Preliminary. Please do not cite/distribute without permission.

Mortality Among Persons Who Self-Identify as American Indian or Alaska Native

Michael Bird (Kewa Pueblo)
Jacob Bor* (Boston University)
Jahn Hakes (U.S. Census Bureau)
David Himmelstein (City University of New York)
Steffie Woolhandler (City University of New York)

*jbor@bu.edu

Prepared for NBER Racial and Ethnic Health Disparities Workshop, April 28, 2023

ABSTRACT

Race is not always reported accurately on death certificates, which can lead to biased estimates of mortality rates when numerators and denominators are taken from different sources. Native American/American Indian/Alaskan Native (AIAN) identity in particular is significantly underreported, leading to underestimates of mortality rates in this population. Using nationally representative, longitudinal data from the Mortality Disparities in American Communities study, 2008-2015, we estimate mortality rates and life expectancy for self-identified AIAN and other racial/ethnic groups without bias due to misclassification on death certificates. We find that the excess mortality risk among AIAN relative to the total U.S. population increases by 53% after accounting for misclassification.

1. Background

Life expectancy in the U.S. varies by race/ethnicity. Official U.S. government estimates rely on matching death certificates with population denominators. However, demographics on death certificates are not always reported accurately, which can lead to biased mortality estimates.[1] [2] Native American/American Indian/Alaskan Native (AIAN) identity in particular is significantly underreported in death certificates,[3] despite rapid growth in the U.S. population identifying as AIAN on Census forms in recent decades.[4] The invisibility of AIAN in Census and health data has been referred to as "statistical erasure"[5] and follows centuries of deliberate policy to erase indigenous culture and identity, e.g. through genocide and forced assimilation.[6] The under-reporting of mortality among AIAN is particularly troubling because, even with the under-reporting, AIAN have among the highest mortality rates of any U.S. racial/ethnic group.

In this paper, we analyze restricted-use data from the Mortality Disparities in American Communities (MDAC) study. MDAC is a collaboration between the U.S. Census and National Vital Statistics System (NVSS). Specifically, the U.S. Census has linked individual respondents in the 2008 American Community Survey (ACS) to U.S. death records through 2015. (MDAC is in the process of being updated to 2019.) The MDAC database offers the opportunity to estimate population mortality rates based on 7 years of prospective follow-up for a representative population of the U.S. with different sociodemographic characteristics. In contrast to the conventional race-stratified mortality rates reported by CDC, the longitudinal mortality estimates from the MDAC are based on self-identified race/ethnicity, are not biased due to misreporting on death certificates.

---

[1] Arias E. The Validity of race and Hispanic origin reporting on death certificates in the United States. CDC 2008. Series: DHHS publication no. (PHS) 2009-1348. https://stacks.cdc.gov/view/cdc/7066

[2] Arias E, Heron MP, Hakes JK. Validity of race and Hispanic origin reporting on death certificates in the United States: an update. CDC. Vital and health statistics. Series 2, Data evaluation and methods research; no. 172, 2016. https://stacks.cdc.gov/view/cdc/45533

[3] *Op Cit*. Arias et al. 2014

[4] Passel JF. The growing American Indian Population: beyond demography. In: National Research Council, Committee on Population. Changing numbers, changing needs: American Indian demography and public health. 1996. https://www.nap.edu/catalog/5355/changing-numbers-changing-needs-american-indian-demography-and-public-health

[5] Small-Rodriguez D. https://www.drdrdesi.com/research

[6] Dunbar-Ortiz R. 2014. *An Indigenous People's History of the United States.*

Prior research has found substantial underreporting of AIAN racial identify on death certificates, ranging from 10% to 40%. Using death certificates augmented by Indian Health Service patient registration records, Arias et al[7] assessed AIAN life expectancy in 2007-2009 among persons living on or near tribal lands. They estimated AIAN life expectancy in the areas they studied at 71.1 years, compared to 78.4 years for the non-Hispanic White population nationally. A study of the AIAN population in the Pacific Northwest yielded similar findings.[8] Analyzing data on a sample of deaths linked to the 2010 Census, Arias and colleagues found that AIAN deaths are underreported by 25% in death certificates.[9] However, no recent longitudinal follow-up study has yet assessed AIAN mortality rates prospectively. The MDAC data enable calculation of unbiased estimates of AIAN age-specific mortality rates and life expectancy; comparison of mortality rates for different groups, e.g. AIAN alone, AIAN in combination with another race, Hispanic AIAN, AIAN on vs. off reservations; and assessment of individual-level socioeconomic factors as mediators of mortality disparities.

