

# Including Self-reported Race to Improve Cancer Surveillance Data for American Indians and Alaska Natives in Washington State

Megan J. Hoopes, MPH<sup>a</sup>; Maile Taulii, PhD, MPH<sup>b</sup>; Thomas M. Weiser, MD, MPH<sup>a,c</sup>;  
Rachel Brucker, RN, MPH<sup>b</sup>; Thomas M. Becker, MD, PhD<sup>a</sup>

**Abstract: Background.** American Indians and Alaska Natives (AI/AN) are frequently misclassified as another race in cancer surveillance systems, resulting in underestimated morbidity and mortality. Linkage methods with administrative records have been used to correct AI/AN misclassification, but AI/AN populations living in urban areas, and those who self-identify as AI/AN race, continue to be under-ascertained. The aim of this study was to evaluate racial misclassification in two cancer registries in Washington State using an urban AI/AN patient roster linked with a list of Indian Health Service (IHS) enrollees. **Methods.** We conducted probabilistic record linkages to identify racial misclassification using a combined demographic dataset of self-identified AI/AN patients of a large, urban Indian health center, and administratively-identified AI/AN enrolled with the IHS. Age-adjusted incidence rates were calculated for 3 linkage populations: AI/AN originally coded in each cancer registry, post-linkage AI/AN identified through the IHS roster alone, and post-linkage AI/AN identified through either the urban or IHS file. **Results.** In the state and regional cancer registries, 11% and 18%, respectively, of matched cases were originally coded as a race other than AI/AN; approximately 35% of these were identified by the urban file alone. Incidence rate estimates increased after linkage with the IHS file, and further increased with the addition of urban records. Matches identified by the urban patient file resulted in the largest relative incidence change being demonstrated for King County (which includes Seattle); the all-site invasive cancer rate increased 8.8%, from 443 to 482 per 100,000. **Conclusions.** Inclusion of urban and self-identified AI/AN records can increase case ascertainment in cancer surveillance systems beyond linkage methods using only administrative sources.

**Key words:** Alaska Native, American Indian, AI/AN, cancer, incidence, misclassification

## Introduction

Racial misclassification in public health data sources has been well documented, and for American Indians and Alaska Natives (AI/AN), results in a systematic underestimation of morbidity and mortality. Previous reports have demonstrated that AI/AN are incorrectly classified in administrative data sources more frequently than other racial/ethnic groups,<sup>1</sup> with racial misclassification ranging from 30% to 70%.<sup>1-3</sup> In the case of AI/AN, such misclassification is almost always unidirectional—that is, AI/AN are incorrectly coded as a race other than AI/AN, rather than another race being misclassified as AI/AN.<sup>1,4-6</sup>

Terminology and methods used for the classification of race have received frequent criticism.<sup>7-9</sup> Inaccuracies may result from fixed-response categories, whereas self-assigned race permits individuals to account for all racial groupings with which they identify and minimizes constraints of traditional categorization.<sup>7,10</sup> A shortcoming of this technique is that individual racial identity is dynamic and context-dependent; self-report allows for people of AI/AN ancestry to self-identify their race differently in different settings.<sup>11</sup>

Additionally, many institutions lack a systematic way to record race, resulting in the potential for racial misclassification in surveillance systems. Specific mechanisms include inconsistencies in the collection of race data (eg, many private health insurance companies do not collect or report race), inaccurate third party form completion, assumptions based on appearance, and hesitation to ask an individual or family member to identify race.

The Indian Health Service (IHS) offers a unique ability to identify individuals of documented AI/AN ancestry, since eligibility for services is tied to enrollment in a federally-recognized tribe. In addition, many tribes and other organizations which serve AI/AN maintain registry systems to help determine eligibility for services. These systems are particularly useful in efforts to correctly identify AI/AN people. Linkage methods with IHS patient records have been successfully used to supplement and improve AI/AN representation in public surveillance systems including cancer,<sup>2,5-6,11-12</sup> death records,<sup>13-14</sup> and sexually transmitted diseases.<sup>3,15</sup> The use of IHS records alone, however, may fail to capture AI/AN people who have not enrolled in a tribal

*"Including Self-reported Race to Improve Cancer Surveillance Data for American Indians and Alaska Natives in Washington State"*

<sup>a</sup>Northwest Tribal Epidemiology Center, Northwest Portland Area Indian Health Board, Portland, OR. <sup>b</sup>Seattle Indian Health Board, Urban Indian Health Institute, Seattle, WA. <sup>c</sup>Portland Area Indian Health Service, Portland, OR.

