Principal Component Analysis of Morbidity and Mortality among the United States Homeless Population: A Systematic Review and Meta-Analysis

Albert Nguessan Ngo, PhD1,2* and David Joseph Turbow, PhD3*

Abstract

Background: Homelessness is a modern social and economic problem of major public health significance. In the United States of America (USA), several studies have identified different leading causes of morbidity and mortality among homeless populations.

For this systematic review, Principal Component Analysis (PCA), a multi-variate technique is applied to elucidate the strength of association of comorbidities and mortality among the U.S homeless population.

Methods: A search was conducted for published data for the 20-year period between 1/1/1998 and 12/31/2018. The keywords consisted of the terms: homeless, mortality, morbidity, United States, and health disparities. Data were drawn from cohort studies, observational studies and surveillance reports. Data were then extracted from eleven studies for meta-analysis using PCA.

Results: The first principal component analysis (PCA1) revealed a strong correlation between the co-morbidities of heart disease, cancer, diabetes, and psychoactive substance use. By contrast, PCA2 and PCA3 were strongly correlated to respiratory disease, and transportation-related injuries, respectively. Diseases correlated to PCA1, PCA2, and PCA3 were the major causes of mortality and morbidity of US homeless people with a prevalence of 53.88% (χ² = 228.941, degree of freedom (DF) = 41.924); 26.03% (χ² = 176.27, DF = 176.274) and 19.85% (χ² = 131.266, DF = 131.266), respectively.

Conclusions: Based on the Principal Component Analysis (PCA), homeless people in the U.S. die mainly from heart disease, cancer, diabetes, and psychoactive substance use. Thus, there is a need to address health disparities and to further promote health education and intervention programs.

Keywords
Homelessness, Mortality, Morbidity, United States, Health disparities, Principal component analysis

Introduction

According to the United States (U.S.) Department of Health and Human Services (HHS), a homeless person is defined in section 330(h)(5)(A) as “an individual who lacks housing (without regard to whether the individual is a member of a family), including an individual whose primary residence during the night is a supervised public or private facility (e.g., shelters) that provides temporary living accommodations, and an individual who is a resident in transitional housing” [1].

By early 1970, homelessness had been recognized as a national problem of major public health and economic significance in the U.S. [2]. However, at that time, morbidity and mortality rates associated with homelessness had not yet been formally documented [3,4].
Homelessness affects all types of people, irrespective of age and gender. There are currently an estimated 554,000 homeless people living in the U.S., representing 0.17% of the total U.S. population [5]. According to the National Health Service, the life expectancy of a homeless person is 30 years less than her/his counterpart who has a home. The life expectancy of a homeless person is less than 47.6-years on average. Thus, homelessness is strongly associated with early death [6]. While the total number of U.S. homeless declined over the period 2007 to 2015, the homeless population increased in several major cities.

Many factors lead to homelessness: including lack of affordable housing [7-10], marriage separation, lawful eviction, post-traumatic stress disorder [11,12], foreclosure [13], accidental fire, and natural disasters such as hurricanes, and earthquakes. Other factors leading to homelessness include: poverty, lack of family support, substance abuse, mental impairment, and reduction in unemployment entitlements [14].

As of 2017, 28 million Americans are without health insurance and lack preventive health care [15]. One out of eight Americans is likely to become homeless and in poor health [16]. Homelessness itself is a contributing factor to human morbidity [17]. Poor health conditions and exposure to extreme weather conditions are cofactors that lessen the life expectancy of homeless people. For instance, many homeless perish during extreme hot weather events due to dehydration [15].

Homeless people are also prone to chronic diseases that are prevalent in the general population. In fact, the prevalence of diabetes and high blood pressure in the adult homeless population is identical to the general population [18]. According to a 2017 CDC report, 30.3 million (9.4%) and 84.1 million Americans had diabetes and prediabetes [19]. Diabetes was the seventh foremost cause of death in the U.S. [19]. The prevalence of severe mental illness and substance abuse addiction is high in the homeless population leading to homelessness [20]. Many homeless people suffer from depression, anxiety, personality disorders, schizophrenia, heroin dependence, benzodiazepine dependence, alcohol dependency, bone fractures from assault, suicidal ideation and self-harm, asthma, gastrointestinal disease, tuberculosis, eczema, hepatitis C, hepatitis B and HIV/AIDS [17]. Several efforts are devoted to homeless such as assertive community treatment in the rehabilitation of homeless persons with severe mental illness [21].

