A

Methods

Approximately 148,042[1] Alaska Native and American Indian (ANAI) people reside in Alaska (individuals reporting ANAI identity alone, or in combination with another racial identity). ANAI people comprise 19.5% of the Alaskan population, and almost 90% of ANAI people living in Alaska identify as Alaska Native; therefore, throughout this report, we refer to all ANAI people resident in Alaska as "Alaska Native (AN) people."

The Alaska Native Tumor Registry collects cancer information for all AN people who are resident in Alaska at the time of cancer diagnosis; who are eligible for Indian Health Service benefits; and who have been diagnosed with cancer from 1969 through the present day. According to ANTR standard case-finding practices, cases were ascertained through a variety of sources, including...

- hospital discharge diagnoses for Tribal and non-Tribal health facilities in Alaska
- tumor registry and pathology files of the ANMC and other in-state healthcare facilities
- linkages with other population-based cancer registries (e.g., the Alaska Cancer Registry and the Washington State Cancer Registry), and
- death certificates (<1% cases were registered solely on the basis of information from a death certificate).

Cancer incidence and mortality data for the US White (USW) population are included in this report to provide a standard point of reference or comparison. Such comparisons help us to describe similarities and differences in cancer patterns, which in turn can help identify areas of strength, and opportunities for improvement. Cancer data for USW were taken from the SEER 9 data set.[3] Population estimates for both the Alaska Native and US White populations originate from the US Census (1970, 1980, 1990, 2000, and 2010) as well as from the National Center for Health Statistics' bridged population series for AN people, 1990–2018, available from SEER.[4,5]

Primary cancer site, pathology, behavior, and grade coding followed the International Classification of Diseases for Oncology, 3rd edition.[6] Cancer sites of origin were grouped according to SEER primary site groups. These are defined in detail in Section A.2.3 (pg. 171) SEER site recode definitions.

[1] Population of ANAI people (race alone or in combination) in Alaska as of July 2019 based on Alaska Department of Labor and Workforce Development, Research and Analysis Section vintage 2019 population estimates. Sources: `https://live.laborstats.alaska.gov/pop/index.cfm` (accessed 2021-04-09), `https://live.laborstats.alaska.gov/pop/estimates/data/AgeBySexByRaceAICHispAK.xls` (accessed 2021-04-09)

Incidence rates, rate ratios (RRs), and confidence intervals were computed using R version 3.6.3 (2020-02-29)[7] and cross-checked against output from SEER*Stat.[8] Rates for AN people were age-adjusted to the US Census 2000 standard population using the direct method for comparison with USW rates (see Section A.1, pg. 170). Rate ratios are expressed as the AN rate divided by the USW rate. Cancer incidence rates for each Tribal health region within Alaska were compared with the average annual cancer incidence rates for all AN people statewide. In this report, cell sizes <5 are censored for patient privacy.

A.1   Equations

Equation (A.1) defines the crude incidence rate while equation (A.5) defines the age-adjusted incidence rate.

A.1.1   Crude

$$rate_{crude} = \frac{count}{population} \times 100,000 \tag{A.1}$$

$$SE_{crude} = \frac{\sqrt{count}}{population} \times 100,000 \tag{A.2}$$

$$CI_{lo} = \frac{\frac{1}{2} \times \text{invchi}(\frac{p}{2}, 2 \times count)}{population} \times 100,000 \tag{A.3}$$

$$CI_{hi} = \frac{\frac{1}{2} \times \text{invchi}(1 - \frac{p}{2}, 2 \times (count + 1))}{population} \times 100,000 \tag{A.4}$$

A.1.2   Adjusted

$$rate_{adj\ x \to y} = \sum_{i=x}^{y} \left[\left(\frac{count_i}{population_i} \right) \times 100,000 \times \left(\frac{stdmil_i}{\sum_{j=x}^{y} stdmil_j} \right) \right] \tag{A.5}$$

$$SE_{adj} = \left[\sum_{i=x}^{y} \left(\frac{stdmil_i}{\sum_{j=x}^{y} stdmil_j} \right)^2 \times \left(\frac{count_i}{population_i^2} \right) \right]^{\frac{1}{2}} \times 100,000 \tag{A.6}$$