Address correspondence to Megan Hoopes, MPH; Northwest Portland Area Indian Health Board, 2121 SW Broadway Dr., Suite 300, Portland, OR 97201. Telephone: (503) 228-4185, fax: (503) 228-8182, email: mhoopes@npaihb.org.

registry, members of tribal groups that are not federally recognized, and other indigenous people of North America who might culturally or ancestrally identify as AI/AN but do not meet IHS eligibility criteria. These methods also exclude AI/AN who haven't accessed care at an IHS or tribal facility, including many residing in urban areas where IHS services are limited or non-existent.

American Indians and Alaska Natives who reside in urban areas are more likely to be misclassified in administrative and other data sets than those living in less populated places.<sup>1,3-4</sup> Contrary to common belief, the majority of AI/AN live in urban areas.<sup>16</sup> Nearly 7 out of 10 AI/AN live in a census-defined metropolitan area, and only an estimated 57% of the nation's AI/AN population are eligible to receive health care services from the IHS.<sup>17</sup> Even among those who may be eligible, IHS is not the main point of care for all AI/AN, especially those living in cities. Few previous studies have sought to include urban AI/AN populations in record linkage efforts,<sup>18</sup> and it is likely that racial misclassification of urban AI/AN persists, even in surveillance systems that have been previously corrected by IHS linkages.

The Seattle/King County area is the largest urban setting in Washington, home to more than 33,000 AI/AN.<sup>16</sup> The objective of the current study was to identify and evaluate the impact of AI/AN racial misclassification in two Washington cancer registries through probabilistic linkage with a combined urban/tribal file. Patient enrollment records from IHS facilities and tribal clinics, and a large urban Indian health care facility were combined to create the most comprehensive demographic dataset of AI/AN individuals in Washington state to date. We sought to obtain more accurate cancer incidence rates for Washington's AI/AN population, particularly those living in the Seattle/King County region.

## Methods

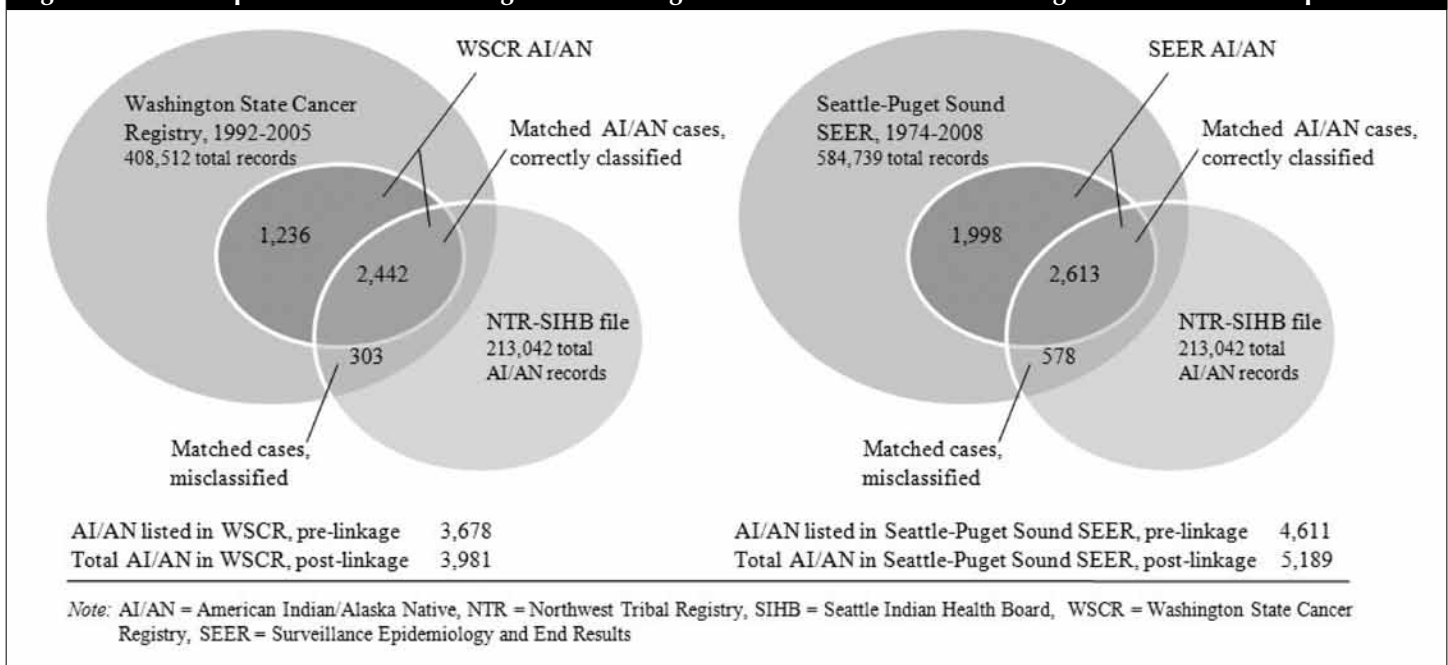
The Northwest Tribal Registry (NTR) is an enumeration of the AI/AN population in Idaho, Oregon, and Washington which has been maintained and regularly updated by the Northwest Portland Area Indian Health Board since 1999. The NTR is derived from the Portland Area Indian Health Service patient file, and includes all AI/AN who received services from an IHS or tribal health care facility in the 3-state region between 1986 and August 2008. Because IHS eligibility is based on tribal membership or descent from an enrolled member of a federally recognized tribe, all patient records in the NTR are of documented AI/AN ancestry.

