HHS Public Access

Author manuscript

Demogr Res. Author manuscript; available in PMC 2015 July 02.

Published in final edited form as:

Demogr Res. 2014 ; 21(2): 27–70. doi:10.4054/DemRes.2014.31.2.

Factors responsible for mortality variation in the United States: A latent variable analysis

Christopher Tencza[1], **Andrew Stokes**[2], and **Samuel Preston**[3]

[1] University of Pennsylvania, U.S.A.

[2] University of Pennsylvania, U.S.A.

[3] University of Pennsylvania, U.S.A.

Abstract

BACKGROUND—Factors including smoking, drinking, substance abuse, obesity, and health care have all been shown to affect health and longevity. The relative importance of each of these factors is disputed in the literature, and has been assessed through a number of methods.

OBJECTIVE—This paper uses a novel approach to identify factors responsible for interstate mortality variation. It identifies factors through their imprint on mortality patterns and can therefore identify factors that are difficult or impossible to measure directly, such as sensitive health behaviors.

METHODS—The analysis calculates age-standardized death rates by cause of death from 2000-2009 for white men and women separately. Only premature deaths between ages 20-64 are included. Latent variables responsible for mortality variation are then identified through a factor analysis conducted on a death-rate-by-state matrix. These unobserved latent variables are inferred from observed mortality data and interpreted based on their correlations with individual causes of death.

RESULTS—Smoking and obesity, substance abuse, and rural/urban residence are the three factors that make the largest contributions to state-level mortality variation among males. The same factors are at work for women but are less vividly revealed. The identification of factors is supported by a review of epidemiologic studies and strengthened by correlations with observable behavioral variables. Results are not sensitive to the choice of factor-analytic method used.

CONCLUSIONS—The majority of interstate variation in mortality among white working-age adults in the United States is associated with a combination of smoking and obesity, substance abuse and rural/urban residence.

tencza@sas.upenn.edu. astokes@sas.upenn.edu. spreston@sas.upenn.edu.

1. Introduction

Understanding the underlying causes of variation in mortality is useful for health policy and intervention design. However, risk factors can be difficult to measure directly, as observed measures are often products of traits or circumstances that are unobserved, partially observed, or complex and multidimensional. An alternative is to infer the effects of risk factors indirectly using a latent variable approach. Because factors that influence mortality typically manifest themselves in several causes of death, associations among causes of death over time or space may provide important information about underlying causal factors. Without explicitly introducing a latent variable model, such reasoning has been used to infer the role of cigarette smoking (Peto et al. 1992; Preston et al. 2011) and the quality of the health care system in explaining variation in US mortality (Nolte and Mckee 2004).

In this paper we take advantage of variation in mortality by cause of death across US states to identify the underlying factors that are creating such variation. The emphasis is on behavioral factors that affect the risk of death. Previous studies have suggested that behavioral factors play a leading role in explaining US mortality (Mokdad et al. 2004; Mokdad et al. 2005; Murray et al. 2006; Danaei et al. 2009; 2010; Mehta and Preston 2012; Murray et al. 2013).

The primary approach to identifying the role of behavioral factors in interstate mortality variation is to apply relative risks derived from epidemiologic studies to the risk factor distribution of populations using the population attributable fraction (PAF) (Danaei et al. 2009; Danaei et al. 2010). In order to provide reliable results, such an approach requires accurate data on both relative risks and on risk factor distributions. Neither is measured with a high degree of accuracy or certainty and in some instances the data are altogether unavailable. For example, the fraction of deaths attributable to obesity in the US varies by a factor of 3-4 depending on which set of national estimates of relative risks is employed (Mehta and Chang 2009). Relative risks from smoking depend on the number of cigarettes consumed per day, inhalation, filtration, tar content, and especially the duration and past intensity of the habit. These elements are not readily captured in a single variable. Data on other behaviors, such as use of illicit drugs and unsafe sex, are often unreliable because of their sensitive nature. Finally, the source of data for most regional analyses of health patterns in the US, the Behavioral Risk Factor Surveillance Survey (BRFSS), is subject to several important limitations related to validity and comparability of data, including reliance on self-reported data, exclusion of households without telephones, and high rates of non-response. The national response rates for BRFSS in 2011 were 53.0% for landlines and 27.9% for cell phones (Centers for Disease Control and Prevention 2013a).

The present study takes an entirely different approach. It treats behavioral factors as latent variables that are identifiable through covariation of causes of death across populations. The operation of a particular risk factor is expected to appear in the form of high correlations across states within the cluster of causes of death for which its relative risks are greatest. Unlike prior studies, the present study is not limited to the subset of risk factors that can be reliably measured; thus it has the potential to uncover previously overlooked patterns of risk in populations. Also, because this approach is independent of the attributable-risk approach, it provides a valuable independent assessment of the contribution of behavioral risk factors to mortality variation.

2. Background and approach

Through straightforward decompositional methods, inter-population differences in death rates or life expectancy can be readily assigned to various causes of death (Preston et al. 2001). Beyond their contributions to such accounting exercises, death rates from a particular cause have also been used as indicators of broader determinants of mortality. One set of applications of this approach has used particular causes of death (variously termed “sentinel”, “tracer” or “amenable” causes) as indicators of the performance of a medical system (Polednak 2000; Nolte and McKee 2008).

A second set of applications has used lung cancer as an indicator of smoking prevalence and intensity (Peto et al. 1992; Preston et al. 2010; Fenelon and Preston 2012). Statistical relations were established between lung cancer mortality and mortality from other causes of death across OECD countries (Preston et al. 2011) and across states of the United States (Fenelon and Preston 2012). The statistical model used in these applications was negative binomial regression in which death rates from other causes of death were regressed on death rates from lung cancer and a set of indicator variables. The causal model assumed in this approach is shown in Figure 1A. But if mortality from both lung cancer and other causes of death are functions of smoking, then smoking should be treated as a latent variable, as shown in Figure 1B. Furthermore, there is no reason to expect that smoking is the only factor at work in fashioning population mortality patterns: the role of smoking should be inferred in a multivariate context.

In the present analysis we assume that behavioral factors are latent variables whose effects manifest themselves in a variety of causes of death. We will employ a list of mutually exclusive causes of death, without prejudging which causes of death, or behavioral risk factors, are to be featured.

The influence of a particular risk factor will be inferred from the significance of variation in causes of death with which that factor is most closely associated. We look, in particular, for causes of death associated with four behavioral risk factors: smoking, obesity, alcohol abuse, and illicit drug use. Obesity is not technically a behavioral risk factor but we refer to it as such because it closely related to behaviors of dietary intake and physical activity. We also consider a non-behavioral factor, the quality of the health care system. In some instances operation of a factor can be identified through explicit terms contained in the International List of Diseases, Injuries, and Causes of Death, e.g., alcohol poisoning. In these cases the behavioral factor (alcohol) has been assigned by the attending physician or coroner. More commonly, we rely on epidemiologic studies, and meta-analyses thereof, to identify the causes of death that demonstrate the highest relative risks for a particular risk factor.

Table 1 presents, in alphabetical order, the set of causes of death that we expect to be most closely associated with the five risk factors that we consider. Amenable causes represent deaths that are considered avoidable through medical services. References to the studies that support such identifications are included.

