Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

ACTUARIAL SCIENCE

MASTER´S FINAL WORK

DISSERTATION

MORTALITY IN THE UNITED STATES' BORDER REGIONS: A CLOSER LOOK AT THE U.S.-MEXICO AND U.S.-CANADA BORDERS

MELANIE JEAN JOERGER

OCTOBER - 2021

Lisbon School of Economics & Management
Universidade de Lisboa

MASTER
ACTUARIAL SCIENCE

MASTER´S FINAL WORK
DISSERTATION

MORTALITY IN THE UNITED STATES' BORDER REGIONS: A CLOSER LOOK AT THE U.S.-MEXICO AND U.S.-CANADA BORDERS

MELANIE JEAN JOERGER

SUPERVISION
ONOFRE ALVES SIMÕES

OCTOBER - 2021

ABSTRACT

The objective of this thesis is to explore the mortality trends in the United States' border regions. Using the Center of Disease Control and Prevention's WONDER database, we examine overall mortality from 1999-2019 through the calculation of standardized mortality ratios for the border region versus the non-border areas. We analyse sub-populations of the border by state, ethnicity, and cause of death, and we use varying combinations of confounders in our standardization including age, gender, and cause of death. The findings confirm significant differences between the border and non-border regions, with opposite results at each border. When accounting for all confounders, the border region at the Mexican border (SMR = 0.958) has lower mortality than the non-border region (SMR = 1.011), and at the Canada border, the border region (SMR = 1.033) has higher mortality than the non-border region (SMR = 0.985). In this manner, the county of residence and proximity to the border could be a useful contributor to mortality estimations.

KEYWORDS: Mortality; U.S.-Mexico border; U.S.-Canada border

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgments

First, I would like to thank my supervisor, Professor Onofre, for all his advice helping me through this process and for his improvement to my work.

I would like to thank my parents for their encouragement and support of my pursuit in this masters program abroad.

I want to thank my dog, Nora, for staying up with me all those late nights I was working on this paper and forgiving me for not taking her to the park as often at the end.

Finally, I would like to express my sincerest appreciation for my fiancé, Katherine, and her unwavering support throughout this masters program, this dissertation, and in life.

CHAPTER 1 - INTRODUCTION

Mortality has long been a topic of interest because mortality rates are a key indicator of the well-being of a population. The more that is known about mortality patterns, the better we can plan for the future and identify areas for improvement. Mortality is also a major factor affecting life insurance and annuity products since they depend on a precise prediction of future mortality trends. Therefore, an understanding of the current and future mortality and longevity of populations is also crucial for insurance companies to accurately price and reserve for life insurance and annuity products.

In addition to estimating mortality rates, researchers are also interested in what factors are associated with, and could potentially be causing, higher or lower mortality. One of these factors of interest is geographic location. Mortality is often viewed from a national level, but a recent study in the United States showed that there are significant differences in death rates at the county level (Dwyer-Lindgren *et al.,* 2016). Geographic disparities in mortality have also been shown to be persistent, meaning that even as the population of an area is constantly changing, the high or low mortality that the area faces can remain (James, Cossman, and Wolf, 2018). This suggests that geographic location could be a somewhat steady explanatory variable in predicting mortality.

Proximity to a country's border is one way of classifying geographic areas. The areas near a country border have interesting properties and demographics compared to non-border areas since they are the area where two cultures collide (Moya *et al.,* 2016). The United States has two land borders, one with Canada to the north and the other with Mexico to the south. Since they are bordering two largely different countries, these borders have dissimilar sets of concerns.

The United States-Mexico border region tends to be of particular interest because of the different levels of development of the two bordering countries (Botero *et al.,* 2015). The U.S.-Mexico border region has documented inequalities and disadvantages when compared to other regions of the United States. For instance, if the border region were considered a state, it would rank last in access to healthcare and per capita income (Moya *et al.,* 2016). Border residents also have lower levels of education and employment than non-border residents (Mills and Caetano, 2016). These disparities have led to the establishment of the bi-national Border Health Commission which aims to improve the health conditions in the border region. They have also led to a growing body of literature on the mortality and morbidity of the border region.

On the other side, it is true that the United States and Canada experience similar standards of living and have many commonalities with regards to culture and ethnic makeup (Feeny *et al*., 2010). However, there are still differences in the countries and the health and mortality of the United States versus Canada has been a topic of research. There has also been an interest in the mortality and health effects of the pollution generated at high-trafficked border crossings along the U.S.-Canada border. While the U.S.-Canada border has received relatively less attention with regards to mortality, it provides another perspective for comparison to the U.S.-Mexico border.

In this thesis, we will take a closer look at the mortality in the United States' border regions, especially when compared to the non-border regions of the border states. To our knowledge there hasn't been a mortality study concerning both the U.S.-Canada and U.S.-Mexico border regions. In Chapter 2, we will review the existing literature on the health and mortality of the U.S.-Mexico border, followed by the U.S.-Canada border. In Chapter 3, we define the border regions, detail the data's source and characteristics, and explain the mortality methods that will be discussed and utilized in this work. In Chapter 4, the U.S.-Mexico border region will be compared to the non-border region, then we will do the same for the U.S.-Canada border, and finally we will compare the two border areas directly. The statistical software R will be used to carry out all calculations and generate figures. Finally, Chapter 5 will conclude the work.

CHAPTER 2 - LITERATURE REVIEW

2.1 U.S.-Mexico Border

A majority of the residents in the U.S.-Mexico border region are of a Hispanic origin. Based on the intercensal populations for years 2001-2009, the U.S.-Mexico border region's population had 53.5% of a Hispanic or Latino origin compared to only 14.9% in the U.S. non-border region (Centers for Disease Control and Prevention, 2020). Given this significant demographic, much of the literature on the region focuses on the Hispanic-American population and most notably what has been known as the Hispanic paradox.

