	Radioisotopes
4	Combination of 1 with 2 or 3
5	Radiation, NOS - method or source not specified
7	Patient or patient's guardian refused
8	Radiation recommended, unknown if administered
9	Unknown if radiation administered

✓	**RX Summary--Rad to CNS** **(Radiation to the Brain and/or Central Nervous System)** **(Radiation Therapy to the Brain or Central Nervous System)**	1	SEER/COC	312 - 312

Description:
For lung and leukemia cases, codes for radiation given to the brain or central nervous system. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0, 1, 7-9

Codes:

For Lung and Leukemia Cases Only:

0	No radiation to the brain and/or central nervous system
1	Radiation
7	Patient or patient's guardian refused
8	Radiation recommended, unknown if administered
9	Unknown

For All Other Cases:

9	Not applicable

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Summary--Surg/Rad Seq (Radiation Sequence with Surgery) (Radiation Therapy Sequence with Surgery)**	1	SEER/COC	313 - 313

Description:
Codes for the sequencing of radiation and surgery given as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0, 2-6, 9

Codes:

0	No radiation and/or cancer-directed surgery
2	Radiation before surgery
3	Radiation after surgery
4	Radiation both before and after surgery
5	Intraoperative radiation
6	Intraoperative radiation with other radiation given before or after surgery
9	Sequence unknown, but both surgery and radiation were given

✓	**RX Summary--Chemo (Chemotherapy)**	1	SEER/COC	314 - 314

Description:
Codes for chemotherapy given as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0-3, 7-9

Codes:

0	None
1	Chemotherapy, NOS
2	Chemotherapy, single agent
3	Chemotherapy, multiple agents (combination regimen)
7	Patient or patient's guardian refused
8	Chemotherapy recommended, unknown if administered
9	Unknown if chemotherapy administered

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Summary--Hormone (Endocrine (Hormone/Steroid) Therapy) (Hormone/Steroid (Endocrine) Therapy)**	1	SEER/COC	315 - 315

Description:
Codes for the type of hormonal treatment given as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0-3, 7-9

Codes:

0	None
1	Hormones (including NOS and antihormones)
2	Endocrine surgery and/or endocrine radiation (if cancer is of another site)
3	Combination of 1 and 2
7	Patient or patient's guardian refused
8	Hormonal therapy recommended, unknown if administered
9	Unknown if hormonal therapy administered

✓	**RX Summary--BRM (Biological Response Modifiers) (Biological Response Modifier Therapy)**	1	SEER/COC	316 - 316

Description:
Codes for the type of biological-response-modifier therapy given as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0, 1, 7-9

Codes:

0	None
1	Biological response modifier
7	Patient or patient's guardian refused
8	Biological response modifier recommended, unknown if administered
9	Unknown if BRM administered

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Summary--Other (Other Cancer-Directed Therapy)**	1	SEER/COC	317 - 317

Description:
Codes for other treatment performed as part of the first course of treatment. Includes treatment given at all facilities as part of the first course.

Allowable Values and Format:
0-3, 6-9

Codes:

0	No other cancer-directed therapy except as coded elsewhere
1	Other cancer-directed therapy
2	Other experimental cancer-directed therapy (not included elsewhere)
3	Double-blind study, code not yet broken
6	Unproven therapy (including laetrile, krebiozen, etc.)
7	Patient or patient's guardian refused therapy which would have been coded 1-3 above
8	Other cancer-directed therapy recommended, unknown if administered
9	Unknown if other cancer-directed therapy administered

✓	**First Course Calc. Method**	1	AACCR	318 - 318

Description:
Code indicating how the time interval for limiting first course of therapy is calculated.