2. Data Description

Study population

The study population consisted of all persons residing in the U.S. in 2008, the sampling frame for the 2008 ACS. The ACS is a national survey conducted by the U.S. Census, with data collected every month of the year in all states, D.C., and Puerto Rico. Each year the survey is sent to a 2.5% sample of the U.S. resident population. Surveys are filled out by mail or online. The Census Bureau follows up with a sample of non-responses by telephone and through in-person visits. The ACS includes both people residing in households and those residing in group quarters, such as schools, prisons, and nursing homes.

---

[7] Arias E, Xu J, Jim MA. Period life tables for the non-Hispanic American Indian and Alaska Native population, 2007–2009. Am J Public Health. 2014;104(S3):S312-9.

[8] Dankovchik J, Hoopes MJ, Warren-Mears V, Knaster E. Disparities in life expectancy of Pacific Northwest American Indians and Alaska natives: analysis of linkage-corrected life tables. Public Health Reports. 2015 Jan;130(1):71-80.

[9] Arias E, Xu J, Curtin S, Bastian B, Tejada-Vera B. 2021. Mortality Profile of the Non-Hispanic American Indian or Alaska Native Population, 2019. NVS Report 70(12).

Data linkage

The U.S. Census Bureau has linked the 2008 ACS at the individual level with mortality records from the National Death Index to construct the MDAC database. The MDAC was established to enable research on mortality disparities by sociodemographic and small-area geographic variables.[10] The MDAC includes information on date of death, place of death, and cause of death (ICD-10 codes). Linkage between the ACS and National Death Index was conducted using an exact match on Social Security Number (SSN) or Name and Date of Birth. 91% of records in the ACS have SSN, enabling highly accurate matches. The remainder are linked on name and birth date.[11] The linkage algorithm is tuned to be conservative such that a match would only be made on name and birth date if there was very high confidence it was the same person.

Access to the MDAC

Access to the MDAC is restricted and we obtained approval to conduct this project from the Census. The team included a Census Statistician in the Census's Mortality Research Group who conducted the analyses and had direct access to the data. The broader team designed the analyses and Dr. Bor, who obtained Special Sworn Status, reviewed intermediate output through the Census's Research Output Direct Access System (RODAS). All output reported in this paper have been reviewed by the Census's Data Review Board (DRB) to ensure compliance with policies to protect confidentiality of respondents. Exact counts have been rounded according to a specific algorithm and small cells suppressed.

How representative is the MDAC?

The MDAC is intended to include all respondents in the 2008 ACS. However, not all respondents had adequate identifiers (Social Security Number [SSN], names, dates of birth) to be linked to the NDI. Of the 4.5 million ACS respondents, 33,000 (0.7%) lacked sufficient identifiers and were not sent to the National Vital Statistics System (NVSS) for linkage. An additional 7.6% were sent to NVSS but without SSN. When we analyzed mortality rates based on the linked data, we found that whereas 6.8% of all respondents died during the follow-up period, only 1.2% of respondents without a SSN died during follow-up (**Table S1**). After

[10] US Census Bureau. Mortality Disparities in American Communities (MDAC): Project overview. https://www.census.gov/topics/research/mdac.html
[11] I*bid.*

consultation with the Census statistician, we concluded that this most likely reflects under-linkage among persons without a SSN. In the absence of a SSN, the NVSS would have required a very close match on names and date of birth, leading to a lower linkage rate. Based on this analysis, we decided to exclude from the sample respondents without a SSN, because their mortality rates would be underestimated. People without a SSN were somewhat more likely to be poor, less educated, rural, immigrants, and to reside in group quarters. Differences were relatively small, however.

We assessed for evidence of a "healthy respondent" effect in the ACS. We find some suggestive evidence that mortality rates are a bit lower in year 1 for working age adults – consistent with survey non-participation of people in poor health. However, the differences were not very large, were not apparent in other age groups, and we decided not to exclude the first year of follow-up in our analysis (**Table S2**). Another limitation of the linked data is that persons who migrate outside of the U.S. remain in the denominator but would not be observed in U.S. vital statistics. These "immortals" could artificially deflate mortality estimates. We note that external migration is relatively low among AIAN, but it may affect mortality estimates of other groups, particularly those with large immigrant populations.