In this review article, we systematically elucidate the strength of association between comorbidities and mortality among the U.S homeless population based on principal component analysis (PCA), a multi-variate technique.

**Methods**

The systematic review was conducted following Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines [22].

Studies were selected for analysis based on the following inclusion criteria: the study must identify causes of mortality and morbidity among the United States homeless population. Only studies published from 1998-2018 were included for this analysis. As an exclusion criterion, studies focusing solely on one cause of disease (e.g., cancer) were excluded from this systematic and meta-analysis.

The literature search was conducted in Medline, PubMed, and Europe PubMed Central electronically. The following search terms were employed alone and in combination: causes, morbidity, mortality, homeless, United States, health disparities.

**Statistical Analysis**

**Principal component analysis and Bartlett’s test**

The aim of Principal Component Analysis is to use an orthogonal transformation procedure to extract and present tabular data as a graph which best explains the variance [23]. In the present study, PCA was used to identify the leading causes of death in the U.S. homeless population based on previously published data from peer-reviewed and review articles related to mortality and morbidity among the U.S. homeless [24-26]. For this study, we drew upon the results of several previously published articles on the causes of mortality/morbidity of US homeless. PCA lends itself to meaningful analysis of health data by reducing the number of variables to a few, readily interpretable linear combinations of the original quantitative variables. Each linear combination of variables corresponds to a principal component [27,28]. The assumption is that each principal component is strongly intercorrelated with the original variable [28]. The correlation coefficient between the first principal component analysis (PCA1) is computed, followed by the second Principal component analysis (PCA2) etc. and the original quantitative variable for an efficient and rational interpretation of the data. The Pearson correlation coefficient is defined as follows:

\[ r = \frac{\text{cov}(X,Y)}{S_X S_Y} \]

Where, \( r \) is the correlation coefficient; \( r \) varies from zero (no correlation) to one (perfect positive correlation) or one (perfect negative correlation); \( \text{cov}(X,Y) \) is the covariance of \( X \) and \( Y \). \( S_X \) and \( S_Y \) are the standard deviation of \( X \) and \( Y \), respectively. The eigenvalues are derived from the correlation matrix and provide useful information of the interpretation of the data. The correlation coefficient and the eigenvalues were computed using JMP software version 9.0.3 (SAS Institute, Cary, North Carolina, U.S.). A Bartlett’s Test was used to determine whether there was homogeneity of variances among the different eigenvalues correlated to the

---

The Preferred Reporting Items for Systematic Review (PRISMA) flow diagram is shown in Figure 1. The selected studies for qualitative and meta-analysis are shown in Table 1.

A summary of the literature used for this systematic review is shown in Table S1.

Statistical analysis for Mortality and morbidity in homeless population

Figure 2A reveals the summary plot of the PCA1, PCA2 and PCA3 accounting for 53.8%, 26.41%, and 19.55% of the variability of the sample respectively. The correlation coefficients between PCA1 and cancer, heart disease, psychoactive substance use, and diabetes
The PCA2 and PCA3 were found to be strongly related and metabolic disease were found to be greater than or equal to 0.948, as shown in Table 2. This suggests that homeless people died primarily from cancer, heart disease, psychoactive substance use, diabetes and metabolic disease. These health conditions reflect the major leading causes of death in the broader U.S. population. The PCA1 analysis also reveals that the original quantitative variables are inter-related and explain the prevalence of commodities and mortality in the US homeless population. For instance, untreated health conditions such as diabetes can lead to heart-related events and conditions such as stroke and diabetic ketoacidosis [39,40]. Likewise, an addiction to psychoactive substances can trigger the release of brain-derived neurotrophic factor, a brain protein associated with tumor growth [41].

