Trends in Rates of Suicide by Decedent Age over Time: United States Mortality Data, 1960-2007

S. Chen [a], S. Suzuki [a], A. Podawitz [b] and C.A. Claassen [b]

[a] University of North Texas Health Science Center, Department of Biostatistics
[b] University of North Texas Health Science Center, Department of Psychiatry
*Email: shande.chen@unthsc.edu*

Abstract: As calculated using health gap summary measures, suicide is predicted to rank as the fourteenth most burdensome global health condition by the year 2020, and suicide rates are still climbing rapidly in many countries of the world. Unfortunately, the unique and challenging methodological issues which confront clinical suicide epidemiologists complicate investigations which could potentially inform suicide prevention efforts. There is, for instance, a popular, albeit largely unexplored, theory that environmental and biological influences impact suicidal behavior differentially by age group. However, neither the validity of the theory itself nor the potential mediators of the age-suicide relationship are well explicated. The usual approach to age group analysis for suicide decedents is to use fixed intervals (e.g., five-year intervals) to define age groups. However, when the data are available for many years, the suicide rate can vary over time for a given age group, and displaying the rate over time for multiple age groups can be confusing. In this paper, we use a method of cluster analysis to derive internally consistent age groups which exhibit distinct changing patterns across time. The method is applied to the US suicide data between 1960 and 2007.

Keywords: Suicide, epidemiology, mortality, decedent age

1. INTRODUCTION

Global Burden of Disease health gap summary statistics suggest that suicide will be the fourteenth most burdensome global health condition by the year 2020, and suicide rates continue to climb rapidly in many countries of the world (Mathers et al., 2006; Brundtland, 1999). However, despite its huge associated burden, suicide is a relatively rare event, and the unique and challenging methodological issues associated with rare event analyses represent substantial barriers to many of the epidemiological investigations which could potentially inform suicide prevention efforts (Geddes, 1999; Moscicki, 1989). For one, capricious fluctuations are sometimes observed over time in the rates of occurrence associated with infrequent events, making it statistically difficult to measure the impact of rate mediators and moderators. Accurate characterization of infrequent and rare events requires data collection among very large, representative groups (Geddes, 1999). In the absence of rigorous sampling techniques, epidemiologically sound conclusions cannot be generated, leaving risk management strategies poorly informed and injury prevention strategies inadequately conceptualized (Geddes, 1999; Salib, 2003; Welch, 2001).

There is a popular, albeit largely unexplored, theory that environmental and biological influences impact suicidal behavior differentially by age group. In particular, neither the validity of the theory itself nor the potential mediators of age-suicide relationships are well explicated for United States (U.S.) data. The conventional perspective is that the relative contribution of various age groupings to the overall suicide rate is stable across time (Figure 1), with the highest crude rates per 100,000 U.S. population always occurring among the elderly (American Foundation for Suicide Prevention, 2011; Lester, 2003). This conclusion, however, may well ignore other important information available from a study of within-age group rate fluctuations across time (Gunnell et al., 1997).

Figure 1. Crude Rate of Suicide by Five-year Age Group, 1981-2007
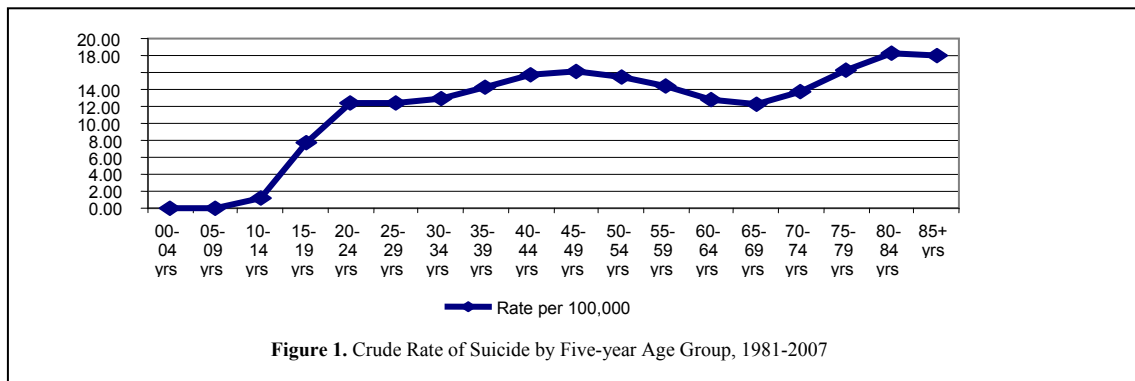
Moreover, the usual approach to age group analysis for suicide decedents is to use fixed intervals (e.g., five-year intervals) to define age groups. However, when the data are available for many years, the suicide rate can vary over time for a given age group, and displaying the rate over time for multiple age groups can be confusing (Figure 2).

