

Analyzing Cause of Death in Midlife in the United States*

Group 6: Sanghyun Cho[†], Jonghwa Do[‡], Hansol Woo[§]

University of Wisconsin-Madison

In the late 20th century, the U.S. saw declining mortality rates for midlife and older adults, but from 1999 to 2013, a troubling trend emerged. Middle-aged non-Hispanic whites experienced a notable increase in all-cause mortality due to suicides, drug addiction, and alcoholism. This phenomenon, known as "deaths of despair," disproportionately affected those with lower socioeconomic status and extended to uneducated whites, as highlighted by Case and Deaton's (2021) research. Our project explores how causes of death change over time and across demographic groups among midlife citizens and examines the extent of "deaths of despair." Our findings align with previous research, showing that cancer and heart disease remained the top two causes of death, but their overall proportion decreased over time due to the emergence of other diseases. Additionally, our research indicates that suicide, drug, and alcohol-induced causes of death are closely linked to the low-education group and non-white populations.

Analysis

1. Data

We analyze data from Midlife in the United States (MIDUS), administered by the Institute On Aging (IOA) at UW Madison, a national longitudinal survey of adults in the United States. The mortality data cover the deaths that occurred by the end of 2022. The analytic sample consists of MIDUS participants who died after MIDUS 1 and the cause of death information was available via NDI Plus (2004–2022). NDI Plus (National Death Index Plus) provides the cause of death information in addition to basic NDI information. Among the 7,108 MIDUS 1 survey respondents 2,459 were identified as deceased through a search of the National Death Index (NDI) files by 2022. Out of these, 2,158 had valid respondent's characteristics and cause of death information.

In addition, we utilize the leading causes of death data from National Center for Health Statistics (NCHS) between 1999 and 2017 to enhance the analysis by assessing the representatives of the MIDUS data through a comparison with another national dataset.

2. Expanding simplified cause of death category

The MIDUS dataset incorporates information on both the underlying and multiple causes of death classified by the International Classification of Diseases (ICD) from the World Health Organization (WHO). Additionally, the mortality data in MIDUS includes simplified categories of cause of death, aligning with the prevalent causes of death in the U.S. These categories currently consist of six pri-

* We thank professor Kim Ruhl and Ashley Swanson for setting profound foundation for the data handling, visualization and data analysis. All errors are on us. [†] corresponding author, scho95@wisc.edu [‡] jdo7@wisc.edu [§] hwoo28@wisc.edu

mary classifications, such as heart disease, cancer, Chronic Lower Respiratory Disease (CLRD), diabetes, COVID-19, and other diseases.

To delve deeper into the analysis of the trend of cause of death as well as deaths of despair, we recognize the need to expand these categories. Specifically, individuals are categorized under "death of despair" if their cause of death, as documented in the National Death Index (NDI), includes suicide, drug addiction, or alcoholism, including alcoholic liver diseases. Consequently, we established twenty-seven simplified categories of cause of death based on ICD codes, covering suicide, drug-induced, and alcohol-induced causes of death (Table 1). The selection of these twenty-seven categories is in line with the leading causes of death reported by the National Center for Health Statistics (NCHS) data for the year 2019. The top underlying causes of death include cancer (26.8%), heart disease (24.1%), CLRD (7.6%), and stroke (5.6%).

3. Time trend of causes of death: MIDUS and NCHS data

We conducted a time trend analysis of causes of death using two datasets, MIDUS and NCHS. Given the significantly smaller sample size of the MIDUS data compared to the NCHS data, we systematically examined both datasets to enrich our understanding of the nationwide trends in leading causes of death. The analysis focused on the underlying cause of death.

