INTRODUCTION TO VOLUME ONE I-1

CANCER IN NORTH AMERICA, 1996-2000

VOLUME ONE: INCIDENCE

PURPOSE

The Data Evaluation and Publication Committee, a standing committee of the North American Association of Central Cancer Registries (NAACCR), produced this monograph. The NAACCR bylaws charge the Data Evaluation and Publication Committee (DEPC) to gather data from member registries, review, evaluate, and compile the information for publication. The year 2003 marks the 13th release of the annual publication of Cancer in North America (CINA) series and the 7th monograph to include cancer mortality data. This monograph reflects the much appreciated participation of NAACCR member registries in this monograph, covering cancer incidence in 12 Canadian provinces and territories, 45 U.S. states, the District of Columbia, and five metropolitan areas in the United States (U.S.). We hope this effort continues to improve the completeness, the timeliness and the quality of data collected by member registries; to promote the use of their cancer registry data; and to provide cancer statistics that are inclusive of all racial/ethnic groups and geographic coverage of North America.

MONOGRAPH FORMAT

The Cancer in North America (CINA), 1996-2000 monograph includes three volumes: Volume One contains cancer incidence data for individual member registries, Volume Two contains cancer mortality data and Volume Three is comprised of NAACCR Combined Incidence Data. Volume One of this monograph includes cancer incidence data for the years 1996 through 2000, unless indicated otherwise, for all participating registries. Volume One has three major sections as described below.

VOLUME ONE, SECTION I: Introduction and Technical Notes

This section states the purpose of the monograph, describes the monograph format, details the various data sources and the NAACCR criteria for the combined cancer statistics for the United States, Canada, and North America, and provides data interpretation. Background information about NAACCR is also provided.

VOLUME ONE, SECTION II: Registry-specific Cancer Incidence by Race and Sex, Age-adjusted to the 2000 U.S. and 1970 U.S. Population Standards

Incidence data are presented in this monograph for 63 central population-based registries: 51 from the United States (45 states, 5 metropolitan areas and the District of Columbia) and 12 from Canada (10 provinces and 2 territories). This represents almost total coverage in Canada (excluding Nunavut) and all but 5 states in the U.S. (Delaware, Kansas, South Dakota, Tennessee, and Vermont). None of the U.S. territories submitted their data this year.

The first page for each participating registry provides descriptive information about the registry, identifies the contact person(s), and summarizes several data quality indicators. This descriptive information about each registry can be used to evaluate the general quality of the registry data and the comparability of incidence rates across the registries. These indicators are the percentage death certificate only (DCO) cases, the rates of duplicate records (from the NAACCR protocol) and the rate of data errors based on the NAACCR EDITS Metafile. Whether or not the registry met the criteria for inclusion in the combined rates is also indicated. Although the data quality indicators presented on this page are for all covered years combined, data quality has been evaluated for each of the single years of data. To assist in data comparisons across registries within the United States, the percentages of total cases that are of races other than white or black and of unknown race are also listed. The sources for case finding are included, with the diagnosis year that each source was implemented (e.g., hospitals, death certificates, pathology laboratories, radiation therapy sites, interstate data exchanges, physician’s offices,
ambulatory surgical centers, and nursing homes or hospices). The “Year case finding began” specifies the first
diagnosis year of cases reported to a registry, not the year of operation in which a registry initiated case finding.

In addition to these general quality indicators, several cancer site-specific data quality indicators are presented.
These cancer sites are selected based on both their frequent occurrence and their importance in cancer control
and prevention activities. For each cancer, four data quality indicators are listed: the percent DCO cases for that
cancer site; the percent microscopically confirmed cases; and the site-specific incidence-to-mortality rate ratios
for whites and for blacks (except for Hawaii and Canadian registries, for which ratios are calculated for all races
combined). The same time periods for incidence rates (numerator) and mortality rates (denominator) are used
in computing the rate ratios, and both rates are adjusted to the 2000 U.S. population standard. Incidence-to-
mortality rate ratios are suppressed when fewer than 6 cases or deaths were reported. For cancer types occurring
in only one gender, e.g., prostate and ovary, the ratios are calculated based on the rates for the specific gender.

