

How Healthy Are Survey Respondents Compared with the General Population?

Using Survey-linked Death Records to Compare Mortality Outcomes

Katherine M. Keyes,^{a,b} Caroline Rutherford,^a Frank Popham,^c Silvia S. Martins,^a and Lindsay Gray^c

Background: National surveys are used to capture US health trends and set clinical guidelines, yet the sampling frame often includes those in noninstitutional households, potentially missing those most vulnerable for poor health. Declining response rates in national surveys also represent a challenge, and existing inputs to survey weights have limitations. We compared mortality rates between those who respond to surveys and the general population over time.

Methods: Survey respondents from 20 waves of the National Health Interview Survey from 1990 to 2009 who have been linked to death records through 31 December 2011 were included. For each cohort in the survey, we estimated their mortality rates along with that cohort's mortality rate in the census population using vital statistics records, and differences were examined using Poisson models.

Results: In all years, survey respondents had lower mortality rates compared with the general population when data were both weighted and unweighted. Among men, survey respondents in the weighted sample had 0.86 (95% confidence interval = 0.853, 0.868) times the mortality rate of the general population (among women, RR = 0.887; 95% confidence interval, 0.879, 0.895). Differences in mortality are evident along all points of the life course. Differences have remained relatively stable over time.

Conclusion: Survey respondents have lower death rates than the general US population, suggesting that they are a systematically healthier source population. Incorporating nonhousehold samples

and revised weighting strategies to account for sample frame exclusion and nonresponse may allow for more rigorous estimation of the US population's health.

(*Epidemiology* 2018;29: 299–307)

Nationally representative surveys serve multiple purposes for health promotion and planning.^{1–3} Reliable estimates of the population prevalence of health-related behaviors are critical to public health infrastructure for multiple purposes, including formulating and evaluating policies aimed at improving and maintaining population health and well-being, national guidelines, percentiles, and clinical decision-making. Two developing dynamics warrant reconsideration of the role of surveys purporting to be nationally representative in describing the nation's health.

First, validity of inference from survey data is dependent on their representativeness of the general population, yet the sampling frame of such surveys is typically restricted to non-institutionalized households and so excluding homeless and institutionalized populations known to have high rates of poor health.⁴ For example, the population of the US state and federal prisons was less than 200,000 from the 1950s through the 1970s when many national surveys began. Since then, both the number and the proportion of individuals in the United States experiencing incarceration has increased dramatically. The number in state and federal prisons doubled by 1983, reached over a million by 1994, and reached 1.6 million by 2010.⁵ As of 2015, approximately 1 in 218 Americans was incarcerated in a state or federal prison,⁶ and the burden of incarceration falls disproportionately on Black and low-income Americans, as well as men.⁷ These trends imply that an increasing number of Americans—particularly those concentrated in certain demographic groups—fall outside the sampling frame of US surveys, which may obscure our ability to assess health disparities by race and socioeconomic status.

Second, survey response has been declining for the last decade.^{8–11} For example, the US National Health and Nutrition Examination Survey achieved household response levels

Submitted December 7, 2016; accepted October 19, 2017.

From the ^aDepartment of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY; ^bDepartment of Psychiatry, Columbia University Medical Center, New York, NY; and ^cMRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Glasgow, United Kingdom.

The authors report no conflicts of interest.

Data and code for replication: Full SAS and Stata code for these analyses are included as an online supplement, and data files are available by request to the corresponding author.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Katherine M. Keyes, Columbia University Department of Epidemiology, Mailman School of Public Health, 722 West 168th Street, Suite 503, New York, NY 10032. E-mail: kmk2104@columbia.edu.

Copyright © 2017 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ISSN: 1044-3983/18/2902-0299

DOI: 10.1097/EDE.0000000000000775

between 80% and 90% throughout the 1980s and 1990s⁸ but have been less than 80% since 2007, and in 2011–2012, the response was 72.6%.¹² The National Health Interview Survey's (NHIS) household response level fell below 80% in 2010 from over 90% in the 1990s and decreased to 77.6% in 2012.¹³

There is growing concern about the resulting nonrepresentativeness of national surveys, including whether use of households as a sampling frame is appropriate and the ensuing potential for unreliable inference to the general population.^{11,14} Two issues are particularly salient to this discussion. First, undercoverage of the general population may result in prevalence estimates that are not an accurate snapshot of the nation's health. Second, associations between demographic characteristics and risk behaviors may not reflect these associations in the general population. For example, we may over- or underestimate the extent of health disparities by such factors as sex and race if there are differential associations by inclusion status. Such potential for inferential bias is nontrivial. Estimates of both smoking and excessive alcohol consumption indicate decreases over the past 10 years.¹⁵ Given that nonrespondents and the incarcerated are more likely to be smokers and excessive alcohol consumers than those who respond,^{10,16–20} at least part of the reported prevalence decreases over time could be attributable to growing nonresponse bias and misrepresentation of the nation based on the sampling frame. As data sources from which population health estimates are generated become larger and with the rise of “big data,”²¹ understanding representativeness, or lack thereof, is growing in importance for epidemiologic investigation.

The most commonly used post hoc method to address nonresponse bias is to use population demographics and selection probabilities and reweight the samples using inverse probability weighting. However, the factors selected for weighting may not be adequate to account for differential selection into the sample if they do not capture or adjust for differences between respondents and nonrespondents in terms of health aspects, such as those related to smoking or alcohol consumption.^{22,23} Further, survey weighting methods applied to most national surveys do not account for factors associated with living in a household (e.g., wealth and criminal justice involvement). Poststratification weighting uses demographically stratified population totals from the US census, which does include those residing outside of households, and as such, poststratification sample weighting should in theory aid in estimating prevalences and associations that are more representative of the general population, despite excluding the institutionalized in the sampling frame. The extent to which weighting may be insufficient to render whole population-representative estimates is unknown, given that weighted estimates are difficult to validate: using weighting, we can ensure that the distributions of the factors for which there is population-level data at the household level (e.g., sex, age, race) match between population and sample, but this does not necessarily resolve representativeness for unmeasured factors.

