A Comparison of Mortality Estimates from Multiple Nationally Representative Surveys and Vital Statistics Data in the United States

Dustin C. Brown[†]
Department of Sociology
Social Science Research Center
Mississippi State University

Joseph T. Lariscy
Department of Sociology
University of Memphis

Lucie Kalousová
Department of Sociology
Department of Health Services Organization and Policy
Population Studies Center
University of Michigan

[†]Please direct questions and comments to Dustin C. Brown, Department of Sociology, P.O. Box C, 207 Bowen Hall, Mississippi State University, Mississippi State, MS 39762. Email: dbrown@soc.msstate.edu.

Abstract

Population surveys prospectively linked with the death records of respondents provide invaluable opportunities for the study of the relationship between social and economic circumstances and mortality outcomes. However, the methods for establishing vital status vary across data sources and this variation may influence the accuracy of their mortality estimates. We conduct the first study that systematically compares mortality estimates from multiple survey-linked mortality files with U.S. vital statistics data. Our results show that mortality rates and life expectancies in U.S. vital statistics data resemble those of the National Health Interview Survey Linked Mortality Files, Health and Retirement Study, and Americans' Changing Lives study. Compared to vital statistics, General Social Survey-National Death Index (GSS-NDI) mortality rates are much higher at younger adult ages, similar in midlife, and much lower at older adult ages. Results from multivariate analyses also imply that the GSS-NDI does not accurately reflect the mortality experience of the U.S. population.

Keywords: Adult mortality • Vital statistics • Record linkage

BACKGROUND

Several nationally representative surveys in the United States are now linked to mortality records in the National Death Index (NDI). Survey-linked mortality files are critically important and increasingly popular sources of information on population health because they allow researchers to examine the links between social factors, disease risk, disability, and mortality (Preston and Taubman 1994; Rogers et al. 2000). Survey–linked mortality files have enormous potential as a source of information on social, economic, and behavioral determinants of U.S. adult mortality risk since researchers are able to assess sociodemographic characteristics and behaviors not available in vital statistics data (Muennig et al. 2011). However, the utility of these data sources depends on whether they accurately reflect the mortality experience of the U.S. adult population.

Although population health researchers increasingly rely upon survey-linked mortality files to explain and monitor adult mortality disparities in the United States, no studies have systematically compared mortality estimates derived from multiple nationally representative survey-linked mortality files with vital statistics data. An important step in any scientific enterprise is critically evaluating their methods and data. Such confirmatory analyses are particularly important for health demographers, who often rely on administrative or survey data that were not initially intended for mortality research. Several prior studies have examined data quality issues in survey-linked datasets, but these studies only focus on one or two survey-linked mortality files at a time (Brown et al. 2012; Lariscy et al. 2015; Lin et al. 2003). Research that systematically cross-validates results between multiple nationally-representative survey-linked mortality files is also sparse. In order to have confidence in the results obtained from analyses of surveys with prospectively-linked mortality follow-up, social scientists need to systematically assess the extent to which the influence of social, economic, and behavioral risk factors on mortality risk varies across data sources.

There are several reasons for why mortality follow-up across multiple nationally-representative surveys might yield varying estimates with respect to social, economic, and behavioral risk factors of death. The record linkage methodologies used to match surveys to death certificate data are probabilistic and thus do not correctly ascertain the vital status for all survey respondents (Harron et al. 2016). Some survey respondents may die during the follow-up period but the personally identifiable (PII) information used to match the records is inaccurate or missing on the survey and/or death certificate so that the two records cannot be matched. Failure to match a survey respondent to their death certificate will produce a false negative such that these respondents will appear “statistically immortal” (Pablos-Méndez 1994:1237). Survey-linked data sources differ in their strategies in ascertaining vital status and this variation could impact the accuracy of their mortality estimates. For instance, some datasets rely exclusively on passive linkage (i.e., linking records based on PII). Other datasets that are longitudinal in design are able to incorporate active follow-up (i.e., re-contact the surviving respondent or collect information regarding the respondent’s death from a surviving family member). Finally, differential record linkage by population subgroup could affect estimates of mortality disparities if members of some groups are more likely to be linked to the death certificates than members of other groups.

Vital statistics data are generally seen as more accurate than survey-linked data since they are larger include information on the entire U.S. population whereas the survey-linked data are samples of the non-institutionalized U.S. population. These differences in study design could produce differences in mortality estimates vital statistics and survey-linked datasets.

Additionally, vital statistics data are drawn from two data sources: death certificate data are used in the numerator of mortality rates and population estimates provided by the U.S. Census Bureau are used in the denominator of mortality rates. Errors in content or coverage in either source could bias the resulting mortality rates. Other studies have demonstrated issues in vital statistics rates among racial/ethnic minority populations, such as ethnic misclassification on death certificates among Hispanics (Arias et al. 2010) and census undercount among blacks (Anderson and Fienberg 1999).

This study takes a two-fold approach to addressing the uncertainty regarding the comparability of the multiple data sources used in the study of U.S. adult mortality. First, we compare mortality rates and life expectancy estimates derived from several commonly used nationally-representative surveys in the United States that have prospective mortality follow-up to mortality estimates in vital statistics data. Second, in order to determine whether the various methods used to establish vital status across data sources influences the accuracy of their estimates of mortality disparities, we contrast coefficients of sociodemographic variables obtained from multivariate hazard models predicting respondent's death in the four surveys as a function of sociodemographic characteristics.

METHODS

Data

We compare mortality estimates from four nationally-representative U.S. surveys that have longitudinal mortality follow-up: National Health Interview Survey Linked Mortality Files (NHIS-LMF), Health and Retirement Study (HRS), Americans' Changing Lives (ACL), and General Social Survey National Death Index (GSS-NDI).

The NHIS is a nationally-representative cross-sectional survey of the non-institutionalized U.S. adult population ages 18 and over. The NHIS is conducted annually by the U.S. National Center for Health Statistics. The NHIS-LMF prospectively links the 1986-2009 NHIS to death records in the 1986-2011 NDI. NHIS-LMF data were obtained from the Integrated Health Interview Series (IHIS) website (Minnesota Population Center 2015). The 1986-2011 NHIS-LMF contains survey records for 1,605,246 respondents who were eligible for mortality follow-up and 256,751 decedents. Additional information about the NHIS-LMF is available elsewhere (Lochner et al. 2008).

The Health and Retirement Study (HRS) is a nationally-representative longitudinal survey of U.S. adults over the age of 50. Respondents were first interviewed in 1992 and follow-up interviews have occurred bi-annually through the present. The HRS sample design has changed slightly over the years to include respondents from a related study (i.e., AHEAD) and it periodically incorporates refresher cohorts. The HRS contains approximately 20,000 respondents in any given wave and over 12,000 deaths as of 2012. We conduct analyses with a harmonized version of the 1992-2010 HRS created by the RAND Corporation (HRS RAND File, Version N). Additional information about the HRS (Sonnega et al. 2014) and HRS RAND File (Chien et al. 2014) is available elsewhere.