Following the registry description, annual age- and sex-specific population estimates for all races, whites, and
blacks are reproduced for each registry in the United States. For Canadian registries, population estimates are
presented for all races combined.

The cancer incidence tables show the total five-year (unless indicated otherwise) incidence counts and the
average annual incidence rates age-adjusted to the 2000 U.S. and the 1970 U.S. population standards, by the
SEER recode groupings for primary site based on the International Classification of Diseases for Oncology,
Second Edition (ICD-O-2) by sex for all races, whites, and blacks for each registry in the U.S. except for Hawaii
(all races only). Canadian registries use ICD-O-2 for their incidence data from 1996-2000. For Canadian
registries, rates are presented only for all races combined because race identifiers are not collected by Canadian
cancer registries.

VOLUME ONE, SECTION III: Registry-specific Cancer Incidence by Sex, Age-adjusted to the 1996
Canadian and World Population Standards

Age-adjusted incidence rates for all races combined were also calculated for males and females separately using
the 1996 Canadian and World population standards for all U.S. and Canadian participating registries.

TECHNICAL NOTES

Data Sources

**Incidence.** Each member registry provided its own incidence data for 1996 to 2000; if all five years were not
available, data were provided for as many of the five years as possible. Cancer incidence data for registries in
the SEER program were obtained from the SEER public use data file (November 2002 submission) produced by
the National Cancer Institute (NCI). Statistics Canada provided data for all Canadian registries except for Alberta
and Ontario that submitted their own data. The 1996-2000 incidence data for Northwest Territories reflects
current geographic boundaries.

**Population Estimates.** Estimates of the population for the U.S., individual U.S. states, and all SEER areas for 1996
through 2000 were obtained from the SEER program, based on U.S. Bureau of Census population estimates for these
years. These population estimates represent a modification of the annual time series of population estimates
produced by the Population Estimates Program of the Bureau of the Census with support from the NCI. Please refer
to the SEER Cancer Statistics Review, 1975-2000 ¹ and its methodologies for specific documentation regarding
modifications made by the NCI to the Census Bureau estimates. The following summarizes these modifications.


1-2 INTRODUCTION TO VOLUME ONE
The initial modification affects only population estimates for the State of Hawaii. The Epidemiology Program of the Hawaii Cancer Research Center has developed its own set of population estimates, based on sample survey data collected by the Hawaii Department of Health. This effort grew out of a concern that the native Hawaiian population has been vastly undercounted in previous censuses. The “Hawaii-adjustment” to the Bureau of the Census estimates has the net result of reducing the estimated white population and increasing the Asian and Pacific Islander population for the state. The Bureau of the Census estimates for the total population, black population, and American Indian and Alaska Native populations in Hawaii are unaffected.

The population estimates now incorporate bridged single-race estimates that are derived from the original multiple-race categories in the 2000 Census. These bridged estimates are consistent with the four race groups enumerated in the 1990 Census and were produced under a collaborative arrangement between the National Center for Health Statistics and the Census Bureau.

A revised set 1990 through 2000 population estimates was recently obtained by NCI from the Census Bureau. This file contains populations by year, county, race, Hispanic origin, sex, and age. The methodology implemented by the Census Bureau to develop these county estimates is comparable to that used to produce national and state 1990-2000 intercensal estimates and is described on the Census Bureau’s website. Thus, statistics published here may differ from those published in previous editions of CINA.

Statistics Canada provided the estimates of the Canadian population for all Canadian provinces and territories, adjusted for census under-coverage and non-permanent residents. The 1996-2000 populations for Northwest Territories reflects current geographic boundaries.