Overall, mortality rates in the MDAC data matched up with the CDC's reported mortality rates for the whole U.S. population (**Table S3**). Initial analyses revealed age-adjusted mortality estimates substantially lower than what CDC had reported nationally. (We extracted national data on deaths and population denominators for the period 2008-2015 from CDC WONDER as a comparison.) Excluding people without SSNs made a substantial difference, bringing the mortality rates in the MDAC much closer to the CDC reported national data. **Table S3** shows that MDAC age-specific mortality rates line up closely with the national CDC data at ages 5-34 and ages 65+, with a modest undercount observed at ages 35-64. Mortality rates in the first year of life are severely underestimated. Whereas the national infant mortality rate was 617 per 100K, the MDAC estimate was 111 per 100K. It is likely that the ACS simply misses newborns who die before they leave the hospital. Additionally small numbers of deaths among children mean that data do not meet Census suppression rules for release. Therefore, we focus on the population ages 15 and older.

Study Sample

After excluding people without SSNs, our sample included 4,136,000 ACS respondents, including 58,000 people who self-identified as American Indian or Alaskan Native (AIAN). We defined AIAN as non-Hispanic persons reporting AIAN alone or in combination with other racial categories. **Table 1** shows the characteristics of the study sample. AIAN had similar poverty and education levels as Black and Hispanic respondents but were much more likely to reside in rural areas than those demographic groups. 24% of AIAN resided in a Census-defined "AIAN area" including an Indian Reservation or other designated area.

3. Misclassification of AIAN on death certificates

Table 1 shows the deaths occurring in the linked MDAC database. We observed 58,000 people who self-identified as AIAN for 423,000 person years of follow-up, during which period 926 self-identified AIAN died. When relying on race/ethnicity as reported in death certificates, however, only 449 of all deaths in the MDAC database had "AIAN" recorded as the decedent's race. In other words, using death certificate-based race/ethnicity instead of self-reported race/ethnicity would yield death counts (and mortality rates) that were 48.5% of what they should be, implying that true AIAN mortality could be as much as twice as high as existing unadjusted estimates. This number (48.5%) is substantially lower than the 75% estimated by Arias et al. using 2010 Census data (Arias 2021) and we are working to resolve this discrepancy.

Table 2 shows the racial classification matrix of deaths observed in the MDAC database. Nearly 60% of deaths occurring to self-identified AIAN were misclassified on death certificates, with 51% coded as non-Hispanic White and 7% coded as non-Hispanic Black. Deaths among self-identified AIAN were much more likely to be misclassified than any other group, with just 0.5% of White decedents misclassified, 3% of Black decedents misclassified, 7% of Asian or Pacific Islander decedents misclassified, and 13% of Hispanic decedents misclassified, compared to 60% of AIAN decedents.

4. Impact of misclassification on death rates

Table 1 shows crude death rates for AIAN when using self-identified vs. death certificate assigned race/ethnicity. When using self-identified race/ethnicity, the crude death rate was 926 per 100K, the second-highest of any racial group. When using death certificate assigned race/ethnicity, the crude death rate was 449, about 50% lower and the second-lowest of any racial group. Of course, difference across racial groups are driven in part by differences in the age distribution. **Table 3** shows results by age group. The table compares death rates for the AIAN population when using self-identified race/ethnicity vs. race/ethnicity from death certificates. The table also reports misclassification ratios by age. Relying on death certificate-assigned race leads to underestimating death rates under 45 years by about one third and over 45 years by over half. To take one example, the estimated death rate for AIAN ages 65-74 using death certificate race/ethnicity was 1180 per 100K. When using self-reported race, the death rate was 2720 per 100K.

One reason for the large observed difference is that the race data on death certificates are provided with "bridged race" categories, in which people who report multiple races are assigned to a single race. Some people who self-reported AIAN and another category may have been bridged into the other category on the death certificate. We therefore conducted additional analyses comparing the MDAC self-report data with CDC Wonder data which use bridged race categories in both the numerator and denominator.