Markides and Coreil (1986) were the first to critically review literature on the health and mortality of southwestern Hispanic population. It was concluded that the health status of Hispanics was closer to that of non-Hispanic Whites than that of non-Hispanic Black Americans, even though their socio-economic status is closer to the latter group. This is significant because it has been well documented that lower socio-economic status typically leads to higher morbidity and mortality in the United States (Adler *et al*., 1994; Sorlie, Backlung, and Keller, 1995; Crimmins, Kim and Seeman, 2009). Markides and Coreil (1986) looked at several health indicators (infant mortality, life-expectancy, cardiovascular diseases, cancer, diabetes and other diseases) and found that Mexican Americans fared better than expected on many of them. This pattern of Hispanics being healthier and having lower mortality than whites despite their lower economic status has been found consistently in literature and has come to be known as the Hispanic Paradox (Shai and Rosenwaike, 1987; Becker *et al*, 1988; Liao *et al*., 1998; Hummer *et al*., 2007).

By dividing the population by age, ethnic subgroup, or immigration status, some studies have found that the paradox only applies to certain subgroups. A study of longitudinal cohorts shows that the Hispanic paradox is most pronounced in older age groups (over 64 years of age) and that in the lower age brackets (18-44 years of age) the mortality of Hispanics may even be higher than that of whites (Liao *et al*., 1998). Hummer *et al*. (2000) divided the Hispanic population into subgroups based on origin: Mexican Americans, Cuban Americans, Puerto Ricans, Central/Southern American, and Other Hispanic. Their analysis of mortality data for the years 1986-1994 found that when accounting for mortality risk factors all Hispanic subgroups had lower mortality overall than non-Hispanic whites. They found that the Mexican American population had the lowest mortality when compared to whites. However, Puerto Ricans between the ages of 18-44 were found to have higher mortality than whites. Both these findings were reinforced by Palloni and Arias (2004) who concluded that the paradox was only true for foreign-born Hispanics excluding Cubans and Puerto-Ricans.

As the existence of this paradox was first emerging, several studies hypothesized that the data must be misleading and that migratory factors were mainly at the source of this seemed paradox. The two main ideas were referred to as the "healthy migrant hypothesis" and the "salmon bias" (Shai and Rosenwaike, 1987; Sorlie *et al*. 1993; Liao *et al*. 1998). The healthy migrant hypothesis postulates that there is healthy selection process when it comes to migration which leads to a migrant population that is healthier than the origin population. The salmon bias suggests that immigrants will migrate back to their home country when they are near death. Since foreign deaths are not included in the U.S. mortality data, this would skew the mortality rates lower.

Abraído-Lanza *et al*. (1999) conducted a study to test these two hypotheses. To test the healthy immigrant hypothesis, they isolated the U.S.-born Hispanics from the foreign-born Hispanics since the U.S.-born population are not immigrants and would not be affected by this bias. This segmentation of nativity status can also cast doubt on the salmon bias since U.S.-born people would be far less likely to leave the country before death than the foreign-born population.

Using Cox proportional hazards models (Kumar and Klefsjö, 1994) with adjustment for age, income, and education, Abraído-Lanza *et al*. (1999) first studied four sub-groups of U.S. Hispanics. The resulting hazard ratios of mortality were Mexican (Men: 0.57, Women: 0.60), Puerto Rican (0.63, 0.45), Cuban (0.53, 0.47) and Central or South American (0.52, 0.46), showing that all subgroups had lower mortality than their non-Hispanic white counterparts. Because the Puerto-Rican and Cuban subgroups also showed lower mortality, it adds evidence that the salmon bias is likely not at fault here. Cuba is difficult to reach and the political conditions that led to migration have not changed, so it is unlikely that a U.S. Cuban would return. Puerto Rico, on the other hand, is very accessible but their deaths are included in U.S. mortality statistics, so the salmon bias cannot affect this group.

Next, the authors used hazard ratios to compare populations based on birth status. When comparing U.S.-born Hispanics and U.S.-born Whites, the U.S.-born Hispanics had significantly lower mortality (Hazard ratios for women ranged 0.49-0.65 and for men 0.59-0.62). Neither the salmon bias nor the healthy immigrant hypothesis could attribute to this difference.

This study shows that the healthy migrant hypothesis and the salmon bias are not the main factors causing the Hispanic paradox and that other factors need to be explored. Abraído-Lanza *et al*. (1999) suggested health behaviours such as smoking and diet, and cultural behaviours like social cohesion may be the reasons to lower mortality. McDonald and Paulozzi (2019) echoed the hypothesis that behavioural factors like the lower prevalence of smoking among Hispanic Americans could be behind their lower rates of major chronic diseases.

Eschbach, *et al*. (2004) studied the morbidity and mortality of Mexican Americans in relation to the percentage of Mexican Americans in their neighbourhood to test if the paradox could be attributed to social support. Typically, neighbourhoods with a high percentage of Mexican Americans or African Americans are economically disadvantaged. It has also been shown that African American communities are correlated with negative health effects. The authors conducted a 7-year longitudinal study on 3050 Mexican Americans over the age of 65. Their results supported their hypothesis that living in high-density Mexican American neighbourhoods is associated with a positive health and mortality effect. This leads to the possibility that social cohesion and support systems may be factors causing the Hispanic Paradox.

There are, however, some health indicators that have been proven to be outside the paradox, meaning that Hispanics don't fare as well as their non-Hispanic white counterparts. The United States-Mexico Border Health Commission, a binational entity created in 2000 to address public health in the border communities, established several health priorities for the region in the Healthy Border 2020 program (US–Mexico Border Health Commission 2020). Some of the notable diseases were diabetes, breast and cervical cancer, and communicable diseases especially tuberculosis and HIV.

Starting from the first study by Markides and Coreil (1986) it has often been found that diabetes rates are higher than expected in Hispanics and along the U.S.-Mexico border region (Salinas, Su and Snih, 2013; McDonald and Paulozzi, 2019). The U.S.-Mexico Border Diabetes Prevention and Control Project conducted a study with 4,027 adults living in 44 border communities in the United States and Mexico. They collected surveys, conducted interviews, and required a health assessment including a blood pressure reading and a fasting blood sample from each participant. Using these survey results, two studies sought to understand the prevalence of diabetes in the border region. Stoddard *et al*. (2010) found that 25.9% of adults with diabetes living along either side of the border were undiagnosed, but that Mexicans and Mexican immigrants were significantly more likely to be undiagnosed than U.S.-born Hispanics or non-Hispanic whites. Early diagnosis is important to reduce diabetes-related mortality and other possible health complications that arise from untreated diabetes. In the second study, it was found that the overall prevalence of diabetes in the border region was 17.6% compared to the U.S. national estimate of 6.3% at the time of this study. They also found an "inverse relationship between diabetes and education and socio-economic level" (Díaz-Apodaca *et al*., 2010).