Allowable Values and Format:
1, 2, 9

Codes:

1	Defined from diagnosis date (COC)
2	Defined from treatment start date (SEER)
9	Other, unknown

IX.F Treatment - 1st Course Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	18		319 - 337

IX.G Treatment - Subsequent Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Subsequent Treatment**	60	COC	338 - 397
	Four areas, 15 positions each, for 4 separate sets or courses of subsequent therapy, with subfields in each area, as follows:			
	Area #1:	15	COC	338 - 352
	Subfields:			
	Subseq RX--Date	8	COC	338 - 345
	Subseq RX--Surg	2	COC	346 - 347
	Subseq RX--Rad	1	COC	348 - 348
	Subseq RX--Chemo	1	COC	349 - 349
	Subseq RX--Hormone	1	COC	350 - 350
	Subseq RX--BRM	1	COC	351 - 351
	Subseq RX--Other	1	COC	352 - 352
	Area #2:	15	COC	353 - 367
	(Subfields as in Area #1)			
	Area #3:	15	COC	368 - 382
	(Subfields as in Area #1)			
	Area #4:	15	COC	383 - 397
	(Subfields as in Area #1)			

Description:

Fields for coding up to four sets or courses of treatments administered subsequent to the first course. Each set has an associated date and 1 to 6 different treatment modalities.

Allowable Values and Format:

Subseq RX--Date: SEE NOTE 1, page 42.

For the six treatment modalities, the codes are identical to those used in the corresponding fields of RX--Summary, columns 308 - 317, except there are no fields in the subsequent treatment area corresponding to "Reason for No Surgery" or "RX Summary--Rad to CNS".)

Special Codes:

For Subseq RX--Date fields:

00000000 No subsequent therapy

99999999 Unknown if any subsequent therapy

IX.G Treatment - Subsequent Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	20		398 - 417

IX.H Follow-up/Recurrence Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Date of Last Contact (Date of Last Follow-Up or of Death)**	8	SEER/COC	418 - 425

Description:
Date of last contact with the patient, or date of death.

Rationale:
Death clearance is required of all incidence registries to complete case identification. This field is used to store the date of death when a match is found. For registries that perform follow-up, the entry in this field is used to calculate survival time, for both living and deceased patients. It is helpful to hospital registries when central registries can report dates of death to them for their own follow-up purposes.

Allowable Values and Format:
SEE NOTE 1, page 42.

✓	**Vital Status**	1	COC	426 - 426

Description:
Vital status of the patient as of the date entered in "Date of Last Contact".

Allowable Values and Format:
0,1

Codes:
0 Dead
1 Alive

✓	**Tumor Status (Cancer Status)**	1	COC	427 - 427

Description:
Records the cancer status for this primary as of the date entered in "Date of Last Contact". If patient has multiple primaries, may be different.

Rationale:
Hospitals use this field to compute survival analysis (disease-free intervals). By maintaining it, central registries can help hospitals share follow-up information between hospitals that see the same case.

Allowable Values and Format:
1, 2, 9

Codes:
1 No evidence of this cancer
2 Evidence of this cancer
9 Unknown, indeterminate whether this cancer is present

IX.H Follow-up/Recurrence Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Quality of Survival (Performance Indicator)**	1	COC	428 - 428

Description:
Records patient's ability to carry on the activities of daily living at the date of last contact.

Allowable Values and Format:
0-4, 8, 9

Codes:

0	Normal Activity
1	Symptomatic and ambulatory
2	Ambulatory more than 50% of the time, occasionally needs assistance
3	Ambulatory less than 50% of the time, nursing care needed
4	Bedridden, may require hospitalization
8	Not applicable, dead
9	Unknown or unspecified

✓	**Last Type of Follow-Up (Follow-Up Method)**	1	COC	429 - 429

Description:
Records the source from which the latest follow-up information was obtained.

Rationale:
For registries performing follow-up, this field helps evaluate the success rates of various methods of follow-up. It also can be used to report to institutions the source of follow-up information that is sent to them. When there is a conflict in follow-up information, knowing the source can help resolve the inconsistency.