We also examine mortality estimates from two smaller surveys: Americans' Changing Lives Study and the General Social Survey National Death Index. The Americans' Changing Lives Study is a survey of U.S. adults ages 25 and older (see House 2014 for details). ACL respondents were first interviewed in 1986 (N = 3,617) and follow-up interviews were conducted with the original respondents in 1989 (N = 2,867), 1994 (N = 2,562), and 2001/2 (N = 1,787), and 2011 (N = 1,427). The ACL has a closed cohort design and is linked to the National Death Index through 2011 (N = 1,832 deaths).

The GSS-NDI links cross-sectional data from the nationally-representative U.S. General Social Survey (1978-2002) to death records in the National Death Index between 1979 and 2008 (Muennig et al. 2011). The 1978-2008 GSS-NDI contains 32,830 respondents and 9,271 deaths. To maximize comparability with the other datasets examined, we limit our GSS-NDI analyses to survey years 1986-2002 (N = 24,133 Respondents, N = 5,519 Decedents).

We compare mortality rates and life expectancies in the NHIS-LMF, HRS, ACL, and GSS-NDI with mortality rates and life expectancies calculated from U.S. vital statistics data. We obtain U.S. life tables online from the Human Mortality Database (HMD 2016). Finally, we take several steps to maximize comparability across the datasets, given the differences in their years of data collection, age groups (i.e., HRS respondents are aged 50 years and older), ascertainment methodologies, etc. For instance, we run separate multivariate analyses for ages 25+ and 50+, so that the HRS data can be compared to the other datasets in the 50+ analyses.

Measures

The dependent variable in the NHIS-LMF, HRS, ACL, and GSS-NDI is all-cause mortality risk (0 = alive, 1 = dead). Covariates include self-reported age in years, gender (0 = male, 1 = female), race (1 = white, 2 = black, 3 = other race), marital status (1 = married, 2 = previously married, 3 = never married), and educational attainment (1 = less than high school, 2 = high school graduate, 3 = some college, 4 = bachelor's degree or more). Exposure to the risk of death is measured in calendar years. Survivors receive a partial year of exposure the first year they are interviewed (calculated based on interview month/quarter) and full years of exposure each year thereafter until they are censored (due to death or the end of follow-up). Decedents receive a partial year of exposure the first year they are interviewed (calculated based on interview month/quarter), a full year of exposure each year they survive, and a partial year of exposure the year they die (calculated based on month/quarter of death).

Note that in the analyses used to create the life tables (see below), age and exposure have slightly different specifications. Exact age on January 1st of year X is calculated based on respondents' self-reported birth dates[1] (birth month and year) and interview dates (GSS, HRS, and ACL: calculated using interview month and year, NHIS-LMF: calculated using interview quarter and year). Based on this information we then calculated age x in January, where age x ranges from exact *age x - 0.05* to exact *age x + 0.49*. This allows us to approximate traditional occurrence-exposure rates in the models used to create the life tables. Notably, results based on reported age were very similar to those based on calculated age x. In the models used to retrieve the mortality rates to create the life tables, respondents were assigned one year of exposure each year they survived (including the first interview year) and decedents were assigned 0.5 years of exposure in the interval in which they died. This is consistent with the common assumption that deaths occur, on average, at mid-year. Results from preliminary analyses with more precise calculations of exposure were virtually identical to the ones presented herein.

Analyses

The NHIS-LMF, HRS, ACL, and GSS-NDI are all reformatted into person-year files. Age in years, mortality status, and exposure to the risk of death vary over time (person-years). Although it is possible in the ACL and HRS to obtain time-varying information for other covariates (education, marital status), we measure these at baseline in all analyses. This is done to maximize

[1] Respondents' birth month was not available in the two most recent waves of the GSS-NDI (2000, 2002). However, the GSS did include respondents' Zodiac sign. We were able to use this information to randomly assign missing birth months.

comparability across each dataset. The analyses proceeded in two stages. The first stage involves comparing mortality rates and life expectancies estimated from the NHIS-LMF, HRS, ACL, and GSS-NDI to ones based on vital statistics data. We accomplish this by estimating parametric hazard models. We determined the optimal specification of the baseline hazard by estimating a series of exploratory analyses with different specifications of the baseline hazard (Exponential, Gompertz, Weibull, Log-Logistic, etc.). After reviewing these results, we decided to present results from Exponential models because this model fit the data best. To limit the amount of structure that we impose on the relationships examined, the models that the life tables are based on are parsimonious and predict adult mortality risk as a function of age in years. In each dataset, we estimate this model for the overall sample and separately for men and women. We use these models to estimate age-specific mortality rates and construct life tables using a multivariate life table approach (Teachman and Hayward 1993). We compare these results with mortality rates and life expectancies based on vital statistics data.

The second stage of analyses involves estimating a series of nested multivariate models within each dataset and comparing the relative mortality estimates (hazard ratios) generated from these analyses across each survey. We estimate semi-parametric hazard models (i.e., Cox models). We estimate two Cox Proportional Hazard models that progressively adjust for relevant covariates. The first model predicts all-cause mortality risk as a function of age in years, gender (reference group: men), and race (reference group: whites). The second model predicts all-cause mortality risk as a function of age in years, gender (reference group: men), race (reference group: whites), marital status (reference group: currently married at interview), and educational attainment (reference group: college education).

RESULTS

Mortality Rates and Life Tables

Tables 1 (Overall), 2 (Men), and 3 (Women) compare mortality rates and life expectancies from each survey with vital statistics data. Results for men and women combined reveal that age-specific all-cause mortality rates based on NHIS-LMF, HRS, and ACL data are very similar (see Tables 1-3, Figures 1-3). Mortality rates (m_x) and life expectancies (e_x) from these surveys and vital statistics data are also very similar. However, mortality rates based on NHIS-LMF, HRS, and ACL data are slightly lower than the ones reported in vital statistics data beginning around exact age 80 years. The actual timing and amount of divergence varies slightly across surveys though.

Yet, our results also imply that these observed differences in mortality rates have modest effects on life expectancy. This fact is most evident in Figures 4-6, which compare life expectancies in each survey with ones based on U.S. vital statistics data. As these figures illustrate, shows, life expectancies calculated from NHIS-LMF and ACL data are only about one year lower in the than mortality rates in U.S. vital statistics data. Life expectancy in the HRS is also generally comparable to vital statistics data, but differences between the HRS and vital statistics data are more pronounced before age 65 or so when compared to the other surveys (Figures 4-6). This especially is true for HRS men (Figure 5). Previous studies have shown similar results in the NHIS-LMF and HRS (Brown et al. 2012; Hummer et al. 1999; Lariscy et al. 2015), but to our knowledge no published studies have shown results of this kind with the ACL data. Differences between mortality rates at older ages based on surveys and vital statistics data primarily are thought to exist because the survey sampling frames exclude institutionalized persons – most notably nursing homes residents – who, on average, have a higher risk of death than non-institutionalized persons, whereas vital statistics data includes mortality information for the entire population (Brown et al. 2012; Lariscy et al. 2015; Lin et al. 2003).