Breast cancer and cervical cancer have also been shown to have higher rates among Hispanic women along the border region. (Herrera *et al*., 2012; McDonald and Paulozzi, 2019). The mortality from both cancers can be reduced with early detection which requires regular preventative screenings. Given the documented higher prevalence of breast cancer at the border, Banegas *et al*. (2011) studied breast cancer knowledge by ethnic group. They found that Hispanic women in the border region had higher knowledge of breast cancer prevention, but they were not receiving the recommended screening procedures. This suggests that it isn't a lack of education that's the problem, but another factor keeping Hispanic women from receiving preventative healthcare. Herrera *et al*. (2012) sought to understand the higher prevalence of cervical cancer in the border region. They collected data from 1,724 U.S. women in border communities and ran multivariable logistic regression analyses to determine which variables could be contributing to lower rates of cervical cancer screenings in the area, and they found that lack of health insurance coverage had a clear correlation. From both these studies, we can see that women in the U.S.-Mexico border region are suffering from barriers to adequate health care, especially preventative screenings, which could be leading to the increased rates of breast and cervical cancer in the region.

Hispanics and residents in the border region have been shown to have higher rates of mortality from infectious diseases (Markides and Coreil, 1986), especially the chronic infections, tuberculosis, and HIV (McDonald and Paulozzi, 2019). These diseases disproportionately affect poor communities which leaves the border region at risk. Additional risk factors active at the border are increased mobility and migration, closeness of social interactions, and limited access to healthcare (Moya *et al*., 2016).

2.2 U.S.-Canada Border

Compared to the U.S.-Mexico border, there is a much smaller body of literature on the health of residents living in the U.S.-Canada border region. This may be because the United States and Canada enjoy similar standards of living and have more in common culturally and ethnically than the United States and Mexico. However, when it comes to health and mortality, there are still a few notable differences between the two countries, namely, their health care systems and inequality (Siddiqi and Hertzman, 2007).

Feeny *et al*. (2010) compared the health and mortality of the U.S. versus Canada on three metrics: health-related quality of life, life expectancy, and health-adjusted life expectancy. Canada had more favourable outcome on all three metrics. The authors explored the potential explanations as "access to health care over the full life span (universal health insurance) and lower levels of social and economic inequality, especially among the elderly." Notably, regarding inequality, when they compared only the white populations of both countries, they found that the health statuses were much more similar.

Krueger, Bhaloo, and Rosenau (2009) added to the conversation regarding how similar the U.S. and Canada are with regards to health lifestyles. The authors complied data for both countries for health behaviours (smoking, obesity, and binged drinking), fertility measures (crude birth rate, total fertility rate, and fertility rate for women aged 15-19), and cause-specific mortality (HIV, diabetes, chronic liver disease and cirrhosis, accidents, suicide, and homicide). They used a variety of general linear models to determine if healthy lifestyles between the two countries converge at the border. Their results confirmed that Canada in general has healthier lifestyles than the U.S., and supported their hypothesis that these healthy lifestyles converge at the border. It was the non-border populations of Canada and the U.S. that showed the more striking outcomes.

For instance, the highest levels of binge-drinking and smoking were in Canada not at the border, and the highest levels of almost all cause-specific mortalities were found in the U.S. not near the Canadian border.

Although not an issue along the entire border, there is some documentation on the air pollution and its effects near Canadian border crossings. In response, the two countries jointly formed the Canada-US Border Air Quality Strategy (BAQS) to reduce the transborder air pollution.

The Ambassador Bridge is the most heavily trafficked border crossing in terms of commercial trade. It connects the Canadian city Windsor, Ontario, to the American city Detroit, Michigan (Bureau of Transportation Statistics, 2021). In line with BAQS, Band *et al*. (2016) studied the mortality and cancer rates from 1979 to 1999 of Windsor (Essex County) compared to those in the province of Ontario as a whole. They calculated the standardised mortality ratios (SMR) for causes of death associated with long term air pollution exposure including lung cancer, bronchitis, and emphysema, and standardised cancer incident ratios (SIR) for lung cancer (Breslow and Day, 1987). Through statistical significance testing, they documented significantly increased incidence of lung cancer and mortality from bronchitis, emphysema, circulatory diseases and lung cancer in Essex County.

The second busiest border crossing is the Peace Bridge crossing connecting Buffalo, New York, to Fort Erie, Ontario (Bureau of Transportation Statistics, 2021). A health study of Erie County sought to study the prevalence of asthma at the Peace Bride border crossing (Oyana, Rogerson and Lwebuga-Mukadsa, 2004). The authors designed a cross-sectional study of asthma case subjects and non-asthma case subjects (control group) and used spatial analysis techniques to conclude that there was a significant cluster of asthma cases in the border crossing proximity. They also found that the closer a resident is to the crossing their probability of having asthma increases; for instance, living within 0.5 km of the crossing increases the risk of asthma by 15 times compared to just 2 km away.

While significant differences are observed on the border with Mexico, for better or worse, this is not the case in the states bordering Canada. Because the U.S. is more culturally different than Mexico compared to Canada, the southern border region has more observable differences than the northern one. There is also a distinction in the number of immigrants from the two bordering countries. By far, the highest number of immigrants in the United States come from Mexico. According to Pew Research Center (2020), the number of Mexican immigrants living in the U.S. as of 2018 was 11.2 million which comprises 25.0% of all immigrants. In addition, over 2 million Mexican immigrants live within the border region (Israel and Batalova, 2020). On the contrary, the number of Canadian immigrants was only 800 thousand or 1.8% of immigrants. Canadian immigrants also tend to live in large metropolitan areas in California, Florida, New York, and Texas, not near the Canadian border (Israel and Batalova, 2021). Therefore, it can be expected that Mexican immigrants will have a larger effect on their respective border region than Canadian immigrants.