The Seattle Indian Health Board (SIHB), established in 1970, is a non-profit, multi-service community health center that targets urban AI/AN in the Puget Sound region of Western Washington. In contrast to IHS facilities, individuals of all races are eligible to receive care at the SIHB and are served on an ability-to-pay sliding fee scale; racial identification of patients is ascertained by self-report at registration. Thus there is no requirement or incentive for patients to report AI/AN race if they do not self-identify as such. The patient registry used for this analysis included all patients who accessed care at one of SIHB's medical or dental clinics and self-identified as AI/AN between 1996 and 2007.

We used LinkPlus (version 2.0; Atlanta, GA), a probabilistic linkage software developed by the Centers for Disease Control and Prevention, to link the NTR to the SIHB patient file, arriving at a combined demographic data set of the tribal and urban Washington AI/AN population. A flag variable was added to the combined file to indicate if each record originated from the tribal registry only, the urban patient file only, or was a match (contained in both files). The combined file contained 213,042 records.

Cancer records for years 1992–2005 were obtained from the Washington State Cancer Registry (WSCR) (released

**Figure 1. AI/AN representation in Washington cancer registries as a result of record linkages with urban/tribal patient file**



January, 2008), and years 1974–2008 from the Surveillance, Epidemiology, and End Results (SEER) Registry of Seattle-Puget Sound (released February, 2009). The Seattle-Puget Sound SEER Registry contracts with the state of Washington to collect cancer case information among residents of the 13 Puget Sound area counties of Clallam, Grays Harbor, Island, Jefferson, King, Kitsap, Mason, Pierce, San Juan, Skagit, Snohomish, Thurston, and Whatcom. These cases are transmitted regularly to the WSCR for inclusion in the state-wide registry. Both of these registries adhere to quality control standards set by the North American Association of Central Cancer Registries (NAACCR).<sup>19</sup>

We then used LinkPlus software and uniform linkage configurations to match the combined urban/tribal AI/AN demographic file to each of the cancer registries. Uncertain matches were clerically reviewed by two project staff members, who together assigned each potential match as a “true” or “false” match. We identified all true matches for which the primary race field in the cancer registry was coded as something other than AI/AN (misclassified AI/AN cases), and reported these records to cancer registry staff for correction. We then removed all personal identifiers from the cancer registry files, and retained all matched and non-matched cases for data analysis.

We calculated age-adjusted incidence rates for the most recent diagnosis years considered complete by NAACCR standards at the time of linkage with each registry: 2002–2005 for WSCR records, and years 2002–2006 for Seattle-Puget Sound SEER records. State-wide rates were calculated from WSCR data, and county-level rates calculated from Seattle-Puget Sound SEER. WSCR, and SEER data were not aggregated, and analyses from each registry are presented separately. Calculated rates include invasive cancers (behavior code 3), plus *in situ* urinary bladder cases. To arrive at the best estimates for cancer incidence among Washington individuals who identify as AI/AN, we included in the analyses all matched cases, plus non-matched cases coded as AI/AN in the cancer registry. Throughout the analyses, we counted records as AI/AN if AI/AN race was indicated in either of the two primary race fields in the cancer registries (Race 1 or Race 2, NAACCR data items 160 and 161).<sup>19</sup> To demonstrate the effect of the linkages on incidence estimates, we calculated rates using three linkage populations: those originally coded as AI/AN in the cancer registries (pre-linkage), those originally coded as AI/AN plus misclassified cases identified by NTR records (post-linkage, NTR only), and those originally coded as AI/AN plus misclassified cases identified by NTR and/or SIHB records (post-linkage, NTR+SIHB). To provide an update on AI/AN cancer statewide and in the Seattle area, incidence rates for selected leading cancer sites were calculated for Washington State, the Seattle-Puget Sound SEER region, and King County. We used the National Center for Health Statistics bridged-race population estimates for 2002–2006 (vintage 2007) for population-at-risk denominators. These census-based estimates employ bridging methods to bridge multiple-race population counts to single-race categories, thus adjusting population estimates for individuals who self-select more than one race.<sup>20</sup> All incidence rates were age-adjusted to the 2000 US standard population with