3. Data and methods

We use factor analysis to identify latent variables underlying mortality variation between US states during the period 2000–2009. Factor analysis describes observed, correlated variables by means of a smaller number of latent, unobserved variables or ‘factors’. These factors are ordered by the amount of variance they explain. Each factor identified has a factor loading for each of the observed variables (i.e., mortality rates for a particular cause of death). The loading is the correlation between that factor and a particular cause of death, and the sum of their correlations across k factors is how much variation in that cause of death can be explained by k factors. Each factor also produces a factor score for each observation (i.e., state), indicating the relative intensity of the operation of that factor in the state.

In mathematical terms, cause-specific death rates (C_1, C_2...C_p) are treated as functions of underlying factors (F_1, F_2...F_m). Each cause of death is linearly related to the chosen number of factors as shown:

$$C_j = B_{j1}F_1 + B_{j2}F_2 + \cdots + B_{jm}F_m + e_j \quad (1)$$

where C_j represents the j^{th} of p causes of death obtained from n independent subjects (states). B_{jm} represent the factor loadings relating cause of death j to the m^{th} factor F, and e_j represents the variance unique to cause of death C_j (Brown 2006).

We examine mortality by cause of death for white persons aged 20-64. Deaths within this age range are conventionally considered ‘premature’. Those dying are less likely to suffer from multiple pathologic processes that can make cause of death assignments more uncertain at older ages. As demonstrated in the electronic supplementary material, results are very similar when the age range is expanded. Attention is confined to the white population in order to minimize the role of sources of variation associated with race.

We obtained counts of deaths by age, state, and cause of death from the National Center of Health Statistics (NCHS) (2010). State identifiers for deaths in years 2007-2009, not available in the public-use mortality file, were obtained through a special request to the NCHS. State level mid-year population estimates are taken from the National Center for Health Statistics bridged-race population estimates (National Center for Health Statistics 2012b). Hawaii is dropped from the analysis because of its very small number of deaths. We calculate the death rates by age, state, and cause of death over the decade 2000–2009. By pooling ten years of data we are able to calculate reliable rates for a larger set of causes of death. We then age-standardize the death rates using the age distribution of the US population in 2000 as the standard (Anderson and Rosenberg 1998). We perform separate analyses for men and women.

We use cause-of-death groupings from the recently published Global Burden of Disease compendia (GBD), which were developed for public health applications from the

International Classification of Diseases versions 9 and 10 (Naghavi et al. 2010). We have made several modifications to the GBD list that are described in the electronic supplementary material. Most of these adjustments disaggregate individual causes of death with epidemiological significance from larger groups of causes (e.g., alcoholic liver cirrhosis is separated from the cirrhosis of the liver category). Causes of death with fewer than 5,000 deaths for either sex during the decade for the US as a whole are eliminated from the analysis.

Analysts using factor analysis face several methodological choices. For our basic analysis we use the procedures that are most conventional: factors are assumed to be uncorrelated with one another, factor scores for each state or state-year are estimated using the regression-based approach, scree plots are used to determine the number of factors selected, and varimax rotation is used after determining the number of factors to retain. As demonstrated in the electronic supplementary material, results are not sensitive to these choices.

This is not the first effort to use factor analysis or its close relative, principal components analysis, in a study of mortality variation (see United Nations (1982) and Gavrilova et al. (2002) for earlier examples). However, to the best of our knowledge it is the first effort to use spatial variation in mortality by cause of death to identify latent variables responsible for mortality variation.

4. Results

We chose to include three factors for each sex, because plots of the proportion of variance explained by the successive addition of factors to the analysis (scree plots) show a rapid fall-off after three for males. For females the drop off is not as abrupt, but we retained three factors to remain consistent across sexes. In the electronic supplementary material we show that factor loadings for the first three factors are not sensitive to the choice of the number of factors retained. The scree plots and eigenvalues of the factor analysis also appear in the electronic supplementary material. The first factor explains 34% of the total variance in causes of death for males and 33% of the total variance for females. The cumulative variance explained by three factors is 61% and 56% for males and females, respectively.

Table 2 presents the factor loadings for males. The table contains only factor loadings that are greater than 0.60 ($p<.00001$). These factor loadings are simply the correlation coefficients between the columns (factors) and rows (causes of death). We chose a relatively high cutoff of 0.60 in our interpretation of factors, but all loadings are reported in the appendix. We have given names to the factors based upon our expectations of the causes of death with which a factor is most closely identified. We label the first factor “Smoking/ Obesity”. The cause of death with which the first factor is most highly correlated is lung cancer ($r = 0.947$). In fact, this is the highest correlation of any cause of death with any factor for either sex. This result is a clear indication of the importance of smoking to explaining interstate variation in mortality.

Most of the other causes of death that are heavily loaded on Factor 1 are also closely associated with smoking: COPD, oral cancers, ischemic heart disease, and cerebrovascular disease. Such a pattern accords with our expectations about patterns of variation by cause of death that should be observed if smoking were an important contributor to interstate variation in mortality. In fact, 7 of the 10 causes of death identified in Table 1 to be most strongly associated with smoking have loadings on Factor 1 above 0.60. The 3 remaining causes of death identified for smoking in Table 1 have moderately high loadings: esophagus cancer (0.594), aneurysm (0.567), and bladder cancer (0.457). Also appearing on the list of heavily loaded causes of death is exposure to smoke and heat, often a consequence of careless smoking (United States Fire Administration 2013).

One cause of death with only mildly elevated risks for smokers also appears on the list of causes with high loadings on Factor 1: diabetes mellitus. That is one of the most important causes of death through which obesity operates. Colorectal cancers are also closely associated with obesity (Table 1) and appear in the list of heavily loaded causes on Factor 1. Cerebrovascular disease and ischemic heart disease appear on the lists of diseases closely associated with both smoking and obesity. In other words, Factor 1 includes representations of both smoking and obesity as causal factors in mortality variation, and their influences are not readily disentangled. This issue is addressed in more detail below.

Figure 2 shows the map of factor scores for Factor 1 among males. Clearly, this factor is heavily concentrated in the Appalachian region down through the Deep South. The lowest factor scores are concentrated in the West, especially Utah and Colorado. This pattern is highly correlated with death rates from lung cancer, but it is also highly correlated with death rates from all causes combined (r = .72). This latter correlation is not a necessary result of using factor analysis because the correlations between causes of death do not reflect the relative magnitude of death rates from various causes, only their patterns of variation across multiple causes.

We have labeled Factor 2 for males “Substance Abuse”, an ascription that appears particularly clear-cut. Hepatitis and HIV/AIDS are associated with drug abuse and unsanitary use of needles, and accidental poisonings consist almost exclusively of drug overdoses. Liver cirrhosis and liver cancer are associated with hepatitis and substance abuse. Six of the nine causes of death hypothesized to be most closely associated with drug abuse in Table 1 have factor loadings above 0.60 in Table 2. Alcohol abuse appears to play a role in Factor 2 as well. Liver cancer, alcoholic liver cirrhosis, and interpersonal violence are closely associated with alcohol abuse (Table 1) and appear on the list of causes of death most heavily loaded on Factor 2. “Other digestive diseases” are also heavily loaded on male Factor 2. This category includes pancreatitis and liver disease, which are often an outcome of alcohol abuse (Rehm 2011). The appearance of substance abuse as a major factor in explaining mortality variation is an important finding, as previous studies have found it difficult to measure substance abuse and to understand its contribution to mortality variation. This appears to be the first instance in which a geographic pattern of mortality from various causes of death associated with substance abuse has been identified.