Mortality estimates based on the NHIS-LMF, HRS, and ACL closely correspond with vital statistics data, but mortality estimates from the GSS-NDI do not (Tables 1-3, Figures 1-3). Our results strongly suggest that the GSS-NDI does not accurately reflect the mortality experience of the U.S. adult population. All-cause mortality rates (m_x) and life expectancies (e_x) generated from GSS-NDI and vital statistics data differ substantially. As Figures 1-3 show, GSS-NDI mortality rates are much higher than vital statistics mortality rates at younger ages, gradually converge with vital statistics mortality rates throughout midlife until the rates crossover at exact age 76, and become much lower than vital statistics mortality rates thereafter. Accordingly, life expectancy estimates in the GSS-NDI and vital statistics life tables are also discordant (Figures 4-6). When compared to life expectancies based on U.S. vital statistics data, life expectancy in the GSS-NDI is approximately 3-4 years lower at younger ages and 3-4 years higher at the oldest ages. The gender-specific results shown in Tables 2 (Men) and 3 (Women) are very similar to those shown in Table 1 for the overall population (also see Figures 2, 3, 5, 6).

Multivariate Results

Tables 4 (Ages 25+) and 5 (Ages 50+) show descriptive statistics for all variables included in the Cox Proportional Hazards models. Proportions in attributes of variables and mean age are comparable across datasets. This suggests thatthat the problems evident in the GSS-NDI mortality estimates possibly arise from NDI linkage process rather than from the collection of the GSS itself. The only exception to this is the proportion of deaths in the ACL. The ACL has a much higher proportion respondents who died during follow-up (Ages 25+ = 35.0%, Ages 50+ =

71.8%) than the other datasets. In contrast, only 20.1% of respondents ages 25 and older and 39.7% of respondents ages 50 and older died over the same period (1986-2011) in the NHIS-LMF. On first glance, the ACL seems to have an excess of deaths, but these differences likely arise due to differences between the ACL sample design and the sample design of the other surveys. Note that the ACL has a closed cohort design while the other surveys are either cross-sectional (NHIS-LMF, GSS-NDI) or contain refresher cohorts (HRS). Thus, these differences in the proportion of deaths in the ACL and the other surveys are reasonable.

Finally, Tables 6 (Ages 25+) and 7 (Ages 50+) show results from two nested Cox models that regress all-cause mortality risk on key covariates. Overall, the results from each datasets are relatively similar. All of the associations displayed in the tables are in the expected directions. The most apparent differences between the hazard ratios for the GSS-NDI and other datasets involve a higher mortality risk among black respondents relative to white respondents in the GSS-NDI and a flatter education-mortality gradient in the GSS-NDI. Interestingly, despite having a relatively small number of deaths in comparison to the other surveys, the ACL performs reasonably well when compared to the other datasets. Hazard ratios in the ACL generally fall somewhere in between those found in the GSS-NDI and the NHIS-LMF.

DISCUSSION

The major objectives of Healthy People 2020 involve the reduction of population health disparities, and surveys with mortality follow-up are an exemplary data source to monitor and evaluate progress toward accomplishing these objectives. Analyses based on surveys that contain longitudinal mortality follow-up have proliferated rapidly in recent years. These studies have become an increasingly important source of information on the influence of various social, economic, and behavioral factors on adult mortality risk. However, no studies have systematically examined whether mortality estimates based on these data sources are comparable across surveys and with national mortality estimates in vital statistics data.

Our findings suggest that the algorithm which links the GSS with the NDI has not accurately matched respondents with their death records. Probabilistic record linkage depends on PII reported in both data sources to correctly match respondents who died during the follow-up period to their death records. Provision of Social Security numbers (SSN) is particularly important for linkage since it is more unique to individuals than the other linkage items such as name and date of birth. Unlike in other surveys, SSN has not been consistently collected by GSS interviewers, with SSN only being solicited in years 1993 forward. Our results suggest that the GSS-NDI is not a reliable source of information on mortality and mortality disparities among U.S. adults.

Limitations

Since the markers of linkage quality (class and score) are not available in the public-use GSS-NDI dataset, we cannot be entirely certain that poorer linkage quality in the GSS-NDI than in the other survey-linked datasets is the source of the unanticipated GSS-NDI mortality estimates. However, Muenning et al. (2011) state that no GSS-NDI cases are placed into class 1 and very few cases are placed into class 2. Of the five classes, classes 1 and 2 are considered to be the most reliably matched decedents whereas the match quality for decedents placed into classes 3 and 4 is less certain. Class 5 matches are assumed to have survived the follow-up period. In contrast to the few GSS-NDI decedents in classes 1 and 2, among US-born white NHIS-LMF respondents who were identified as having died during follow-up, 63% of matches are in class 1 and 15% are in class 2 (Lariscy 2011).

Conclusion

Linked-mortality files are an increasingly important source of information on population health disparities, but few studies have examined how well these types of data perform relative to vital statistics data. Our results suggest that not all survey-linked mortality are created equally. Mortality estimates based on the NHIS-LMF, HRS, ACL are all very comparable to U.S. vital statistics data, but this is not the case for the GSS-NDI. The GSS-NDI is a clear outlier in terms of performance. Although we cannot be certain without detailed information on linkage quality, we are reasonably certain that the observed differences between the GSS-NDI and the other surveys arise because the GSS-NDI does not accurately capture mortality follow-up among its respondents The GSS-NDI does not accurately reflect the mortality experience of U.S. adults. Caution is warranted when using the GSS-NDI to examine adult mortality disparities.

References

Anderson, M. J., & Fienberg, S. E. (1999). *Who counts? The politics of census-taking in contemporary America*. New York: Russell Sage Foundation.

Arias, E., Eschbach, K., Schauman, W. S., Backlund, E. L., & Sorlie, P. D. (2010). The Hispanic mortality advantage and ethnic misclassification on U.S. death certificates. *American Journal of Public Health, 100*(Supplement 1), S171-S177.

Brown, D. C., Hayward, M. D., Montez, J. K., Hummer, R. A., Chiu, C.-T., & Hidajat, M. M. (2012). The significance of education for mortality compression in the United States. *Demography, 49*(3), 819-840.

Chien, S., Campbell, N., Hayden, O., Hurd, M., Main, R., Mallett, J., Martin, C., Meijer, E., Moldoff, M., Rohwedder, S., & St. Clair, P. (2014). *RAND HRS data documentation, version N*. Santa Monica, CA: RAND Center for the Study of Aging.

Harron, K., Goldstein, H., & Dibben, C. (Eds.). (2016). *Methodological developments in data linkage*. West Sussex, UK: Wiley.

House, J. S. (2014). *Americans' changing lives: Waves I, II, III, IV, and V, 1986, 1989, 1994, 2002, and 2011. [Computer file] ICPSR04690-v7*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Human Mortality Database. (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). www.mortality.org.

Hummer, R. A., Rogers, R. G., Nam, C. B., & Ellison, C. G. (1999). Religious involvement and U.S. adult mortality. *Demography, 36*(2), 273-285.

Lariscy, J. T. (2011). Differential record linkage by Hispanic ethnicity and age in linked mortality studies: Implications for the epidemiologic paradox. *Journal of Aging and Health, 23*(8), 1263-1284.