5-year age groupings.<sup>21</sup> Confidence intervals were calculated by the gamma method described by Fay and Feuer.<sup>22</sup> All data management and statistical analyses were conducted using SAS software (version 9.1; SAS Institute Inc, Cary, NC).

At all stages, the study was planned, designed, and implemented collaboratively by the Northwest Portland Area IHB and the SIHB. We established data sharing and confidentiality agreements prior to beginning work on the project. The protocol for the present study was approved by the institutional review boards of Portland Area Indian Health Service, Washington State Department of Health, and Fred Hutchinson Cancer Research Center (where the Seattle-Puget Sound SEER registry is housed), as well as the privacy board of SIHB.

## Results

Results of the two record linkages are presented in Figure 1. Approximately 11% of 2,745 matches to WSCR and 18% of 3,191 Seattle-Puget Sound SEER matches were misclassified as a race other than AI/AN. As a result of the linkages, the representation of AI/AN cases in the cancer registries increased by 8.2% for WSCR and 12.5% for Seattle-Puget Sound SEER. Race classification had been corrected through previous linkages between IHS and each cancer registry, thus these numbers represent only newly identified AI/AN misclassification. Characteristics of misclassified cancer cases are presented in Table 1. In each registry, over 84% of misclassified cases were originally coded as White; the linkage with Seattle-Puget Sound SEER

**Table 1. Characteristics of Misclassified AI/AN Cases by Cancer Registry**

|                             | Washington State<br>Cancer Registry,<br>1992–2005 | Seattle-Puget<br>Sound SEER,<br>1974–2008 |
|-----------------------------|---|---|
|                             | n (%)   | n (%)                                     |
| Total misclassified records | 303   | 578                                       |
| Race 1 as coded             |   |   |
| White                       | 272 (89.8)  | 489 (84.6)                                |
| Black                       | 5 (1.7)   | 15 (2.6)                                  |
| Unknown                     | 15 (5.0)  | 60 (10.4)                                 |
| All others combined         | 11 (3.6)  | 14 (2.4)                                  |
| Linkage source              |   |   |
| NTR alone                   | 189 (62.4)  | 352 (60.9)                                |
| SIHB alone                  | 104 (34.3)  | 203 (35.1)                                |
| Both                        | 10 (3.3)  | 23 (4.0)                                  |
| Sex                         |   |   |
| Male                        | 135 (44.6)  | 189 (32.7)                                |
| Female                      | 168 (55.5)  | 389 (67.3)                                |

Note: AI/AN = American Indian/Alaska Native, NTR = Northwest Tribal Registry, SIHB = Seattle Indian Health Board, SEER = Surveillance, Epidemiology, and End Results

also identified a substantial number of cases for which race had been unknown ( $n=60$ , 10.4% of misclassified cases). In each registry, approximately 35% of misclassified cases were identified by SIHB records alone, and less than 4% of misclassified cases in either cancer registry were identified by records originally contained in both NTR and SIHB files. Among misclassified AI/AN cases in WSCR, a slightly greater proportion were female (55.5%,  $p$ -value = 0.06); in the SEER registry, misclassified cases were significantly more likely to be female (67.3%;  $p$ -value <0.001).