Figure 2 shows a map of factor scores for male Factor 2. The highest scores occur in states along the southern perimeter from Florida to California, with the highest scores in New Mexico and Nevada. Lowest values are observed in the upper Midwest, especially in North and South Dakota. Unlike other factors, we have a limited understanding of the expected geographic variation in substance abuse, and therefore are more speculative in our interpretation of the factor scores. One hypothesis is that the geographic pattern could reflect a greater relative ease of drug trafficking along the southern border.

Factor 3 for males is a vivid pattern associated with injuries. The five causes of death that correlate the highest with this factor are all some form of injury, including self-inflicted injuries, transport injuries, exposure to mechanical forces, other accidents, and drowning (Table 2). The death rate from falls is the next most highly correlated with Factor 3 at 0.586. While deaths from drowning and falls are not numerous, their geographic distribution is quite similar to that of several more important sources of fatal injury.

Figure 2 presents the map of factor scores for male Factor 3. Montana and Wyoming lead the way in injury mortality, with other states in the Mountain time zone also exhibiting high scores. States with large urban populations like New York, Massachusetts, New Jersey, Illinois, and California have the lowest scores on this factor. Below, we demonstrate more formally the connection between rurality and Factor 3, a connection that supports our designation of this factor as a rural/urban factor. That mortality from injuries is higher in rural areas has been clearly demonstrated (Myers et al. 2013). A high rate of mass transit use may help keep mortality from transport injuries low in urban states, while lower population density in rural states may increase driving exposure. An individual’s greater proximity to trauma centers in urban areas may also contribute (Myers et al. 2013). Another factor may be increased exposure to occupational hazards in rural areas. That suicide is also heavily loaded on this factor is intriguing, since transport injuries, falls, and drownings represent additional ways of ending one’s life. Residents of rural areas have traditionally had higher suicide rates than urban residents (McCarthy et al. 2012; Singh and Siapush 2002). Furthermore, the map may reflect injury prevention laws, since mountain states tend to have the fewest injury prevention laws and northeastern states the most (Levi et al. 2012).

Table 3 presents the factor loadings for women. Factor 1 is labeled “substance abuse” because three of the five highest loadings are associated with substance abuse, including the clearest marker, accidental poisoning. The other two causes most heavily loaded on female Factor 1 are other digestive diseases and interpersonal violence, also plausibly linked to substance abuse. The first two factors for men and women are the same two factors, but reversed in order. Like male factor 2, female Factor 1 contains high loadings for liver cirrhosis, liver cancer, accidental poisoning, hepatitis, and other digestive diseases. Suicide and transport injuries are also plausibly associated with substance abuse and appear in the list of the causes most highly correlated with Factor 1. Cerebrovascular disease has the second highest loading, and is an important condition through which substance abuse operates (Kaku and Lowenstein 1990). But there are a total of 15 causes of death loaded at 0.60 or better with female Factor 1, including causes closely associated with smoking (COPD, respiratory diseases) and obesity (diabetes mellitus). As a result, this factor is not as coherently identified with substance abuse as is Factor 2 for males. Nevertheless, female Factor 1 scores and male Factor 2 scores are correlated at a modest 0.60, indicating that they are tapping into somewhat the same sources.

Figure 3 presents a map of factor scores for Factor 1 for women. As in the case of substance abuse for men (Figure 2), scores for women are lowest in the upper Midwest and high in the Southwest, especially in New Mexico and Nevada. However, high scores are also found in Appalachia for women, whereas male scores in Appalachia were closer to zero.

Factor 2 for women is labeled “Smoking/Obesity”, because out of the five causes of death with loadings over 0.60, two are strongly associated with smoking (lung cancer and ischaemic heart disease) and three are strongly associated with obesity (colorectal cancer and breast cancer, as well as ischaemic heart disease). In contrast to male Factor 1, neither COPD nor oral cancers makes the list of most highly correlated causes, thus failing to support a primary smoking interpretation for the factor. Males have historically and currently smoked more heavily than females, while obesity levels are similar for the two sexes (National Center for Health Statistics 2012a). So it may not be surprising that smoking is more strongly represented in male Factor 1 than in female Factor 2. In further support of the interpretation of female Factor 2 as a “Smoking/Obesity” factor, female Factor 2 scores and male Factor 1 scores are highly correlated at 0.82, suggesting that the two factors are responding to the same underlying sources of variation. Figure 3 also shows that the factors scores for this factor are highest in a band that runs from West Virginia through Kentucky, Arkansas, and Oklahoma, with high scores in the Midwest as well. The lowest factor scores can be found in the West, especially the Four Corners. This map is quite similar to that of male Factor 1, as implied by the high correlation between the factors. Alcohol poisoning and alcoholic liver cirrhosis are two negatively loaded causes on this factor (see Appendix Table S3), implying a negative spatial association between alcohol abuse and smoking/obesity.

Factor 3 for females does not emerge as a coherent set of causes of death since only one cause, HIV/AIDS, has a factor loading above 0.60. This result is consistent with the fact that female Factor 3 explains less variance (8.4%) than male Factor 3 (13.4%). However, as shown below, Factor 3 for females has an important rural/urban dimension, as it does for males. The death rate from HIV/AIDS can be interpreted as a (negative) indicator of rurality. In fact, death rates from injury, which dominate Factor 3 for males, are negatively correlated with Factor 3 for females: falls (−0.416), transport injuries (−0.411), other accidents (−0.309) and self-inflicted wounds (−0.276). Male and female factor scores for Factor 3 have a fairly strong inverse relationship (r= −0.60). The highest scores on Factor 3 for females is found in states with the largest urban areas: New York, New Jersey, California, Texas, Illinois, and Florida. The lowest factor scores are concentrated in the Northwest.

4.1 Relation of factors to all-cause mortality

The factors have been identified through patterns of correlation between causes of death. These correlations take no account of the magnitude of a particular death rate, so there is no necessary relationship between a factor and the proportion of all-cause mortality that it accounts for. Nevertheless, the factors do account for a high degree of variation in all-cause mortality. The three male factors correlate with the age-standardized death rate from all causes, at 0.72, 0.45, and 0.47, respectively. The three female factors correlate with all-cause mortality, at 0.68, 0.65, and 0.20, respectively. In an Ordinary Least Squares regression, the three factors explain 94% of the variance in all-cause mortality for men and 93% for women.

So far we have used only age-standardized rates in the analysis. It would be reassuring if the factors identified had a plausible age pattern of correlation with all-cause mortality, one that reflected the age patterns of mortality by cause. Figure 4 for males and 5 for females show that this expectation is realized. The correlation between all-cause age-specific death rates and the smoking/obesity factor grows stronger with age, its correlation with the rural/urban factor grows weaker with age, and its correlation with the substance abuse factor is an inverted U-shape. These age patterns mimic the age patterns of mortality for the main causes of death associated with each of the factors (National Center for Health Statistics 2012a).