Definitions

Primary Cancer Sites. SEER rules were used to define cancer sites (see Appendix) for all ages combined. Numbers and rates include invasive cancers only, with the exception of cancer of the bladder, for which invasive and in situ cases were included in the counts and rates. Numbers and rates for carcinoma in situ of the breast were listed separately, when supplied by the registry. In situ cases of the breast were not included in the “All Sites” category. Squamous and basal cell carcinomas of the skin were excluded, except those of the lip and genital organs (see Appendix). Cancers among non-residents of the area and cases with unknown sex or age were omitted from all calculations, but cases with unknown race were included in the computation of “all races” cancer counts and rates.

---


Incidence Rates. Rates were calculated per 100,000 population and age-adjusted by the direct method to the 2000 U.S., the 1970 U.S., the 1996 Canadian (Cdn.) and the World population standards. The incidence rates are annual averages for the period 1996 through 2000. (Note: Not all registries submitted data for all five years; their rates are annual averages for the years submitted.) The age distributions of the four population standards are presented below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>6,913.5</td>
<td>8,441.6</td>
<td>6,623.5</td>
<td>12,000</td>
</tr>
<tr>
<td>5-9</td>
<td>7,253.3</td>
<td>9,820.4</td>
<td>6,798.5</td>
<td>10,000</td>
</tr>
<tr>
<td>10-14</td>
<td>7,303.2</td>
<td>10,230.4</td>
<td>6,771.5</td>
<td>9,000</td>
</tr>
<tr>
<td>15-19</td>
<td>7,216.9</td>
<td>9,384.5</td>
<td>6,784.1</td>
<td>9,000</td>
</tr>
<tr>
<td>20-24</td>
<td>6,647.8</td>
<td>8,056.1</td>
<td>6,776.1</td>
<td>8,000</td>
</tr>
<tr>
<td>25-29</td>
<td>6,452.9</td>
<td>6,632.0</td>
<td>7,291.4</td>
<td>8,000</td>
</tr>
<tr>
<td>30-34</td>
<td>7,104.4</td>
<td>5,624.9</td>
<td>8,703.0</td>
<td>6,000</td>
</tr>
<tr>
<td>35-39</td>
<td>8,076.2</td>
<td>5,465.6</td>
<td>8,851.0</td>
<td>6,000</td>
</tr>
<tr>
<td>40-44</td>
<td>8,185.1</td>
<td>5,895.8</td>
<td>8,005.5</td>
<td>6,000</td>
</tr>
<tr>
<td>45-49</td>
<td>7,211.8</td>
<td>5,962.2</td>
<td>7,184.7</td>
<td>6,000</td>
</tr>
<tr>
<td>50-54</td>
<td>6,271.6</td>
<td>5,464.3</td>
<td>5,581.2</td>
<td>5,000</td>
</tr>
<tr>
<td>55-59</td>
<td>4,845.4</td>
<td>4,907.7</td>
<td>4,486.8</td>
<td>4,000</td>
</tr>
<tr>
<td>60-64</td>
<td>3,879.3</td>
<td>4,240.3</td>
<td>4,070.5</td>
<td>4,000</td>
</tr>
<tr>
<td>65-69</td>
<td>3,426.4</td>
<td>3,440.6</td>
<td>3,785.8</td>
<td>3,000</td>
</tr>
<tr>
<td>70-74</td>
<td>3,177.3</td>
<td>2,678.9</td>
<td>3,258.9</td>
<td>2,000</td>
</tr>
<tr>
<td>75-79</td>
<td>2,699.9</td>
<td>1,887.1</td>
<td>2,323.1</td>
<td>1,000</td>
</tr>
<tr>
<td>80-84</td>
<td>1,784.2</td>
<td>1,124.1</td>
<td>1,542.4</td>
<td>500</td>
</tr>
<tr>
<td>85+</td>
<td>1,550.8</td>
<td>743.5</td>
<td>1,161.7</td>
<td>500</td>
</tr>
<tr>
<td>Total</td>
<td>100,000.0</td>
<td>100,000.0</td>
<td>100,000.0</td>
<td>100,000</td>
</tr>
</tbody>
</table>

