

# Effect of Racial/Ethnic Misclassification of American Indians and Alaskan Natives on Washington State Death Certificates, 1989–1997

Paul Stehr-Green, DrPH, MPH, James Bettles, BS, and L. Dee Robertson, MD, MPH

Previous reports have shown that American Indian and Alaskan Native race is underreported on death certificates and other health-related data sets in Washington State and elsewhere.<sup>1–7</sup> In 1999, the Northwest Tribal Epidemiology Center, Northwest Portland Area Indian Health Board, established the Northwest Tribal Registry to provide a means of correcting for racial/ethnic misclassification and missing race/ethnicity data in public and other health-related data sets. In this article we describe the results of a linkage study of the Registry and Washington State death files that examined racial/ethnic misclassification.

## METHODS

The Northwest Tribal Registry includes all American Indians/Alaskan Natives who received health care services at American Indian facilities in Idaho, Oregon, or Washington (excluding the Seattle Indian Health Board and 2 other smaller Puget Sound–area tribal clinics) between 1994 and mid-1998 and an unknown percentage served during the mid-1980s through 1993 ( $n=135\,060$ ). The Washington State death files contain data for persons who died between 1989 and 1997 ( $n=363\,760$ ).

We used AutoMatch, Version 4.2 (Matchware Technologies, Inc, Kennebunk, Me, 1998), which employs a probabilistic record linkage algorithm, to link the Registry to the death files. This software not only identifies exact matches of records but also calculates the probability of a correct match in situations in which minor differences exist in the selected characteristics between the 2 different data files for a prospective match.

The record linkage was conducted in a 5-pass run, rotating primary matching variables and allowing for errors on some data fields but not others in each pass. On completion,

**Objectives.** This study examined effects of racial/ethnic misclassification of American Indians and Alaskan natives on Washington State death certificates.

**Methods.** Probabilistic record linkage were used to match the 1989–1997 state death files to the Northwest Tribal Registry.

**Results.** We identified matches for 2819 decedents, including 414 (14.7%) who had been misclassified as non-American Indians and Alaskan natives on the death certificates. The likelihood of being correctly classified increased 3-fold for each higher level of American Indian and Alaskan native ancestry (odds ratio = 2.88; 95% confidence interval [CI] = 2.51, 3.30) and decreased by 6.9% per calendar year (95% CI = 2.0, 11.5).

**Conclusions.** Systematic biases on death certificates in Washington State persist. Methods to reduce misclassification can improve data quality and enhance efforts to measure and reduce racial/ethnic health disparities. (*Am J Public Health.* 2002;92:443–444)

all exact matches were identified, and we reviewed questionable matches to further distinguish true matches. The final set of matched records, with all identifiers removed, comprised a data set of deaths in Washington State during 1989 through 1997 among persons known to be American Indians/Alaskan Natives. For our analyses of misclassification, we excluded all unmatched records because we had no authoritative basis (e.g., the Registry) for verifying the correct racial/ethnic classification of these decedents.

For analyses that examined place of residence at time of death, we used the US Office of Management and Budget–defined metropolitan statistical areas, which comprised Benton, Clark, Franklin, Island, King, Kitsap, Pierce, Snohomish, Spokane, Thurston, Whatcom, and Yakima counties.

## RESULTS

We identified matches for 2819 decedents, including 2405 (85.3%) who had been identified as American Indians and Alaskan natives in both data sets and 414 (14.7%) who had been misclassified as non-American Indians/Alaskan Natives on the death certificates.

Women were equally as likely as men to be misclassified (14.6% vs 14.7%,  $P>.10$  by  $\chi^2$  test). The youngest (i.e., <21 years) and oldest (i.e.,  $\geq 65$  years or older) decedents were statistically significantly more likely to be misclassified (19.8% and 17.7%, respectively) than were those aged 21 to 64 years (10.9%), but there was no statistically significant linear trend of misclassification by age group. Persons living in metropolitan statistical areas at the time of their death were slightly, but statistically significantly, more likely to be misclassified (15.8% of the metropolitan statistical area residents vs 13.0% of those living in rural areas,  $P<.05$  by  $\chi^2$  test).

A strong, statistically significant inverse association was found between blood quantum (i.e., a measure of American Indian/Alaskan Native ancestry) and misclassification: persons with a 100% blood quantum were less than one tenth as likely to be misclassified (4.0%) as persons with a less than 25% blood quantum (43.6%) ( $P<.001$  by linear test for trend). There was also a statistically significant trend of increasing misclassification over the 9-year time span of this study ( $P<.001$  by linear test for trend).

All of these factors were controlled simultaneously in a logistic regression model. We included, in addition to the variables listed

above, 3 dummy variables in the model indicating whether the death was due to any of the 3 causes known to be more prevalent among American Indians/Alaskan Natives (i.e., diabetes, injuries, and mental health disorders) and was therefore more likely to be correctly classified. Independent of all other factors, the strongest predictor of correct racial/ethnic classification was blood quantum: for each successively higher level of blood quantum, the likelihood of being correctly classified as American Indians/Alaskan Natives increased by almost 3-fold (odds ratio [OR]=2.88; 95% confidence interval [CI]=2.51, 3.30). Living in a non-metropolitan statistical area at the time of death also statistically significantly increased the likelihood that a decedent would be correctly classified (OR=1.55; 95% CI=1.18, 2.03). The probability of correct classification decreased for every year of death after 1989 by 6.9% per year (95% CI=2.0, 11.5).