Figure 2. Suicide rates by age in the United States from 1960 to 2007
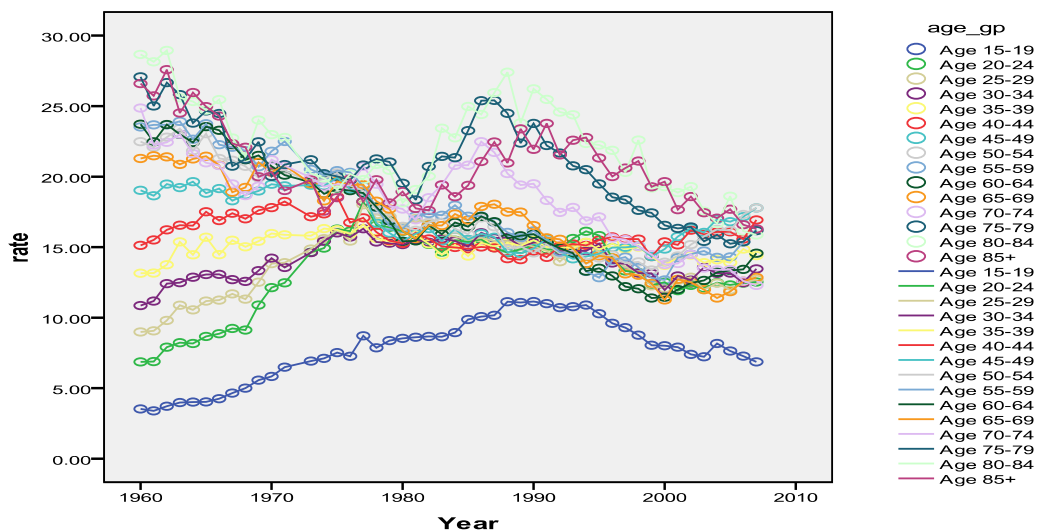
Cluster analysis is a common statistical technique that groups similar observations into clusters. Hierarchical clustering creates a tree diagram so one can determine a good criterion to select optimal clusters. Using cluster analysis, one can develop subgroups within data that would otherwise be confusing. This technique is useful in many fields, such as data mining, pattern recognition, image analysis and bioinformatics.

The purpose of this manuscript is to apply cluster analysis to US suicide data in order to differentiate age groups with different suicide rates and change patterns over time period. By doing so, we will be able to further study the environmental influences by age.

2. METHODS

2.1. Data

The data used in this IRB-approved study consist of annual population-based, crude suicide rates per 100,000 US citizens broken down into five year age groups for the years 1960-2007. Death data were obtained from annual volumes of Statistical Abstracts, a national mortality data source compiled by Centers for Disease Control's (CDC) National Center for Health Statistics (U.S. Census Bureau, 2011). This dataset contains information from death certificates filed in state vital-statistics offices and includes causes of death reported by attending physicians, medical examiners, and coroners. It also includes demographic information about decedents reported by funeral directors, who obtain that information from family members and other informants. The denominator population data is intercensal and postcensal census data were obtained from the Bureau of the Census (U.S. Census Bureau, 2011; CDC, 2004). These data are based on information gathered in censuses and on estimation procedures conducted in non-census years. All data to be used in these analyses are in the public domain.