Figure 1 illustrate the temporal trends in the top 10 causes of death in the MIDUS and NCHS datasets. Cancer and heart disease consistently emerged as primary causes of death from the late 1990s to the 2020s in both datasets. However, their proportions gradually declined, offset by the increasing prominence of diseases like Alzheimer's, suicide, and COVID-19. In the MIDUS data, the percentages attributed to cancer and heart disease consistently decreased from 33.0% before 2010 to 23.2% after 2010 (Figure 2). Meanwhile, Alzheimer's and Parkinson's disease notably increased from 1.0% and 0.4% before 2010 to 4.7% and 1.7% after 2010 (Figure 2). A similar trend is observed in the NCHS data, with significant decrease in the proportion of cancer and heart disease (Figure 3) and notable increases in Alzheimer's and suicide from 1.9% and 1.2% in 1999 to 4.3% and 1.7% in 2017, respectively (Figure 3).

Upon comparing both datasets, we have identified differences attributable to the unique characteristics of each dataset. Firstly, the NCHS dataset spans the years 1999 to 2017, excluding data for COVID-19, which emerged as a primary cause of death in 2020 according to the MIDUS dataset. Additionally, the MIDUS data primarily represent the middle and older population, influencing the distribution of causes of death. In contrast, the NCHS dataset encompasses all age groups, including the younger generation. Taking cancer, heart disease, and alcohol-related causes into consideration, these tend to be relatively rare in the younger population. Consequently, the proportions of these causes are higher in the MIDUS dataset compared to the NCHS dataset due to the distinctive age composition of the MIDUS data.

4. Descriptive analysis of Death of Despair

Our previous analysis focused solely on the underlying causes of death. However, due to limited sample sizes, further analysis was challenging. To address this, we extended our investigation to include multiple causes of death, as provided in MIDUS. As a result, individuals were categorized into the "death

of despair” group if their cause of death in NDI Plus, in either the multiple causes of death or underlying cause of death measures, included suicide, drug addiction, or alcoholism, including alcoholic liver diseases.

Table 2 displays descriptive statistics for all sample in MIDUS and decedents who died from despair-related causes (suicide, drug addiction, alcoholism). Notably, the number of such decedents gradually increased from 1995 to 2022. We also conducted an analysis by the socioeconomic status, and distinctions were observed in the age, sex, education, income, race, and marital status of the decedents. Those with despair-related causes of death were more prevalent in the low-education, non-married, and male compared to the whole sample (Table 2).

We expanded our analysis to include a geographical perspective at the state level. Due to privacy regulations and the absence of geographic information in MIDUS, we turned to the NCHS, which provides both suicide cause of death and geographical data. Figure 4 visually represents the proportion of decedents whose cause of death is suicide across entire state-level populations. Alaska stands out with the highest suicide rate, and there is also a notably higher proportion of suicides in the Midwest. Specifically, Utah, Wyoming, and Colorado exhibit elevated suicide mortality rates. In contrast, New York, New Jersey, and the District of Columbia show lower rates of suicide.

5. Regression analysis of Death of Despair

To find an association between socioeconomic status and the risk of death of despair, we conducted a regression analysis. Table 3 provides a comprehensive overview of the results from Ordinary Least Square (OLS), Logit, and Probit models. They outline the predicted relative risks of death of despair based on race, education, marital status, age, sex, and income. The dependent variable is assigned a value of 1 if the decedent’s cause of death is related to the death of despair. The findings in Table 3 indicate statistically significant associations. Specifically, individuals with lower education levels and lower income exhibit a significantly greater risk of death of despair. In contrast, race and marital status did not reach statistical significance in predicting death of despair. These results align with previous research trends, reinforcing the notion that the middle-aged, non-college-educated demographic group is experiencing a concerning rise in mortality rates associated with the death of despair.

Conclusions and directions for future research

”Deaths of despair” describe a notable rise in mortality rates attributed to suicides, drug addiction, and alcoholism. This project is driven by a better understanding of the ’death of despair’ phenomenon. Our findings align with prior research, revealing a complex interplay of social and economic factors contributing to a pervasive sense of hopelessness and despair in specific populations. However, it’s essential to acknowledge the limitations of our study, such as the constraint of sample size and the use of simple logistic regression. Future research should consider expanding the sample size and employing advanced statistical techniques to yield more generalizable results and employ more rigorous methods. This will enhance the reliability and robustness of our insights into the intricate dynamics surrounding the ”deaths of despair” phenomenon.