Increasingly, national surveys are being individually record-linked to routine mortality data,^{24,25} providing the ability to assess death rates among survey respondents. Because we also have the information for the entire population, we can compare death rates between survey samples and the general population. If there is no nonresponse bias and household dwellers are representative of the whole population, these estimates should be equivalent, especially when the data are weighted for demographics, as well as nonresponse. To the extent that the survey respondents are healthier, however, we may conclude that current approaches are not capturing individuals mostly likely to have important risk factors for illness. Such survey to general population death comparisons has revealed insufficiencies in survey coverage and weighting in various countries,^{22,26–32} but to date, no attempt has been made, to our knowledge, to assess whether such results generalize to the United States where survey recruitment, sampling strategies, and population size and characteristics differ greatly. The present study compares weighted and unweighted mortality rates in US survey samples to comparable cohorts in the general population.

METHODS

Data Sources

National Health Interview Survey

The NHIS is an annually conducted in-person household survey comprising a multistage probability sample of noninstitutionalized respondents. Total household response rates ranged from ≈82% (2009) to 96% (1991 and 1992). We used data on adults in the NHIS surveys from 1990 to 2009 that have been individual-level record-linked to the National Death Index (NDI) through 31 December 2011, with successful linkage of ≈94% of eligible respondents (N = 1,309,449). Study respondents were probabilistically^{24,33} linked to the NDI through at least one of seven matching criteria, including some combination of social security number, first and last name, middle initial, date of birth, and father's surname. Those respondents who were 18+ at the time of interview were eligible to be linked; the present study includes those eligible respondents aged 18 to 79 years (ceiling coding in the NHIS precluded precise designation of birth cohort for those older than 80 years). Table 1 provides the weighted and unweighted sample size for each year of the NHIS, and the proportion of each sample decreased by 31 December 2011. The Columbia Institutional Review Board approved analyses of the NHIS-linked mortality data.

Details of the NHIS sampling frame and the estimation of sample weights are found elsewhere.^{34,35} Briefly, NHIS uses a multistage area probability design, which divides the United States into geographically defined Primary Sampling Units and groups those into strata to ensure broad geographic representation and, within those Primary Sampling Units, area and permit segments. The data on individuals for each

TABLE 1. NHIS Sample Sizes by Year and the Number and Percentage of Those Participants Who Died as of December 31, 2011

| Survey Year | Weighted | | Unweighted | | Difference Weighted (%) – Unweighted (%) |
|-------------|-----------------|--------------------|-----------------|----------------|---|
| | Sample Size (n) | Deaths (%) | Sample Size (n) | Deaths (%) | |
| 1990 | 175,411,048 | 38,165,833 (21.8) | 82,170 | 18,349 (22.3) | 0.5 |
| 1991 | 176,976,027 | 36,089,935 (20.4) | 81,907 | 17,241 (21.1) | 0.7 |
| 1992 | 182,508,639 | 34,590,922 (19.0) | 86,716 | 16,856 (19.4) | 0.4 |
| 1993 | 180,216,525 | 32,527,886 (18.1) | 74,696 | 13,706 (18.4) | 0.3 |
| 1994 | 182,824,384 | 30,425,708 (16.6) | 78,429 | 13,791 (17.6) | 1 |
| 1995 | 184,246,906 | 28,463,384 (15.5) | 68,477 | 10,641 (15.5) | 0 |
| 1996 | 185,946,202 | 25,704,369 (13.8) | 42,392 | 5,795 (13.7) | –0.1 |
| 1997 | 185,371,933 | 23,910,926 (12.9) | 65,528 | 8,494 (13.0) | 0.1 |
| 1998 | 189,454,603 | 22,027,315 (11.6) | 60,959 | 7,218 (11.8) | 0.2 |
| 1999 | 191,683,613 | 20,031,054 (10.5) | 60,028 | 6,299 (10.5) | 0 |
| 2000 | 193,572,176 | 18,063,618 (9.3) | 62,141 | 5,742 (9.2) | –0.1 |
| 2001 | 195,617,711 | 15,619,887 (8.0) | 61,692 | 4,887 (7.9) | –0.1 |
| 2002 | 197,322,085 | 14,053,099 (7.1) | 56,937 | 4,077 (7.2) | 0.1 |
| 2003 | 204,137,203 | 12,431,310 (6.1) | 55,009 | 3,394 (6.2) | 0.1 |
| 2004 | 205,545,112 | 10,276,640 (5.0) | 57,569 | 2,931 (5.1) | 0.1 |
| 2005 | 208,256,306 | 9,159,404 (4.4) | 59,314 | 2,672 (4.5) | 0.1 |
| 2006 | 210,844,397 | 7,323,039 (3.5) | 48,991 | 1,623 (3.3) | –0.2 |
| 2007 | 213,460,098 | 5,796,466 (2.7) | 48,900 | 1,323 (2.7) | 0 |
| 2008 | 215,226,246 | 4,627,273 (2.2) | 48,833 | 1,027 (2.1) | –0.1 |
| 2009 | 217,067,952 | 3,197,955 (1.5) | 59,838 | 818 (1.4) | –0.1 |
| Total | 3,895,689,166 | 392,486,023 (10.1) | 1,260,526 | 146,884 (11.7) | 1.6 |

NHIS indicates National Health Interview Survey.