4.2 Validation

Table 4 presents correlations between factor scores and various measures related to the factors that we have identified. Most of these measures are drawn from the Behavioral Risk Factor Surveillance System (BRFSS), an annual survey conducted by the National Center for Health Statistics that provides data on risk factor distributions by state. Despite the flaws in BRFSS data as described in the introduction, it is the best source of information on how risky behaviors vary from state to state. For most measures we have been able to tabulate micro-level data in order to examine the distributions for the white population aged 20-64 by sex.

The correlations in Table 4 provide solid support for the interpretations that we have given to the factors. Male Factor 1 (Smoking/Obesity) scores are correlated across states at 0.75 or above with measures of the prevalence of smoking and of obesity. These are among the highest correlations in Table 4. The highest correlation involving Male Factor 2 (Substance Abuse) is with self-reported nonmedical prescription drug use (0.53). Male Factor 3 is most highly correlated with indicators of rural residence: miles driven (0.77), gun ownership (0.87), and living outside a Metropolitan Statistical Area (0.63). The correlation between gun ownership and the male injury factor is the highest of any of the 72 correlations in Table 4. Gun ownership is also highly correlated at −0.79 with female Factor 3.

Consistent with the interpretation of female Factor 1 as one of substance abuse, the factor has moderately high loadings with painkiller abuse (0.63) and illicit drug abuse (0.53). However, the highest correlations for female Factor 1 are with the two measures of health care access. Health care access may play a strong role in the spatial patterning of this factor, which would be consistent with the broad range of causes of death highly loaded on Factor 1 for females. On the other hand, the ‘amenable causes’ that are thought to be most sensitive to the quality of medical services are not prominent in female Factor 1. Female Factor 2 (Smoking/Obesity) is most highly correlated with the prevalence of smoking (0.56), obesity I (0.38), and obesity II (0.42). The correlations of other measures with female Factor 2 are very low. Female Factor 3 correlations are consistent with a rural/urban interpretation, as all three indicators of rural residence are negatively correlated: miles driven (−0.54), gun ownership (−0.79), and living outside a Metropolitan Service Area (MSA) (−0.72).

5. Discussion

The application of factor analysis to state-specific mortality by cause of death has revealed a set of factors with a substantial degree of epidemiologic coherence. Smoking/obesity, substance abuse, and rural/urban residence emerge as the most prominent sources of mortality variation among men. Similar but weaker patterns are observed in females. Confidence in the identification of these factors is increased by high correlations between the factors and survey-based estimates of exposure to various risks.

Of note, we found that three factors were able to explain the majority of US mortality variation among working-age adults in both men and women. The finding of a large contribution of behavioral risk factors to mortality variation is consistent with previous studies (Danaei et al. 2010; Fenelon and Preston 2012). One prior study estimated that targeting just four of the leading modifiable risk factors could increase overall life expectancy in the US by more than four years and reduce differences in life expectancy across population sub-groups by as much as 20% (Danaei et al. 2010).

The two leading behavioral risk factors in the United States, smoking and obesity, did not emerge as separate factors in the analysis. The identification of smoking and obesity as separate factors may have been prevented by the fact that these two behaviors are highly correlated at the population level. The correlation of smoking and obesity prevalence across states as measured in the 2005 BRFSS (variables in Table 4) was .63 for men and .52 for women. The fact that both behaviors have significant pathways operating through ischaemic heart disease and cerebrovascular disease further impedes their distinction. With respect to male Factor 1, however, the preeminence of lung cancer and the list of highly loaded causes of death clearly suggest a stronger role for smoking than for obesity.

Substance abuse, including alcohol and particularly drug-use, emerged as important sources of mortality variation for both men and women in our analysis. The role of drug-use in American mortality patterns is underappreciated, likely because it is not easily amenable to measurement. However, our findings are consistent with recent literature, which shows rising death rates from prescription medication abuse, particularly among white women (Centers for Disease Control and Prevention 2013b).

Our analysis also revealed an important role for rural/urban status in US mortality patterns. This role was particularly prominent among males, for whom rurality was associated with higher injury death rates, including transport injuries and suicide. Our findings suggest that gun ownership and increased exposure to driving may be two factors underlying the rural penalty for males; however, future work should seek to identify a more complete set of risk factors associated with rurality.

The set of causes associated with a particular factor has greater coherence for men than for women, and the factors explain more variance for men. This gender difference is likely a result of the fact that risk factors like smoking and drug abuse, as well as the constellation of risk factors that give rise to injury, are more prominent among men. The clearer results for males may also be caused by the fact that there are many more male deaths for statistical analysis than female deaths (2,308,783 deaths for the male analysis compared to 1,215,719 deaths for the female analysis).

Causes of death that we expect to be most sensitive to the availability and quality of health services did not emerge in a latent variable for either sex. Uncertainty is added about the role of health services because the percent uninsured and the percent who cannot afford health care in a state are highly correlated with female Factor 1. It may be that health services manifest themselves in mortality over a wider variety of causes of death than implied by previous efforts to identify the set of sentinel or amenable causes of death.

Latent variable analysis of causes of death is complementary to the more standard attributable-risk approach in which a set of relative risks associated with a particular risk factor is applied to the risk factor distribution of the population. The latent variable approach does not require prevalence or relative risk estimates and is not sensitive to errors therein. Population exposure data and relative risks are typically available for only a small subset of risk factors and data errors may be especially prominent for data on sensitive health topics commonly elicited through self-report (e.g., body mass index, unsafe sex, and illicit drug-use). The present approach uses US vital statistics, a rich source of data with temporal depth and fine geographic detail.

Our latent variable analysis of US cause-of-death data reveals that the majority of mortality variation in white working-age adults can be explained by three readily interpretable factors. These factors are smoking/obesity, substance abuse, and rural/urban status. Although the patterns were stronger in men, findings were generally consistent across the sexes. Future analyses using the present approach should include extensions to older ages, to other racial/ethnic groups, to time series data, and to finer geographic detail (e.g., US counties). The present approach may also prove valuable for identifying the contribution of modifiable risk factors to mortality patterns in other countries that have complete or near-complete vital registration and in which data to implement the standard attributable risk approach are lacking. More than 60 countries around the world met this criterion as of 2003 (Mathers et al. 2005).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

6. Acknowledgements

This research received support from the Demography of Aging grant (NIH T32 AG000177 & P30 AG012836) from the National Institute of Health's (NIH)'s National Institute of Aging awarded to the Population Studies Center at the University of Pennsylvania. This project was also supported by Grant Number R01AG040212 from the National Institute on Aging.

References

Anderson RN, Rosenberg HM. Age standardization of death rates: implementation of the year 2000 standard. National Vital Statistics Report. 1998; 47(3):1–16. http://www.cdc.gov/nchs/data/nvsr/nvsr47/nvs47_03.pdf.

Brown, TA. Confirmatory factor analysis for applied research. Guilford Press; 2006.