CHAPTER 3 - DATA AND METHODS

3.1 Defining the Border Areas

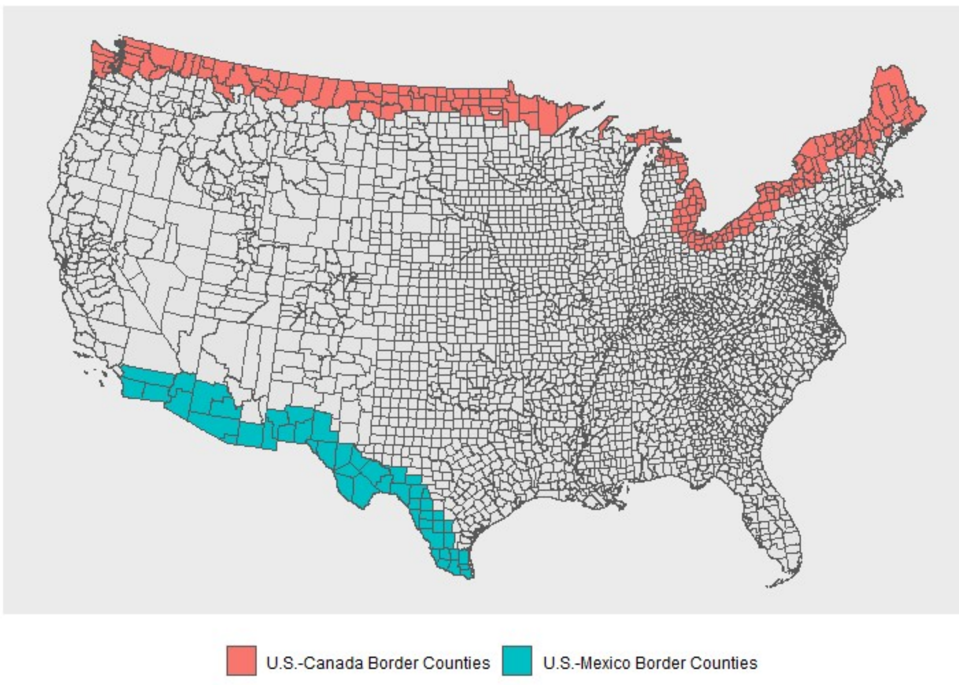
The boundary line between the United States and Mexico spans 3,145 km (1,950 mi) and touches four U.S. states: California, Arizona, New Mexico, and Texas. The border between the United States and Canada is 8,891 km (5,525 mi) and borders 13 states: Alaska, Washington, Idaho, Montana, North Dakota, Minnesota, Michigan, Ohio, Pennsylvania, New York, Vermont, New Hampshire, and Maine. For the purpose of this study, we will only work with the contiguous United States so we will exclude Alaska's eastern border with Canada and only look at the contiguous states' northern border.

To our knowledge, there is no work that looks at both the Mexican and Canadian border regions, which leaves us to determine a definition of the border region that can be used for both. Because there is a much greater body of work on the U.S.-Mexico border region, we will examine the definitions of this border region and choose one that can be reproduced to define the U.S.-Canada border region.

Because mortality data is available at the county level, we need a definition of the border region that includes whole counties. For the U.S.-Mexico border region, there is still not a true consensus on what counties are included in the border region. The U.S.-Mex Political Analysis Tool project by Arizona State University uses a definition including 37 border counties, the Pan-American Organization uses the definition of 48 counties, and even when the border region is defined as only the counties directly touching the border, there are still differing definitions including either 23, 24 or 25 border counties (Payan and Cruz, 2017). The Border Health Commission defines the border region as the 48 counties that lie within 100 km (62.5 mi) of the border citing the La Paz Agreement, a binational agreement signed in 1983 with the goal to improve the environment in the border area. However, their studies only include 44 of those counties (excluding Maricopa, Pinal, and La Paz in Arizona and Riverside County in California) without stating a reason. This definition of 44 counties is the most common definition used in the government and literature. (Salinas *et al*., 2013; Moya *et al*., 2016; Herrera *et al*., 2012).

With the constraint of replication, we used the R package "sf" (Simple Features) to standardize the measurement between country and county borders. When we include any county that has any portion within 100km of the Mexico border, we agreed on the same 48 counties as the Border Health Commission. We ran other scenarios, such as requiring the county to have at least 25% of its area within 100km of the border or requiring the county's centroid to be within a certain distance of the border, but we were unable to find a rule that would include only the 44 counties most commonly used. Therefore, we will move forward with the definition of 48 U.S.-Mexico border counties, which when replicated to the Canadian border, defines 182 U.S.-Canada border counties.

FIGURE 1 – THE CHOSEN U.S.-MEXICO AND U.S.-CANADA BORDER COUNTIES.



Source: United States Census Bureau (author's colouring)

3.2 Data Source

The main data source in this work is from the Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research (CDC WONDER) database. We used the Underlying Cause of Death dataset which identifies one cause of death from U.S. residents' death certificates, classified by the Tenth Revision of the International Classification of Death (ICD-10). The dataset covers deaths from the years 1999 to 2019 compiled by the National Center for Health Statistics (NCHS) through the Vital Statistics Cooperative Program. Death counts lower than 10 and their subsequently calculated mortality rates are "suppressed" to avoid identifiable information and counts lower than 20 are marked "unreliable". The accompanying population estimates are always the mid-interval estimate from the U.S. Census Bureau reports from 1990, 2000, and 2010 (Centers for Disease Control and Prevention, 2020). The variables that will be used in this study are age, gender, Hispanic origin, and cause of death. For cause of death, we will use the nine ICD-10 chapters with the highest national death counts and aggregate the remaining chapters into an "other cause" category, as detailed in Table I.

TABLE I – CAUSE OF DEATH CATEGORIES (1999-2019)

ICD-10 Chapter	Description	Percent of U.S. Deaths
I00-I99	Diseases of the circulatory system	33.4%
C00-D48	Neoplasms	23.1%
J00-J98	Diseases of the respiratory system	9.7%
V01-Y89	External causes of morbidity and mortality	7.5%
G00-G98	Diseases of the nervous system	5.7%
E00-E88	Endocrine, nutritional, and metabolic diseases	4.2%
F01-F99	Mental and behavioural disorders	4.1%
K00-K92	Diseases of the digestive system	3.7%
A00-B99	Certain infectious and parasitic diseases	2.6%
Other	Other cause	5.9%

Source: CDC WONDER database (author's calculations)

When cause of death is being used to segment the population in our analysis, some counties have too small of a population to have reliable death counts in each category. Therefore, we removed these small counties, defined as counties with total populations from 1999-2019 less than 100,000 (~ less than 4,800 average per year), from the analysis. In Table II, the effect of removing these counties is shown to be minimal and the percentage of the population removed is evenly distributed among regions.