NHIS sample has a set of sample weights that are based on the design, nonresponse, and poststratification adjustments. Poststratification adjustments are based on age, sex, and race/ethnicity such that the distributions of these demographics in the weighted sample approximate the US census population for the closest year of Census data collection. In the NHIS Linked Mortality Files, the sampling weights are further adjusted based on ineligibility or insufficient identifying data for linkage. Of the 1,953,298 survey respondents from 1990 to 2009, 1,309,449 (67%) were eligible for linkage, 547,290 (28%) were under age 18 years and not available for public release, and 96,559 (5%) were ineligible. Respondents >79 years old at the time of survey (N=48,923) were removed from the analytic sample because of differences across years in age categorization. For the present analysis, we utilized the sample weight for the linked mortality files that incorporates these weights.³⁶

Vital Statistics

The National Center for Health Statistics maintains the National Vital Statistics System for all deaths in the United States, coordinating and processing US Standard Death Certificates obtained from each state registrar.³⁷ Death data are collected and reported by trained medical certifiers. The present study examines all-cause mortality.

US Census

Population totals stratified by sex and age are derived from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program; they represent a modification of the intercensal and Vintage 2014 annual time series of July 1, county population estimates by age, sex, race, and Hispanic origin produced by the US Census Bureau's Population Estimates Program, in collaboration with the National Center for Health Statistics, and with support from the National Cancer Institute through an interagency agreement.³⁸ SEER population totals provide the most granular population estimates that are publicly available. Population census estimates are not directly linked to vital statistics regarding death.

Analytic Strategy

The general approach was to first estimate mortality numerators and denominators for the linked NHIS sample and for comparably aged individuals in the general population. We determined the mortality among individuals in the entire United States who would have been the same age at the time of the NHIS survey in any given year. This figure comprises those who would have been eligible for the survey (including nonrespondents) and those ineligible but living in the United States (e.g., the institutionalized). We detail our process for these estimations below.

Among adult survey respondents aged 18 to 79 years who were linked ($N = 1,260,526$), we first determined the numbers in each survey cohort who had died by 31 December 2011, overall and by age and sex. We then estimated the person time at risk among both those who died and those who did not die to inform the mortality rate. Among those in NHIS who died during follow-up, person time was estimated as survey year subtracted from death year. For those who had not died, person time was estimated as survey year subtracted from 2012 (given that respondents were linked through 31 December 2011).

We then obtained corresponding information on death and person-time at risk in the population. We did so by using age of death from vital statistics to determine the age of a deceased person in each year in which they would have been, based on age, eligible to be in the NHIS. For each death, we calculated a “pseudo age at survey” for each year in which the deceased individual would have been eligible and used that pseudo-age to estimate their person-time as the difference between death age and pseudo-age. For example, an individual who died at age 70 years in 2005 would have been eligible to be in the 1990 NHIS survey at age 55, and eligible for the 1991 survey at age 56, and so forth. Thus, the pseudo age at survey is 55 years in 1990 and 56 years in 1991. We then calculated the person-time for each decedent by subtracting would-be survey year from real death year, for each year they would have been eligible for survey in. We then take the complement of this person time by subtracting it from the total possible person time—2012 minus the survey year. Then we collapse the data, summing the complement of person time variable by single year age and survey year. Next, we merge these data with the census population totals, by year and age. To calculate the total synthetic person time sum at each year and age, we multiply the census population by the person time for those who did not die (2012 minus year of potential survey) and then subtract the complement of the person time from the mortality data—that is, person time lost to death.

Once the data on deaths and person-time were collected, we then combined survey and population data and estimated age-standardized, as well as age-specific, mortality rates and rate ratios from Poisson models. Race (for those listed as White and Black, as other racial groups did not have sufficient sample size in the NHIS to produce reliable estimates) and sex stratifications were performed given well-documented mortality differences and potential for differential exclusion from the sampling frame due to institutionalization as well as nonresponse, with existing evidence of sex-specific effects.²² Age-standardization was conducted with reference to the US 2000 Standard Population, as recommended by the US Centers for Disease Control and Prevention.^{39,40} These models estimated predicted mortality rates overall and by year, rate ratios comparing NHIS participants to the general population, and corresponding 95% confidence intervals (CI). We additionally included single year of age as a covariate in these models to account for possible differences in age

distribution within 5-year age groups. Mortality rates in the NHIS and the general population were estimated in a single model with a dichotomous indicator for survey versus population.

We also estimated the mortality rate for each single year of age in the data from Poisson models. To estimate age-specific mortality rates by year, we included a three-way interaction between survey/population (1 versus 0), age (in single years), and year. We then used the marginal predicted rates from this model to estimate the mortality rate from the survey and the mortality rate from the population by age by year.

Modeling was done both without sample weighting for NHIS and including the NHIS sample weight for participants in the death record linkage (general population weights were assigned the value 1). Modeling was conducted in Stata Version 13. Full SAS and Stata code for these analyses are included as an online supplement, and data files are available by request to the corresponding author.

RESULTS

Figure 1 shows the mortality rate per 100,000 for NHIS respondents through 2014, by year, using both weighted and unweighted NHIS samples, as well as the estimated mortality rates for comparably aged contemporaneous general population cohorts. Results are stratified by sex. Confidence intervals for these estimates are provided in eTable 1; <http://links.lww.com/EDE/B288>. Sample weighting had little effect on the mortality estimates, and in all years, mortality among NHIS respondents was lower than the general population. Mortality declined over time as more recent survey respondents were followed for a shorter time and, therefore, over a younger age span than the respondents to earlier surveys.

Table 2 shows the age-standardized mortality rate ratios and confidence intervals between survey respondents and the general population for men and women, respectively. Across all years, among men, survey respondents in the weighted sample have 0.86 times the rate of mortality as the general population (95% CI = 0.853, 0.868). Among men, the rate ratio by year ranged from 0.798 (95% CI = 0.759, 0.838) in 2007 to 0.893 (95% CI = 0.874, 0.911) in 1995. Across all years, among women, survey respondents in the weighted sample have 0.887 times the rate of mortality as the general population (95% CI = 0.879, 0.895). Among women, the rate ratio by year ranged from 0.724 (95% CI = 0.66, 0.787) in 2009 to 0.925 (95% CI = 0.912, 0.938) in 1994.