The PCA2 and PCA3 were found to be strongly related and metabolic disease were found to be greater than or equal to 0.948, as shown in Table 2. This suggests that homeless people died primarily from cancer, heart disease, psychoactive substance use, diabetes and metabolic disease. These health conditions reflect the major leading causes of death in the broader U.S. population. The PCA1 analysis also reveals that the original quantitative variables are inter-related and explain the prevalence of commodities and mortality in the US homeless population. For instance, untreated health conditions such as diabetes can lead to heart-related events and conditions such as stroke and diabetic ketoacidosis [39,40]. Likewise, an addiction to psychoactive substances can trigger the release of brain-derived neurotrophic factor, a brain protein associated with tumor growth [41].

The PCA2 and PCA3 were found to be strongly related and metabolic disease were found to be greater than or equal to 0.948, as shown in Table 2. This suggests that homeless people died primarily from cancer, heart disease, psychoactive substance use, diabetes and metabolic disease. These health conditions reflect the major leading causes of death in the broader U.S. population. The PCA1 analysis also reveals that the original quantitative variables are inter-related and explain the prevalence of commodities and mortality in the US homeless population. For instance, untreated health conditions such as diabetes can lead to heart-related events and conditions such as stroke and diabetic ketoacidosis [39,40]. Likewise, an addiction to psychoactive substances can trigger the release of brain-derived neurotrophic factor, a brain protein associated with tumor growth [41].

The PCA2 and PCA3 were found to be strongly related and metabolic disease were found to be greater than or equal to 0.948, as shown in Table 2. This suggests that homeless people died primarily from cancer, heart disease, psychoactive substance use, diabetes and metabolic disease. These health conditions reflect the major leading causes of death in the broader U.S. population. The PCA1 analysis also reveals that the original quantitative variables are inter-related and explain the prevalence of commodities and mortality in the US homeless population. For instance, untreated health conditions such as diabetes can lead to heart-related events and conditions such as stroke and diabetic ketoacidosis [39,40]. Likewise, an addiction to psychoactive substances can trigger the release of brain-derived neurotrophic factor, a brain protein associated with tumor growth [41].
to respiratory disease and injuries from transportation. This suggests that respiratory disease and injury from transportation contribute significantly to the death of homeless people in U.S. with a prevalence of 26% and 19.95%, respectively as shown in Figure 2C. PCA1, PCA2 and PCA3 were statistically significant and explain the major cause of deaths in U.S. population with a cumulative percentage 99.766% as shown in the scree plot Figure 2B. The other eigenvalues were found to be negligible. Thus, the scree plot suggests the use of three principal components is sufficient to yield meaningful analysis and interpretation of the data [42] (Table 2).

**Sensitivity Analysis**

A sensitivity analysis was conducted undertaken to determine the extent to which uncertainty in the PCA output could be allocated to different input sources thus corroborating the findings. In this second PCA analysis, studies reporting less than five causes of disease were removed from the meta-analysis. As shown in Figure 2C and Figure 2D, the PCA values were identical for PCA1 based on the sensitivity analysis. Likewise, PCA2 and PCA3 values for the meta-analysis ranges are [26.415-26.794] and [19.167-19.550] showing a narrow variation of the principal components. In addition, PCA1, PCA2 and PCA3 are still strongly correlated to the same original variable as shown in (Table S2).

**Discussion and Interventions**

As a method of analysis, PCA is advantageous in extracting the maximum of information explaining the primary cause of death of homeless people in the USA.
As revealed by the PCA, untreated chronic diseases are the major contributing causes to early death among U.S. homeless people. Our analyses thus support previously published research on the topic [43]. For instance, Eyre, et al. reported that collectively cancer, heart disease, diabetes contribute to two-third of the deaths of US general population [44]. This corroborates our finding that in the U.S. homeless die mainly from diseases correlated to PCA1 with a probability of 53.9%. Unfortunately, many homeless people do not have access to preventive and affordable health care.