TABLE II – COUNTY SIZE DISTRIBUTION (1999-2019)

Border Area	County Size	Count	Population	Percentage
Mexico Border Area	Large	37	284,874,362	99.8%
	Small	11	569,199	0.2%
Mexico Non-Border Area	Large	267	1,180,517,467	99.8%
	Small	45	2,450,741	0.2%
Canada Border Area	Large	167	449,247,625	99.8%
	Small	15	960,037	0.2%
Canada Non-Border Area	Large	387	1,054,940,675	99.7%
	Small	50	2,762,587	0.3%

Source: CDC WONDER database (author's calculations)

3.3 Mortality Analysis Methods

The purpose of this section is to comprehensively explain the different mortality methods to be discussed in this work. We are focusing on single indices used to summarize and compare mortality, in our case from different geographical areas. For each index, we will offer its mathematical formula, using the notation utilized by Kim *et. al* (2020) in a study of small area mortality with some modifications to fit this study, and a practical example that will be followed throughout the section.

We will breakdown the calculations for each method by looking at one county in detail, Webb County, TX, which is one of the U.S.-Mexico border counties. Webb County was chosen because of its unique mortality data which provides insight into the different mortality indices. The death and population data for Webb County from the CDC WONDER database is in Table III.

TABLE III – WEBB COUNTY DEATHS AND POPULATION 1999-2019

Ten-Year Age Group	Deaths	Percentage of Deaths	Population	Percentage of Population
<1 years	611	2.5%	109,920	2.2%
1-4 years	99	0.4%	413,044	8.2%
5-14 years	129	0.5%	962,072	19.0%
15-24 years	617	2.5%	845,784	16.7%
25-34 years	702	2.9%	721,399	14.3%
35-44 years	1,018	4.2%	670,634	13.3%
45-54 years	1,956	8.0%	542,312	10.7%
55-64 years	3,137	12.9%	375,678	7.4%
65-74 years	4,228	17.4%	236,915	4.7%
75-84 years	5,877	24.1%	129,668	2.6%
85+ years	5,969	24.5%	47,906	0.9%
Total	24,343		5,055,332	

Source: CDC WONDER database (author's calculations)

3.3.1 Crude Mortality Rate

The crude mortality rate (CMR) is the most basic mortality index and is therefore the easiest to calculate. It is a ratio where the numerator is the number of deaths in a particular population over a specified time period, and the denominator is the total exposed to the risk (that is, the total estimated population) in the same population over the same period:

(1) $$CMR_c = \frac{Observed\ Deaths}{Total\ Exposed\ to\ Risk} = \frac{\sum_i d_{i,c}}{\sum_i t_{i,c}}.$$

where $d_{i,c}$ is the number of observed deaths in county c and in ten-year age group i and $t_{i,c}$ is the corresponding population. The CMR is typically multiplied by some factor of ten; for instance, the CDC WONDER's mortality rates are multiplied by 100,000 to represent mortality per 100,000 population. For the county data, this is a simple calculation of the total deaths divided by the total population:

$$CMR_{Webb} = \frac{24{,}343}{5{,}055{,}332} \text{ x } 100{,}000 = 481.531$$

This result can be interpreted as the total deaths in Webb County per 100,000 residents. Although the CMR is beneficial for its simplicity, it doesn't consider the age distribution of the population. This can be problematic when attempting to compare the mortality of two areas with potentially different age distributions because the younger population is going to have a lower CMR than the older population by design. This is why there have been other methods developed to standardize the mortality rate for comparison especially with regards to the population's age distribution.

3.3.2 Directly Standardized Mortality Rate

The first method of standardization is referred to as the directly standardized mortality rate (DSMR) which is essentially a weighted average. In the case of age standardization, the weights are the age proportion in the standard population which is multiplied by the age-specific crude mortality rate observed in the target population. This can also be expressed as:

$$(2) \qquad DSMR_c = \frac{\sum_i T_i \left(\frac{d_{i,c}}{t_{i,c}}\right)}{\sum_i T_i}.$$

where T_i is population of the standard population in ten-year age group i. In this paper, we will utilize ten-year age groups for the standardized rates, also known as age-adjusted rates. For the standard population in this exercise, we will use the entire United States population as detailed in Table IV.

TABLE IV – UNITED STATES STANDARD POPULATION 1999-2019

Ten-Year Age Group	Population
< 1 year	83,178,746
1-4 years	333,004,097
5-14 years	860,231,167
15-24 years	893,239,038
25-34 years	874,019,823
35-44 years	889,151,096
45-54 years	887,210,087
55-64 years	724,021,170
65-74 years	477,908,943
75-84 years	282,052,886
85+ years	112,855,471
Total	6,416,872,524

Source: CDC WONDER database

Next, the steps of the calculation of the DSMR are shown in Table V for the example of Webb County.

TABLE V – WEBB COUNTY DSMR CALCULATION

(1) Ten-Year Age Group	(2) Webb County Deaths	(3) Webb County Population	(4) Age-Specific CMR (2) / (3)	(5) Webb County Population Percentage	(6) Standard Population Percentage	(7) DSMR Component (4) x (6)
< 1 year	611	109,920	0.00556	2.2%	1.3%	0.00007
1-4 years	99	413,044	0.00024	8.2%	5.2%	0.00001
5-14 years	129	962,072	0.00013	19.0%	13.4%	0.00002
15-24 years	617	845,784	0.00073	16.7%	13.9%	0.00010
25-34 years	702	721,399	0.00097	14.3%	13.6%	0.00013
35-44 years	1,018	670,634	0.00152	13.3%	13.9%	0.00021
45-54 years	1,956	542,312	0.00361	10.7%	13.8%	0.00050
55-64 years	3,137	375,678	0.00835	7.4%	11.3%	0.00094
65-74 years	4,228	236,915	0.01785	4.7%	7.4%	0.00133
75-84 years	5,877	129,668	0.04532	2.6%	4.4%	0.00199
85+ years	5,969	47,906	0.12460	0.9%	1.8%	0.00219
Total	24,343	5,055,332	0.20888			0.00750

Source: CDC WONDER database (author's calculations)

$$DSMR_{Webb} = \sum DMSR\ Components \ \ x \ 100{,}000 = \ 750.04$$

The DSMR for Webb County is 56% higher than its CMR calculated in the last section. This is due to the younger population in the county. In columns (5) and (6) in Table V, it is shown that the population proportions for ages 55 years and above are all at least 50% lower than the national average and the age groups of 24 years and below are all in higher proportions than the standard population. This younger age distribution skewed the CMR to be lower, showing the necessity of an age-standardized rate.