Figure 2 shows the predicted 5-year age group-specific mortality rate ratios for each year from a Poisson model with survey/population by age by year interaction; confidence intervals are provided in eTable 2; <http://links.lww.com/EDE/B288> ($P < 0.001$ for the interaction, both when age was considered in single years and in 5-year age groups). Mortality was consistently lower for survey respondents than for the general population across all ages through 2002, with those at younger ages consistently showing more divergence in mortality between survey and population than those in older ages;

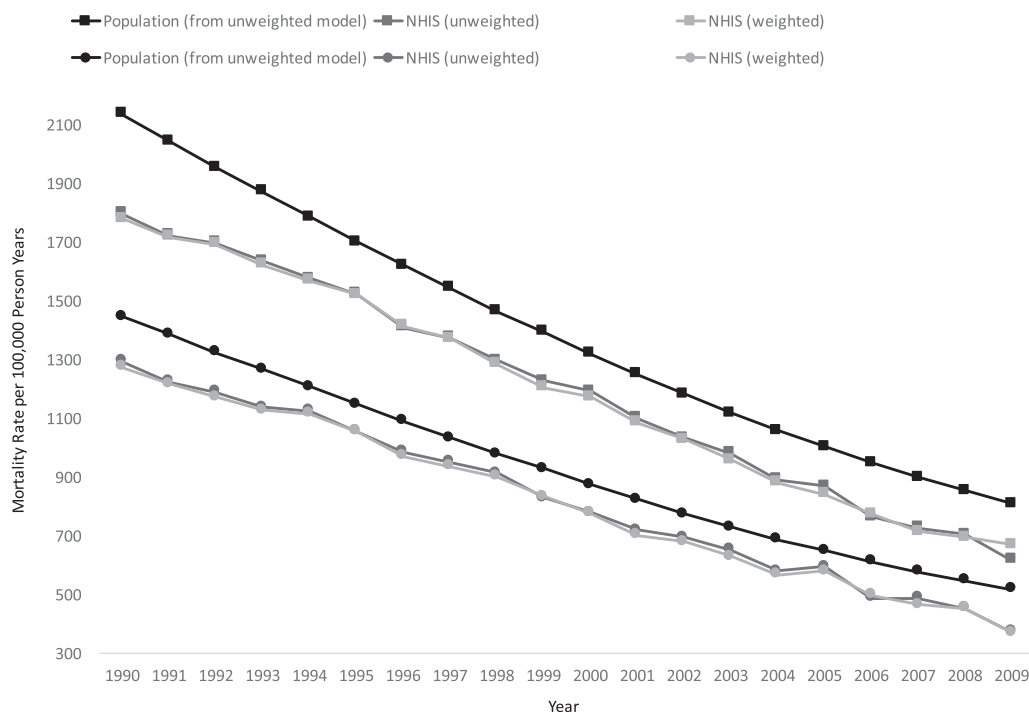


FIGURE 1. Comparison of age-standardized mortality rates in National Health Interview Survey respondents versus the population, among men (square markers) and among women (circular markers).

after 2002, there was more variation in the correspondence between survey and general population samples with mortality, likely due to smaller numbers of deceased for the more recent years.

Supplementary Analyses

In eTables 3 and 4; <http://links.lww.com/EDE/B288>, we show the associations between sex (women compared with men) and race (Black compared with White), respectively, for each year in the NHIS and the general population. In NHIS, rate ratios for the mortality difference between women and men ranged from 0.556 (95% CI = 0.494, 0.618) in 2009 to 0.702 (95% CI = 0.684, 0.72) in 1990. In the general population, mortality rates among women were 0.657 times those of men in 1990 and increased to 0.633 times those of men by 2009. Generally, the trend toward growing disparity in mortality favoring women over the survey years was largely consistent between NHIS and the general population.

Comparing mortality rates among Black and White individuals in NHIS, mortality rate ratios ranged from 1.195 in 1991 to 1.767 in 2003, tending to be larger for more recent survey years. In contrast, in the general population, there was little evidence for trends over time in rate ratios, which ranged from 1.431 to 1.459 in all years from 1990 to 2009.

DISCUSSION

The present study documents that NHIS survey respondents have lower mortality rates than the general population. We draw this conclusion by using information on survey

respondents from a 20-year period who have been linked to death records. We compared their mortality rates to the estimated mortality rates among the contemporaneously aged general population. Overall, survey respondents die at 0.86 (men) to 0.887 (women) times the rate of the general population, and survey weighting did not have an impact on the results. The differential was evident for older and younger respondents, though with smaller magnitude in the earlier years.

Further, there is little evidence for systematic change in the relative magnitude of differences between survey respondents and the general population over time. As survey response levels have declined,⁸⁻¹¹ we might have expected to see growing differences between survey respondents and the general population, but this is not evident at present. Continued surveillance of these mortality rates as younger cohorts age, as well as examinations of subgroup differences, will be important next steps in survey research.

Finally, our results indicate that the association between sex and mortality is similar in survey respondents and the general population, indicating that even if prevalence estimates from surveys by sex may not represent the general population, associations with mortality are consistent. On the other hand, associations between race (Black versus White) and mortality increased over time in the NHIS participants but did not change through the course of this study in the general population, indicating that surveys may be increasingly misrepresenting racial disparities in health in the United States as a whole. Other work suggests that disparities in health between Black and White Americans are decreasing nationally, though studies have not followed