Standard Errors. Standard errors (S.E.) of the rates were calculated using the formula:

\[
S.E. = \sqrt{\sum w_j^2 n_j / p_j^2}
\]

where \(w_j\) = the fraction of the standard population in age group (5-year age interval), \(n_j\) = number of cases in that age group, and \(p_j\) = person-years denominator. For many registries, the standard error of the rates are small, as the population covered is large. However, for registries that cover a small population, the standard error may be substantial.
**Comparison of Rates.** The standard error of adjusted rates can be used to evaluate the statistical significance of rate differences among comparable populations. For example, if the adjusted rates in two populations are $R_1$ and $R_2$ and their standard errors are $S.E._1$ and $S.E._2$, an approximate confidence interval for the rate ratio can be calculated using the following formula:

$$(R_1/R_2)^{z^2/2}$$

where $x = (R_1 - R_2) / \sqrt{(S.E._1^2 + S.E._2^2)}$ and $z = 1.96$ for 95% confidence limits. If this interval does not include one, the two rates are statistically significantly different at a $p$ value of 0.05. This test can be inaccurate for rates based on fewer than 16 cases, and it should not be used for rates based on fewer than six cases.

It should be emphasized that this kind of comparison of adjusted rates must be undertaken with caution as misleading conclusions may be drawn if the ratios of the age-specific rates in the two populations are not constant in all age groups. In these circumstances, the ratios of the adjusted rates will vary according to the standard populations used.

While it is possible to compare incidence rates among populations in various individual registries, it is important to consider whether the groups are comparable by race (i.e., percent other races or percent unknown race from Section II of Volume One). One should also consider the registries’ quality and completeness as differences can be related to both true underlying differences and differences in reporting completeness and data quality. Interpretation without consideration of these factors may contribute to misleading or inaccurate conclusions.

**Combined Rates for the United States, Canada, and North America.** To be included in combined rates, a registry’s data had to meet or exceed six quality criteria for each single year, 1996, 1997, 1998, 1999, and 2000. This standard, equivalent to meeting NAACCR’s silver registry certification criteria for five consecutive years, is stricter than those used in the past. The six quality criteria are:

1. Data were submitted for each of the five years, 1996 through 2000.
3. 97% of cases pass all internal consistency checks defined by NAACCR’s EDITS metafile, Version 9.1.8
4. The code for “unknown” was used to describe:
   - sex in less than 3% of the cases;
   - age in less than 3% of the cases;
   - county of residence in less than 3% of the cases;
   - race in less than 5% of the cases;
5. Cases registered with information abstracted from death certificates only (DCOs) comprise less than 5% of all cases for each of the five years’ data submitted.
6. Case ascertainment was estimated to be 90% or higher for each of the five years’ data submitted.

NAACCR uses the incidence-to-mortality rate ratio method to measure completeness of case ascertainment. The method assumes that cancer death data are complete, and that the ratio of age-adjusted cancer incidence rates to

---


Race-specific completeness of case ascertainment in jurisdiction $s$ ($C_{sk}$) was computed by dividing the observed race-specific (white; black) age-adjusted (2000 U.S.) incidence rate for both sexes and all cancer sites combined7 (“Observed T”) by the expected race-specific (white; black) age-adjusted (2000 U.S.) incidence rate for both sexes and all cancer sites combined7 (“Expected T”):

$$C_{sk} = \frac{ObservedT_{sk}}{ExpectedT_{sk}}$$


$$ExpectedI_{skij} = \left( M_{skij} \right) \left( \frac{I_{SEER_{skij}}}{M_{U.S., kij}} \right)$$

$$ExpectedT_{sk} = \sum_{i=1}^{2} \sum_{j=1}^{N} ExpectedI_{skij}$$
where:

\[ I = \text{Age-adjusted (2000 U.S.) incidence rate for race } k, \text{ sex } i, \text{ site } j, \text{ 1996 to 2000} \]
\[ M = \text{Age-adjusted (2000 U.S.) mortality rate for race } k, \text{ sex } i, \text{ site } j, \text{ 1996 to 2000} \]
\[ s = \text{State, SEER area, province, or territory} \]
\[ \text{SEER} = \text{Combined eleven SEER areas}^9 \]
\[ \text{U.S.} = \text{United States} \]
\[ T = \text{Age-adjusted (2000 U.S.) incidence rate for total sites}^{10} \]