Lariscy, J. T., Hummer, R. A., & Hayward, M. D. (2015). Hispanic older adult mortality in the United States: New estimates and an assessment of factors shaping the Hispanic paradox. *Demography*, *52*(1), 1-14.

Lin, C. C., Rogot, E., Johnson, N. J., Sorlie, P. D., & Arias, E. (2003). A further study of life expectancy be socioeconomic factors in the National Longitudinal Mortality Study. *Ethnicity & Disease*, *13*(2), 240-247.

Lochner, K., Hummer, R. A., Bartee, S., Wheatcroft, G., & Cox, C. (2008). The public-use National Health Interview Survey Linked Mortality Files: Methods of reidentification risk avoidance and comparative analysis. *American Journal of Epidemiology, 168*(3), 336-344.

Minnesota Population Center. (2015). *Integrated Health Interview Series: Version 6.11*. Minneapolis, MN: University of Minnesota.

Muennig, P., Johnson, G., Kim, J., Smith, T. W., & Rosen, Z. (2011). The general social survey-national death index: An innovative new dataset for the social sciences. *BMC Research Notes, 4*, 385.

Pablos-Méndez, A. (1994). Mortality among Hispanics. *Journal of the American Medical Association, 271*(16), 1237.

Preston, S. H., & Taubman, P. (1994). Socioeconomic differences in adult mortality and health status. In L. G. Martin & S. H. Preston (Eds.), *Demography of Aging*. Washington, DC: National Academies Press.

Rogers, R. G., Hummer, R. A., & Nam, C. B. (2000). *Living and dying in the USA: Behavioral, health, and social differentials of adult mortality*. San Diego, CA: Academic Press.

Sonnega, A., Faul, J. D., Ofstedal, M. B., Langa, K. M., Phillips, J. W. R., & Weir, D. R. (2014). Cohort profile: The Health and Retirement Study (HRS). *International Journal of Epidemiology, 43*(2), 576-585.

Teachman, J. D., & Hayward, M. D. (1993). Interpreting hazard rate models. *Sociological Methods & Research, 21*(3), 340-371.

Table 1. Mortality Rates and Life Expectancies at Selected Ages for Men and Women Combined in U.S. Vital Statistics, NHIS-LMF, HRS, ACL, and GSS-NDI

	Vital Statistics[a]	NHIS-LMF[b]	HRS[c]	ACL[d]	GSS-NDI[e]
Age-Specific Mortality Rate (m_x)					
Age 25	0.0010	0.0006	---	0.0005	0.0026
Age 35	0.0015	0.0014	---	0.0012	0.0045
Age 45	0.0030	0.0032	---	0.0028	0.0079
Age 55	0.0070	0.0073	0.0046	0.0067	0.0137
Age 65	0.0164	0.0168	0.0123	0.0159	0.0239
Age 75	0.0379	0.0382	0.0324	0.0376	0.0417
Age 85	0.1000	0.0873	0.0856	0.0891	0.0727
Life Expectancy (e_x)					
Age 25	52.9	53.3	---	53.8	48.3
Age 35	43.5	43.8	---	44.2	39.8
Age 45	34.3	34.6	---	34.9	31.8
Age 55	25.6	26.0	27.6	26.2	24.7
Age 65	17.9	18.3	19.3	18.4	18.5
Age 75	11.3	12.0	12.2	11.9	13.3
Age 85	6.2	7.2	6.9	7.0	9.3

Vital statistics is for years 1986-2010

Mortality rates and life expectanices are based on weighted analyses. Exponential hazard models were estimated to obtain mortality rates. To construct the life tables (not shown), age-specific mortality rates were estimated for exact ages 25 to 100+ (HRS: 50-100+). The predicted mortality rates used to create the life tables are equivalent to exponentially smoothed central death rates (i.e., the mx column of a life table). Additional information about this approach is available elsewhere (Teachman and Hayward 1993). Other functional forms were considered, but exploratory analyses indicated that the exponential models fit best

Table 2. Mortality Rates and Life Expectancies at Selected Ages for Men in U.S. Vital Statistics, NHIS-LMF, HRS, ACL, and GSS-NDI

	Vital Statistics[a]	NHIS-LMF[b]	HRS[c]	ACL[d]	GSS-NDI[e]
Age-Specific Mortality Rate (m_x)					
Age 25	0.0015	0.0008	---	0.0006	0.0034
Age 35	0.0021	0.0017	---	0.0015	0.0057
Age 45	0.0039	0.0040	---	0.0036	0.0097
Age 55	0.0089	0.0090	0.0059	0.0087	0.0165
Age 65	0.0209	0.0206	0.0152	0.0209	0.0279
Age 75	0.0486	0.0472	0.0393	0.0502	0.0472
Age 85	0.1266	0.1079	0.1014	0.1207	0.0799
Life Expectancy (e_x)					
Age 25	50.2	50.9	---	50.7	45.6
Age 35	40.9	41.4	---	41.1	37.4
Age 45	31.9	32.3	---	32.0	29.8
Age 55	23.5	23.9	25.8	23.5	23.1
Age 65	16.1	16.6	17.8	16.0	17.3
Age 75	9.9	10.6	11.1	10.0	12.5
Age 85	5.4	6.2	6.2	5.6	8.8

Vital statistics is for years 1986-2010

Mortality rates and life expectanices are based on weighted analyses. Exponential hazard models were estimated to obtain mortality rates. To construct the life tables (not shown), age-specific mortality rates were estimated for exact ages 25 to 100+ (HRS: 50-100+). The predicted mortality rates used to create the life tables are equivalent to exponentially smoothed central death rates (i.e., the mx column of a life table). Additional information about this approach is available elsewhere (Teachman and Hayward 1993). Other functional forms were considered, but exploratory analyses indicated that the exponential models fit best.

Table 3. Mortality Rates and Life Expectancies at Selected Ages for Women in U.S. Vital Statistics, NHIS-LMF, HRS, ACL, and GSS-NDI

	Vital Statistics[a]	NHIS-LMF[b]	HRS[c]	ACL[d]	GSS-NDI[e]
Age-Specific Mortality Rate (m_x)					
Age 25	0.0005	0.0004	---	0.0003	0.0019
Age 35	0.0010	0.0010	---	0.0008	0.0035
Age 45	0.0022	0.0024	---	0.0019	0.0063
Age 55	0.0051	0.0057	0.0035	0.0048	0.0115
Age 65	0.0125	0.0135	0.0097	0.0120	0.0208
Age 75	0.0302	0.0319	0.0272	0.0301	0.0378
Age 85	0.0870	0.0753	0.0759	0.0754	0.0685
Life Expectancy (e_x)					
Age 25	55.6	55.7	---	56.6	50.6
Age 35	45.9	46.0	---	46.8	41.8
Age 45	36.5	36.6	---	37.3	33.5
Age 55	27.5	27.8	29.1	28.3	26.0
Age 65	19.3	19.8	20.5	20.1	19.4
Age 75	12.2	13.1	13.0	13.1	13.8
Age 85	6.6	7.9	7.4	7.7	9.5