TABLE 2. Rate Ratios Comparing Death Rates in NHIS to Comparably Aged Cohorts in the General Population by Year, by Sex

| Survey Year | Rate Ratio, Weighted (95% CI) | Rate Ratio, Unweighted (95% CI) |
|-------------|----------------------------------|------------------------------------|
| Men Only | | |
| 1990 | 0.834 (0.822, 0.845) | 0.842 (0.826, 0.857) |
| 1991 | 0.841 (0.824, 0.857) | 0.843 (0.823, 0.864) |
| 1992 | 0.86 (0.844, 0.877) | 0.868 (0.844, 0.892) |
| 1993 | 0.868 (0.85, 0.886) | 0.874 (0.849, 0.898) |
| 1994 | 0.878 (0.861, 0.894) | 0.883 (0.863, 0.904) |
| 1995 | 0.893 (0.874, 0.911) | 0.896 (0.875, 0.918) |
| 1996 | 0.867 (0.842, 0.891) | 0.87 (0.839, 0.901) |
| 1997 | 0.885 (0.865, 0.905) | 0.89 (0.862, 0.918) |
| 1998 | 0.876 (0.856, 0.896) | 0.885 (0.859, 0.912) |
| 1999 | 0.866 (0.844, 0.887) | 0.879 (0.852, 0.907) |
| 2000 | 0.889 (0.868, 0.91) | 0.904 (0.879, 0.93) |
| 2001 | 0.865 (0.84, 0.889) | 0.881 (0.85, 0.911) |
| 2002 | 0.866 (0.842, 0.889) | 0.874 (0.842, 0.906) |
| 2003 | 0.857 (0.829, 0.886) | 0.877 (0.84, 0.914) |
| 2004 | 0.831 (0.797, 0.866) | 0.844 (0.796, 0.892) |
| 2005 | 0.843 (0.806, 0.88) | 0.865 (0.813, 0.918) |
| 2006 | 0.822 (0.781, 0.864) | 0.806 (0.748, 0.865) |
| 2007 | 0.798 (0.759, 0.838) | 0.809 (0.755, 0.863) |
| 2008 | 0.815 (0.764, 0.867) | 0.824 (0.752, 0.896) |
| 2009 | 0.824 (0.762, 0.887) | 0.764 (0.695, 0.834) |
| Total | 0.86 (0.853, 0.868) | 0.868 (0.86, 0.876) |
| Women only | | |
| 1990 | 0.882 (0.867, 0.898) | 0.895 (0.874, 0.915) |
| 1991 | 0.879 (0.868, 0.891) | 0.884 (0.867, 0.901) |
| 1992 | 0.885 (0.872, 0.899) | 0.898 (0.878, 0.918) |
| 1993 | 0.89 (0.874, 0.907) | 0.898 (0.878, 0.919) |
| 1994 | 0.925 (0.912, 0.938) | 0.932 (0.915, 0.95) |
| 1995 | 0.921 (0.905, 0.938) | 0.92 (0.899, 0.94) |
| 1996 | 0.889 (0.863, 0.915) | 0.905 (0.868, 0.941) |
| 1997 | 0.907 (0.884, 0.929) | 0.92 (0.89, 0.949) |
| 1998 | 0.921 (0.896, 0.947) | 0.932 (0.895, 0.969) |
| 1999 | 0.9 (0.878, 0.922) | 0.893 (0.859, 0.927) |
| 2000 | 0.889 (0.855, 0.922) | 0.893 (0.85, 0.936) |
| 2001 | 0.853 (0.819, 0.887) | 0.874 (0.826, 0.923) |
| 2002 | 0.878 (0.85, 0.906) | 0.895 (0.857, 0.933) |
| 2003 | 0.865 (0.829, 0.901) | 0.895 (0.84, 0.95) |
| 2004 | 0.827 (0.792, 0.861) | 0.845 (0.802, 0.888) |
| 2005 | 0.896 (0.862, 0.929) | 0.918 (0.868, 0.968) |
| 2006 | 0.824 (0.758, 0.89) | 0.798 (0.712, 0.884) |
| 2007 | 0.807 (0.758, 0.857) | 0.846 (0.773, 0.918) |
| 2008 | 0.833 (0.767, 0.899) | 0.829 (0.754, 0.905) |
| 2009 | 0.724 (0.66, 0.787) | 0.723 (0.636, 0.81) |
| Total | 0.887 (0.879, 0.895) | 0.899 (0.89, 0.908) |

CI indicates confidence interval; NHIS, National Health Interview Survey.

pseudo-cohorts in a similar methodology as the present study,⁴¹ suggesting that a greater focus on longitudinal research would benefit overall understanding of disparities.

One source of this difference is known and well described: as outlined earlier, the sampling frame of US surveys is most often households.^{25,42} The omission of individuals who are incarcerated, otherwise institutionalized, homeless, or serving in the military, provides an over-optimistic picture of our nation's health. Further, exclusion of institutionalized populations from the sample frame of surveys under-represents the extent of health disparities by race.⁷ Such exclusion may be increasingly problematic, as a disproportionate number of African American men continue to be incarcerated at high rates. At least a portion of the difference between survey respondents and general population documented in our analyses is certain to be attributable to incarcerated and institutionalized individuals being systematically excluded from national survey sampling frames.

Other sources of this difference are less well understood. While men as well as racial/ethnic minorities (e.g. African American and Hispanics) are less likely to respond to surveys,^{11,19,43} such nonresponse should, in theory, be accounted for with sample weighting. While oversampling of harder-to-reach groups may enable nonresponse weighting to perform better, currently the design of NHIS does not oversample racial/ethnic minorities except those over 65 years.²⁵ Heavy alcohol consumers and smokers, on the other hand, are also less likely to respond to surveys,^{10,16–19} and while such health behaviors are associated with demographics, standard weighting schemes cannot account for variation in these health behaviors between respondents and nonrespondents within demographic subgroups. Given that heavy alcohol consumers and smokers have higher mortality rates than the general population,^{1,2,27} such health-related selection likely accounts for a proportion of the differences observed in the present study as well.