Allowable Values and Format:
0-5, 7-9

Codes:

0	Reported hospitalization
1	Readmission
2	Physician
3	Patient
4	Department of Motor Vehicles
5	Medicare/Medicaid file
7	Death Certificate
8	Other
9	Unknown

IX.H Follow-up/Recurrence Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

Note: The nonconfidential parts of the follow-up contact address are in this section of the document, and the confidential parts are in IX.M, the Patient-Confidential Section. The follow-up contact address is the current address that will be used to send a follow-up inquiry, usually the patient's current address, but may be that of another contact. It should correspond to the Follow-up Contact Name, col. 991 - 1020. The following rationale applies to all of the follow-up contact fields:

Rationale:
Sometimes central registries carry out follow-up by contacting the patients via letter or phone call to ascertain their vital status. The most current reported address and phone number are needed. This information is also useful for conducting interviewing studies.

✓ **Follow-up Contact City** | **20** | **AACCR** | **430 - 449**

Description:
Name of city to be used in the follow-up contact address. If the patient has multiple tumors, the city of residence should be the same.

Allowable Values and Format:
Alpha only, no special characters, mixed case left justified, blank filled; except if unknown, leave blank.

✓ **Follow-up Contact State** | **2** | **AACCR** | **450 - 451**

Description:
U.S. Postal Service abbreviation for the state, commonwealth or country to be used in the follow-up contact address. If a commonly accepted two-letter abbreviation does not exist for the country, leave blank. If the patient has multiple tumors, the states should be the same. No standard has yet been set for Canadian Provinces.

Allowable Values and Format:
Alpha only, upper case, or all blank

✓ **Follow-up Contact Zip** | **9** | **AACCR** | **452 - 460**

Description:
United States Postal Service zip code of the address in the follow-up contact address. If the patient has multiple tumors, the zip codes should be the same. No standard has yet been set for entry of Canadian postal codes.

Allowable Values and Format:
5- or 9-digit zip code, or blanks. Blanks follow the 5 digit code.

Special Codes:
blank Unknown, or resident of foreign country

IX.H Follow-up/Recurrence Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Recurrence--Date (First)** (Date of First Recurrence)	8	COC	461 - 468

Description:
The date of the first recurrence of this reportable tumor.

Allowable Values and Format:
SEE NOTE 1, page 42.

✓	**Recurrence--Type (First)** (Type of First Recurrence)	1	COC	469 - 469

Description:
Code for the type of first recurrence.

Allowable Values and Format:
0-4, 9

Codes:

0	No recurrence
1	Local recurrence
2	Regional recurrence
3	Distant recurrence
4	Never disease-free
9	Unknown

✓	**Recurrence--Distant Sites** (Distant Site or Sites of First Recurrence)	3	COC	470 - 472

Description:
Codes for the distant site or sites in which the reportable tumor has recurred. This is treated as 3 separate 1-digit fields.

Allowable Values and Format:
For each character position: 0-9. Left-justify, zero fill if fewer than 3 distant sites of recurrence.

Codes:

0	None
1	Peritoneum
2	Lung
3	Pleura
4	Liver
5	Bone
6	Central Nervous System
7	Skin
8	Lymph Nodes (Distant)
9	Other, generalized, NOS

IX.I Death Information Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Cause of Death (Underlying Cause of Death)**	4	SEER/COC	473 - 476

Description:
Official cause of death as coded on the death certificate.

Rationale:
Cause of death is used for calculation of adjusted survival rates by the life table method. The adjustment corrects for deaths from competing causes. (See reference 5, pages 16 ff.)

Allowable Values and Format:
Valid ICDA-8, ICD-9, or ICD-10 cause-of-death codes. See SEER Code Manual or DAM for additional instructions.

Special Codes:

0000	Patient alive at last contact
7777	State death certificate not available
7797	State death certificate available but underlying cause of death is not coded

✓	**ICD Revision Number (ICD Code Revision Used for Cause of Death) (ICD Code Revision)**	1	SEER/COC	477 - 477

Description:
Indicator for the coding scheme used to code the cause of death.

Allowable Values and Format:
0, 1, 8, 9

Codes:

0	Patient Alive at Last Follow-Up
1	ICD-10
8	ICDA-8
9	ICD-9

✓	**Autopsy**	1	AACCR	478 - 478

Description:
Code indicating whether or not an autopsy was performed.

Allowable Values and Format:
0, 1, 2, or 9

Codes:

0	Alive
1	Dead, with autopsy
2	Dead, no autopsy
9	Dead, unknown if autopsy

IX.I Death Information Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Place of Death**	3	AACCR	479 - 481

Description:
State or country where patient died.