2.2. Analysis

Age at suicide in this dataset is calculated directly from birth and death dates as found on death certificates. Graphical descriptions of suicide trends by five-year age group over time were developed for this project. We first used five-year intervals to create initial age groups. For each year, we computed the suicide rate for a given age group based on the number of suicides and the population of individuals within that age group. Next we used cluster analysis to merge groups with similar rates and changing pattern. Ward's minimum-variance method was applied for the clustering (Ward, 1963). We used an R^2 value of 0.95 as the cutting value and created 7 clusters. Because of the nature of the eventual analyses that will use these data (e.g., to identify moderators and mediators of rates across time), we allowed no gap in age within each cluser. Finally, we created seven age groups based on these seven clusters.

3. RESULTS

The correlational tree created during the cluster analysis is displayed in Figure 3. The seven age groups derived from this analysis included: 'Age 15-19', 'Age 20-34', 'Age 35-44', 'Age 45-49', 'Age50-69', Age 70-74' and 'Age 75+'. In other words, these age groups tended to be internally consistent in terms of their rate fluctuations across time.
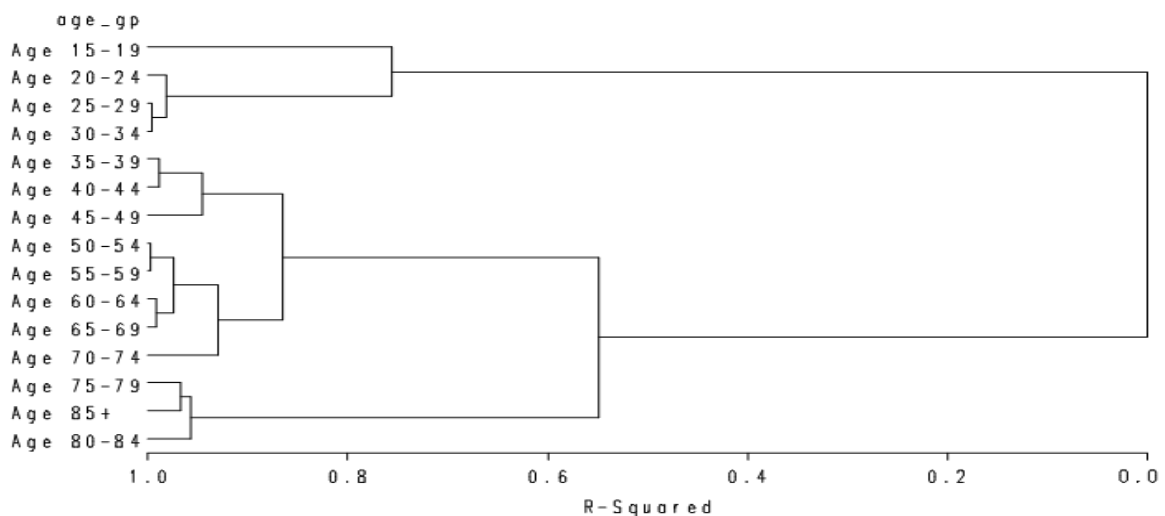