The most common cancer sites for AI/AN by number of incident cases were female breast, lung and bronchus, colorectal, and prostate. State-wide incidence rate estimates

by linkage population are presented in Table 2 and estimates for the Seattle-Puget Sound SEER region and King County in Table 3. Incidence rate estimates increased as a result of linkage with NTR records, and further increased with the addition of SIHB records. The magnitude and direction of this change was similar across all selected cancer sites, geographic regions, and genders. A significant increase was demonstrated between pre- and post-linkage incidence estimates for the state-wide AI/AN rates calculated from WSCR (371 vs. 439 per 100,000 person-years, respectively, Table 2). For the Seattle-Puget Sound SEER region, the all-site incidence rate increased a total of 8.4%, from 475 to 501 per 100,000 after linkage with NTR records alone (5.5%

**Table 2. Selected Age-adjusted Invasive<sup>a</sup> Cancer Incidence Rates (per 100,000)<sup>b</sup> by Linkage Population, Washington AI/AN, 2002–2005**

|                   | Pre-linkage |                      | Post-linkage, NTR only |                             | Post-linkage, NTR+SIHB |                             |
|-------------------|-------------|----------------------|------------------------|-----------------------------|------------------------|-----------------------------|
|                   | n           | Rate (95% CI)        | n                      | Rate (95% CI)               | n                      | Rate (95% CI)               |
| All sites         | 1,035       | 370.5 (345.9, 396.8) | <b>1,195</b>           | <b>426.0 (399.6, 454.0)</b> | <b>1,237</b>           | <b>439.2 (412.4, 467.6)</b> |
| Colon and rectum  | 122         | 47.1 (38.3, 57.6)    | 142                    | 55.1 (45.6, 66.5)           | 146                    | 56.1 (46.5, 67.5)           |
| Lung and bronchus | 151         | 62.1 (51.8, 74.1)    | 162                    | 66.4 (55.8, 78.8)           | 166                    | 67.6 (56.9, 80.1)           |
| Breast (female)   | 161         | 93.5 (78.5, 111.1)   | 189                    | 110.2 (93.9, 129.2)         | 194                    | 113.2 (96.6, 132.3)         |
| Prostate (male)   | 116         | 98.9 (79.8, 122.6)   | 133                    | 113.4 (92.9, 138.7)         | 138                    | 116.4 (95.7, 141.8)         |

<sup>a</sup>Excluding non-urinary bladder *in situ* cases

<sup>b</sup>Adjusted to 2000 US standard population

Note: AI/AN = American Indian/Alaska Native, CI = confidence interval, NTR = Northwest Tribal Registry, SIHB = Seattle Indian Health Board, **bold** = rate is significantly different from pre-linkage rate

**Table 3. Selected Age-adjusted Invasive<sup>a</sup> Cancer Incidence Rates (per 100,000)<sup>b</sup> by Linkage Population, Seattle-Puget Sound SEER Region, AI/AN, 2002–2006**

|                   | Pre-linkage |                      | Post-linkage, NTR only |                      | Post-linkage, NTR+SIHB |                      |
|-------------------|-------------|----------------------|------------------------|----------------------|------------------------|----------------------|
|                   | n           | Rate (95% CI)        | n                      | Rate (95% CI)        | n                      | Rate (95% CI)        |
| All sites         |             |                      |                        |                      |                        |                      |
| SEER counties     | 1,032       | 475.1 (443.3, 509.2) | 1,091                  | 500.7 (468.0, 535.5) | 1,126                  | 514.8 (481.7, 550.0) |
| King County       | 289         | 443.0 (388.4, 504.4) | 293                    | 446.8 (392.8, 508.4) | 319                    | 482.0 (426.0, 546.5) |
| Colon and rectum  |             |                      |                        |                      |                        |                      |
| SEER counties     | 111         | 56.4 (45.3, 69.9)    | 115                    | 58.2 (46.9, 71.8)    | 116                    | 58.4 (47.2, 72.1)    |
| King County       | 23          | 39.6 (23.9, 63.2)    | 23                     | 39.6 (23.9, 63.2)    | 24                     | 40.5 (24.7, 64.2)    |
| Lung and bronchus |             |                      |                        |                      |                        |                      |
| SEER counties     | 160         | 87.0 (72.9, 103.4)   | 163                    | 88.1 (73.9, 104.5)   | 169                    | 90.3 (76.1, 106.9)   |
| King County       | 56          | 105.3 (78.0, 140.5)  | 56                     | 105.3 (78.0, 140.5)  | 60                     | 109.8 (82.0, 145.1)  |
| Breast (female)   |             |                      |                        |                      |                        |                      |
| SEER counties     | 167         | 122.9 (103.5, 145.8) | 181                    | 133.8 (113.4, 157.6) | 182                    | 135.3 (114.7, 159.4) |
| King County       | 55          | 133.2 (98.2, 179.0)  | 58                     | 139.0 (103.4, 185.2) | 59                     | 143.4 (106.9, 190.7) |
| Prostate (male)   |             |                      |                        |                      |                        |                      |
| SEER counties     | 101         | 112.7 (89.2, 142.2)  | 112                    | 124.4 (99.8, 155.0)  | 117                    | 129.6 (104.5, 160.7) |
| King County       | 23          | 79.4 (47.1, 135.4)   | 23                     | 79.4 (47.1, 135.4)   | 27                     | 95.5 (59.0, 155.6)   |