Rationale:
This field also helps carry out death clearance. When a hospital reports a place of death, the information can help in death certificate matching. It also can signal an out-of-state death for which the death certificate is to be requested.

Allowable Values and Format:
Numeric. SEER 3-digit place-of-birth codes. See Appendix B of *SEER Program Code Manual* or the COC DAM Part Four.

Special Codes:
000 Alive
999 Place of death unknown

IX.J Edit Overrides/Conversion History Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Over-ride Flags**	9	SEER	482 - 490
	Nine Separate Fields:			
	Over-ride Age/Site (Age/Site/Histology Interfield Review)	1	SEER	482 - 482
	Over-ride SeqNo/DxConf (Sequence Number/Diagnostic Confirmation Interfield Review)	1	SEER	483 - 483
	Over-ride Site/Lat/SeqNo (Site/Histology/Laterality/Sequence Interrecord Review)	1	SEER	484 - 484
	Over-ride Surg/DxConf (Surgery/Diagnostic Confirmation Interfield Review)	1	SEER	485 - 485
	Over-ride Site/Type (Site/Type Interfield Review)	1	SEER	486 - 486
	Over-ride Histology (Histology/Behavior Interfield Review)	1	SEER	487 - 487
	Over-ride Report Source (Type of Reporting Source/Sequence Number Interfield Review)	1	SEER	488 - 488
	Over-ride Ill-define Site (Sequence Number/Ill-defined Site Interfield Review)	1	SEER	489 - 489
	Over-ride Leuk, Lymphoma (Leukemia or Lymphoma/Diagnostic Confirmation Interfield Review)	1	SEER	490 - 490

Description:
Nine flags used to override certain interfield and interrecord edits defined by SEER.

Allowable Values and Format:
Blank, 1

Codes:
blank Not reviewed
1 Reviewed

IX.J Edit Overrides/Conversion History Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Site (1973-91)** **(Primary Site (1973-91))**	4	SEER	491 - 494

Description:
Area for retaining the primary site code entered before a conversion to ICD-O-2.

Allowable Values and Format:
Numeric or all blank

Codes:
For cases diagnosed before 1992, contains the ICD-O-1 site code as originally coded, if available. Blank for cases coded directly into ICD-O-2, i.e., 1992 and later cases.

✓	**Morph (1973-91)** **(Morphology (1973-91))**	6	SEER	495 - 500

Description:
Area for retaining the morphology code entered before a conversion to ICD-O-2. Includes 4 digits for histologic type, 1 digit for behavior code, and 1 digit for grade code.

Allowable Values and Format:
Numeric or all blank

Codes:
For cases diagnosed before 1992, contains the ICD-O-1 or field trial 6-digit morphology code as originally coded, if available. Blank for cases coded directly into ICD-O-2, i.e., 1992 and later cases.

✓	**ICDO-2 Conversion Flag** **(Review Flag for 1973-91 Cases)**	1	SEER	501 - 501

Description:
Code specifying how the conversion of site and morphology codes from ICD-O-1 and the field trial editions to ICD-O-2 was accomplished.

Allowable Values and Format:
0-4

Codes:

0	Primary site and morphology originally coded in ICD-O-2
1	Primary site and morphology converted without review
2	Primary site converted with review; morphology machine converted without review
3	Primary site machine converted without review; morphology converted with review
4	Primary site and morphology converted with review

IX.J Edit Overrides/Conversion History Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	1		502 - 502

IX.K System Administration Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Date Case Completed**	8		503 - 510

Description:
The date that 1) the abstractor decided that the case report was complete, and 2) the case passed all edits that were applied.

Allowable Values and Format:
See NOTE 1, page 42. Local use field.

✓	**Date Case Last Changed**	8		511 - 518

Local use field, but see NOTE 1, page 42.

✓	**General Coding Procedure**	2		519 - 520

Local use field.

✓	**SEER Coding Sys. - Current**	1	AACCR	521 - 521

Description:
This shows the SEER coding system best describing the majority of SEER items as they now are in the record (after conversion).