Table 1: Expanding Simplified Categorized Causes of Death

Leading Cause of Death	ICD-9	ICD-10	Count(Share) : N(%)
Cancer	C00-C97	140-208	578(26.8)
Heart Disease	I00-I09,I11,I13,I20-I41, I42.0-I42.5,I42.7-I42.9, I44-I51	390-398,402,404,410-424, 425.0-425.4,425.6-425.9, 426-429	520(24.1)
Chronic Lower Respiratory Disease (CLRD)	J40-J47	490-494,496	165(7.6)
Cerebrovascular Disease (Stroke)	I60-I69	430-434,436-438	120(5.6)
Alzheimer's Disease	G30	331.0	72(3.3)
Diabetes	E10-E14	250	69(3.2)
Unintentional Injuries	V01-V99,W00-W99, X00-X39,X46-X59, Y85-Y86	E800-E848,E880-E928, E929.0,E929.1, E929.3-E929.5, E929.8,E929.9	70(3.2)
Influenza and Pneumonia	J09-J18	480-487	34(1.6)
Kidney Disease	N00-N07,N17-N19, N25-N28	580-589	28(1.3)
Septicemia (Blood Poisoning)	A40-A41	038	23(1.1)
Parkinson's Disease	G20-G21	332	26(1.2)

Table 1: Expanding Simplified Categorized Causes of Death (cont.)

Leading Cause of Death	ICD-9	ICD-10	Count(Share) : N(%)
	F11.0-F11.5,F11.7-F11.9, F12.0-F12.5,F12.7-F12.9, F13.0-F13.5,F13.7-F13.9, F14.0-F14.5,F14.7-F14.9, F15.0-F15.5,F15.7-F15.9, F16.0-F16.5,F16.7-F16.9, F17.0,F17.3-F17.5, F17.7-F17.9,F18.0-F18.5, F18.7-F18.9,F19.0-F19.5, F19.7-F19.9,X40-X44, X60-X64,X85,Y10-Y14, Y45,Y47,Y49	292,304,305.2-305.9, E850-E858,E935,E937, E939,E950.0-E950.5, E962.0,E980.0-E980.5	16(0.7)
Drug Induced Causes			
Hypertension	I10,I12,I15,I16	401,403	19(0.9)
Chronic Liver Disease	K73-K74	571.4-571.9	15(0.7)
COVID-19	U07.1	n/a	40(1.9)
Suicide	U03,X60-X84,Y87.0	E950-E959	12(0.6)
Alcohol Induced Causes	E24.4,F10,G31.2,G62.1, G72.1,I42.6,K29.2,K70, K85.2,K86.0,R78.0,X45, X65,Y15	291,303,305.0,357.5, 425.5,535.3,571.0-571.3, 790.3,E860	9(0.4)
Aortic Aneurysm	I71	441	8(0.4)
Assault (Homicide)	U01-U02,X85-X99, Y00-Y09,Y87.1	E960-E969	5(0.2)
Cholelithiasis and other Gallbladder Disorder	K80-K82	574-575	5(0.2)

Table 1: Expanding Simplified Categorized Causes of Death (cont.)

Leading Cause of Death	ICD-9	ICD-10	Count(Share) : N(%)
Human Immunodeficiency Virus (HIV) Disease	B20-B24	042-044	5(0.2)
Complications of Medical Care	Y40-Y84, Y88	E870-E879, E930-E949	5(0.2)
Congenital Malformations	Q00-Q99	740-759	1(0.0)
Legal Intervention	Y35, Y89.0	E970-E978	0(0)
Tuberculosis	A16-A19	010-018	0(0)
Pregnancy, Childbirth and Puerperium	O00-O99	630-676	0(0)
Other Disease			313(14.5)