<sup>a</sup>Excluding non-urinary bladder *in situ* cases

<sup>b</sup>Adjusted to 2000 U.S. standard population

Note: AI/AN = American Indian/Alaska Native, CI = confidence interval, NTR = Northwest Tribal Registry, SIHB = Seattle Indian Health Board, SEER = Surveillance, Epidemiology, and End Results

increase), and to 515 per 100,000 based on linkage with the NTR+SIHB file. Within the Seattle-Puget Sound SEER catchment area, the largest relative changes between pre- and post-linkage rate estimates were seen for King County, although low case counts resulted in wide confidence intervals and none of the differences reached statistical significance. All-site incidence for King County AI/AN increased 8.8%, from 443 to 482 per 100,000 person-years.

The effect of the linkages on incidence estimates varied by primary site. In WSCR, the greatest relative changes between pre- and post-linkage incidence were seen for female breast cancer (21.1% increase), colorectal (19.1% increase), and prostate cancer (17.7% increase). Similarly, in the Seattle-Puget Sound SEER region, female breast and prostate cancer rates were most affected by racial misclassification (10.1% and 15.0% increase, respectively). These relative changes, however, did not achieve statistical significance.

## Discussion

We describe racial misclassification of AI/AN in two cancer registries in Washington State based on a collaborative linkage between a large urban clinic and an IHS patient roster. As a result of urban patient inclusion, linkage-corrected AI/AN cancer incidence rates increased in state-wide estimates, the Seattle-Puget Sound SEER catchment area, and notably, King County. The current study is the first to our knowledge to demonstrate that inclusion of urban AI/AN patient rolls can increase cancer case ascertainment above and beyond annual cancer linkages conducted with IHS records.

We also present one of the first studies to include a linkage population that self-reports race as AI/AN (ie, the urban patient roster). Self-reported race/ethnicity is considered by researchers and the US government as the scientific gold standard,<sup>1,9-10,23</sup> though past studies that have examined racial misclassification in health surveillance systems have been largely limited by the absence of self-reported race data. Boehmer and colleagues<sup>1</sup> reported very low agreement between self-reported and administrative sources of AI/AN race data in a large sample of Department of Veterans Affairs dental patients. When assessing outcomes by race, some studies have noted differential diagnostic and treatment results according to the source of race data (administrative or self-report), suggesting that racially misclassified individuals may systematically differ with regard to outcomes.<sup>1,24</sup>

In the current report, the high proportion of misclassified AI/AN cases identified by the urban (self-identified) file alone highlights two important implications: first, that record linkage with administrative data sources, such as IHS patient lists, omits individuals who are either not eligible for or have not received care at an IHS facility but may still self-identify as AI/AN; and second, that the use of the IHS patient file alone fails to capture many AI/AN who reside and access health care in urban areas. Both of these factors result in continued under-ascertainment of AI/AN in surveillance data sources. Through this research we were able to more closely approximate the total AI/AN population by including a large patient population that

self-identifies as AI/AN. This allowed for closer congruence in collection methods between numerator (cancer registry) and denominator (Census-based, self-reported) data, a limitation that has been noted by several other researchers.<sup>4,13,24</sup>

Unlike some other IHS areas, all IHS facilities in the Portland Area (which includes Idaho, Oregon, and Washington) provide only outpatient care; there are no IHS-operated inpatient centers where AI/AN cancer patients could receive cancer-related diagnostic, treatment, or surgical procedures. Additionally, the Seattle-Puget Sound SEER region does not contain any large tribes or areas where AI/AN make up a substantial proportion of the population.<sup>5</sup> Past research has demonstrated that the likelihood of racial misclassification of AI/AN is greater in urban areas and locations where the minority population constitutes a small proportion of the population.<sup>3,5,11,13,24</sup> Record linkages with IHS patient files<sup>12</sup> and some tribal enrollment lists<sup>11</sup> have been ongoing in the Pacific Northwest for over 5 years, with linkage-derived AI/AN race information being regularly reported back to WSCR for correction. Thus we expected that the baseline AI/AN racial classification in WSCR would be relatively accurate compared to other regions and registries without this history. The proportions of matched AI/AN cases coded as another race in Seattle-Puget Sound SEER and WSCR were similar, despite the SEER registry having been updated less frequently through linkages with the NTR. We do note that the Seattle-Puget Sound SEER registry has been previously linked with a national IHS patient file,<sup>12</sup> resulting in partial correction of AI/AN racial misclassification through this effort. In both registries, approximately 35% of misclassified cases were identified through urban patient records alone; these would have remained misclassified had the urban file not been included. A very small proportion of matched cases (<4% in either cancer registry) matched a record in both the IHS and urban file. This observation reaffirms the notion that the SIHB largely serves a population that does not additionally access IHS clinical services. These populations are also likely to differ in other characteristics related to cancer risk factors, diagnosis, and treatment.