Allowable Values and Format:
0-4

Codes:

0	No SEER Coding
1	1987 SEER Coding Manual
2	May 1988 SEER Coding Manual
3	Jan 1989 SEER Coding Manual
4	Jan 1992 SEER Coding Manual

✓	**SEER Coding Sys. - Original**	1	AACCR	522 - 522

Description:
This shows the SEER coding system best describing the way the majority of SEER items in the record were originally coded.

Allowable Values and Format:
0-4

Codes:

0	No SEER Coding
1	1987 SEER Coding Manual
2	May 1988 SEER Coding Manual
3	Jan 1989 SEER Coding Manual
4	Jan 1992 SEER Coding Manual

IX.K System Administration Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**ACoS Coding Sys. - Current**	1	AACCR	523 - 523

Description:
Code for the ACoS CoC coding system currently used in the record.

Allowable Values and Format:
0-4

Codes:

0	No ACoS Coding
1	Pre-1988 (Cancer Program Manual Supplement)
2	1988 Data Acquisition Manual
3	1989 Data Acquisition Manual revisions
4	1992 Data Acquisition Manual

✓	**ACoS Coding Sys. - Original**	1	AACCR	524 - 524

Description
Code for the ACoS CoC coding system used to originally code the items.

Allowable Values and Format:
0-4

Codes:

0	No ACoS Coding
1	Pre-1988 (Cancer Program Manual Supplement)
2	1988 Data Acquisition Manual
3	1989 Data Acquisition Manual revisions
4	1992 Data Acquisition Manual

✓	**Subseq Report for Primary**	1		525 - 525

Description:
Indicator for known duplicate case reports.

AACCR has not adopted standards for this item.

✓	**Vendor Name**	10	AACCR	526 - 535

Description:
System-generated. Name of the computer services vendor who programmed the system submitting this data. Abbreviate as necessary, and keep a consistent name throughout all submissions. Include software version number where available.

Rationale:
This is used to track which vendor and which software version submitted the case. It helps define the source and extent of a problem discovered in data submitted by a software provider.

Allowable Values and Format:
Alphanumeric, spaces

IX.K System Administration Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**AACCR Record Version**	1	AACCR	536 - 536

Description:
Code for the AACCR record definition which was used to create the record.

Allowable Values and Format:
Blank, 1

Codes:

blank	September 1989 version
1	1992 version

IX.K System Administration Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	14		537 - 550

IX.L Special Use Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Site-Specific Studies**	200		551 - 750
	Description: Reserved for special studies. COC intends to use this area for Patient Care Evaluation Studies.			
✓	**State-Specific Items**	100		751 - 850
	Description: Defined by individual states or central registries.			

IX.M Patient - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

Note: The Patient-Confidential Section contains all fields that can be used to identify the individual. This includes the patient's name and identifying numbers, and also the most specific parts of the address at diagnosis and the follow-up contact address. The other fields needed to complete the addresses are in IX.B, Demographic Section and IX.H, Follow-up Recurrence Section.

✓ **Patient's Last Name** 15 COC 851 - 865
(Last Name)

Description:
Last name of the patient.

Allowable Values and Format:
Alpha only, no embedded spaces, no special characters, left-justified, blank filled

---

✓ **Patient's First Name** 14 COC 866 - 879
(First Name)

Description:
First name of the patient.

Allowable Values and Format:
Alpha only, no embedded spaces, no special characters, left-justified, blank filled.

✓ **Patient's Middle Initial** 1 COC 880 - 880
(Middle Initial)

Description:
Middle initial of the patient.

Allowable Values and Format:
Alpha, or blank if none or unknown.

✓ **Alias** 15 COC 881 - 895

Description:
Records an alternate name or "AKA" used by the patient, if known. Note that maiden name is entered in columns 1062 - 1076.

Rationale:
This is used to match reports on the same patient using different aliases.

Allowable Values and Format:
Alpha, left-justified and blank filled, or all blank if none or unknown

---

✓ **Spouse/Parent Name** 50 896 - 945

AACCR has not adopted standards for this item.