Overall completeness of case ascertainment in jurisdiction \( s \) (\( C_s \)) was calculated by adding weighted estimates of race-specific completeness of case ascertainment in jurisdiction \( s \) (\( C_{sk} \)), using the proportion of the population in each of the race groups (\( P_{sk} \)) as weights:

\[ C_s = \sum_{k=1}^{2} C_{sk} \times P_{sk} \]

This method of estimating completeness assumes that race-sex-site-specific incidence-to-mortality rate ratios are relatively stable (within 20% limits). The incidence-to-mortality rate ratio standard to which all registries were adjusted, using SEER incidence rates and U.S. death rates, is the current NAACCR standard for this purpose.

The same methods were applied to Hawaii and all Canadian registries, except that jurisdiction-specific data were not race specific, and SEER-incidence-to-U.S.-mortality rate ratios were computed for whites only.

\( C_s \) was adjusted for the presence of duplicate records in the data of jurisdiction \( s \) (\( CA_s \)) thus:

\[ CA_s = C_s \times U_s \]

where:

\( CA = \text{Adjusted overall completeness of ascertainment} \]
\( C = \text{Unadjusted overall completeness of ascertainment} \]
\( s = \text{State, SEER area, province, or territory} \]
\( U = \text{Proportion of unduplicated records, based on NAACCR’s Protocol for Assessing Duplicate Cases.} \]


Every registry included in the combined rates had an adjusted completeness estimate of at least 90 percent for each year of the five years’ data submitted. In the United States, 34 registries met all the criteria for inclusion in the combined rates. These were Alaska, Arizona, California, Greater Bay Area, Los Angeles, Colorado,
Connecticut, Florida, Atlanta, Hawaii, Idaho, Illinois, Iowa, Kentucky, Louisiana, Michigan, Detroit, Minnesota, Montana, Nebraska, New Jersey, New Mexico, New York, North Carolina, Ohio, Oregon, Pennsylvania, Rhode Island, Utah, Washington, Seattle, West Virginia, Wisconsin, and Wyoming. This year combined incidence statistics for Canada are not included in this monograph due to concerns about the appropriateness of NAACCR inclusion criteria for Canadian registries.

**Impact of the Modified Population Estimates on the NAACCR Completeness Estimates.** Recently, the United States Bureau of the Census revised the U.S. population estimates for the 1990s by using 2000 decennial census data to adjust the original post-1990 census population projections. The revised population estimates have an effect on both the incidence and death rates differentially across cancer site and region. The completeness estimates for all cancer registries have also been affected. Despite this revision, the number of registries meeting the NAACCR combined inclusion criteria has increased compared to last year's monograph. The population represented by these registries has also increased this year from 55 percent to 68 percent of the United States population.

**Data Interpretation**

**Race-specific Rates.** Race-specific (either white or black) incidence rates are presented for an aggregate only when more than five cases are included in that aggregate. Canadian data are presented for all races only. To facilitate comparisons among registries of similar race distribution, the proportion of *unknown race* and *races other than black or white* is reported in the Registry Description at the beginning of each registry's section.

**Percent DCOs.** The proportion of cases identified by Death Certificate Only (DCO) has been used as a rough guide to assess completeness of case ascertainment. Only invasive cancer cases are included in the denominator of this proportion. The percent of DCO cases in a registry’s data set may be reduced by intensive “follow-back” to identify other sources of information on DCO cases.

Many new registries postpone the collection of DCO cases until the registry has at least five years’ data, because deaths which occur prior to the five-year mark are very likely to have been diagnosed before the registry’s date of establishment. Were these deaths to be collected and registered as DCO cases, they would inflate cancer incidence during the first several years of registry operations, because they are registered in the year of death. Whether or not a registry used death certificates as a source for case ascertainment in the period 1996 to 2000 is indicated in the Registry Description at the beginning of each registry section.