Centers for Disease Control and Prevention, (CDC). Behavioral Risk Factor Surveillance System Survey Data. Department of Health and Human Services, Centers for Disease Control and Prevention; Atlanta, GA: 2001. http://www.cdc.gov/brfss/annual_data/annual_2001.htm

Centers for Disease Control and Prevention, (CDC). Behavioral Risk Factor Surveillance System Survey Data. Department of Health and Human Services, Centers for Disease Control and Prevention; Atlanta, GA: 2005. http://www.cdc.gov/brfss/annual_data/annual_2005.htm

Centers for Disease Control and Prevention, (CDC). Behavioral Risk Factor Surveillance System: 2011 Summary Data Quality Report. Department of Health and Human Services, Centers for Disease Control and Prevention; Atlanta, GA: 2013a.

Centers for Disease Control and Prevention, (CDC). Prescription Painkiller Overdoses: a growing epidemic, especially among women. Department of Health and Human Services, Centers for Disease Control and Prevention; Atlanta, GA: 2013b. http://www.cdc.gov/vitalsigns/PrescriptionPainkillerOverdoses/

Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. Preventive Medicine. 2004; 38(5):613–619. doi:10.1016/j.ypmed.2003.11.027. [PubMed: 15066364]

Danaei G, Ding EL, Mozaffarian D, Taylor B, Rehm J, Murray CJL, Ezzati M. The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Medicine. 2009; 6(4) e1000058. doi:10.1371/journal.pmed.1000058.

Danaei G, Rimm EB, Oza S, Kulkarni SC, Murray CJL, Ezzati M. Correction: The promise of prevention: the effects of four preventable risk factors on national life expectancy and life expectancy disparities by race and county in the United States. PLoS Medicine. 2010; 7(3):1–12. doi:10.1371/journal.pmed.1000248.

Ezzati M, Henley SJ, Lopez AD, Thun MJ. Role of smoking in global and regional cancer epidemiology: current patterns and data needs. International Journal of Cancer. Journal International du Cancer. 2005a; 116(6):963–971. doi:10.1002/ijc.21100. [PubMed: 15880414]

Ezzati M, Henley SJ, Thun MJ, Lopez AD. Role of smoking in global and regional cardiovascular mortality. Circulation. 2005b; 112(4):489–497. doi:10.1161/CIRCULATIONAHA.104.521708. [PubMed: 16027251]

Federal Highway, Adminstration. Highway Statistics. U.S. Department of Transportation, Federal Highway Administration; 2001. http://www.fhwa.dot.gov/ohim/ohimstat.htm

Fenelon A, Preston SH. Estimating smoking-attributable mortality in the United States. Demography. 2012; 49(3):797–818. doi:10.1007/s13524-012-0108-x. [PubMed: 22610474]

Gavrilova, NS., Semyonova, VG., Evdokushkina, GN. Mortality Crisis in Russia: New Health Threats. Center on Aging, University of Chicago; 2002. http://health-studies.org/PAA-2002-Russia.pdf

Howden, LM., Meyer, JA. Age and sex composition: 2010. 2010. 2010 Census Briefs, US Department of Commerce, Economics and Statistics Administration. US CENSUS BUREAU

Hughes, A., Sathe, N., Spagnola, K. State estimates of substance use from the 2006-2007 National Surveys on Drug Use and Health. Office of Applied Studies, Substance Abuse and Mental Health Services Administration,NSDUH Series H-35, HHS Publication No. SMA 09-4362; Rockville, MD: 2009. http://www.samhsa.gov/data/2k7state/2k7state.pdf

Kaku DA, Lowenstein DH. Emergence of recreational drug abuse as a major risk factor for stroke in young adults. Annals of Internal Medicine. 1990; 113(11):821–827. doi: 10.7326/0003-4819-113-11-821. [PubMed: 2240897]

Levi, J., Segal, LM., Kohn, D. The Facts Hurt: A State-By-State Injury Prevention Policy Report. Trust for America's Health; Washington, D.C.: 2012. http://www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf72972

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2013; 380(9859):2224–2260. doi:10.1016/S0140-6736(12)61766-8.

Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bulletin of the World Health Organization. 2005; 83(3):171–177. http://www.who.int/bulletin/volumes/83/3/171.pdf. [PubMed: 15798840]

McCarthy JF, Blow FC, Ignacio RV, Ilgen MA, Austin KL, Valenstein M. Suicide among patients in the Veterans Affairs health system: rural-urban differences in rates, risks, and methods. American Journal of Public Health. 2012; 102(S1):S111–S117. doi:10.2105/AJPH.2011.300463. [PubMed: 22390583]

Mehta NK, Chang VW. Mortality attributable to obesity among middle-aged adults in the United States. Demography. 2009; 46(4):851–872. doi:10.1353/dem.0.0077. [PubMed: 20084832]

Mehta N, Preston S. Continued increases in the relative risk of death from smoking. American Journal of Public Health. 2012; 102(11):2181–2186. doi:10.2105/ AJPH.2011.300489. [PubMed: 23050582]

Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. JAMA: the Journal of the American Medical Association. 2004; 291(10):1238–1245. doi:10.1001/ jama.291.10.1238. [PubMed: 15010446]

Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Correction?: Actual Causes of Death in the United States, 2000. JAMA: the Journal of the American Medical Association. 2005; 293(3):293–294. doi:10.1001/jama.293.3.293. [PubMed: 15657315]

Murray CJL, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandiorio TJ, Ezzati M. Eight Americas: Investigating Mortality Disparities across Races, Counties, and Race-Counties in the United States. PLoS Medicine. 2006; 3(9):1513–1524. doi:10.1353/tam.2006.0127.

Murray CJ, Abraham J, Ali MK, Alvarado M, Atkinson C, Baddour LM, Gutierrez HR. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. JAMA. 2013; 310(6):591–606. doi:10.1001/jama.2013.13805. [PubMed: 23842577]

Myers SR, Branas CC, French BC, Nance ML, Kallan MJ, Wiebe DJ, Carr BG. Safety in Numbers: Are Major Cities the Safest Places in the United States? Annals of Emergency Medicine. 2013 doi: 10.1016/j.annemergmed. 2013.05.030.

Naghavi M, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. Population Health Metrics. 2010; 8(9):1–14. [PubMed: 20181218]

National Center, for Health Statistics. Mortality data, multiple cause-of-death public-use data files. National Center for Health Statistics; Hyattsville, MD: 2010. http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm#Mortality_Multiple

National Center, for Health Statistics. Health, United States, 2012: With Special Feature on Emergency Care. Hyattsville, MD: 2012a. http://www.cdc.gov/nchs/data/hus/hus12.pdf

National Center, for Health Statistics. Bridged-race intercensal estimates of the resident population of the United States for July 1, 2000-July 1, 2009, by year, county, single-year of age (0, 1, 2, .., 85 years and over), bridged race, hispanic origin, and sex. 2012b. http://www.cdc.gov/nchs/nvss/bridged_race.htm

Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. Health affairs. 2008; 27(1):58–71. doi:10.1377/hlthaff.27.1.58. [PubMed: 18180480]

Nolte, E., Mckee, M. Does health care save lives? Avoidable mortality revisited. The Nuttfield Trust; 2004.