However, the DSMR does have some constraints. When the geographical areas under consideration are small or sparsely populated, the DSMR can be incalculable. This is because the observed deaths by age group are needed from the population which can be too low to be reliable or may not even be reported.

3.3.3 Indirectly Standardized Mortality Rate

The next method of standardization is the indirectly standardized mortality rate (ISMR) which is an approximation of the DSMR and is formulated as follows:

$$(3) \qquad ISMR_c = CMR_{standard} \times \frac{Observed\ Deaths}{Expected\ Deaths} = \frac{\sum_i D_i}{\sum_i T_i} \times \frac{\sum_i d_{i,c}}{\sum_i t_{i,c}\left(\frac{D_i}{T_i}\right)},$$

where D_i is the number of observed deaths in the standard population in ten-year age group, i. The advantage of the ISMR is that the observed deaths by age group in the population of interest is not necessary for the calculation, instead only the age distribution of the population and the total observed deaths are needed. This is an important benefit when the population of interest is too small to have reliable death data by age group. The ISMR for the county of Webb is calculated in Table VI.

TABLE VI – WEBB COUNTY ISMR CALCULATION

(1) Ten-Year Age Group	(2) Webb County Population	(3) Deaths in Standard Population	(4) Standard Population	(5) Standard Age-Specific Rate (3) / (4)	(6) Expected Deaths (2) x (5)
< 1 year	109,920	538,305	83,178,746	0.00647	711
1-4 years	413,044	93,287	333,004,097	0.00028	116
5-14 years	962,072	127,143	860,231,167	0.00015	142
15-24 years	845,784	660,663	893,239,038	0.00074	626
25-34 years	721,399	969,557	874,019,823	0.00111	800
35-44 years	670,634	1,678,859	889,151,096	0.00189	1,266
45-54 years	542,312	3,678,540	887,210,087	0.00415	2,249
55-64 years	375,678	6,444,490	724,021,170	0.00890	3,344
65-74 years	236,915	9,359,781	477,908,943	0.01958	4,640
75-84 years	129,668	13,901,138	282,052,886	0.04929	6,391
85+ years	47,906	15,966,195	112,855,471	0.14147	6,777
Total	5,055,332	53,417,958	6,416,872,524	0.23403	27,062

Source: CDC WONDER database (author's calculations)

$$ISMR_{Webb} = \frac{53,417,958}{6,416,872,524} \text{ x } 100,000 \text{ x } \frac{24,343}{27,062} = 748.82.$$

The ISMR is very close to the DSMR calculated in the previous section, 750.04, demonstrating that the ISMR can be a fair approximation of the DSMR.

3.3.4 Standardized Mortality Ratio

The standardized mortality ratio compares the ISMR of the population of interest with the crude mortality rate in the standard population:

$$(4) \qquad SMR_c = \frac{Observed\ Deaths}{Expected\ Deaths} = \frac{\sum_i d_{i,c}}{\sum_i t_{i,c}\left(\frac{D_i}{T_i}\right)}.$$

Another explanation is that the SMR is the observed deaths in a population of interest divided by the expected deaths in that population. The SMR is preferable to the ISMR because it is a ratio which allows for the comparison of the population of interest and the standard population to be more intuitive. A SMR value greater than 1 means that the population has higher mortality than the standard population, sometimes referred to as "excess deaths", and a value less than one means the population has lower mortality than the standard population. Using the same calculations done in section 3.3.3 for the ISMR, the SMR for Webb County is:

$$SMR_{Webb} = \frac{24,343}{27,062} = 0.8995.$$

In this case we can see that Webb County has a lower mortality rate than the standard population.

It is important to note that since the SMR uses an indirect method of standardization, it is using the age-group weights of the study population, not the standard population. This means it is very dependent on the age distribution of the study population. This creates a caveat for the use of the SMR such that the SMR of two different study populations should not be used the compare the mortality between those two populations (Schoenbach and Rosamond 2000). Instead, the SMR can only be used to compare the mortality of the study population versus the standard population used in its calculation. If there is enough death data per age group in the study populations, it is best to use the direct method of standardization because the DSMR can be used for the direct comparison of two study populations. For this study, there is the constraint that many counties are small which leads to suppressed or unreliable death counts when divided into age groups. Therefore, we will use indirect standardization via the SMR, and we will only be able to draw conclusions on the comparison to the standard population used.

For the SMR, we will also want to test for significance to see if the difference in SMR is significant compared to the standard population or if it can be attributed to random variability. We will do this by calculating a 95% confidence interval for the SMR, which in context means that if the study were replicated many times, the calculated SMR would lie in the interval 95% of the time. There are several methods of approximating the bounds of the confidence interval, but we will utilize the one designated in a recent small area mortality study shown in equation (5) (Kim *et. al* 2020). If the confidence interval does not include 1, the SMR will be significant when compared to the standard population.

$$(5) \qquad 95\%\ Confidence\ Interval = \left[\frac{SMR}{\exp\left(\frac{1.96}{\sqrt{d_c}}\right)}, \qquad SMR * \exp\left(\frac{1.96}{\sqrt{d_c}}\right)\right].$$