Figure 1 displays the age-specific mortality rates shown in Table 3 alongside the equivalent estimates from CDC Wonder (2008-2015), calculated using separate numerator and denominator data. The yellow line in Figure 1 shows the ratio of AIAN mortality as estimated in CDC Wonder data vs. our MDAC estimates. The ratio declines with age, indicating that CDC Wonder data are more complete for younger age groups and the impact of misclassification is larger at older age groups.

Figure 2 shows mortality rate ratios comparing AIAN age-specific mortality rates with age-specific mortality rates for the whole U.S. population regardless of race/ethnicity. The line with solid circles shows CDC Wonder data. The line with hollow circles shows MDAC data, which lie above the CDC Wonder data for all age groups. AIAN death rates are more than twice the national average for persons ages 15-54 and remain above the national average at all age groups. The survival advantage previously reported among older AIAN based on CDC Wonder data may be an artifact of misclassification.

Figure 3 summarizes the results across age groups in age-standardized mortality rates. Across all racial/ethnic groups, age standardized mortality in the MDAC sample was 1063 deaths per 100,000. The mortality rate among AIAN was 1324 per 100,000 when using race as reported on death certificate and 1463 per 100,000 when using self-reported race. The gap between AIAN age-standardized mortality and total age-standardized mortality increases by 53% after accounting for misclassification.

5. Explaining the AIAN mortality disadvantage

A final aim of our analysis was to assess what mediating factors are associated with excess mortality among AIAN relative to other U.S. population groups. **Table 4** shows preliminary results. We fit Cox Proportional Hazards models to assess whether being self-identified Non-Hispanic AIAN is associated with higher (or lower) mortality. The models all control flexibly for age. Columns (1) and (3) compare non-Hispanic AIAN to non-Hispanic Whites. Columns (2) and (4) compare non-Hispanic AIAN to all other groups. Columns (1) and (2) show that being AIAN was associated with about a 40% increased mortality risk across follow up. Columns (3) and (4) show that adjusting for economic, social, and demographic factors (see Table for list of controls included) reduced this association by about half to around 20% increased risk. In other words, about half of the excess mortality risk is statistically explained by factors such as living in a rural area and having lower income / education / employment.

6. Conclusion

The AIAN mortality disadvantage relative to other U.S. racial/ethnic groups is substantially larger than it appears in traditional CDC estimates. Addressing high mortality rates among AIAN – a legacy of past and present discrimination and disadvantage – is an urgent policy priority.