The present study would be improved by the addition of tribal enrollment lists, which have been shown to supplement case ascertainment above IHS enrollment records, though to varying degrees.<sup>11,24</sup> This could be a particularly useful approach in the northwest, where the majority of tribes have chosen to take their allotted health care dollars and operate their own health programs rather than using federally-operated IHS facilities.<sup>25</sup> Under this option, some tribes have elected to purchase a private health insurance plan for each tribal member,<sup>26</sup> thus precluding the routine use of an Indian health care facility by its members and further reducing the likelihood that members of those tribes would appear in IHS patient rosters. An additional limitation is that we were precluded from evaluating the cumulative amount of AI/AN racial misclassification in the two cancer registries because they had been previously updated based on past linkages with the NTR and IHS. The proportions reported here reflect only misclassified AI/AN cases newly identified through the current linkages.

Ensuring the accuracy and completeness of surveillance data continues to be an important strategy to inform resource allocation and advocate for improvement in the health of AI/AN people. American Indians and Alaska Natives residing in urban areas are frequently undercounted in surveillance systems, and previous linkage efforts to correct racial misclassification have largely failed to address this omission. In the future, we will continue to pursue opportunities to include urban AI/AN populations in record linkages with northwest disease surveillance systems. Collaborative activities between tribal and urban Indian health organizations can improve data quality for all AI/AN, and result in more accurate estimates of cancer incidence.

### Acknowledgements

The authors are grateful to the staff of Seattle Indian Health Board, Washington State Cancer Registry, and Fred Hutchinson Cancer Research Center Cancer Surveillance System, for providing linkage-related support and critical review of the manuscript. The Washington State Cancer Registry is funded through the Centers for Disease Control and Prevention's National Program of Central Cancer Registries and the Cancer Surveillance System receives funding from National Cancer Institute's SEER Program. Grant support for this project was received from National Cancer Institute's Surveillance Research Program, contract #HHSN261200900180P, and the Indian Health Service Division of Epidemiology and Disease Prevention. We wish to thank Judith Swan, MHS and Brenda Edwards, PhD for their ongoing support. The findings and conclusions in this report are those of the authors and do not necessarily represent official positions of NCI or IHS.