Oza S, Thun MJ, Henley SJ, Lopez AD, Ezzati M. How many deaths are attributable to smoking in the United States? Comparison of methods for estimating smoking-attributable mortality when smoking prevalence changes. Preventive Medicine. 2011; 52:428–433. doi:10.1016/j.ypmed. 2011.04.007. [PubMed: 21530575]

Peto R, Lopez AD, Boreham J, Thun M, Heath C. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. The Lancet. 1992; 339(8804):1268–1278. doi: 10.1016/0140-6736(92)91600-D.

Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. Lancet. 2012; 381(9861):133–141. doi:10.1016/S0140-6736(12) 61720-6. [PubMed: 23107252]

Polednak AP. Black-white differences in sentinel causes of death: counties in large metropolitan areas. Journal of Urban Health Bulletin of the New York Academy of Medicine. 2000; 77(3):501–507. [PubMed: 10976621]

Preston, S., Glei, D., Wilmoth, J. Contribution of Smoking to International Differences in Life Expectancy. In: Crimmins, E.Preston, SH., Cohen, B., editors. International Differences in Mortality at Older Ages: Dimensions and Sources. National Academies press; Washington, D.C.: 2011. p. 105-131.

Preston SH, Glei DA, Wilmoth JR. A new method for estimating smoking-attributable mortality in high-income countries. International Journal of Epidemiology. 2010; 39(2):430–438. doi: 10.1093/ije/dyp360. [PubMed: 20032265]

Preston, S., Heuveline, P., Guillot, M. Demography: Measuring and Modeling Population Processes. Blackwell Publishing; London: 2001.

Rehm J. The risks associated with alcohol use and alcoholism. Alcohol Research and Health. 2011; 34(2):135–143. http://www.researchgate.net/publication/221828484_The_risks_associated_with_alcohol_use_and_alcoholism. [PubMed: 22330211]

Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet. 2008; 371:569–578. doi:10.1016/S0140-6736(08)60269-X. [PubMed: 18280327]

Singh GK, Siahpush M. Increasing Rural – Urban Gradients in US Suicide. American Journal of Public Health. 2002; 92(7):1161–1167. doi:10.2105/AJPH.92.7.1161. [PubMed: 12084702]

Single E, Robson L, Rehm J, Xie X, Xi X. Morbidity and mortality attributable to alcohol, tobacco, and illicit drug use in Canada. American Journal of Public Health. 1999; 89(3):385–390. doi:10.2105/AJPH.89.3.385. [PubMed: 10076491]

Taylor B, Irving HM, Kanteres F, Room R, Borges G, Cherpitel C, et al. The more you drink, the harder you fall: a systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. Drug and Alcohol Dependence. 2010; 110(1):108–116. doi:10.1016/j.drug alcdep.2010.02.011. [PubMed: 20236774]

United States, Fire Administration. 2013. www.usfa.fema.gov/campaigns/smoking

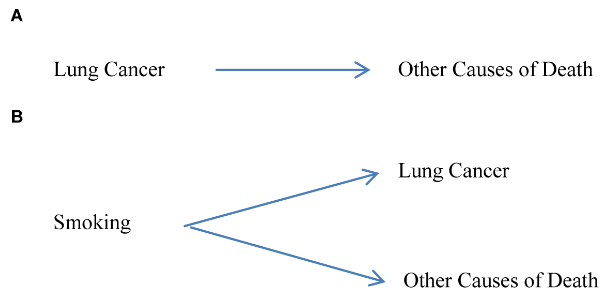
United, Nations. Model Life Tables for Developing Countries. United Nations Publication; New York: 1982. Sales No. E.81.XIII.7

United States, Census Bureau. 2011 American Community Survey. 2013. Tabulations supplied by Alemayehu Bishaw of the U.S. Census Bureau

Van Den Berg C, Smit C, Van Brussel G, Coutinho R, Prins M. Full participation in harm reduction programmes is associated with decreased risk for human immunodeficiency virus and hepatitis C virus: evidence from the Amsterdam Cohort Studies among drug users. Addiction (Abingdon, England). 2007; 102(9):1454–1462. doi:10.1111/j.1360-0443.2007.01912.x.

Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. Lancet. 2009; 373(9669):1083–1096. doi:10.1016/S0140-6736(09)60318-4. [PubMed: 19299006]

Wijetunga M, Seto T, Lindsay J, Schatz I. Crystal methamphetamine-associated cardiomyopathy: tip of the iceberg? Clinical Toxicology. 2003; 41(7):981–986. doi:10.1081/CLT-120026521. [PubMed: 14705845]



Notes: In Preston et al. 2011, lung cancer was treated as a marker of smoking and used to predict the effects of smoking on other causes of death. The causal model inherent to this approach is portrayed in (A). This model is incorrect because lung cancer and other causes of death are not causally related. In the present analysis smoking is treated as a latent variable and its effects are inferred from the correlation structure among causes of death with which smoking is associated. The causal model inherent in this approach is portrayed in (B). This model correctly characterizes the relationship between the variables.

Figure 1.

Directed acyclic graphs portraying the (A) incorrect causal model; (B) correct causal model