Differences between survey respondents and the general population have implications not only for estimates of prevalence but also for estimates of associations and disparities. It is well established that women consume less alcohol and cigarettes than men^{44,45} but are more likely to respond to surveys compared with men.^{11,19,43} Given that heavy alcohol and cigarette consumers are also less likely to respond to surveys,^{10,16–19} heavy alcohol consuming and heavy smoking men may be those least likely to respond to surveys across gender and substance use subgroups, and those who do respond may be atypical. If this is the case, surveys are *underestimating* the gender disparities in the general population. Conversely, African Americans and Hispanics in the United States are less likely to consume alcohol and smoke compared with non-Hispanic whites^{46,47} but are also less likely to respond to surveys.^{11,19,43} Thus, racial/ethnic disparities in alcohol consumption and smoking based on surveys may be *overestimates* of the disparity in the general population. There is evidence of such overestimation in our analysis, as racial disparities in mortality tend to be higher and growing over time in NHIS participants compared with the general population. Development and implementation of survey weighting or imputation

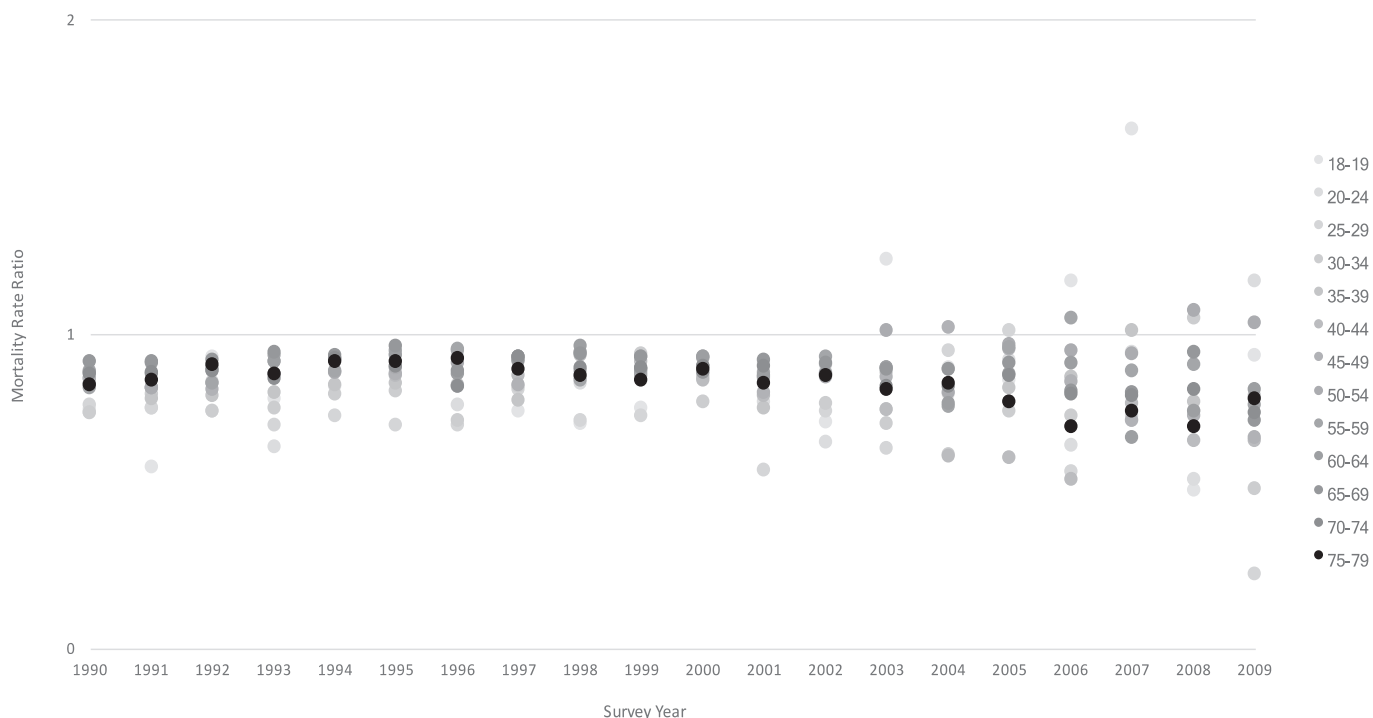


FIGURE 2. Mortality rate ratios comparing National Health Interview Survey (NHIS) to the general population by age group and year.

procedures³² to reflect potential health-related disparities are necessary.

The present study should be considered in light of limitations. We did not have direct estimates of the mortality rates of those individuals who were eligible for NHIS but did not participate. Rather, we compared responders to the general population, estimating their mortality rates using vital statistics and census population totals. Thus, individuals who were not in the sampling frame of NHIS are included in the general population estimate, and we cannot disentangle the proportion of the difference in the NHIS mortality rates and the general population that are due to nonresponse versus institutionalization. Population estimates may not be accurate as the census can miss hard to reach individuals such as homeless and transient populations among whom mortality rates are high. These population estimates should be interpreted with those cautions in mind, although such undercounting of our general population denominators may actually be inflating the mortality rates; thus, the actual differences between survey respondents and the general population could be lower than those we have identified. Further, a small proportion of survey respondents (6% of eligible sample) could not be linked to the NDI, and thus, mortality follow up is incomplete. Additionally, linkages were done probabilistically, and there may be error in the linkage. To the extent that these errors may favor certain sociodemographic groups is unknown. All-cause mortality provides one important metric of health but certainly not the only metric; information on underlying and contributing causes of death are also available on this cohort, and

continued analyses will allow for testing of specific causes of death that may differentiate respondents and the general population. Finally, death record-linked NHIS through 2009 with follow-up to 2011 have been released to date; updated analyses with more recent surveys and years of linkage are important to continue to assess potential divergence between general population and survey representation.

In conclusion, the issue of survey weighting will continue to gain importance in epidemiology. The rise of “big data” and electronic health records promise masses of data,^{21,48,49} many of which are highly likely to be quite unrepresentative of underlying population distributions of health and illness. The *potential* for misrepresentation of population distributions is not necessarily realized, though inferential bias is increasingly being identified across a number of surveys.^{22,28,50–53} However, the potential could be carefully assessed. The issues we highlight in the present article, including systematic exclusion of nonhousehold dwelling individuals from sampling frames, and continuing declines of response rates, are among the multiple challenges that are faced by survey researchers. Continued innovation in methods including sample design and imputation techniques³² or sample weight formulation that allows for broader representation in ways that are fully replicable and valid may improve the utility of large data sources for understanding population health. While some suggest that representative sampling is unnecessary for epidemiologic inquiry,^{54,55} the present results and others^{56,57} continue to demonstrate that our understanding of the distribution and determinants of disease and other health outcomes

vary by population characteristics, and attention to the source populations from which our cases are drawn will likely grow in importance in coming years rather than diminish.

