Racial misclassification is a well-documented weakness of mortality data taken from death certificates. As a result, mortality statistics for American Indians and Alaska Natives (AI/ANs) present, at best, an inaccurate and misleading assessment of mortality in this population.

Studies evaluating the quality of race/ethnicity reporting on death certificates have linked data from death certificates to other data sources collected when the decedent was still alive (e.g., Census, Current Population Survey). Such studies have shown substantial misclassification of AI/AN decedents.

Despite limitations, linking mortality data from death certificates with data from other sources collected when decedents were living provides opportunities to evaluate and correct misclassification of populations such as AI/AN persons and facilitates the calculation and presentation of more accurate mortality statistics. (Am J Public Health. 2014;104: S258–S262. doi:10.2105/AJPH.2013.301647)

MORTALITY DATA PROVIDE critical information for measuring health in the United States. Knowing why and how people die helps researchers understand the nature and magnitude of health problems and is important for the planning, implementation, and evaluation of public health strategies and programs.4–13 Mortality data are widely used and relied on to provide health information at the national, state, and local levels.4 Of the 10 great public health achievements of the 20th century enumerated by the Centers for Disease Control and Prevention, most have been measured wholly or in part using mortality data.5

Addressing racial and ethnic disparities in health and mortality is an important priority in US public health efforts.5 A large body of research has been done to understand why some racial/ethnic groups have better health and mortality outcomes than others.6 The reduction of racial/ethnic disparities in health was an important focus of Healthy People 2000,8 and this goal was strengthened with the introduction of Healthy People 2010 to entail the elimination of racial/ethnic disparities in health.9 Eliminating disparities remains an important goal of Healthy People 2020.10 Nevertheless, racial/ethnic differences in health and mortality persist and remain pervasive.

Accurate assessment of racial/ethnic disparities and tracking progress on these goals depends on the quality, validity, and reliability of proper measurement of the data that are reported. However, well-documented problems with reporting of race and ethnicity in mortality data complicate researchers’ ability to accurately measure and thereby assess disparities for some racial/ethnic groups.11–14 Misclassification has been particularly pronounced for American Indians/Alaska Natives (AI/ANs). The result is that death data for AI/AN populations at the national level and within most states presents, at best, an inaccurate and misleading assessment of mortality in this group.

Mortality data in the US National Vital Statistics System allow for the production of mortality statistics for small populations and small areas because of the uniformity of death certificate content across the country and the virtual completeness of the data. This has important implications for the production of mortality statistics for population subgroups such as AI/AN communities that are too small to be adequately represented in survey data based on sampling. The fact that the racial and ethnic information in the National Vital Statistics System is not sufficiently accurate greatly weakens an otherwise powerful resource for the assessment of AI/AN health.

RACE/ETHNICITY REPORTING ON DEATH CERTIFICATES

Death certificates filed in the 50 states and the District of Columbia provide the basis for US national...
mortality statistics derived from the National Vital Statistics System. Death certificate information is derived from two sources.\textsuperscript{15} Cause-of-death information is reported by physicians (or medical examiners or coroners for injury-related, sudden, or unattended deaths). Demographic and personal information is typically reported by funeral directors. Funeral directors generally obtain this information from an informant, usually a close relative of the decedent.

The reporting of race and ethnicity on death certificates is complicated for two reasons. First, racial identification is a social construct and may change over time.\textsuperscript{16} This is particularly true for those with a multiracial ancestry, which describes much of the US AI/AN population.\textsuperscript{17-21} Indeed, which describes much of the US AI/AN population with a multiracial ancestry, or survey self-report.\textsuperscript{12-14,22}

Linkage studies have shown significant misclassification of Hispanic, AI/AN, and Asian/Pacific Islander decedents. For AI/AN decedents in particular, the misclassification reported in these studies has been quite substantial. Most are usually classified as White.\textsuperscript{14} Because the AI/AN population is small, even a relatively small number of misclassified decedents can have a substantial impact on mortality statistics. The large misclassification rate for the AI/AN population also reflects the complex and personal nature of multiracial identity. It may not be apparent to those completing the death certificate (i.e., funeral directors) or those acting as informants how decedents would have identified themselves. This may be true even for close family members, particularly if the decedent and the family informant are of different generations.