## DISCUSSION

Health assessment involves estimating the burden of disease in a community or group, then using this information to inform public decision making. The validity and accuracy of health assessments can be adversely affected when health or demographic data are not reported accurately or completely in vital records and disease registries (e.g., cancer registries), thereby diminishing the usefulness of this health information. In this case, the misclassification of American Indian/Alaskan Native race/ethnicity on death certificates led to an underestimation of overall and disease-specific mortality burden for American Indians/Alaskan Natives.

Misclassification of race/ethnicity occurs in health-related databases most often because the recorded information is based on observation by physicians, coroners or medical examiners, or other health care workers rather than on patient self-reports or reports by close relatives. This analysis corroborates earlier reports that showed considerable misclassification of American Indian/Alaskan Native race/ethnicity on Washington State death certificates and, furthermore, that this misclassification is systematic (i.e., it varies directly with year of death and urban residency and inversely with blood quantum) and that these biases may be worsening over time.

Improving identification of race/ethnicity from an authoritative external data source has the potential to greatly reduce the problem of misclassification, as demonstrated by this record linkage study that used the Northwest Tribal Registry. In this study, we identified 414 more American Indian/Alaskan Native deaths than the number officially recorded in Washington from 1989 through 1997 (an increase of 9.8% over the official tally), thereby mitigating the effect of the biases and permitting better estimates of the burden of cause-specific deaths for Washington American Indians/Alaskan Natives.

To achieve optimal utility of the Registry, we must improve its completeness and, at the least, ensure its representativeness. Although the Registry was representative of the Northwest American Indian/Alaskan Native population alive at the time of the 1990 census, it is not known with certainty whether the Registry currently includes a representative sample of American Indians/Alaskan Natives in Washington, Oregon, and Idaho (P.A. Stehr-Green, DrPH, MPH, J. Bettles, BS, L.D. Robertson, MD, MPH, unpublished data, 2000); until this is known with confidence, we are limited in our ability to generalize to all Northwest American Indians/Alaskan Natives.

We also know that many urban-dwelling Northwest American Indians/Alaskan Natives who have received services from the Seattle Indian Health Board have not accessed health care services at any other Northwest Indian health care facility in the previous 10 years and thus are not included in the Registry (J. Bettles, Northwest Portland Area Indian Health Board, unpublished data, 1999). Although the total number is sizable (n=3401), it represents less than 3% of the total listed in the Registry, so it seems unlikely that the absence of these individuals from the Registry resulted in a substantial bias in the results of this record linkage study. Their continued absence prevents us, however, from deriving a complete understanding of the burden of disease and death among urban American Indian/Alaskan Native populations and in achieving a complete ascertainment of Northwest American Indians/Alaskan Natives for use in future linkage studies.

Nonetheless, this analysis and others that preceded it show the potential of record link-

age to improve our understanding of the burden of disease among American Indians/Alaskan Natives and of concomitant risk and etiologic factors. Ultimately, this improved understanding of American Indian/Alaskan Native health status can provide the foundation for programmatic activities to improve health and well-being in Northwest American Indian/Alaskan Native communities. ■

## About the Authors

The authors are with the Northwest Tribal Epidemiology Center (The EpiCenter), Northwest Portland Area Indian Health Board, Portland, Ore.

Requests for reprints should be sent to Chandra Wilson, Northwest Portland Area Indian Health Board, 527 SW Hall, Suite 300, Portland, OR 97201 (e-mail: cwilson@npaihb.org).

This article was accepted January 17, 2001.

Note. This study was approved by the Portland Area Indian Health Service's and the Indian Health Service's Headquarters-Albuquerque Area Combined Institutional Review Boards.

## Contributors

P. Stehr-Green, J. Bettles, and L.D. Robertson planned the study. J. Bettles conducted the probabilistic linkages. P. Stehr-Green analyzed the linked data and wrote the paper. L.D. Robertson contributed to the writing of the paper.

## References

1. 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.
2. 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. *Provider.* 1992;July:113-115.
3. Frost F, Taylor V, Fries E. Racial misclassification of Native Americans in a Surveillance, Epidemiology, and End Results Registry. *J Natl Cancer Inst.* 1992;84:957-962.
4. Sugarman JR, Holliday M, Ross A, Castorina J, Hui Y. Improving American Indian cancer data in the Washington State cancer registry using linkages with the Indian Health Service and tribal records. *Cancer.* 1996;78(suppl 7):1564-1568.
5. Sugarman JR, Lawson L. The effect of racial misclassification on estimates of end-stage renal disease among American Indians and Alaska Natives in the Pacific Northwest, 1988 through 1990. *Am J Kidney Dis.* 1993;21:383-386.
6. Sugarman JR, Soderberg R, Gordon J, Rivara FP. Racial misclassification of American Indians: its effect on injury rates in Oregon, 1989 through 1990. *Am J Public Health.* 1993;83:681-684.
7. Epstein M, Moreno R, Bacchetti P. The underreporting of deaths of American Indian children in California, 1979 through 1993. *Am J Public Health.* 1997;87:1363-1366.