Vital statistics is for years 1986-2010

Mortality rates and life expectanices are based on weighted analyses. Exponential hazard models were estimated to obtain mortality rates. To construct the life tables (not shown), age-specific mortality rates were estimated for exact ages 25 to 100+ (HRS: 50-100+). The predicted mortality rates used to create the life tables are equivalent to exponentially smoothed central death rates (i.e., the mx column of a life table). Additional information about this approach is available elsewhere (Teachman and Hayward 1993). Other functional forms were considered, but exploratory analyses indicated that the exponential models fit best

Table 4. Descriptive Statistics for Respondents Ages 25+ in the NHIS-LMF, HRS, ACL, and GSS-NDI Samples

	NHIS-LMF[a]	ACL[b]	GSS-NDI[c]
	N = 1,357,326	*N = 3,614*	*N = 21,339*
Age (Mean)	48.4	47.1	47.2
Female (%)	52.2	52.9	54.9
Race-Ethnicity (%)			
White	83.7	83.5	82.7
Black	10.9	10.9	12.3
Other	5.4	5.6	5.0
Marital Status (%)			
Currently Married	67.9	69.4	64.1
Previously Married	19.8	20.4	22.1
Never Married	12.3	10.2	13.8
Education (%)			
< High school	18.4	25.6	19.0
High school	34.9	31.4	31.0
Some college	22.8	23.3	25.0
College	23.9	19.7	25.0
Deaths (%)	20.1	35.0	22.5

[a] National Health Interview Survey Linked Mortality File (NHIS-LMF, NHIS: 1986-2009, NDI: 1986-2011).

[b] Americans' Changing Lives Survey (ACL, 1986-2011).

[c] General Social Survey-National Death Index (GSS-NDI, GSS: 1986-2002, NDI: 1986-2008).

Notes: The percentages (means) shown in the table are weighted. The analyses are based on individual-level data.

Table 5. Descriptive Statistics for Respondents Ages 50+ in the NHIS-LMF, HRS, ACL, and GSS-NDI Samples

	NHIS-LMF[a]	HRS[b]	ACL[c]	GSS-NDI[d]
	N = 559,903	*N = 34,985*	*N = 2,070*	*N = 8,553*
Age (Mean)	64.2	59.0	55.8	63.6
Female (%)	54.2	53.3	65.0	56.2
Race-Ethnicity (%)				
White	86.8	83.2	86.9	85.8
Black	9.4	10.9	10.4	11.3
Other	3.8	5.9	2.7	3.0
Marital Status (%)				
Currently Married	66.7	73.4	66.6	65.4
Previously Married	28.6	21.5	29.5	30.1
Never Married	4.7	5.1	3.9	4.6
Education categories				
< High school	26.7	22.6	42.2	29.1
High school	35.5	31.0	30.5	32.0
Some college	18.6	22.4	16.0	19.0
College	19.2	24.1	11.4	20.0
Deaths (%)	39.7	26.1	71.8	39.1

[a] National Health Interview Survey Linked Mortality File (NHIS-LMF, NHIS: 1986-2009, NDI: 1986-2011).

[b] Health and Retirement Study (HRS, 1992-2010).

[c] Americans' Changing Lives Survey (ACL, ACL: 1986-2011, NDI: 1986-2011).

[d] General Social Survey-National Death Index (GSS-NDI, GSS: 1986-2002, NDI: 1986-2008).

Notes: The percentages (means) shown in the table are weighted. The analyses are based on individual-level data.

Table 6. Cox Proportional Hazards Models Predicting All-Cause Mortality Risk as a Function of Selected Covariates in the NHIS-LMF, ACL, and GSS-NDI: Men and Women Ages 25 and older

	Model 1			Model 2		
	NHIS-LMF[a]	ACL[b]	GSS-NDI[c]	NHIS-LMF[a]	ACL[b]	GSS-NDI[c]
Age in years	1.09***	1.10***	1.06***	1.08***	1.09***	1.05***
	(0.0003)	(0.0030)	(0.0011)	(0.0003)	(0.0030)	(0.0012)
Female	0.68***	0.59***	0.75***	0.63***	0.56***	0.72***
	(0.0044)	(0.0377)	(0.0235)	(0.0047)	(0.0365)	(0.0234)
Race-Ethnicity						
Black	1.28***	1.43***	1.49***	1.12***	1.25***	1.36***
	(0.0120)	(0.0973)	(0.0706)	(0.0115)	(0.0893)	(0.0657)
Other	0.75***	1.27	1.12	0.73***	1.15	1.07
	(0.0176)	(0.2715)	(0.1059)	(0.0178)	(0.2568)	(0.1015)
Marital Status						
Previously Married				1.24***	1.24***	1.19***
				(0.0063)	(0.0800)	(0.0411)
Never Married				1.40***	1.19	1.19***
				(0.0116)	(0.1764)	(0.0694)
Educational Attainment						
Less than High School				1.81***	1.79***	1.48***
				(0.0097)	(0.2105)	(0.0741)

High School	1.50***	1.35***	1.25***
	(0.0083)	(0.1660)	(0.0604)
Some College	1.33***	1.32***	1.19***
	(0.0090)	(0.1744)	(0.0622)

Notes: * p < .10, ** p < .05, *** p < .001. The table displays hazard ratios and standard errors are in parentheses. The results are weighted. Estimates are based on individual-level analyses. The reference categories are as follows: men, white race-ethnicity, currently married, and college education.

[a] National Health Interview Survey Linked Mortality File (NHIS-LMF, NHIS: 1986-2009, NDI: 1986-2011).

[b] Americans' Changing Lives Survey (ACL, ACL: 1986-2011, NDI: 1986-2011).

[c] General Social Survey-National Death Index (GSS-NDI, GSS: 1986-2002, NDI: 1986-2008).

Table 7. Results from Cox Proportional Hazards Models Predicting All-Cause Mortality Risk as a Function of Selected Covariates in the NHIS-LMF, HRS, ACL, and GSS-NDI, aged 50 years and older

	Model 1				Model 2			
	NHIS-LMF[a]	HRS[b]	ACL[c]	GSS-NDI[d]	NHIS-LMF[a]	HRS[b]	ACL[c]	GSS-NDI[d]
Age in years	1.08***	1.08***	1.10***	1.06***	1.08***	1.07***	1.09***	1.06***
	(0.0002)	(0.0010)	(0.0051)	(0.0021)	(0.0002)	(0.0009)	(0.0051)	(0.0022)
Female	0.68***	0.81***	0.56***	0.77***	0.63***	0.78***	0.53***	0.72***
	(0.0045)	(0.0099)	(0.0378)	(0.0295)	(0.0048)	(0.0110)	(0.0354)	(0.0294)
Race-Ethnicity								
Black	1.19***	1.06	1.30***	1.32***	1.07***	0.98*	1.14*	1.21***
	(0.0075)	(0.0056)	(0.0919)	(0.0782)	(0.0077)	(0.0414)	(0.0851)	(0.0718)
Other	0.70***	0.68***	1.65***	0.98	0.68***	0.65***	1.53*	0.94
	(0.0148)	(0.0526)	(0.3988)	(0.1466)	(0.0148)	(0.0460)	(0.3658)	(0.1415)
Marital Status								
Previously Married					1.19***	1.08***	1.24***	1.21***
					(0.0052)	(0.0205)	(0.0820)	(0.0485)
Never Married					1.27***	0.97	1.36*	1.10
					(0.0105)	(0.0443)	(0.2210)	(0.0973)
Educational Attainment								
Less than High School					1.60***	1.68***	1.59***	1.41***
					(0.0075)	(0.0558)	(0.1971)	(0.0861)
High School					1.35***	1.54***	1.28*	1.23***
					(0.0075)	(0.0493)	(0.1698)	(0.0754)
Some College					1.24***	1.23***	1.24	1.16**
					(0.0088)	(0.0351)	(0.1750)	(0.0810)

Notes: * p < .10, ** p < .05, *** p < .001. The table displays hazard ratios and standard errors are in parentheses. The results are weighted. Estimates are based on individual-level analyses. The reference categories are as follows: men, white race-ethnicity, currently married, and college education.

