The Effects of Racial Misclassification
Part two of a two-part series.

Part one of this series examined problems associated with multiple race categories in calculating race-specific disease rates (see Health News and Notes, July 2003). This problem is secondary to the overall issue of racial misclassification however, in that the allocation of multiple race categories into single categories is meaningless if the race does not accurately reflect an individuals' heritage. This article will illustrate the effects of racial misclassification on statistical measures commonly used to characterize disease burden in a population.

Racial misclassification can be described as the incorrect coding of an individuals' race or ethnicity, for example, an AI/AN individual incorrectly coded as white. For smaller race groups, misclassification generally occurs in one "direction", that is, AI/ANs are often misclassified as white, however white individuals are rarely misclassified as AI/AN.

Correct racial classification of individuals is essential for the accurate characterization of disease within a population. Race-specific measures such as disease rates which are used to express the force of a disease in a given population can increase or decrease significantly based on the number of cases or persons (correctly or incorrectly) included in either the numerator or denominator. The potential for error and ultimate impact is greatest among small populations and race groups.

An example of the effect of racial misclassification on disease rates is shown in table 1. In this example, for disease "X", the true number of cases among the AI/AN population is 650. As illustrated, the resulting rate of disease (with no racial misclassification) would be 867.0 per 100,000 person-years. As racial misclassification increases, (i.e. fewer records are correctly classified as AI/AN in the state disease registry), the disease rates decrease, suggesting (falsely), a lesser overall burden of disease.

At first glance, low disease rates may be interpreted as an indication of a healthy population, however, it is necessary to consider the data source and possible limitations prior to drawing any meaningful conclusions. As with the example presented here, low disease rates may actually be the result of an introduced error or bias (e.g. low screening rates, racial misclassification, etc.) rather than an actual decrease or difference in the number of cases of disease in a population.

It is also important to recognize that because the burden of disease in a given population is often used in the allocation of funding for, and planning and management of, public health programs, it is especially important that racial misclassification and other potential sources of error be addressed.

For additional information, please contact:
Emily Puukka, MS
Project Director, Northwest Tribal Registry Project
503.228.4185 Ext. 285
epuukka@npaihb.org

Table 1. Effect of Racial Misclassification on Disease Rates - Example

True Number of Cases of Disease "X" Among AI/ANs (No Racial Misclassification)	Percent Racial Misclassification in State Disease Registry	Number of Cases of Disease "X" in State Registry Coded as AI/AN	Denominator (From 2000 US Census)	Resulting Disease Rate (per 100,000)
650	0%	650	75,000	(650/75,000)*100,000 = **867.0**
650	20%	520	75,000	(520/75,000)*100,000 = **693.0**
650	40%	390	75,000	(390/75,000)*100,000 = **520.0**
650	60%	260	75,000	(260/75,000)*100,000 = **347.0**