CHAPTER 4 - RESULTS AND DISCUSSION

4.1 U.S.-Mexico Border Analysis

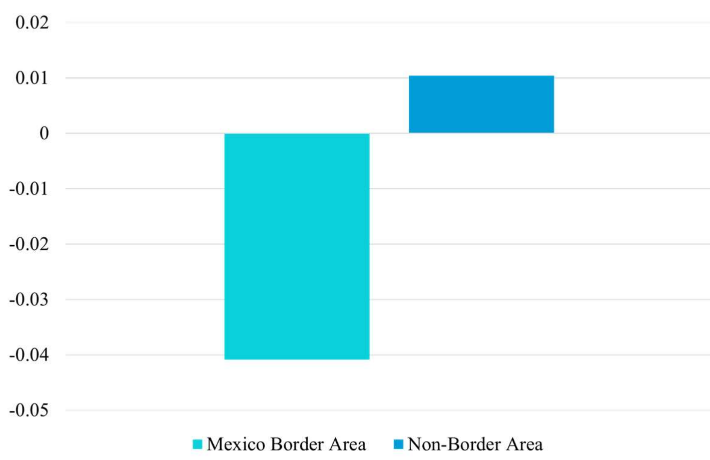
The first approach was to calculate the standardized mortality ratios for the U.S.-Mexico border area (the 48 counties defined in Chapter 3.1) and for the non-border area (the balance of the border states). The full area of the four border states was used as the standard population, age was used as the only confounder of standardization, and a 95% confidence interval was calculated. The results are shown in Table VII. Since the 95% confidence interval for the SMR of the U.S.-Mexico border area does not include 1, we can conclude that the mortality in the border region is significantly lower than in the standard population. On the other hand, the non-border area has an SMR significantly higher.

TABLE VII – SMR BY U.S.-MEXICO BORDER AREA

Border Area	Observed Deaths	Expected Deaths	SMR	95% Confidence Interval	Significant*
Mexico Border Area	1,950,101	2,033,177	0.9591	[0.9578, 0.9605]	Low
Non-Border Area	8,088,962	8,005,886	1.0104	[1.0097, 1.0111]	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 2 – GRAPH OF SMR BY U.S.-MEXICO BORDER AREA



Source: CDC WONDER database (author's calculations)

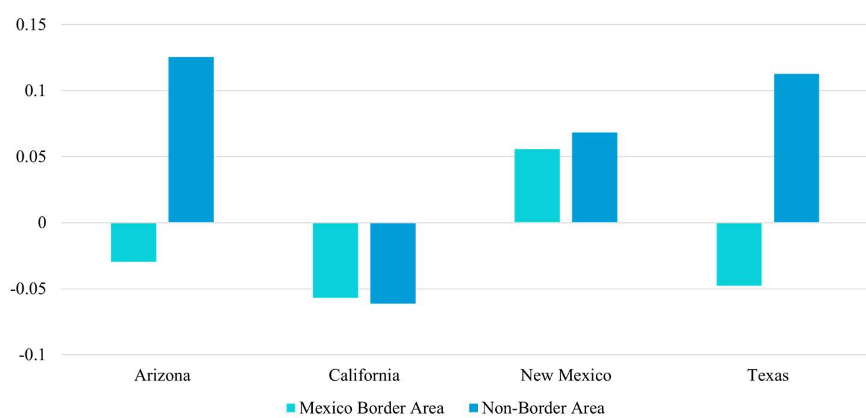
To investigate the distribution further, the SMRs by state and border area were calculated (Table VIII) with the same standard population as above. The SMRs in the border region are most different in Arizona and Texas. In both cases, the Mexico border area has significantly lower mortality (SMR = 0.970 and 0.952, respectively) and the non-border area has significantly higher mortality (SMR = 1.126 and 1.1127). In New Mexico, both areas have significantly higher mortality (SMR = 1.056 and 1.0683) than the standard population, though the border area's is slightly lower. On the other hand, both areas in California had the lowest SMR of any group (SMR = 0.9431 and 0.9386), so they were both significantly lower than the standard population. These results can also be seen geographically in Figure 4. The magenta shades show the areas with SMRs greater than 1 and the teal shaded areas have SMRs under 1.

TABLE VIII – SMR BY STATE AND U.S.-MEXICO BORDER AREA

State	Mexico Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Arizona	0.9702	Low	1.1256	High
California	0.9431	Low	0.9386	Low
New Mexico	1.0557	High	1.0683	High
Texas	0.9521	Low	1.1127	High

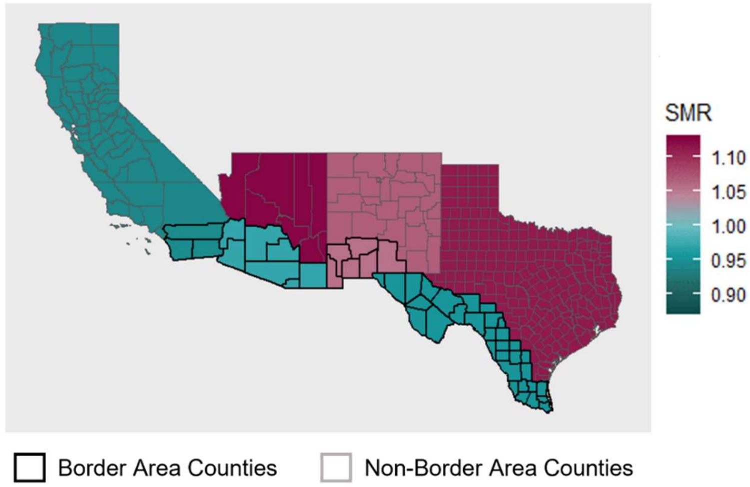
*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 3 – GRAPH OF SMR BY STATE AND U.S.-MEXICO BORDER AREA



Source: CDC WONDER database (author's calculations)

FIGURE 4 – MAP OF SMR BY STATE AND U.S.-MEXICO BORDER AREA



Source: CDC WONDER database (author's calculations)