**Percent of Microscopic Confirmation.** The proportion of total cases with microscopic confirmation can also be used as an indicator of the quality of data collection. From 92 to 96 percent of all SEER cases are confirmed microscopically. A proportion of microscopically confirmed cases that is higher or lower may suggest problems in case ascertainment or abstracting. However, this proportion varies by cancer site. For sites that rely mostly on a clinical diagnosis, e.g., cancers of the pancreas and brain, confirmation rates that are too high may suggest that some clinically-diagnosed cases are missing. Also, registries that do not use death certificates for case finding have an artificially high proportion of microscopically confirmed cases.

**Site-specific incidence-to-mortality rate ratios.** These ratios can be used comparatively to determine completeness of registry data. A registry with a low incidence to mortality ratio may have incomplete ascertainment of cases. Other factors can also influence the site-specific incidence to mortality ratios, such as distribution of stage of disease at diagnosis or differences in coding of death certificates. Because the death rates reflect cases that may have been diagnosed prior to the years included in the incidence rates, the site-specific incidence-to-mortality rate ratios are not intended to be used as an indicator for survival. These ratios may be interpreted in the same way that overall incidence-to-mortality rate ratios are interpreted, adjusting for the larger relative standard errors associated with the incidence and death rates underlying the ratios.
**Cell Suppression, Values of Zero (“0”) and “-” in the Tables.** Although all cases from qualified registries were included in the computation of combined rates, counts and incidence rates were suppressed in the summary tables of selected major cancer sites if they had fewer than six cases. Suppression rules were also applied to the tables of pediatric cancer.

When the incidence rate or count for a specific group is presented as 0.0 or 0, this indicates either 1) that the rate is less than 0.05 per 100,000, or 2) that no cases were reported for the group. “-” is used when the count and rate were suppressed because fewer than six cases were found for the group in question. “-” is also used when data were not available for analysis or to signify “not applicable”, as in the case of female cancer of the prostate, an impossibility.

**Comparisons Among Registries.** All registries responding to the 2003 *Call for Data* are included in Sections II and III of Volume One. In making valid comparisons among registries, it is important to review the data quality indicators for each registry before attributing rate differences to regional variation. Data quality can be an important contributor to observed differences in rates. In addition to data quality, it is important to consider differences in the racial composition of the populations being compared before conclusions are drawn about variations in regional rates.\(^\text{11}\)

**NAACCR MISSION**

The North American Association of Central Cancer Registries, Inc. (NAACCR) is a professional organization that develops and promotes uniform data standards for cancer registration; provides education and training; certifies population-based registries; aggregates and publishes data from central cancer registries; and promotes the use of cancer surveillance data and systems for cancer control and epidemiologic research, public health programs, and patient care to reduce the burden of cancer in North America.

Please address all comments and suggestions about the monograph to:

NAACCR  
Attention: Data Evaluation and Publication Committee  
2121 West White Oaks Drive, Suite C  
Springfield, IL 62704-6495  
(217) 698-0800  
(217) 698-0188 (FAX)

For a copy of this monograph, please contact NAACCR at the above address. The monograph can also be downloaded or viewed from the NAACCR web site ([http://www.naaccr.org](http://www.naaccr.org)).

**Data Evaluation and Publication Committee Members, 2002-2003**

John P. Fulton, Rhode Island, *Chair*  
Vivien W. Chen, Louisiana  
Greer Gay, American College of Surgeons  
Carol Kosary, National Cancer Institute  
Andrew Lake, Information Management Services, Inc.  
Steven D. Roffers, Georgia  
David Roney, Information Management Services, Inc  
Maria Schymura, New York  
Anne-Marie Ugnat, Health Canada  
Ghislaine Villeneuve, Statistics Canada

Dee West, California
William Wright, California
Xiao-Cheng Wu, Louisiana
Royale Anne Hinds, NAACCR
Holly L. Howe, NAACCR
Joellyn Hotes, NAACCR