Figure 1: Change in Leading Causes of Death over Time

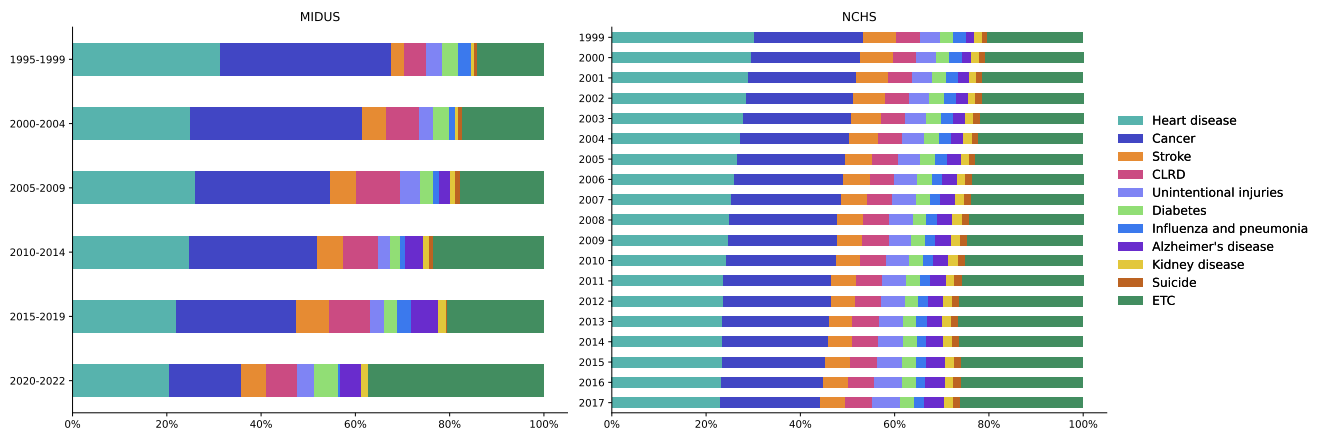


Figure 2: Changes in Leading Causes of Death: MIDUS

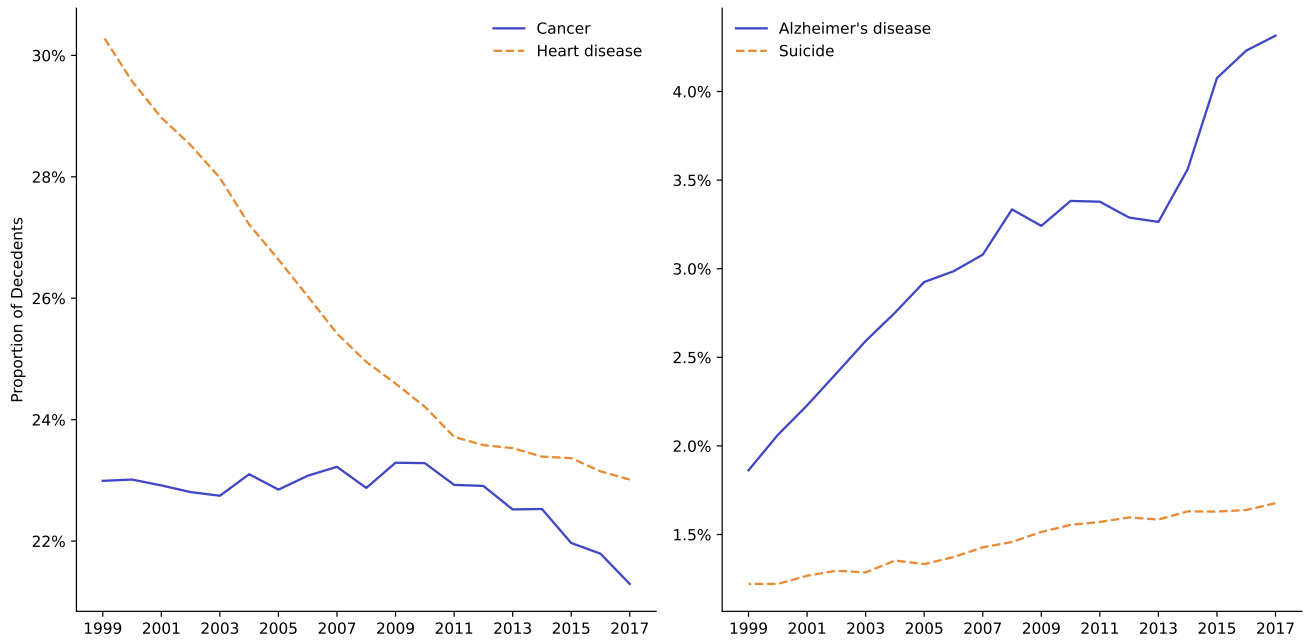


Figure 3: Changes in Leading Causes of Death: NCHS

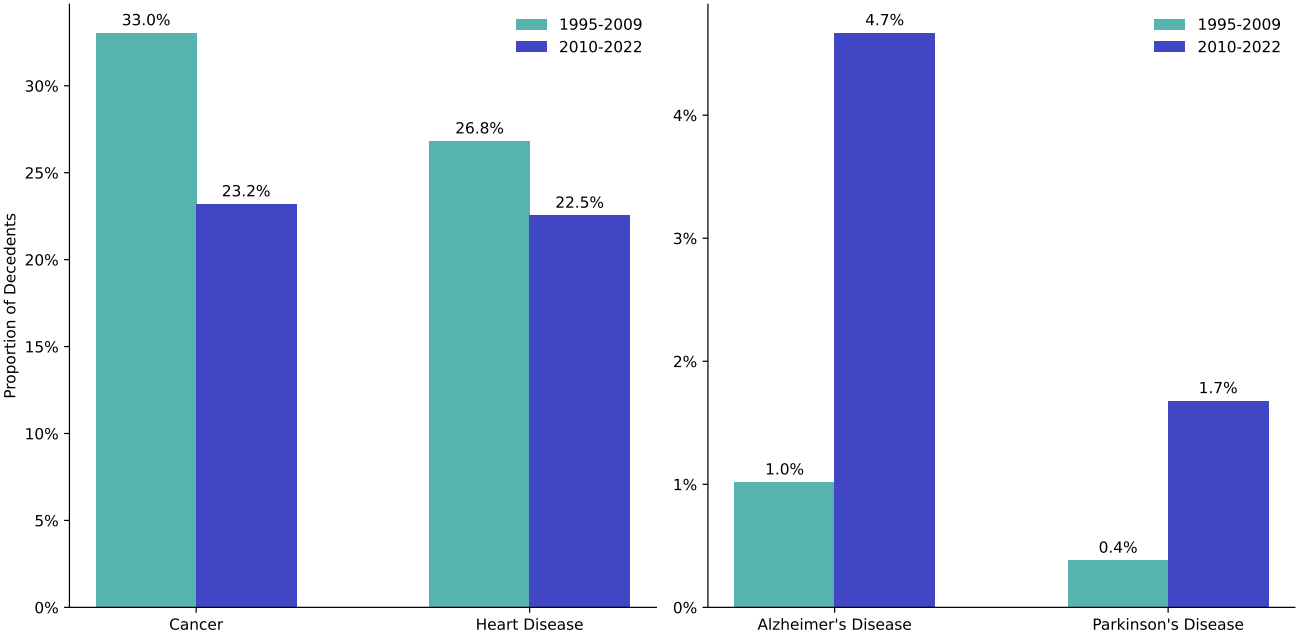


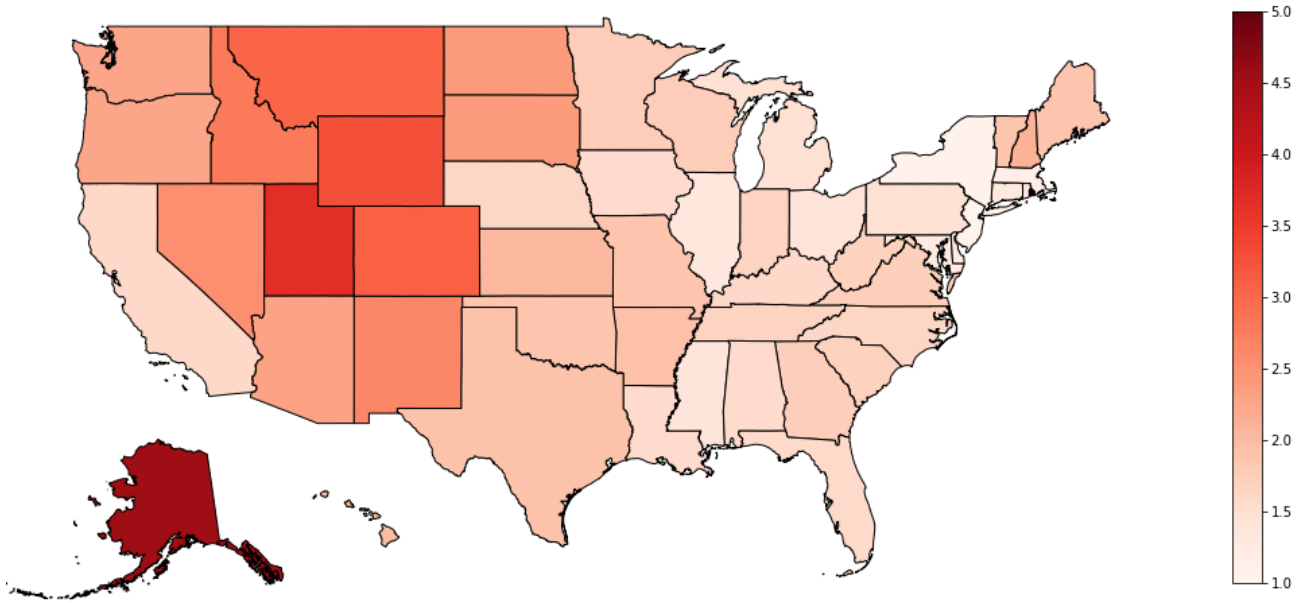
Table 2: Descriptive Statistics

	Total (n = 1,859)	Death of Despair (n = 264)
White ¹	90.7%	75.8%
College degree or above ¹	29.9%	18.9%
Married ¹	65.2%	54.5%
Age ²	58.2	51.6
Male ¹	52.1%	59.8%
Income(\$) ²	58,489	64,797

¹ Each value indicates the portion of the total decedents that correspond to the indicator.

² The values indicate the mean of total decedents.