Misclassification causes particular problems in the calculation of death rates because an inconsistency in reporting occurs between numerator and denominator. The numerator is the number of deaths based on proxy-reported death certificate data and the denominator is the mid-year population based on estimates derived from self-reported census data. The result of this inconsistency is, in the case of AI/AN persons, death rates that are much too low.

**EVALUATING AND CORRECTING FOR MISCLASSIFICATION**

Efforts to evaluate the quality of reporting of race/ethnicity on death certificates date back to the 1960s. Hambright matched death certificates with corresponding records from the 1960 Census to examine social and economic characteristics, including race, not reported on the death certificate and compare responses to the same questions on both records.\textsuperscript{12} Race reported in the census record is believed to be more accurate than race reported on the death certificate because it is self-reported or reported by a household member while the person is still alive. For the AI/AN population, Hambright found relatively low (79.2\%) agreement between the census record and death certificates.\textsuperscript{12}

More recent efforts to evaluate race/ethnicity misclassification have used the National Longitudinal Mortality Study (NLMS). The NLMS is an ongoing data linkage project; it combines data from the Current Population Survey Annual Social and Economic Supplement with a sample of the 1980 decennial census and links these records to death certificate data from the National Vital Statistics System using the National Death Index (NDI).\textsuperscript{13,14,22,24} The first study using the NLMS to evaluate race/ethnicity misclassification found results similar to those of Hambright (73.6\%).\textsuperscript{12,13} Subsequent analyses by Rosenberg et al.\textsuperscript{22} and Arias et al.\textsuperscript{14} using expanded NLMS data (including additional Current Population Survey data and years of mortality follow-up) found much lower agreement (57.4\% and 55.2\%, respectively). The large decline in the level of agreement likely reflects the substantial and well-documented growth in the number of persons identifying themselves as an AI/AN individual between 1960 and 1990, especially among those with multiracial ancestry.\textsuperscript{17-19,21}

The wide differences in the levels of misclassification documented suggest additional study is needed to better understand the factors that affect misclassification if a reliable approach to adjusting the data is to be developed.

Information on the quality of race/ethnicity reporting derived from linkage studies can be used in three different ways. First, the results from these studies allow calculation of adjustment factors or classification ratios that can be applied to death rates to adjust for misclassification. For example, using the NLMS, one can calculate a classification ratio by comparing race or ethnicity collected for individuals on the death certificate with race or ethnicity collected for the same individuals (while still alive) in the Current Population Survey. Death rates for AI/AN persons that are adjusted using a classification ratio based on the most recent NLMS study rise by 30\%.\textsuperscript{14} Although unadjusted death rates for AI/AN persons are lower than those for Whites, adjustment using a classification ratio of 1.30 raises death rates for AI/AN persons such that they are higher than those for Whites. Hispanic life tables are also currently produced using this method.\textsuperscript{24,25}

Second, linkage data are often used directly to generate mortality statistics for racial/ethnic groups.\textsuperscript{26-33} Several studies using NLMS data have used this approach\textsuperscript{26-30} which is especially useful when the death certificate is lacking information that is available in the linked data set (e.g., Current Population Survey, census), such as income, poverty status, occupation, and household composition.

Third, if the linkage is comprehensive enough, corrections
can be made to or flags indicating a reporting discrepancy can be placed on individual records within the mortality files. Using the corrected or flagged records, more accurate statistics can be calculated. For example, the Indian Health Service (IHS)–mortality linkage described in this supplement resulted in an AI/AN flag being placed on mortality records identified as AI/AN in the IHS data, allowing for the calculation of more accurate death rates for AI/AN persons living in counties in which IHS has more of a presence.