Health education programs such as Homeless Empowering Learning Program (HELP) may prove to be effective in reducing morbidity in U.S. homeless population. Those programs promote free educational training on both physical and mental preventive health care to homeless people [45]. Affordable and accessible preventive care for homeless people will be important programs needed to reduce the health disparities in homeless population. In this perspective, the U.S. Preventive Services Task Force (USPSF) approves several guidelines related to preventive care. Those preventive cares include but are not limited to disease screening, mental and behavioral screenings and safety screening and immunizations [39]. For instance, the creation of ambulatory care opportunities for homeless people and creation of health care facilities funded by the federal government or non-profit organization dedicated only to homeless population free of cost may be very beneficial to reduce the health disparities and improve the well-being of homeless people. A recent report of the Office of National Drug Control Policy revealed that 30% of people experiencing chronic homeless exhibits serious mental illness and substance used disorder and other chronic health disorder [21]. Interventions provided by the Substance Abuse and Mental Health Services Administration (SAMHSA) can be pivotal in the battle against mental illness. SAMHSA supports different behavioral health treatment and recovery-oriented services such as outreach, case management, treatment for mental and/or substance use disorders, peer support services, Medicaid and employment readiness services [46].

Here we note the limitations of our systematic review. First, the data inputs are generalized by major types of diseases and injuries types and are pooled for specific cities. This represents an important opportunity to conduct extensive research on health education, promotion and prevention program among U.S. homeless populations. Indeed, those health programs such as homeless empowering learning programs can be beneficial for homeless in many health-related categories: health fostering behavior change, health knowledge, health literacy physical markers of health risk factors, actual physical changes in health and improved food security, and health assessment [47].

Conclusions

Homelessness is a societal, economic and moral problem in U.S. The results of this review with PCA reveal that the major causes of death in homeless population mirror leading causes of death in U.S. general population: Heart disease, cancer, and diabetes/metabolic disease.

Given this stronger understanding of the proximate causes of homelessness, policy levers should be implemented accordingly. Much remains to be done to reduce the health disparities and higher rate of morbidity and mortality in homeless population. Health education and interventions programs such as Homeless Empowering Learning Program (HELP) are indeed very crucial to improve the well-being and life expectancy of USA homeless people. This systematic review sheds new light on an issue of major public health significance and points to the need for leveraged research on improving the health status of homeless people.

Acknowledgments and Disclosure

This work was supported by the American University of Health Sciences Foundation, Signal Hill, CA, USA, grant# IG020719C

Authors Contributions

ANN designed and analyzed the data and drafted the first manuscript. DT generated additional idea and helped and validated the interpretation of the Principal Component Analysis by ANN. All authors edited the manuscript, reviewed and approved the final version, and agreed to be accountable for all aspects of the work.

The views expressed in this article are those of the authors and do not necessarily represent the views of the American university of Health Sciences.

Human Participant Protection

Institutional review board approval was not required because the study scrutinized and synthesized published data only which informed consent has already been obtained by different investigators.

References

planning and development.