Figure 4: Proportion¹ of Death of Despair by States²



¹ Proportion of suicide decedents to all-cause decedents.

² All numbers are from NCHS data.

Table 3: Estimated Coefficients for OLS, Logit and Probit Models^{1,2,3}

	OLS	Logit	Probit
White	0.030 (0.026)	0.297 (0.259)	0.177 (0.139)
Education	-0.059*** (0.017)	-0.656*** (0.184)	-0.347*** (0.095)
Married	-0.041** (0.016)	-0.407** (0.162)	-0.222** (0.087)
Age	-0.005*** (0.001)	-0.046*** (0.007)	-0.026*** (0.004)
Male	0.019 (0.015)	0.178 (0.153)	0.099 (0.081)
Income	0.025* (0.014)	0.269** (0.128)	0.150** (0.070)
constant	0.407*** (0.047)	0.460 (0.420)	0.175 (0.232)

¹ We used the following measures to analyze death of despair. Education: no college degree(e.g., high school graduation, some college education without degree) vs. Bachelor's degree or more. Education level was assessed by the highest level of education completed at MIDUS 1. Several other variables that are found to be associated with the prevalence of specific causes of death in existing literature were included as covariates, including age (in years), gender, race (white vs. others), household income (in terms of \$100,000), and marital status (currently married vs. unmarried [divorced, widowed, never married]).

² ***, **, * indicate the significance level of 1%, 5%, and 10%, respectively.

³ Inside the parentheses are standard errors.