ACKNOWLEDGMENTS

We acknowledge the Columbia University Calderone Prize for Junior Faculty, which provided funding for this project to K.M. Keyes. F. Popham and L. Gray are funded by the Medical Research Council, UK, and Chief Scientist's Office, Scottish Government (MC_UU_12017/13 and SPHSU13), as part of the core funding for the MRC/CSO Social & Public Health Sciences Unit, University of Glasgow, UK.

REFERENCES

- Centers for Disease Control and Prevention. Excessive Alcohol Use: Addressing A Leading Risk for Death, Chronic Disease, and Injury. Available at: <https://www.cdc.gov/chronicdisease/resources/publications/aag/alcohol.htm>. Accessed 21 April 2017.
- U.S. Department of Health and Human Services. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Atlanta: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. 2014. Accessed 21 April 2017.
- Center for Disease Control and Prevention. Trends in Current Cigarette Smoking Among High School Students and Adults, United States, 1965–2011. Available at: https://www.cdc.gov/tobacco/data_statistics/tables/trends/cig_smoking/. 2011.
- Binswanger IA, Krueger PM, Steiner JF. Prevalence of chronic medical conditions among jail and prison inmates in the USA compared with the general population. *J Epidemiol Community Health*. 2009;63:912–919.
- Guerino P, Harrison PM, Sabol WJ. Prisoners in 2010. 2012. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, Bulletin. Available at: <https://www.bjs.gov/content/pub/pdf/p10.pdf>. Accessed 21 April 2017.
- Carson EA, Anderson E. Prisoners in 2015. 2016. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, Bulletin. Available at: <https://www.bjs.gov/content/pub/pdf/p15.pdf>. Accessed 21 April 2017.
- Pettit B. *Invisible Men: Incarceration and the Myth of Black Progress*. New York: Russell Sage Foundation; 2012.
- Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol*. 2007;17:643–653.
- Tolonen H, Helakorpi S, Talala K, Helasoja V, Martelin T, Prättälä R. 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *Eur J Epidemiol*. 2006;21:409–415.
- Ahacic K, Käreholt I, Helgason AR, Allebeck P. Non-response bias and hazardous alcohol use in relation to previous alcohol-related hospitalization: comparing survey responses with population data. *Subst Abuse Treat Prev Policy*. 2013;8:10.
- Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. *JAMA*. 2012;307:1805–1806.
- Centers for Disease Control and Prevention. NHANES Response Rates and Population Totals. *CDC/National Center for Health Statistics*. 2012 Available at: https://www.cdc.gov/nchs/nhanes/response_rates_cps.htm. Accessed 21 April 2017.
- Centers for Disease Control and Prevention. 2012 National Health Interview Survey (NHIS): NHIS Survey Description. Hyattsville, MD: Division of Health Interview Statistics, National Center for Health Statistics. 2013 Available at: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2012/srvydesc.pdf. Accessed 21 April 2017.
- Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics, Committee on National Statistics, Division on Behavioral and Social Sciences and Education, Council NR. *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, DC: The National Academies Press; 2013.
- Centers for Disease Control and Prevention. Current cigarette smoking Among Adults - United States, 2005–2012. *Morb Mortal Wkly Rep*. 2014;63:29–34.
- Osler M, Kriebaum M, Christensen U, Holstein B, Nybo Andersen AM. Rapid report on methodology: does loss to follow-up in a cohort study bias associations between early life factors and lifestyle-related health outcomes? *Ann Epidemiol*. 2008;18:422–424.
- Catto S. *How Much Are People in Scotland Really Drinking? A Review of Data From Scotland's Routine National Surveys*. Glasgow: Public Health Observatory Division, NHS Health Scotland. 2008.
- Roberts H, Bali B, Rushton L. Non-responders to a Lifestyle Survey: A study using telephone interviews. *J Instit Health Edu*. 1996;34(2).
- Wild TC, Cunningham J, Adlaf E. Nonresponse in a follow-up to a representative telephone survey of adult drinkers. *J Stud Alcohol*. 2001;62:257–261.
- Mumola C. Substance Abuse and Treatment, State and Federal Prisoners, 1997. Washington, DC: US Department of Justice (Ed). 1999:1–16.
- Mooney SJ, Westreich DJ, El-Sayed AM. Commentary: Epidemiology in the era of big data. *Epidemiology*. 2015;26:390–394.
- Gorman E, Leyland AH, McCartney G, et al. Assessing the representativeness of population-sampled health surveys through linkage to administrative data on alcohol-related outcomes. *Am J Epidemiol*. 2014;180:941–948.
- Meiklejohn J, Connor J, Kypri K. The effect of low survey response rates on estimates of alcohol consumption in a general population survey. *PLoS One*. 2012;7:e35527.
- Data linkage team. Comparative analysis of the NHIS public-use and restricted-use linked mortality files: 2010 public-use data release. Hyattsville, MD; National Center for Health Statistics March 2010. 2010. Available at: https://www.cdc.gov/nchs/data/datalinkage/nhis_mort_compare_2010_final.pdf.
- National Health Interview Survey. <https://www.cdc.gov/nchs/nhis/>.
- Gray L, McCartney G, White IR, et al. Use of record-linkage to handle non-response and improve alcohol consumption estimates in health survey data: a study protocol. *BMJ Open*. 2013;3:e002647.
- Gray L, McCartney G, White IR, Rutherford L, Katikireddi SV, Leyland AH. A novel use of record-linkage: resolving non-representativeness in health surveys and improving alcohol consumption estimates to inform strategy evaluation [conference abstract]. Public Health Science: A National Conference Dedicated to New Research in Public Health; 12 Nov 2012. London. *Lancet* 2012;380:S42.
- Christensen AI, Ekholm O, Gray L, Glümer C, Juel K. What is wrong with non-respondents? Alcohol-, drug- and smoking-related mortality and morbidity in a 12-year follow-up study of respondents and non-respondents in the Danish Health and Morbidity Survey. *Addiction*. 2015;110:1505–1512.
- Goldberg M, Chastang JF, Leclerc A, et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. *Am J Epidemiol*. 2001;154:373–384.
- Mäkelä P, Paljärvi T. Do consequences of a given pattern of drinking vary by socioeconomic status? A mortality and hospitalisation follow-up for alcohol-related causes of the Finnish Drinking Habits Surveys. *J Epidemiol Community Health*. 2008;62:728–733.
- Vinther-Larsen M, Riegels M, Rod MH, Schiøtz M, Curtis T, Grønbaek M. The Danish Youth Cohort: characteristics of participants and non-participants and determinants of attrition. *Scand J Public Health*. 2010;38:648–656.
- Gorman E, Leyland AH, McCartney G, et al. Adjustment for survey non-representativeness using record-linkage: refined estimates of alcohol consumption by deprivation in Scotland. *Addiction*. 2017;112:1270–1280.
- Data linkage team. Comparative analysis of the NHANES III public-use and restricted-use linked mortality files: 2010 public-use data release. 2010.
- National Center for Health Statistics, Office of Analysis and Epidemiology. Use of Survey Weights for Linked Data Files – Preliminary Guidance. 2013; https://www.cdc.gov/nchs/data/datalinkage/use_of_survey_weights_for_linked_data_files.pdf. Accessed 21 April 2017.
- National Center for Health Statistics, Variance Estimation Guidance, NHIS 2006–2015 (Adapted from the 2006–2015 NHIS Survey Description Documents). 2016; <https://www.cdc.gov/nchs/data/nhis/2006var.pdf>. Accessed 21 April 2017.
- National Center for Health Statistics, Data Linkage, National Health Interview Survey (1986–2004) Linked Mortality Files, Analytic guidelines. https://www.cdc.gov/nchs/data/datalinkage/nhis_mort_analytic_guidelines.pdf. Accessed 21 April 2017.