IX.M Patient - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Medical Record Number**	11	COC	946 - 956

Description:
Records medical record number used by the facility to identify the patient.

Rationale:
This number identifies the patient in a facility. It can be used by a central registry to point back to the patient record, and it helps identify multiple reports on the same patient.

Allowable Values and Format:
Alphanumeric plus blanks, right-justify, or all blank

✓	**Social Security Number**	9	COC	957 - 965

Description:
Records patient's social security number. The number is entered without dashes and without any letter suffix. This is not always identical to the Medicare claim number.

Allowable Values and Format:
Numeric (dashes are not entered)

Special Codes:
999999999 Unknown

✓	**Addr at Dx--No & Street (Number and Street)**	25	COC	966 - 990

Description:
The number and street address or the rural mailing address of the patient's residence at the time the reportable tumor was diagnosed. If multiple tumors, data may differ.

Allowable Values and Format:
Alphanumeric plus spaces, left justified

IX.M Patient - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Follow-up Contact Name**	30	AACCR	991 - 1020

Description:
Name that will be used to generate a follow-up inquiry. Usually the patient's name, but may be that of another contact. Must correspond to the Follow-up Contact Address fields.

Allowable Values and Format:
Free text

✓	**Follow-up Contact Address**	25	AACCR	1021 - 1045

Description:
The number and street address or the rural mailing address to be used to generate a follow-up inquiry. Must correspond to the other fields in the contact address. Usually the patient's current street address, but may be that of another contact. If multiple tumors, data should be the same.

Allowable Values and Format:
Free text

✓	**Patient Phone Number (Telephone)**	10	COC	1046-1055

Description:
Current telephone number with area code for the patient, or sometimes for another contact.

Allowable Values and Format:
All numeric or all blank.

✓	**DC State File Number**	6	State	1056 - 1061

Description:
Death certificate identification number as assigned by the state's vital statistics office.

✓	**Maiden Name**	15	AACCR	1062 - 1076

Description:
Maiden name of the patient.

Rationale:
This is used to link reports on a woman who changed her name between reports. It also is critical when using Spanish surname algorithms to categorize ethnicity.

Allowable Values and Format:
Alpha only, no embedded spaces, no special characters, left-justified, blank filled; or all blank if unknown or not applicable.

IX.M Patient - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion			1077 - 1081

IX.N Hospital - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Hospital Referred From**	8	COC	1082-1089

Description:
Records hospital where the patient was diagnosed, or received any initial treatment for this reportable tumor before admission to the reporting hospital.

Allowable Values and Format:
6-digit number, right-justified, blank-filled; or all blanks

Codes:
Codes assigned by COC.

✓	**Hospital Referred To**	8	COC	1090-1097

Description:
Records hospital where the patient referred for definitive treatment for this reportable tumor after discharge from the reporting hospital.

Allowable Values and Format:
6-digit number, right-justified, blank-filled; or all blanks

Codes:
Codes assigned by COC.

✓	**Last Follow-Up Hospital**	8	AACCR	1098-1105

Description:
Records hospital where the patient was last followed.

Allowable Values and Format:
6-digit number, right-justified, blank-filled; or all blanks

Codes:
Codes assigned by COC.

✓	**Next Follow-Up Hospital**	8	AACCR	1106-1113

Description:
Records hospital where the patient will next be followed.

Allowable Values and Format:
6-digit number, right-justified, blank-filled; or all blanks

Codes:
Codes assigned by COC.

IX.O Physician - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Physician (Attending)**	8	AACCR	1114-1121

Description:
Generally used to identify the physician who is primarily responsible for the management of the patient's cancer. Registry may use physicians' medical license numbers or may create individual numbering systems.

Allowable Values and Format:
Alphanumeric, left-justified, blank-filled, or all blank

✓	**Follow-up Physician**	8	AACCR	1122-1129

Description:
Code for the physician following the patient. This may be a physician of any specialty. Registry may use physicians' medical license numbers or may create individual numbering systems.