**PRESENTING AI/AN MORTALITY STATISTICS**

When presenting unadjusted mortality indicators for AI/AN individuals, it is important to qualify the data and to caution readers and data users as to the extent of and reasons for misclassification. Because of the substantial misclassification of AI/AN decedents, it is important that unadjusted mortality indicators for AI/AN persons not be directly compared with those from other racial/ethnic groups. Unadjusted data are not appropriate for the presentation or analysis of racial/ethnic disparities in mortality. Analysis of unadjusted trends, however, may be acceptable if one can assume that the misclassification is roughly constant over time. Data from the NLMS have shown that this is likely the case, at least for more recent data. Consumers of mortality data need to be alerted to these issues, and care should be taken not to facilitate improper comparisons.

The current linkage data available have important limitations that must be considered when using them to evaluate, generate, or present mortality statistics for small populations such as AI/AN persons. Small numbers and large sampling error in survey linkages such as the NLMS make it difficult to make an accurate assessment of the quality of mortality statistics for AI/AN persons or to generate reliable mortality statistics. This is especially true when looking at subgroups, for example, tribes, within the AI/AN population at the national level or when presenting any data at the state or local level.

The IHS linkage, although not limited by small numbers and sampling error as is the NLMS, is limited in its representativeness of the AI/AN population. Only those AI/AN decedents with records in the IHS patient database are included in the linkage and only death records for these decedents are correctable. Although still very useful, the IHS linkage provides a picture for only a subset of the AI/AN population, and it is also geographically limited in its coverage, primarily to Contract Health Service Delivery Area (CHSDA) counties. CHSDA counties are those in which contract health services are made available by the IHS or tribe to members of an identified Indian tribe who reside in the area. CHSDA counties have a higher proportion of AI/ANs in relation to the total population than do non-CHSDA counties, with 64% of the US AI/AN population residing in the 637 counties designated as CHSDAs (20% of 3141 US counties).

From the standpoint of producing official national, state, or local mortality statistics for AI/AN persons, existing information is largely inadequate or insufficient to comprehensively adjust mortality indicators for misclassification of AI/AN persons. The NLMS sample size is too small to calculate reliable classification ratios by demographic subgroups or subnational geographic areas, and the IHS linkage only allows for correction of a subset of AI/AN deaths. One could simply assume a constant level of misclassification for the AI/AN population across all demographic subgroup and geographic areas and apply a single classification ratio (i.e., 1.30) to all death rates for AI/AN persons. However, Arias et al. showed variation across age groups (although confidence intervals were quite large, especially for the younger population). They also showed substantial geographic variation; misclassification is less likely in areas, such as CHSDA counties, with a greater concentration of AI/ANs.

Limiting presentation and analysis to CHSDA counties mitigates to some extent the effects of misclassification because AI/AN decedents in CHSDA counties are less likely to be misclassified. This strategy has been used in some studies. Now, with the publication of data based on the IHS–NDI linkage study, deaths among AI/AN persons living in CHSDA areas can be largely corrected for misclassification. Of course, AI/AN persons living (and dying) in CHSDA areas are not necessarily nationally representative. Consumers of mortality data need to understand the limitations of this strategy.

**A NEW MODEL**

Despite limitations, data linkage is the future for interpreting and presenting mortality for populations that are often misclassified on death certificates. Linked data are particularly important for the AI/AN population because of the substantial number of AI/AN decedents for whom race is misclassified in mortality data. Linking mortality data with data from other sources collected when decedents were living provides opportunities to evaluate and correct for misclassification and facilitates the calculation and presentation of more accurate mortality statistics. In addition, such linkages present opportunities to add to what is reported on the death certificate, which tends to be limited, especially with regard to socioeconomic indicators.

The effort, described in this supplement, to link IHS data with mortality data via the NDI is an important step forward in improving mortality statistics for AI/AN persons. Although limited in terms of coverage and representativeness, these data are still an important resource for researchers studying AI/AN mortality. The limitations of these very useful data also serve to highlight the potential benefits of more comprehensive linkages.