Figure 3. Tree diagram of clusters for age groups

The suicide rates over time for the seven final age groups are displayed together in Figure 4. This figure illustrates that advancing age is generally equated with higher risk, and that between-group differences in absolute rates of suicide have become smaller across time. The suicide rates for middle-aged individuals appear to be the most intractible across time.
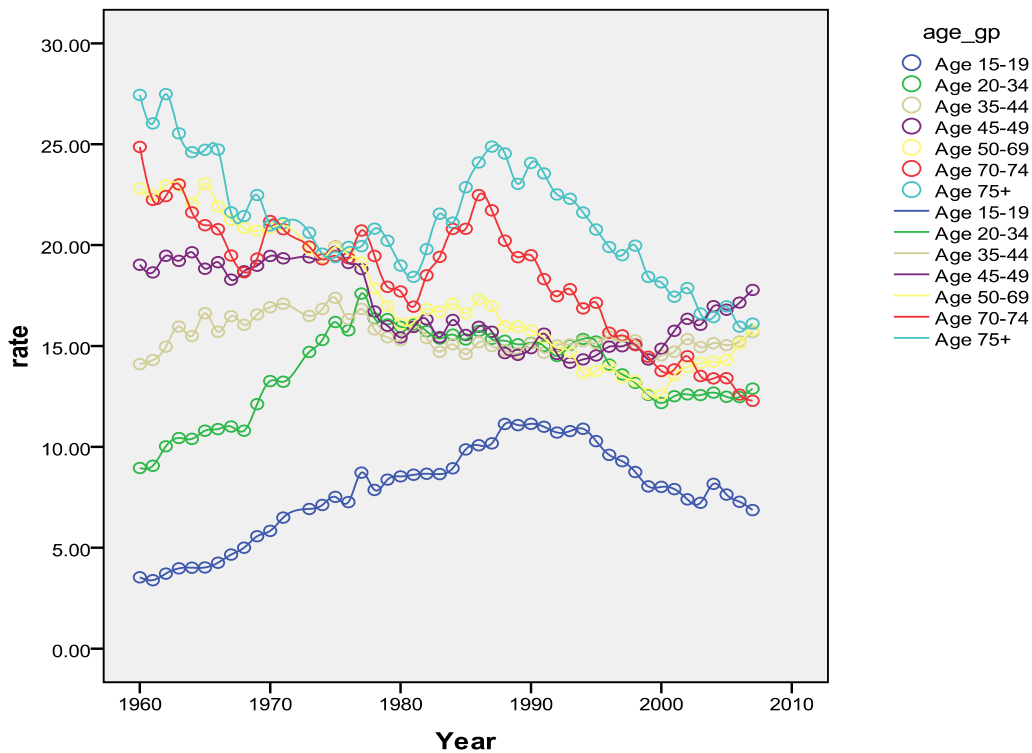

Figure 4. Suicide rates for the seven cluters

Distinct patterns of rate fluctuation over time were observed for each of these seven age clusters. By investigating the initial five-year interval age groups within each cluster, we can demonstrate the within-cluster similarity in longitudinal rate fluctuation across time. Such similarities are shown for clusters 2, 3, 5, and 7 in Figures 5 through 8 respectively. Note that, the other clusters are not shown because they only contain one five-year interval age group.
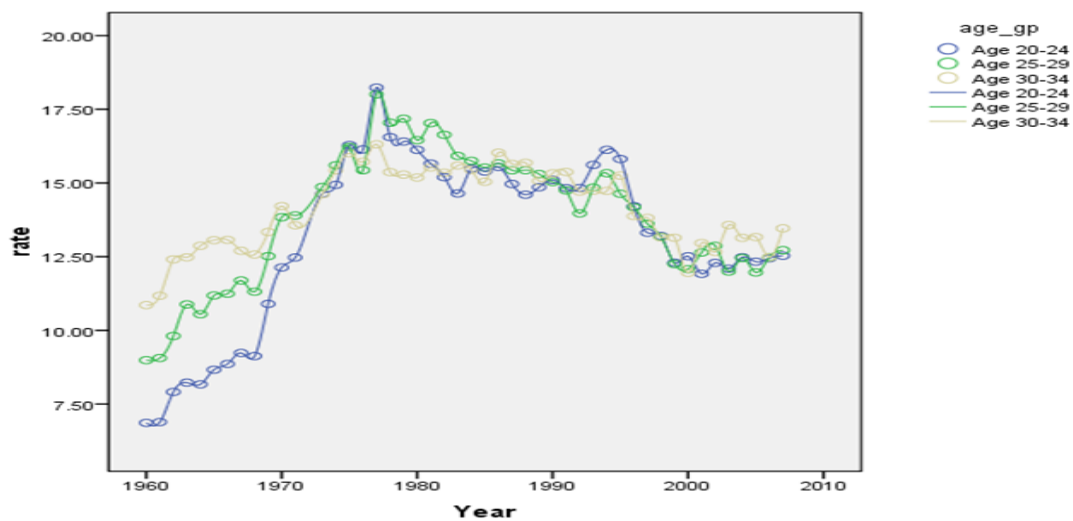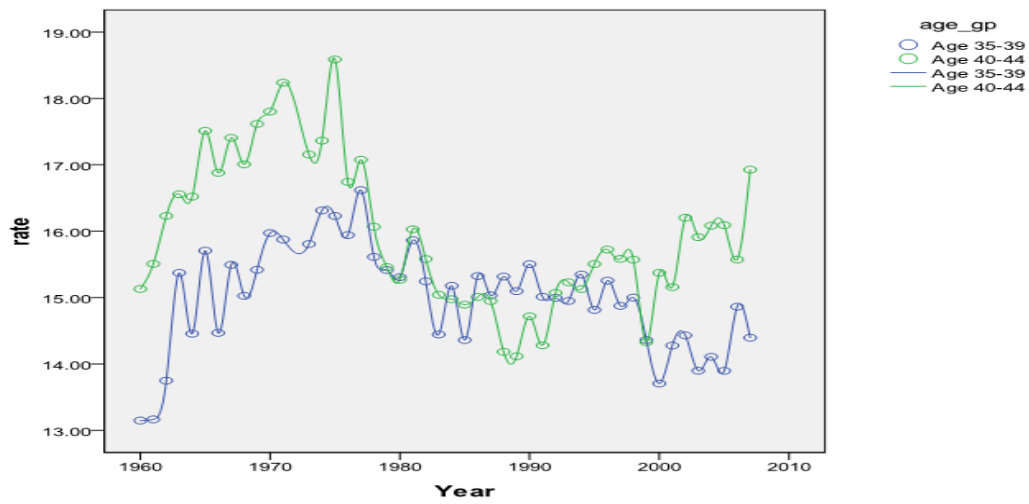