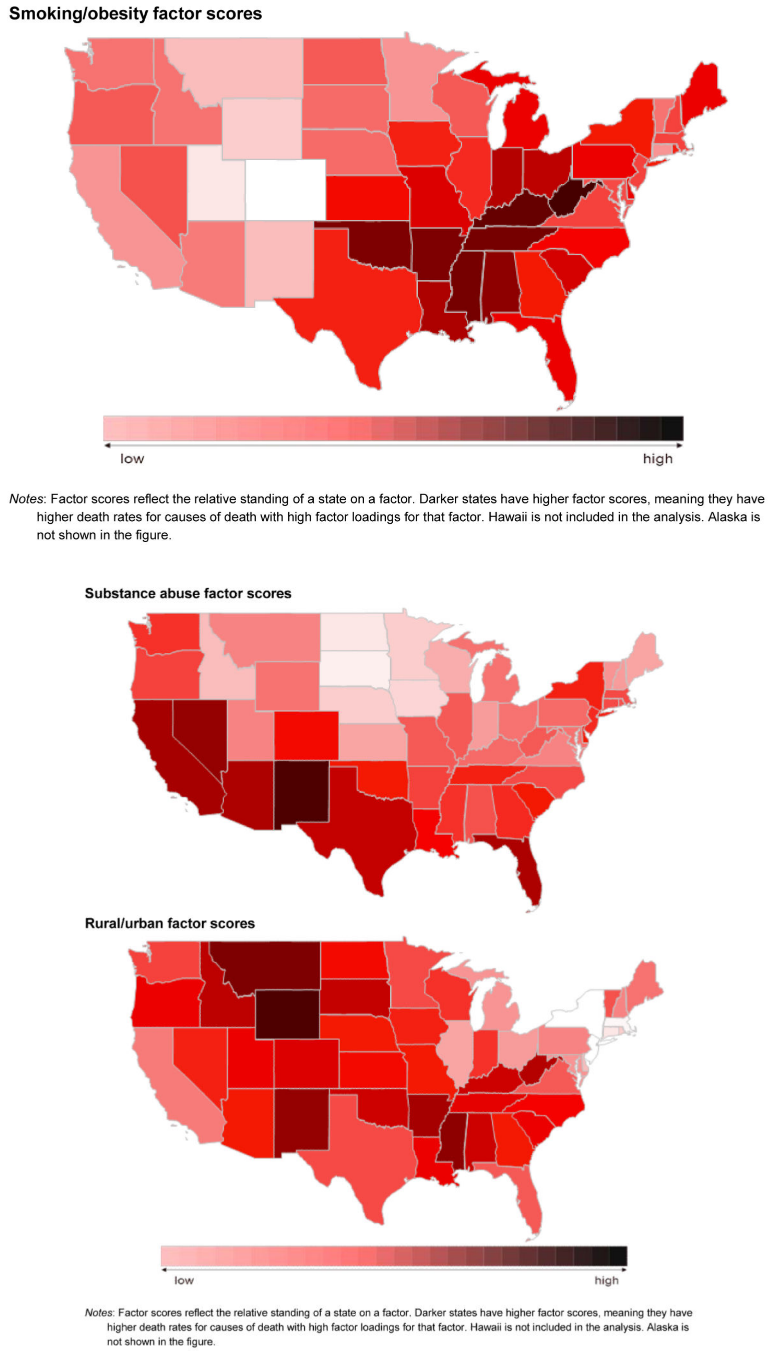
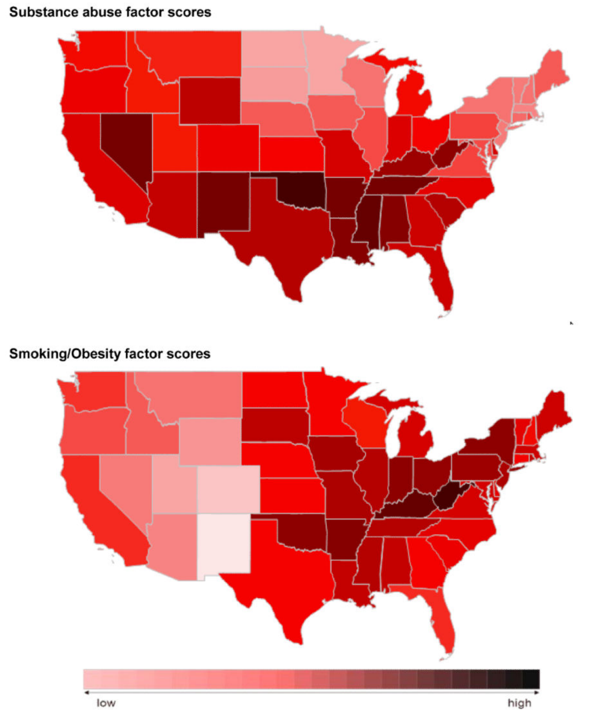
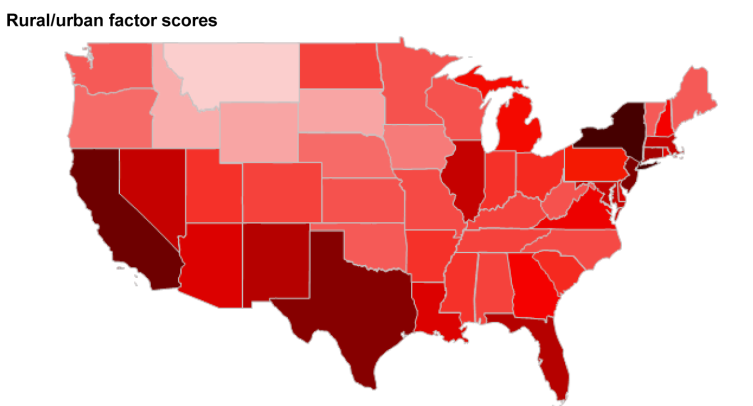


Figure 2.
Male factor score maps



Notes: Factor scores reflect the relative standing of a state on a factor. Darker states have higher factor scores, meaning they have higher death rates for causes of death with high factor loadings for that factor. Hawaii is not included in the analysis. Alaska is not shown in the figure.



Notes: Factor scores reflect the relative standing of a state on a factor. Darker states have higher factor scores, meaning they have higher death rates for causes of death with high factor loadings for that factor. Hawaii is not included in the analysis. Alaska is not shown in the figure.

Figure 3.
Female factor score maps

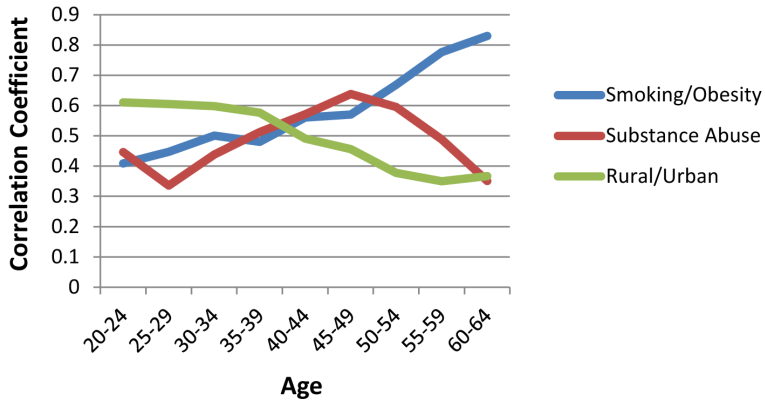


Figure 4.

Correlations between factor scores and age-specific death rates from all causes combined, for men

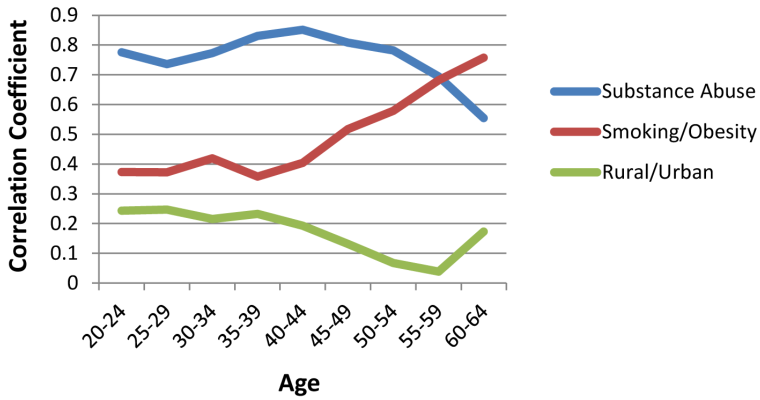


Figure 5.
Correlations between factor scores and age-specific death rates from all causes combined, for women

Table 1

Risk factors and associated causes of death

Smoking	Obesity	Alcohol Use	Drug Use	Amenable Causes
Aneurysm	Breast cancer	Alcohol poisoning	Accidental poisonings	Breast cancer
Bladder cancer	Cerebrovascular disease	Breast cancer	Cardiomyopathy	Cervical cancer
Cerebrovascular disease COPD	Colorectal cancer Corpus Uteri cancer	Esophageal cancer Interpersonal violence	Cerebrovascular disease Hepatitis	Colorectal cancer Diabetes mellitus
Esophageal cancer	Diabetes mellitus	Larynx cancer	HIV/AIDS	Respiratory diseases
Ischemic heart disease	Esophageal cancer	Liver cancer	Hypertensive Heart Disease	Skin cancer
Larynx cancer	Ischemic heart disease	Liver cirrhosis	Liver cirrhosis	
Lung Cancer	Kidney cancer	Oral cancer	Liver Cancer	
Oral Cancer	Kidney disease	Self-inflicted harm	Self-Inflicted harm	
Respiratory diseases	Liver cirrhosis	Transport injuries		

Abbreviations: COPD, Chronic Obstructive Pulmonary Disease.