37. National Center for Health Statistics. National Vital Statistics System. Public-use data file and documentation. https://www.cdc.gov/nchs/nvss/mortality_public_use_data.htm. Accessed 21 April 2017.
38. National Cancer Institute. Surveillance, Epidemiology, and End Results Program. U.S. Population Data. <https://seer.cancer.gov/popdata/download.html>. Accessed 21 April 2017.
39. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People Statistical Notes, no. 20. Hyattsville, Maryland: National Center for Health Statistics. January 2001.
40. National Center for Health Statistics, Continuous NHANES Web Tutorial, Age Standardization and Population Estimates. 2014; https://www.cdc.gov/nchs/tutorials/NHANES/NHANESAnalyses/AgeStandardization/age_standardization_intro.htm. Accessed 21 April 2017.
41. Harper S, MacLehose RF, Kaufman JS. Trends in the black-white life expectancy gap among US states, 1990-2009. *Health Aff (Millwood)*. 2014;33:1375-1382.
42. National Center for Health Statistics. National Health and Nutrition Examination Survey. Questionnaires, Datasets, and Related Documentation. <https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>. Accessed on 21 April 2017.
43. Keeter S, Kennedy C, Dimock M, Best J, Craighill P. Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. *Public Opin Q*. 2006;70(5):759-779. Accessed 21 April 2017.
44. Keyes KM, Li G, Hasin DS. Birth cohort effects and gender differences in alcohol epidemiology: a review and synthesis. *Alcohol Clin Exp Res*. 2011;35:2101-2112.
45. Keyes KM, Martins SS, Blanco C, Hasin DS. Telescoping and gender differences in alcohol dependence: new evidence from two national surveys. *Am J Psychiatry*. 2010;167:969-976.
46. Keyes KM, Vo T, Wall MM, et al. Racial/ethnic differences in use of alcohol, tobacco, and marijuana: is there a cross-over from adolescence to adulthood? *Soc Sci Med*. 2015;124:132-141.
47. Keyes KM, Miech R. Age, period, and cohort effects in heavy episodic drinking in the US from 1985 to 2009. *Drug Alcohol Depend*. 2013;132:140-148.
48. Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *N Biotechnol*. 2012;29:613-624.
49. Khoury MJ, Gwinn ML, Glasgow RE, Kramer BS. A population approach to precision medicine. *Am J Prev Med*. 2012;42:639-645.
50. Gray L, Martins SS, Hamilton A, et al. Correcting non-response bias of US health survey estimates using administrative record-linkage. Presented at the 2015 Society for Epidemiologic Research Meeting, Denver, CO. 2015.
51. Ebrahim S, Davey Smith G. Commentary: Should we always deliberately be non-representative? *Int J Epidemiol*. 2013;42:1022-1026.
52. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider Scope: How selection bias can induce spurious associations. *bioRxiv Available at: http://biorxiv.org/content/early/2016/10/07/079707*. 2016.
53. Howe LD, Tilling K, Galobardes B, Lawlor DA. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. *Epidemiology*. 2013;24:1-9. Accessed 21 April 2017.
54. Rothman K, Hatch E, Gallacher J. Representativeness is not helpful in studying heterogeneity of effects across subgroups. *Int J Epidemiol*. 2014;43:633-634.
55. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42:1012-1014.
56. LeWinn KZ, Sheridan MA, Keyes KM, Hamilton A, K.A. M. The representative developing brain: Does sampling strategy matter for neuroscience? Under review.
57. Falk EB, Hyde LW, Mitchell C, et al. What is a representative brain? Neuroscience meets population science. *Proc Natl Acad Sci U S A*. 2013;110:17615-17622.