Figure 5. Suicide rates for cluster 2

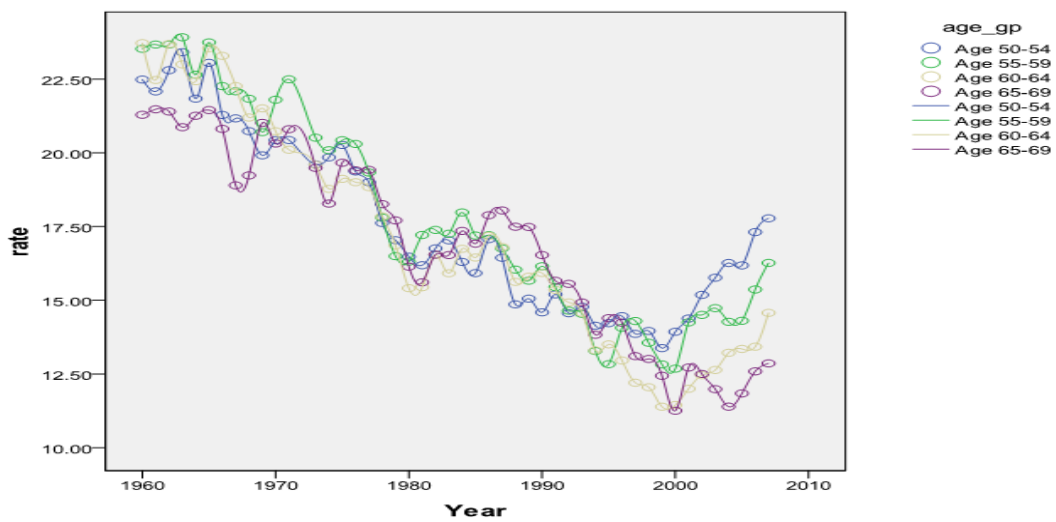**Figure 6.** Suicide rates for cluster 3