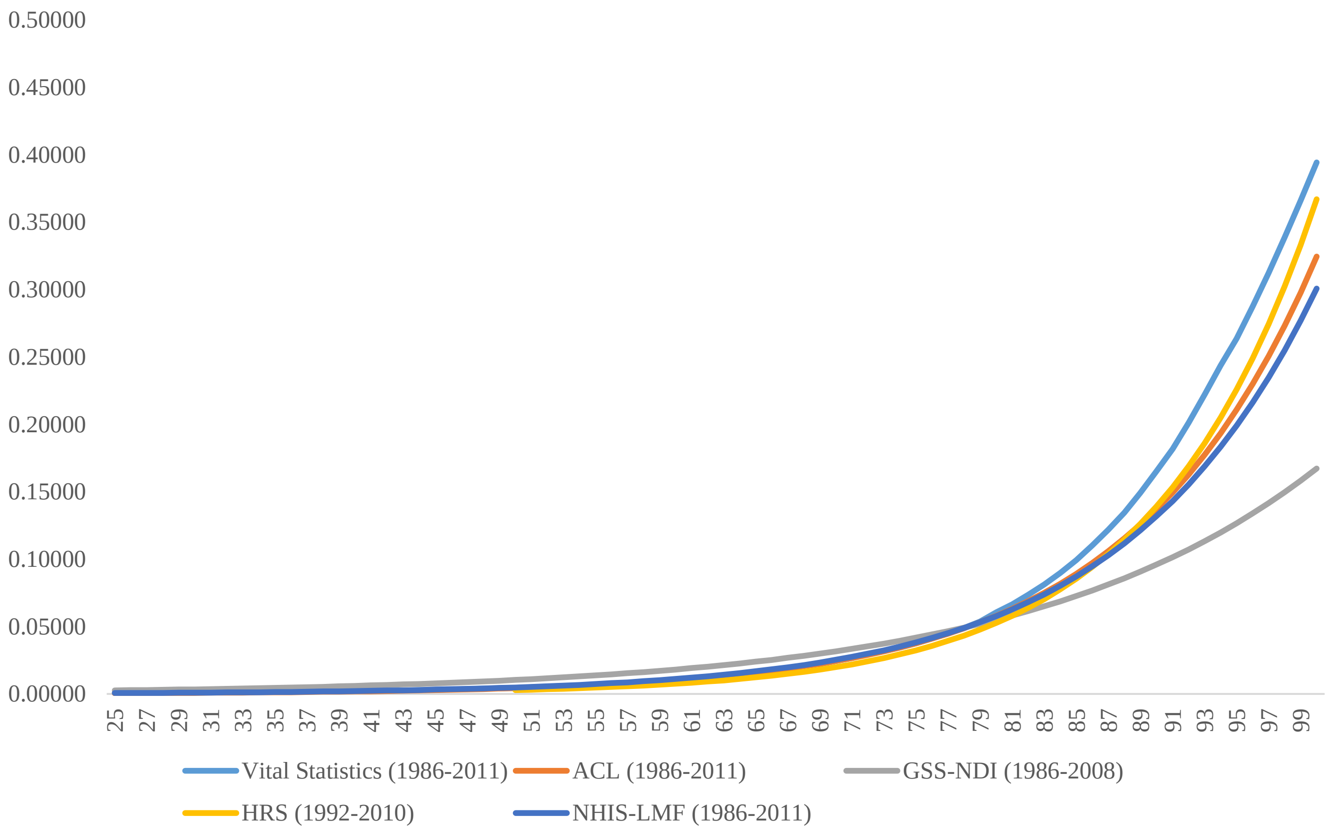
[a] National Health Interview Survey Linked Mortality File (NHIS-LMF, NHIS: 1986-2009, NDI: 1986-2011).

[b] Health and Retirement Study (HRS, 1992-2010).

[c] Americans' Changing Lives Survey (ACL, ACL: 1986-2011, NDI: 1986-2011).

[d] General Social Survey-National Death Index (GSS-NDI, GSS: 1986-2002, NDI: 1986-2008).

Figure 1. Mortality Rates in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Men and Women