FIGURES

Figure 1. Age-specific mortality rates for AIAN: CDC Wonder, MDAC, ratio

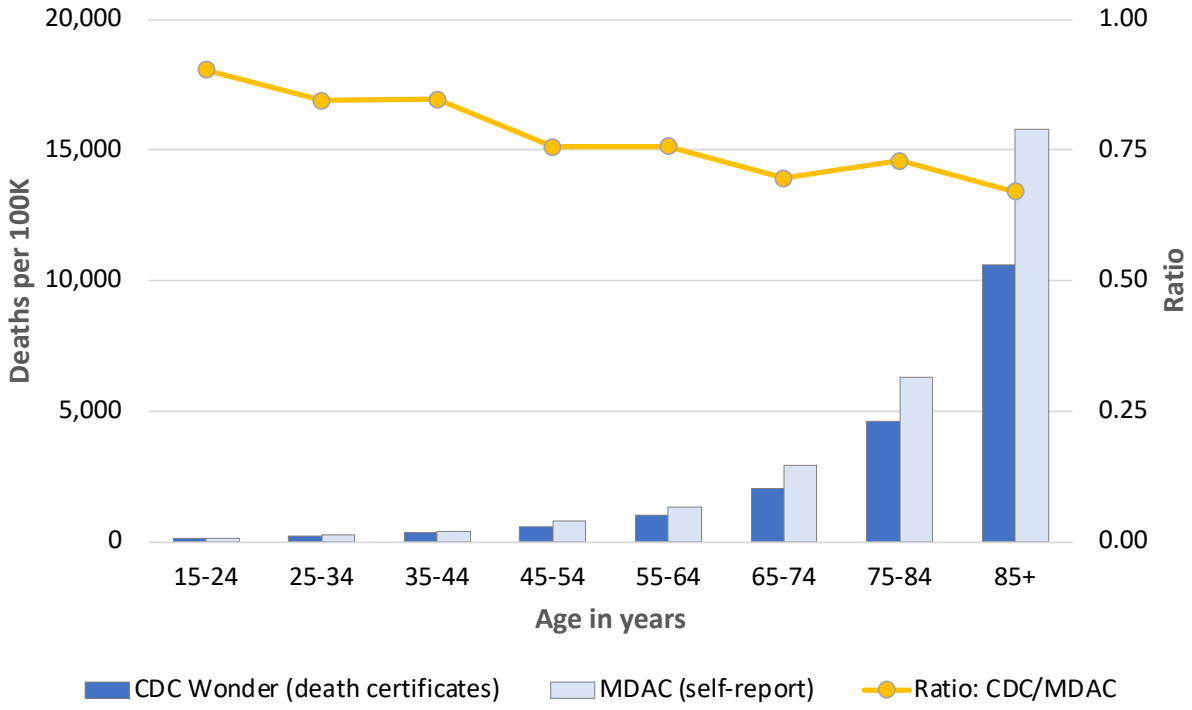


Figure 2. AIAN mortality relative to total U.S. mortality by age: CDC Wonder vs. MDAC

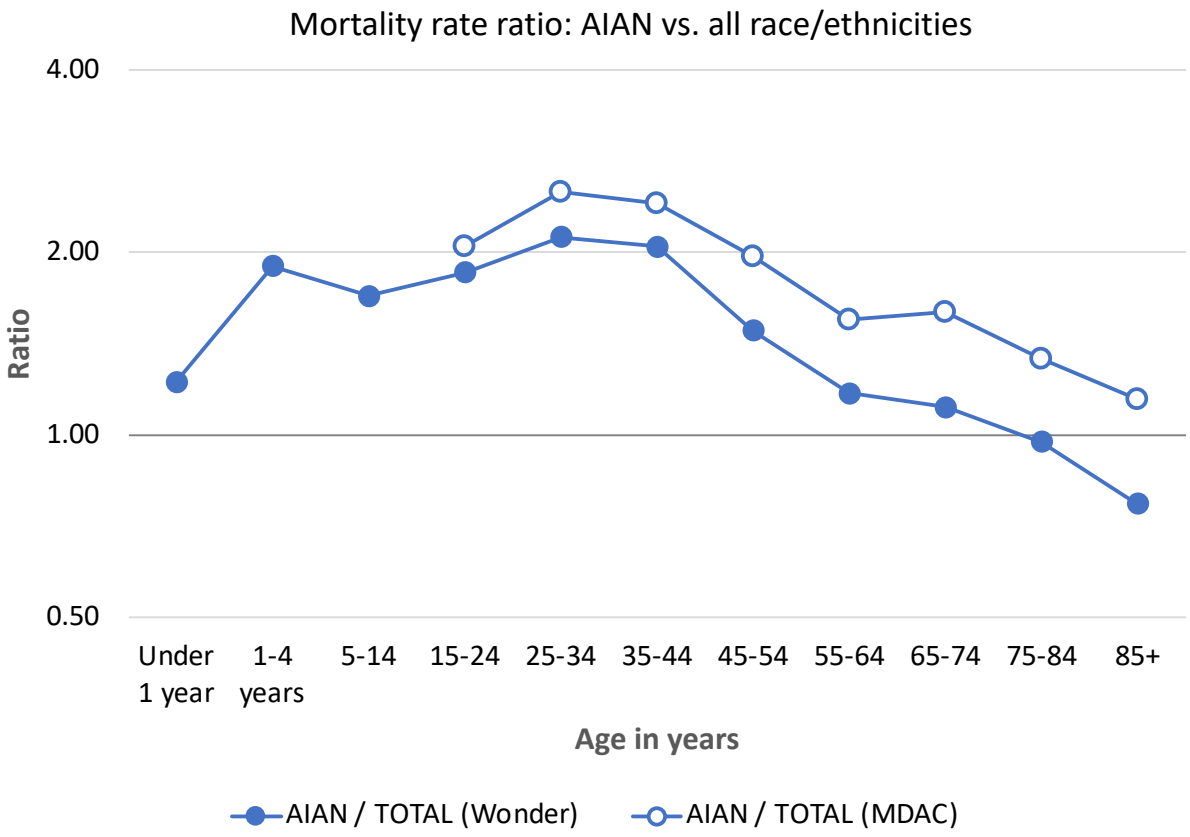


Figure 3. Age-standardized mortality rates (ages 15+)