Notes: Causes of death are compiled from the following sources. Smoking: Pirie et al. 2012, Ezzati et al. 2005a, Ezzati et al. 2005b, Oza et al. 2011; Obesity: Whitlock et al. 2009, Renehan et al. 2008. Alcohol Use: Lim et al. 2013, Corrao et al. 2004, Taylor et al. 2010. Drug Use: Single et al. 1999, Van Den Berg et al. 2007, Wijetunga et al. 2003. Amenable Causes: Nolte and McKee 2004, 2008, Polednak 2000.

Table 2

Male factor loadings

Smoking/Obesity		Substance Abuse		Rural/Urban	
Cause of Death	**Factor Loading**	**Cause of Death**	**Factor Loading**	**Cause of Death**	**Factor Loading**
Lung, Trachea, and Bronchus Cancer	0.947	Liver Cancer	0.844	Self-Inflicted Injuries	0.849
Colorectal Cancer	0.884	Interpersonal Violence	0.834	Transport Injures	0.794
Ischemic Heart Disease	0.880	Hepatitis	0.804	Exposure to Mechanical Forces	0.770
Oral Cancer	0.837	Non-Alcoholic Liver Cirrhosis	0.771	Other Accidents	0.753
Larynx Cancer	0.820	HIV/AIDS	0.758	Accidental Drowning	0.611
Cerebrovascular Disease	0.757	Accidental Poisoning	0.706		
Chronic Obstructive Pulmonary Disease	0.749	Other Digestive Diseases	0.670		
Exposure to Smoke or Heat	0.744	Hypertensive Heart Disease	0.643		
Pancreatic Cancer	0.729	Alcoholic Liver Cirrhosis	0.640		
Non-Hodgkins Lymphoma	0.717				
Leukemia	0.706				
Respiratory Disease	0.690				
Diabetes Mellitus	0.635				

Notes: Death registration data is taken from the NCHS and includes white men ages 20-64 and years 2000-2009. Factor analysis is run on a state by age-standardized cause of death matrix that generates three factors. Factor loadings represent the correlations between causes of death and these factors. All correlations over 0.6 ($p<.00001$) are reported in this table.

Table 3

Female factor loadings

Substance Abuse		Smoking/Obesity		Rural/Urban	
Cause of Death	**Factor Loading**	**Cause of Death**	**Factor Loading**	**Cause of Death**	**Factor Loading**
Interpersonal Violence	0.879	Colorectal Cancer	0.743	HIV/AIDS	0.728
Cerebrovascular Disease	0.876	Breast Cancer	0.740		
Other Digestive Diseases	0.869	Lung, Trachea, and Bronchus Cancer	0.725		
Non-Alcoholic Liver Cirrhosis	0.837	Non-Hodgkins Lymphoma	0.651		
Accidental Poisoning	0.778	Ischaemic Heart Disease	0.618		
Chronic Obstructive Pulmonary Disease	0.769				
Transport Injures	0.761				
Respiratory Diseases	0.751				
Cervical Uteri Cancer	0.749				
Ischaemic Heart Disease	0.726				
Self-inflicted Injuries	0.718				
Diabetes Mellitus	0.690				
Hepatitis	0.663				
Liver Cancer	0.654				
Other Accidents	0.613				

Notes: Death registration data is taken from the NCHS and includes white women ages 20-64 and years 2000-2009. Factor analysis is run on a state by age-standardized cause of death matrix that generates three factors. Factor loadings represent the correlations between causes of death and these factors. All correlations over 0.6 (p<.00001) are reported in this table.

Table 4

Factor score correlations with variables of interest

		Men			Women		
	Measure	**Smoking/ Obesity**	**Substance Abuse**	**Rural/ Urban**	**Substance Abuse**	**Smoking/ Obesity**	**Rural/ Urban**
Health Behaviors	Percent Smoke Everyday (2005)[a]	0.765	0.040	0.325	0.553	0.556	−0.188
	Percent Obese I (BMI>30)[b]	0.759	−0.291	0.220	0.502	0.380	−0.369
	Percent Obese II (BMI>35)[c]	0.752	−0.319	0.085	0.557	0.415	−0.308
	Percent Nonmedical Painkiller Use, Past Month[d]	0.239	0.445	0.269	0.63	−0.095	−0.016
	Percent Drug Use, Past Month (Not Marijuana)[e]	0.199	0.526	0.072	0.528	−0.052	0.171
Health Behaviors	Percent Binge Drinker Last Month[f]	−0.333	−0.290	−0.355	−0.610	0.075	−0.007
Health Care Access	Percent Uninsured[g]	0.148	0.430	0.495	0.765	−0.235	−0.117
	Percent Cannot Afford Health Care Last 6 Months[h]	0.378	0.349	0.547	0.841	−0.118	−0.079
Poverty	Percent in Poverty[i]	0.419	0.470	0.419	0.254	0.280	0.024
Ecological Exposures	Yearly Miles Driven Per Person[j]	0.077	−0.090	0.765	0.438	−0.181	−0.543
	Percent Gun Ownership[k]	0.208	−0.347	0.873	0.345	−0.010	−0.787
	Percent Lives Outside MSA[l]	0.101	−0.464	0.630	0.084	0.082	−0.720

[a]2005 Behavioral Risk Factor Surveillance System (BRFSS) for white men/women ages 20-65 (Centers for Disease Control and Prevention 2005).

[b]2005 Behavioral Risk Factor Surveillance System (BRFSS) for white men/women ages 20-65 (Centers for Disease Control and Prevention 2005).

[c]2005 Behavioral Risk Factor Surveillance System (BRFSS) for white men/women ages 20-65 (Centers for Disease Control and Prevention 2005).

[d]2006-2007 Substance Abuse and Mental Health Services Administration survey data (Hughes et al. 2009). These measures are for the entire United States population

[e]2006-2007 Substance Abuse and Mental Health Services Administration survey data (Hughes et al. 2009). These measures are for the entire United States population

[f]Binge drinking is defined as having 5 or more drinks on at least one occasion in the past 30 days.

[g]Binge drinking is defined as having 5 or more drinks on at least one occasion in the past 30 days.

[h]Binge drinking is defined as having 5 or more drinks on at least one occasion in the past 30 days.

[i]2011 American Community Survey for whites.

[j]Federal Highway Administration (2001) data for all Americans with drivers licenses.

[k]2001 BRFSS measure for all Americans. (Centers for Disease Control and Prevention 2001)

[l]Binge drinking is defined as having 5 or more drinks on at least one occasion in the past 30 days.