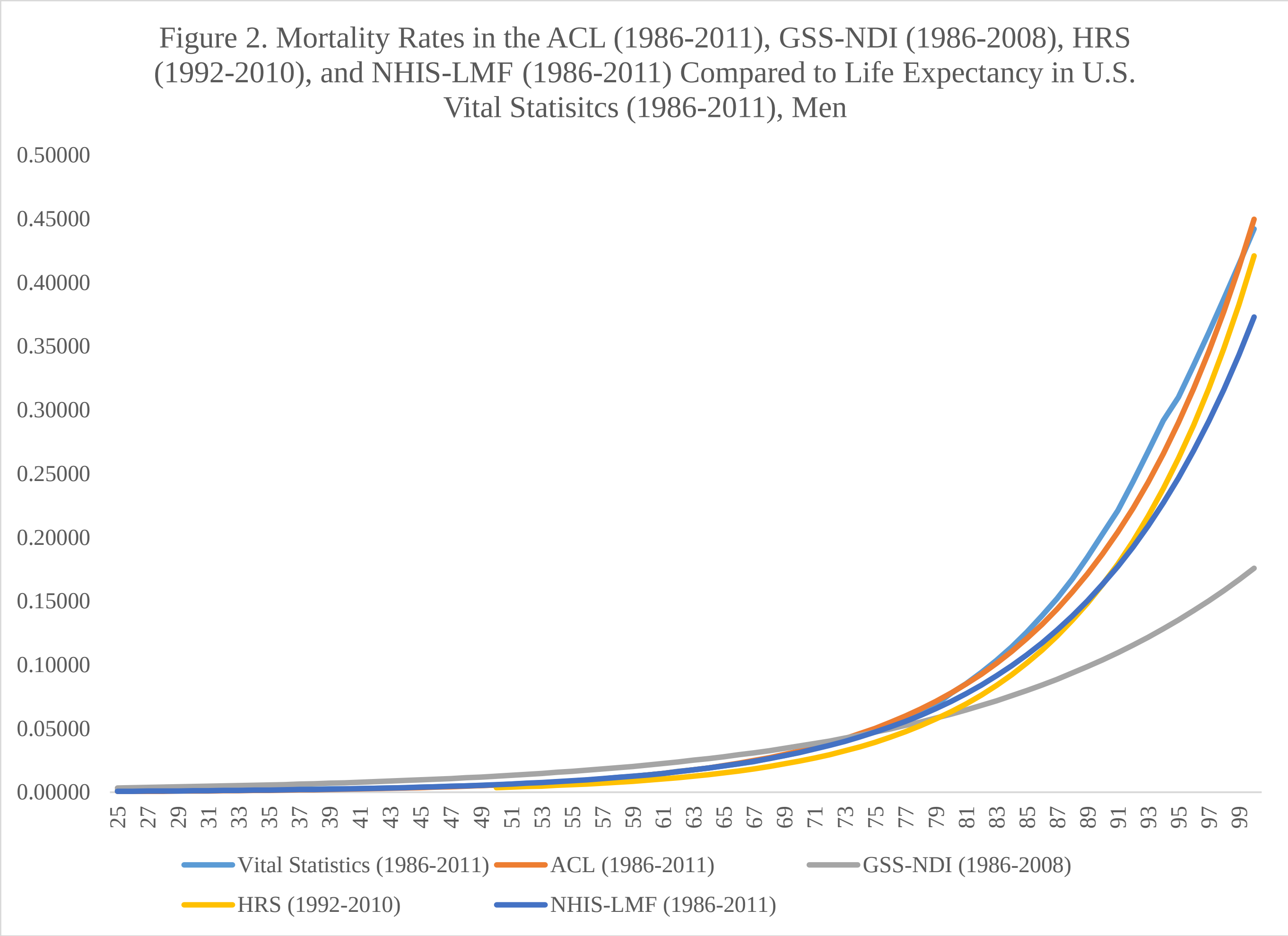
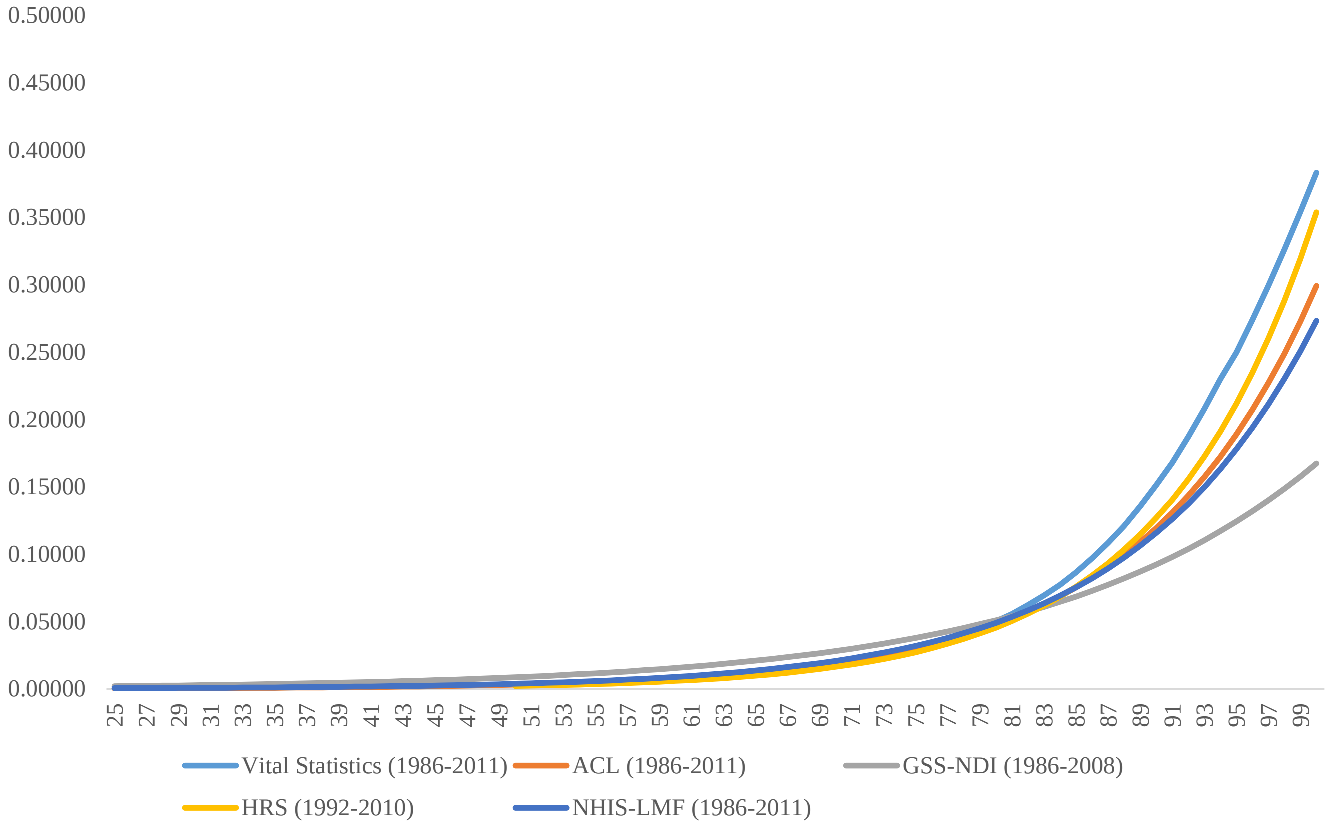


Figure 2. Mortality Rates in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Men
0.50000
0.45000
0.40000
0.35000
0.30000
0.25000
0.20000
0.15000
0.10000
0.05000
0.00000
25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97 99
Vital Statistics (1986-2011)
ACL (1986-2011)
GSS-NDI (1986-2008)
HRS (1992-2010)
NHIS-LMF (1986-2011)

Figure 3. Mortality Rates in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Women