### References

- Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis LE, Jones JA. Self-reported vs. administrative race/ethnicity data and study results. *Am J Pub Health.* 2002;92(9):1471-1473.
- Becker TM, Bettles J, Lapidus J, Campo J, Johnson CJ, Shipley D, et al. Improving cancer incidence estimates for American Indians and Alaska Natives in the Pacific Northwest. *Am J Pub Health.* 2002;92(9):1469-1471.
- Bertolli J, Lee LM, Sullivan PS. AI/AN Race/Ethnicity Data Validation Workgroup. Racial misclassification of American Indians/Alaska Natives in the HIV/AIDS Reporting Systems of five states and one urban health jurisdiction, U.S., 1984-2002. *Public Health Rep.* 2007;122:382-392.
- Arias E, Schauman WS, Eschbach K, Sorlie PD, Backlund E. The validity of race and Hispanic origin reporting on death certificates in the United States. *Vital Health Stat 2.* 2008 Oct; (148):1-23.
- Frost F, Taylor V, Fries E. Racial misclassification of Native Americans in a Surveillance, Epidemiology, and End Results Cancer registry. *J Nat Cancer Inst.* 1992;84(12):957-961.
- Puukka E, Stehr-Green P, Becker TM. Measuring the health status gap for American Indians/Alaska Natives: Getting closer to the truth. *Am J Pub Health.* 2005;95(5):838-843.
- Bradby H. Describing ethnicity in health research. *Ethn Health.* 2003;8(1):5-13.
- Amaro H, Zambrana RE. Criollo, Mestizo, Mulato, LatiNegro, Indigena, white, or black? The US Hispanic/Latino population and multiple responses in the 2000 census. *Am J Pub Health.* 2000;90(11):1724-1727.
- Hahn RA, Stroup DF. Race and ethnicity in public health surveillance: Criteria for the scientific use of social categories. *Public Health Rep.* 1994;109:7-15.
- Office of Management and Budget. Federal Register Notice: Revisions to the standards for the classification of federal data on race and ethnicity. Available at: [http://www.whitehouse.gov/omb/fedreg\\_1997standards/](http://www.whitehouse.gov/omb/fedreg_1997standards/). October 30 1997. Accessed September 29, 2009.
- Sugarman JR, Holliday M, Ross A, Castorina J, Hiu Y. Improving American Indian cancer data in the Washington State Cancer Registry using linkages with Indian Health Service and tribal records. *Cancer.* 1996;78(7 suppl):1564-1568.
- Espey DK, Wiggins CL, Jim MA, Miller BA, Johnson CJ, Becker TM. Methods for improving cancer surveillance data in American Indian and Alaska Native populations. *Cancer.* 2008;113(5 suppl):1120-1130.
- Frost F, Tollestrup K, Ross A, Sabotta E, Kimball E. Correctness of racial coding of American Indians and Alaska Natives on the Washington State death certificate. *Am J Prev Med.* 1994;10:290-294.
- Stehr-Green P, Bettles J, Robertson LD. Effect of racial/ethnic misclassification of American Indians and Alaska Natives on Washington State death certificates, 1989-1997. *Am J Pub Health.* 2002; 92(3):443-444.
- Puukka E, Jackson S, Stehr-Green P. Sexually transmitted diseases among American Indians and Alaska Natives in Washington State, 1995-2000. Unpublished report. Northwest Portland Area Indian Health Board. 2003.
- US Census Bureau. 2005-2007 American Community Survey 3-Year Estimates. Available at: [http://factfinder.census.gov/home/saff/main.html?\\_lang=en](http://factfinder.census.gov/home/saff/main.html?_lang=en). Accessed August 12, 2009.
- Espey D, Paisano R, Cobb N. Regional patterns and trends in cancer mortality among American Indians and Alaska Natives, 1990-2001. *Cancer.* 2005;103(5):1045-1053.
- Sugarman JR, Hill G, Forquera R, Frost FJ. Coding of race on death certificates of patients of an urban Indian health clinic, Washington, 1973-1988. *IHS Prim Care Provid.* 1992 Jul;113-115.
- North American Association of Central Cancer Registries. Cancer Data Standards. Available at: [http://www.naacr.org/index.asp?Col\\_SectionKey=7&Col\\_ContentID=73](http://www.naacr.org/index.asp?Col_SectionKey=7&Col_ContentID=73). 2007. Accessed April 30, 2009.
- National Center for Health Statistics. Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2007, by year, county, age, bridged race, Hispanic origin, and sex (vintage 2007). Prepared under a collaborative arrangement with the US Census Bureau; released September 5, 2008. Available at: [http://www.cdc.gov/nchs/nvss/bridged\\_race.htm](http://www.cdc.gov/nchs/nvss/bridged_race.htm). Accessed October 7, 2009.
- Washington State Department of Health. Guidelines for using and developing rates for public health assessment. 2006. Available at: <http://www.doh.wa.gov/Data/guidelines/Rateguide.htm#standpop>. Accessed July 7, 2009.
- Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: A method based on the gamma distribution. *Stat Med.* 1997;16:791-801.
- Pinheiro PS, Sherman R, Fleming LE, et al. Validation of ethnicity in cancer data: Which Hispanics are we misclassifying? *J Reg Mgmt.* 2009;36(2):42-46.
- Johnson JC, Soliman AS, Tadgerson D, et al. Tribal Linkage and race data quality for American Indians in a state cancer registry. *Am J Prev Med.* 2009;36(6):549-554.
- Indian Health Service, Portland Area Office. Available at: <http://www.ihs.gov/FacilitiesServices/AreaOffices/Portland/>. 2009. Accessed September 9, 2009.
- Northwest Portland Area Indian Health Board. Suquamish Tribe. Available at: [http://www.npaihb.org/member\\_tribes/tribe/suquamish\\_tribe/](http://www.npaihb.org/member_tribes/tribe/suquamish_tribe/). 2009. Accessed August 12, 2009.