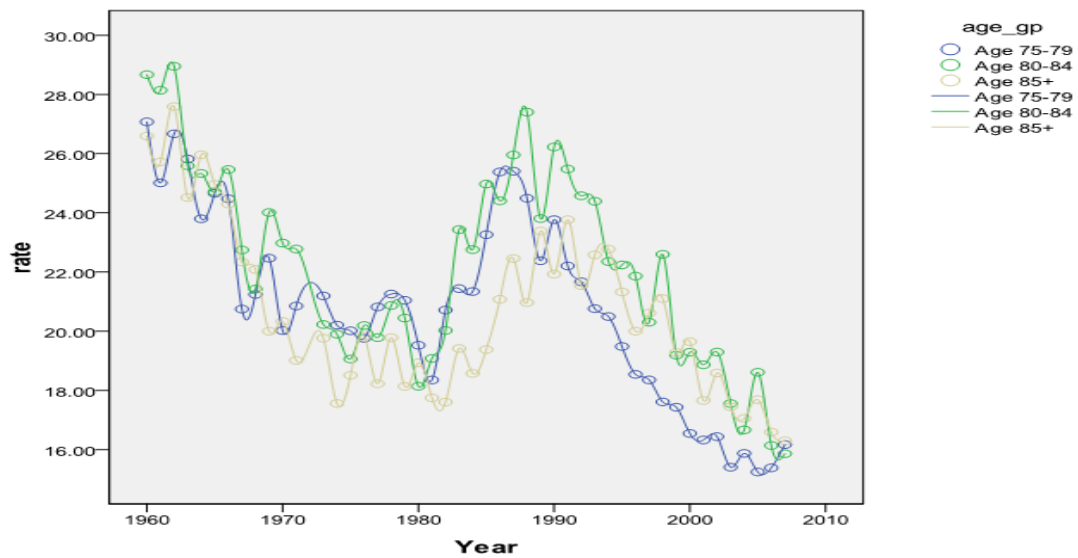**Figure 7.** Suicide rates for cluster 5

Figure 8. Suicide rates for cluster 7

4. DISCUSSION AND CONCLUSIONS

The overall crude US rate of suicide per 100,000 population has varied by less than 5 persons per 100,000 during the past sixty years (U.S. Census Bureau, 2011). However, this analysis demonstrates that the age-related suicide rate has varied to a much greater degree over time. Through cluster analysis, we obtained seven age groups that have different longitudinal trajectories in terms of both absolute annual rate and change over time. In general, the older are have a higher risk to commit suicide (Figure 10). The suicide rates were much sperate out among different ages in early 60's but the differences got smaller in recent years (2000 to 2007). Suicide rates peaked between late 70's and early 90's. The suicide rates for elderly people (70 or older) had a tendency to decrease, while the rates for middle age people (age 35- 49) the rates were similar over the time.

Limitations of this analysis include the following: suicide decedent age data is only available in five-year age groups before 1967 in the United States, which is why this grouping was used for analyses. A finer-grained analysis might have discovered that the patterns fit more exactly with single-year age groupings. Also, low actual numbers of suicide in the very old – as well as low numbers of elderly citizens the population may give the impression that these rates represent larger numbers of suicides than is actually the case. Nonetheless, the crude population rate does suggest the degree to which period effects within a culture can impact behaviors for a given age group.

REFERENCES

American Foundation for Suicide Prevention (2011). Risk Factors for Suicide, from http://www.afsp.org/index.cfm?page_id=05147440-E24E-E376-BDF4BF8BA6444E76.

Brundtland, G. (1999). Global Strategies for Mental Health, from: http://www.who.int/dg/brundtland/en/.

Centers for Disease Control (2011). Web-based Injury Statistics Query and Reporting System (WISQARS), from www.cdc.gov/ncipc/wisqars.

Geddes, J. (1999). Suicide and homicide by people with mental illness: We still don't know how to prevent most of these deaths. *Brithish Medical Journal*, 318(7193), 1225-1226.

Gunnell, D., Shepherd, M., and Evans, M. (1997). Are recent increases in deliberate self-harm associated with changes in socio-economic conditions? An ecological analysis of patterns of deliberate self-harm in bristol 1972-3 and 1995-6. *Psychological Medicine*, 30(5), 1197-1203.

Lester, D. (2003). The variation of the mean age of suicides in the United States. *Crisis: Journal of Crisis Intervention and Suicide*, 24(4), 173-174.

Mathers, C. and Loncar, D. (2006) Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Medicine* from www.plosmedicine.org, **3**(11), e442.

Moscicki, E. (1989) Epidemiological surveys as tools for studying suicidal behavior: A review. *Suicide and Life-Threatening Behavior*, 19, 131-146.

Salib, E. (2003). Lies...damned lies...and suicide statistics! *Medicine, Science and the Law*, 43(4), 315-316.

U.S. Census Bureau (2011), Statistical Abstracts of the United States.

Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association*, 58, 236-244.

Welch, S. (2001). A review of the literature on the epidemiology of parasuicide in the general population. *Psychiatric Services*, 52(2), 368-375.