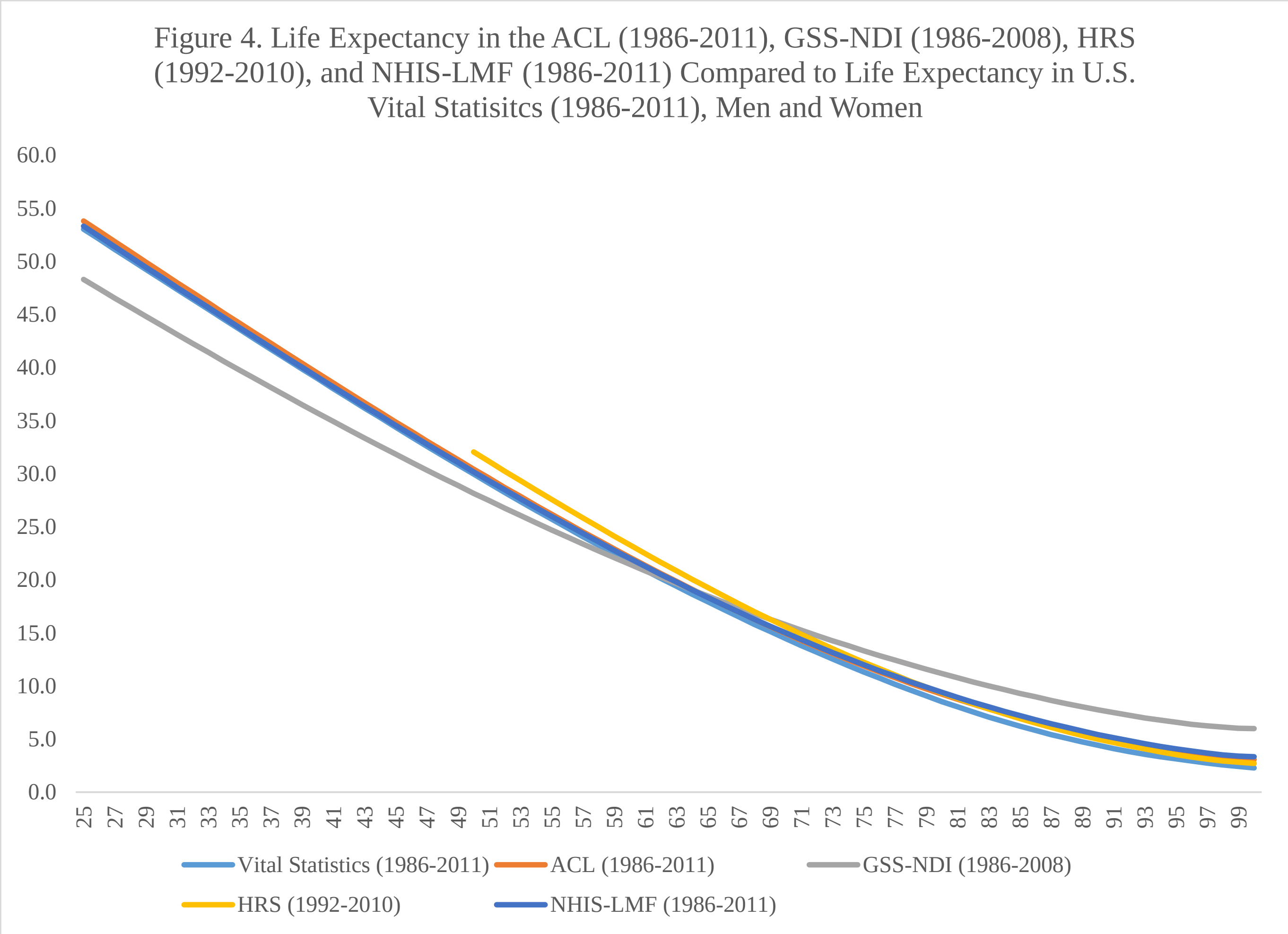


Figure 4. Life Expectancy in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Men and Women
60.0
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97 99
Vital Statistics (1986-2011)
ACL (1986-2011)
GSS-NDI (1986-2008)
HRS (1992-2010)
NHIS-LMF (1986-2011)

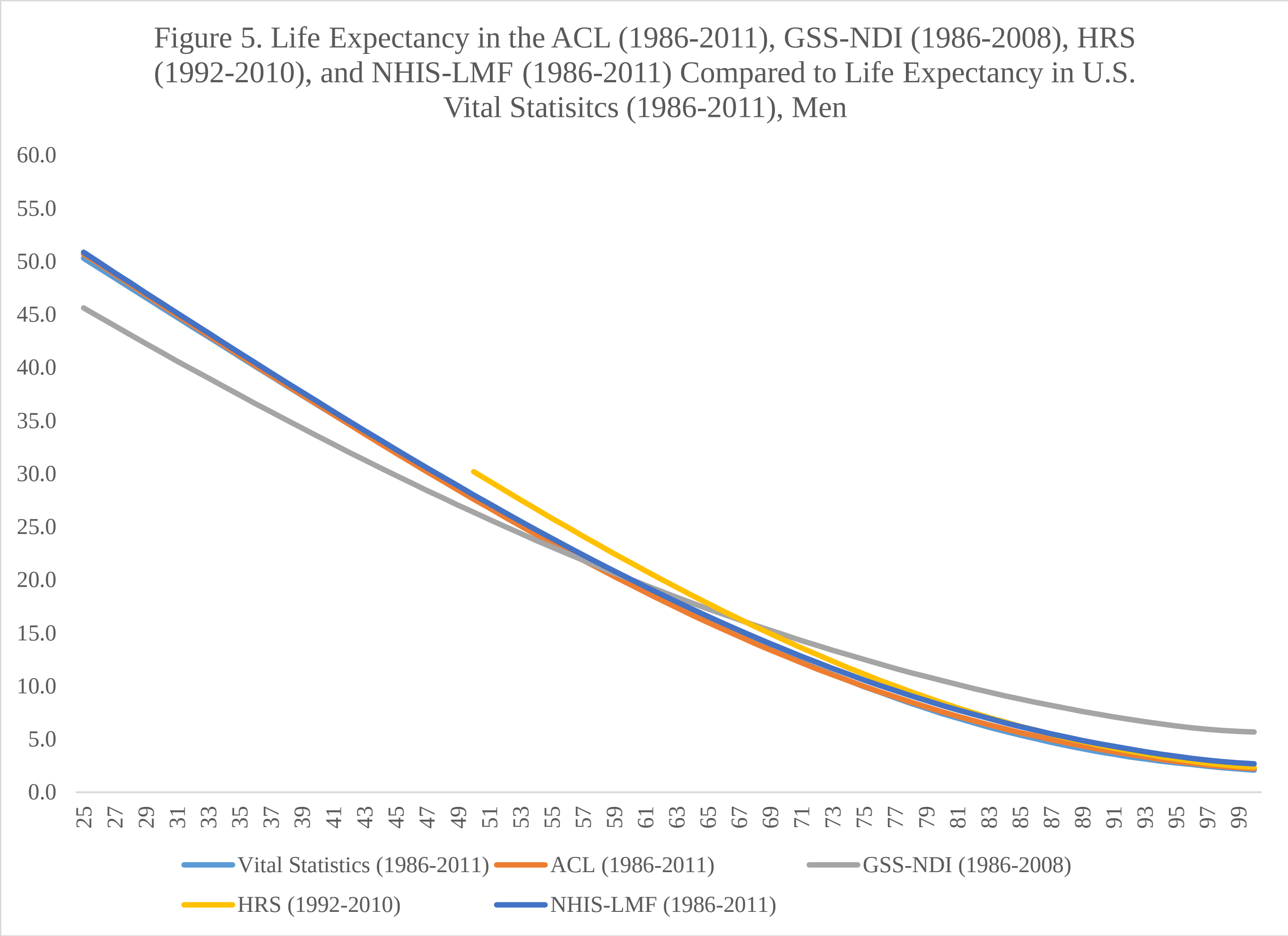


Figure 5. Life Expectancy in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Men
60.0
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97 99
Vital Statistics (1986-2011)
ACL (1986-2011)
GSS-NDI (1986-2008)
HRS (1992-2010)
NHIS-LMF (1986-2011)

Figure 6. Life Expectancy in the ACL (1986-2011), GSS-NDI (1986-2008), HRS (1992-2010), and NHIS-LMF (1986-2011) Compared to Life Expectancy in U.S. Vital Statisitcs (1986-2011), Women