From the standpoint of evaluation and correction of AI/AN mortality, opportunities exist to link decennial census data for AI/AN respondents to mortality data. Ideally, this would involve annual mortality follow-up using the NDI, which would allow for the correction of race reporting for all AI/AN decedents (i.e., all those who reported themselves as such in the decennial census) in annual mortality files at the national, state, and local levels. This would ultimately result in much more accurate information for mortality statistics for AI/AN persons, although the challenge of delayed availability of data would persist.

However, as is often the case, the ideal solution is also an expensive one. The initial effort requires searching at least 2.9
Another approach to correction of mortality statistics for AI/AN populations, though limited in scope to populations that make up federally recognized domestic dependent nations or tribes, is establishing a standardized and ongoing death certificate and tribal membership data linkage process that would be managed by the national Tribal Epidemiology Center (TEC) network. Such an approach has been used with cancer registry data and would involve correcting death certificate data with misclassified race to ensure that all deaths of tribal members are correctly classified as of AI/AN race. Funded by the IHS, the nation’s 12 TECs work in partnership with federally recognized tribal governments, federally recognized tribal government coalitions, and urban Indian health organizations to conduct a variety of epidemiologic activities to improve health.

Each TEC is designated to serve the federally recognized tribes within one of the 12 IHS administrative areas, although one TEC serves two IHS areas and another TEC serves urban Indian health organizations throughout the nation. Because of these unique designations, TECs could be ideal hubs for collecting, linking, protecting, analyzing, and communicating mortality statistics for AI/AN persons.

These approaches are obviously not without limitations. For the census and NDI approach, reporting of race/ethnicity in the census, although superior to what is reported on the death certificate, is certainly not perfect. Under-counting of AI/AN persons in the census potentially limits the ability to detect misclassification in some cases. Challenges associated with changing racial identification and multiple-race reporting also remain. Clear decision rules need to be applied to appropriately classifying persons reporting AI/AN race/ethnicity in combination with other races. The TEC approach would only improve the quality of mortality statistics on domestic dependent nation populations and would not include those who are not tribally affiliated. Furthermore, for such a TEC-based approach to succeed, recognition and acceptance at multiple government levels, a standardized denominator method, and additional investment in TEC infrastructure would be required.

As already stated, linking IHS data with mortality data via the NDI has been an important step in improving the quality of mortality statistics for AI/ANs, as evidenced by this supplemental issue of the Journal. Building on this work, an in-depth exploration of the viability of other approaches will help provide AI/AN leaders and their constituents with improved mortality data and statistics for measuring health. Furthering this work is essential to establishing the ability to routinely assess health disparities in the AI/AN population. Knowing why and how their people are dying will help tribal nations and the United States as a whole to better plan, implement, and evaluate public health strategies and programs focused on alleviating AI/AN health disparities.

**About the Authors**

Robert N. Anderson is with the Division of Vital Statistics, National Center for Health

Statistics, Centers for Disease Control and Prevention, Hyattsville, MD. Glenn Copeland is with the Division for Vital Records and Health Statistics, Michigan Department of Community Health, Lansing. John Mosely Hayes is with the United South and Eastern Tribes, Tribal Epidemiology Center, Nashville, TN.

Correspondence should be sent to Robert N. Anderson, Division of Vital Statistics, 3311 Toledo Road, Room 7318, Hyattsville, MD 20782 (e-mail: R.N.Anderson@cdc.gov). Reprints can be ordered at http://www.aph.html by clicking the "Reprints" link.

This article was accepted August 22, 2013.

**References**


---


Anderson et al. | Peer Reviewed | Commentary | S261
This article has been cited by:

1. Ursula E. Bauer, Marcus Plescia. 2014. Addressing Disparities in the Health of American Indian and Alaska Native People: The Importance of Improved Public Health Data. *American Journal of Public Health* **104**:S3, S255-S257. [Citation] [Full Text] [PDF] [PDF Plus]