$$w_i = \frac{stdmil_i}{population_i \times \sum_{j=x}^{y} stdmil_j} \tag{A.7}$$

$$w_m = \text{max}(w_i) \tag{A.8}$$

$$z = w_m^2 \tag{A.9}$$

$$v = \sum_{i=x}^{y} \left(w_i^2 \times n_i \right) \tag{A.10}$$

$$CI_{lo} = \frac{v}{2 \times rate} \times \texttt{invchi}(\frac{p}{2}, \frac{2 \times rate^2}{v}) \times 100,000 \qquad \text{(A.11)}$$

$$CI_{hi} = \frac{v+z}{2 \times (rate + w_m)} \times \texttt{invchi}(1 - \frac{p}{2}, \frac{2 \times (rate + w_m)^2}{v+z}) \times 100,000$$
$$\text{(A.12)}$$

A.2   Additional Considerations

A.2.1   US White Data

Incidence data for USW people come from the SEER 9 database,[3] which includes data from the following cancer registries: Connecticut, Detroit, Georgia, San Francisco-Oakland, Hawaii, Iowa, New Mexico, Seattle-Puget Sound, and Utah. Data from these registries are only available for the years 1975–2017. Therefore, in five-year data, data for USW include only the years 2014–2017. For longer term trends, data for USW include the years 1975–2017.

A.2.2   Childhood Cancers

In this report the term "men" includes boys (younger than adults) and the term "women" includes girls (younger than adults). Cancer is predominantly an adult disease; however, we do acknowledge that children are also diagnosed with cancers and are not excluded from the data.

A.2.3   Mortality Data

All mortality data for both AN and USW people were sourced from SEER*Stat[9] for the 25-year period from 1994 through 2018 and are included in this publication.

A.2.4   Stage at Diagnosis

Directly corresponding stage variables were not available for ANAI and USW data sets. While the available stage variables use the same general classifications for stage at diagnosis (e.g., local, regional, distant, unknown, etc.) the specific rules used to make those classifications vary slightly. For ANAI data several stage classifications were available; we chose to use the two classifications that were most complete and covered the desired date range: "Derived SS2000" and "Derived Summary Stage 2018". For USW data we chose to use "SEER Combined Summary Stage (2004+)" because it is largely compatible with "Derived SS2000." The following description of the "SEER Combined Summary Stage (2004+)" variable is provided in the *Dictionary of SEER*Stat Variables*.[10]

> Combination of NAACCR Item #s 759 and 3020. For 2004–2015 cases it is copied from Derived SS2000 and for 2016+ it is the best available between Derived SS2000 and SEER Summary Stage 2000. All benign/borderline cases set to NA.

A.2.5   Miscellaneous

- Error bars on charts represent the 95% confidence interval.
- The "in situ" stage is not included in the data used for this report.

A.3   Software Resources

The software resources listed in this section are those that were used directly in production of this report. These listings are provided strictly for informational purposes and do not constitute an endorsement or approval of any of the listed resources.

A.3.1   Applications

- SEER*Stat[8]
- SEER*Prep[11]
- R: A language and Environment for Statistical Computing[7]
- RStudio: Integrated Development Environment for R[12]
- JabRef—An open-source, cross-platform citation and reference management software[13]
- Microsoft Access 2016[14]
- Microsoft Excel 2016[15]
- Microsoft SQL Server 2012[16]
- Microsoft SQL Server Management Studio[17]

A.3.2   R Packages

- bookdown (0.20)[18,19]
- brew (1.0-6)[20]
- clipr (0.7.0)[21]
- DBI (1.1.0)[22]
- dplyr (1.0.0)[23]
- dsr (0.2.2)[24]
- extrafont (0.17)[25]
- ggplot2 (3.3.2)[26]
- kableExtra (1.2.1)[27]
- knitr (1.29)[28–30]
- lubridate (1.7.9)[31]
- purrr (0.3.4)[32]
- readxl (1.3.1)[33]
- rio (0.5.16)[34]
- rlang (0.4.7)[35]
- stringr (1.4.0)[36]
- tidyverse (1.3.0)[37]
- tinytex (0.25)[38,39]

- tufte (0.6)[40]

A.3.3   Citation Styles

- International Journal of Epidemiology Citation Style[41]