### Table S1: Summarizing literature uses in the systematic review and meta-analysis.

<table>
<thead>
<tr>
<th>Study and publication date</th>
<th>Location</th>
<th>Study characteristic</th>
<th>Sampling Method</th>
<th>Diagnostic Method</th>
<th>Special population</th>
<th>No. of Participants</th>
<th>Year of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roncarati, et al. [29]</td>
<td>Boston, MA</td>
<td>Prospective Cohort Study</td>
<td>Homeless Program, Street team</td>
<td>Medical records, clinical, national death index</td>
<td>Unsheltered homeless</td>
<td>445</td>
<td>2000-2009</td>
</tr>
<tr>
<td>Pribish, et al. [31]</td>
<td>Tampa, Florida</td>
<td>Retrospective Cohort Study</td>
<td>Tampa Bay Street medicine</td>
<td>clinic</td>
<td>Homeless in Tampa</td>
<td>183</td>
<td>2015-2016</td>
</tr>
<tr>
<td>Henwood, et al. [33]</td>
<td>Philadelphia, PA</td>
<td>Observational Study</td>
<td>Housing First (HF) programs</td>
<td>HF staff records</td>
<td>Sheltered Homeless</td>
<td>292</td>
<td>2008-2013</td>
</tr>
<tr>
<td>Gambatese, et al. [34]</td>
<td>New York City, NJ</td>
<td>Retrospective cohort study</td>
<td>Homeless surveillance system</td>
<td>Medical examiner</td>
<td>Sheltered and unsheltered homeless</td>
<td>3262</td>
<td>2005-2010</td>
</tr>
<tr>
<td>Bharel, et al. [36]</td>
<td>Boston, MA</td>
<td>Cohort Study</td>
<td>Boston Health Care for the Homeless Program (BHCHP)</td>
<td>BHCHP care</td>
<td>Boston homeless</td>
<td>6494</td>
<td>2010</td>
</tr>
<tr>
<td>Yamane, et al. [37]</td>
<td>Hawaii</td>
<td>Retrospective cohort Study</td>
<td>Hawaii’s Homeless Outreach and Medical Education (H.O.M.E)</td>
<td>Record of H.O.M.E.</td>
<td>Hawaii’s</td>
<td>1182</td>
<td>2009</td>
</tr>
</tbody>
</table>

### Table S2: Principal component Analysis (PCA) on correlations matrix for sensitivity analysis.

<table>
<thead>
<tr>
<th></th>
<th>Cancer</th>
<th>Heart Disease</th>
<th>Infectious and parasitic disease</th>
<th>Respiratory disease</th>
<th>Psychoactive substance use</th>
<th>Transportation Injuries</th>
<th>Diabetes/Metabolic Disease</th>
<th>PCA1</th>
<th>PCA2</th>
<th>PCA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>1.000</td>
<td>0.986</td>
<td>0.399</td>
<td>-0.047</td>
<td>0.991</td>
<td>-0.147</td>
<td>0.976</td>
<td>0.947</td>
<td>-0.348</td>
<td>0.018</td>
</tr>
<tr>
<td>Heart Disease</td>
<td>0.986</td>
<td>1.000</td>
<td>0.528</td>
<td>0.099</td>
<td>0.998</td>
<td>-0.201</td>
<td>0.998</td>
<td>0.986</td>
<td>-0.207</td>
<td>-0.030</td>
</tr>
<tr>
<td>Infectious and parasitic disease</td>
<td>0.399</td>
<td>0.528</td>
<td>1.000</td>
<td>0.871</td>
<td>0.483</td>
<td>-0.118</td>
<td>0.565</td>
<td>0.658</td>
<td>0.712</td>
<td>0.002</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>-0.047</td>
<td>0.099</td>
<td>0.871</td>
<td>1.000</td>
<td>0.046</td>
<td>-0.163</td>
<td>0.146</td>
<td>0.261</td>
<td>0.948</td>
<td>-0.117</td>
</tr>
<tr>
<td>Psychoactive substance use</td>
<td>0.991</td>
<td>0.998</td>
<td>0.483</td>
<td>0.047</td>
<td>1.000</td>
<td>-0.207</td>
<td>1.000</td>
<td>-0.225</td>
<td>-0.243</td>
<td>-0.057</td>
</tr>
<tr>
<td>Transportation Injuries</td>
<td>-0.147</td>
<td>-0.201</td>
<td>-0.118</td>
<td>-0.163</td>
<td>-0.207</td>
<td>1.000</td>
<td>-0.225</td>
<td>-0.243</td>
<td>-0.057</td>
<td>0.985</td>
</tr>
<tr>
<td>Diabetes/Metabolic Disease</td>
<td>0.976</td>
<td>0.998</td>
<td>0.565</td>
<td>0.146</td>
<td>0.995</td>
<td>-0.225</td>
<td>1.000</td>
<td>0.993</td>
<td>-0.161</td>
<td>-0.054</td>
</tr>
</tbody>
</table>