Rationale:
This is used in central registry follow-up, and also may identify a physician to contact to gain approval to conduct patient interviews.

Allowable Values and Format:
Alphanumeric

IX.O Physician - Confidential Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion			1130 - 1150

IX.P Text - Diagnosis Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #

Note: The next three sections contain the text portion of the abstract. The following rationale applies to all of the text areas:

Rationale:
Text is needed to justify the codes selected for the data items and to allow recording information that is not coded at all. It is a component of a complete electronic abstract, and allows for the full abstract to be printed or reviewed on the screen as needed. The text is used for quality control and special studies.

✓ **Text--DX Proc--PE** — **200** — **1151 - 1350**

Description:
Text area for information from history and physical examinations.

Allowable Values and Format:
Free text

✓ **Text--DX Proc--X-ray/scan** — **250** — **1351 - 1600**

Description:
Text area for information from diagnostic imaging reports.

Allowable Values and Format:
Free text

✓ **Text--DX Proc--Scopes** — **250** — **1601 - 1850**

Description:
Text area for information from endoscopic examinations.

Allowable Values and Format:
Free text

✓ **Text--DX Proc--Lab Tests** — **250** — **1851 - 2100**

Description:
Text area for information from laboratory examinations other than cytology or histopathology.

Allowable Values and Format:
Free text

✓ **Text--DX Proc--Op** — **250** — **2101 - 2350**

Description:
Text area for information from operative reports.

Allowable Values and Format:
Free text

IX.P Text - Diagnosis Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Text--DX Proc--Path**	**250**		**2351 - 2600**

Description:
Text area for information from cytology and histopathology reports.

Allowable Values and Format:
Free text

✓	**Text--Primary Site Title**	**40**		**2601 - 2640**

Description:
Text area for description of primary site in natural language.

Allowable Values and Format:
Free text

✓	**Text--Histology Title**	**40**		**2641 - 2680**

Description:
Text area for description of histologic type, behavior, and grade in natural language.

Allowable Values and Format:
Free text

✓	**Text--Staging**	**300**		**2681 - 2980**

Description:
Additional text area for staging information not already entered in the Text--DX Proc areas.

Allowable Values and Format:
Free text

IX.Q Text - Treatment Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**RX Text--Surgery**	150		2981 - 3130

Description:
Text area for information about surgical procedures performed as part of treatment.

Allowable Values and Format:
Free text

✓	**RX Text--Radiation (Beam)**	150		3131 - 3280

Description:
Text area for information about beam radiation given for cancer treatment.

Allowable Values and Format:
Free text

✓	**RX Text--Radiation Other**	150		3281 - 3430

Description:
Text area for information about non-beam radiation given for cancer treatment.

Allowable Values and Format:
Free text

✓	**RX Text--Chemo**	200		3431 - 3630

Description:
Text area for information about chemotherapy treatment.

Allowable Values and Format:
Free text

✓	**RX Text--Hormone**	200		3631 - 3830

Description:
Text area for information about hormonal cancer-directed treatment.

Allowable Values and Format:
Free text

✓	**RX Text--BRM**	100		3831 - 3930

Description:
Text area for information about biological-response-modifier treatment for the cancer.

Allowable Values and Format:
Free text

IX.Q Text - Treatment Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	RX Text--Other	100		3931 - 4030

Description:
Text area for information about other cancer-directed treatment.

Allowable Values and Format:
Free text

IX.R Text - Miscellaneous Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	**Text--Remarks**	350		4031 - 4380

Description:
Text area for information not elsewhere provided for, and for overflow from other text areas.

Allowable Values and Format:
Free text

✓	**Text--Occup/Indus**	100		4381 - 4480

Description:
Text area for information about the patient's occupation and industry.

Allowable Values and Format:
Free text

✓	**Place of Diagnosis**	50		4481 - 4530

Description:
Text area for information about the facility, city, state, or country where the diagnosis was made.

Allowable Values and Format:
Free text

IX.R Text - Miscellaneous Section

	Data Item Name (AKA)	Characters	Source of Standard	AACCR Column #
✓	Reserved for Expansion	770		4531 - 5300

INDEX