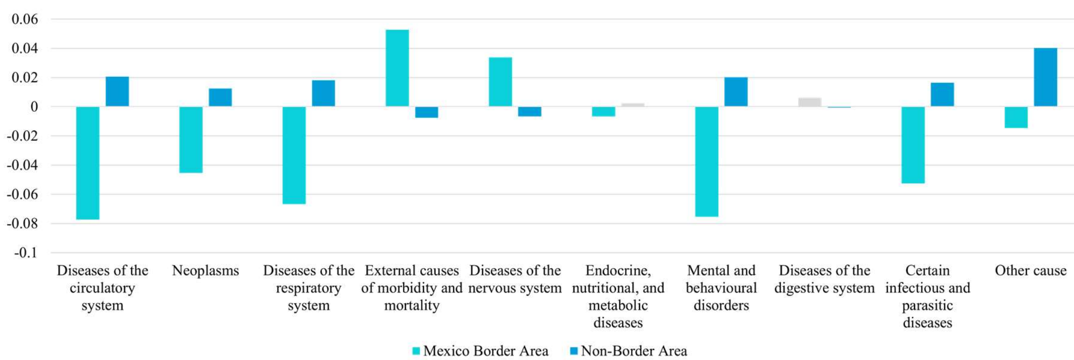
Next, the cause of death was explored by standardizing with respect to age (Table IX). In the U.S.-Mexico border region, most of the causes of deaths SMRs are significantly lower than the standard, but there are two causes of death that are significantly higher: 'External causes of morbidity and mortality' (SMR = 1.053) and 'Diseases of the nervous system' (SMR = 1.034). To investigate the causes of death further, we ran the analysis on the county level and included the maps for each cause of death chapter in Appendix A. For the 'Diseases of the nervous system', the individual border counties which have the highest rates of death are Maricopa County, AZ (SMR = 1.220) and San Diego County, CA (SMR = 1.152). These are the two most populous counties in the border area, so they have a great effect on the overall SMR. Regarding the ICD chapter 'External causes of morbidity and mortality', the non-border areas of Arizona and New Mexico have consistently high SMRs, but these are counteracted by the low rates in California and Texas, as seen in Figure A.4.

TABLE IX – SMR BY CAUSE OF DEATH AND U.S.-MEXICO BORDER AREA

Cause of Death	Mexico Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Diseases of the circulatory system	0.9226	Low	1.0205	High
Neoplasms	0.9547	Low	1.0125	High
Diseases of the respiratory system	0.9333	Low	1.0182	High
External causes of morbidity and mortality	1.0528	High	0.9924	Low
Diseases of the nervous system	1.0338	High	0.9934	Low
Endocrine, nutritional, and metabolic diseases	0.9933	Low	1.0024	-
Mental and behavioural disorders	0.9245	Low	1.0202	High
Diseases of the digestive system	1.0063	-	0.9994	-
Certain infectious and parasitic diseases	0.9475	Low	1.0165	High
Other cause	0.9855	Low	1.0402	High

Source: CDC WONDER database (author's calculations)

FIGURE 5 – GRAPH OF SMR BY CAUSE OF DEATH AND U.S.-MEXICO BORDER AREA



Source: CDC WONDER database (author's calculations)

There has been a significant amount of research regarding the higher levels of diabetes mellitus in the border region, so it was expected that 'Endocrine, nutritional, and metabolic diseases' might be significantly higher in the border region but that was not the case. When this was examined further, we found that the border region in our analysis is benefiting from the inclusion of the four counties that are often excluded from the definition of the Mexico border region even though they are within 100 km of the border. These counties are Riverside County, CA (80 km), La Paz County, AZ (36 km), Maricopa County, AZ (49 km), and Pinal County, AZ (85 km). When these four counties are moved from the border area to the non-border area, the cause of death category 'Endocrine, nutritional, and metabolic diseases' changes entirely. The border area has an SMR of

1.092 making it significantly high, and the non-border area has an SMR of 0.989 which is significantly low. Figure A.6 shows the low mortality rates in the border region of California and Arizona which are offsetting the primarily high rates in the border regions of New Mexico and Texas. It can also be seen that the non-border area of Texas tends to have higher rates of 'Endocrine, nutritional, and metabolic diseases' which causes the SMRs to be more even.

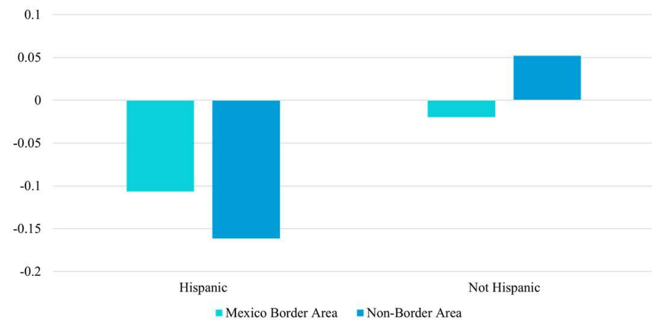
As discussed in Chapter 2, much of the research on the U.S.-Mexico region focuses on the Hispanic or Latino population due to their large proportion in the area and the interesting findings with the Hispanic paradox. Therefore, the next division we examined was dividing the border regions into "Hispanic or Latino" and "Not Hispanic of Latino". On the small percentage (871 / 646,169 = 0.13%) of death certificates where this field was not completed, the category is marked as "Not Stated" and these were removed from the analysis. The results are shown in Table X. The Hispanic populations in the border area (SMR = 0.893) and the non-border area (SMR = 0.839) have significantly lower mortality than the standard population, which is in line with the existing research of the Hispanic paradox. In addition, the mortality of the non-Hispanic population in the border area (SMR = 0.980) is significantly lower than the mortality of the population, compared to the non-border area (SMR = 1.052) which suggests that it is not only the Hispanic population that is causing the border area to have lower mortality.

TABLE X – SMR BY HISPANIC ORIGIN AND U.S.-MEXICO BORDER AREA

State	Mexico Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Hispanic or Latino	0.8932	Low	0.8385	Low
Not Hispanic or Latino	0.9801	Low	1.0521	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 6 – GRAPH OF SMR BY HISPANIC ORIGIN AND U.S.-MEXICO BORDER AREA



Source: CDC WONDER database (author's calculations)

After the variable age, gender is the next most common confounder to add to the standardization. Women have a longer expected lifespan than men and therefore lower mortality rates than men. Consequently, it would be expected that a population with a higher proportion of women would have lower mortality. However, the distribution of gender doesn't vary between different geographies as much as age does, so we wouldn't expect it to have as great of an impact on the SMRs. When gender is added as a variable in the standardization for the U.S.-Mexico border, the border area SMR decreases slightly from 0.9591 to 0.9575, but both are significantly low, and the non-border area increases by an even smaller differential from 1.0104 to 1.0108. When we standardize with either gender or cause of death as the only confounder, or with both as confounders, the expected deaths are strikingly similar and there is no difference in the SMRs in the border and non-border regions, as seen in Table XI.

TABLE XI – SMR BY U.S.-MEXICO BORDER AREA AND STANDARDIZATION METHOD

Confounders in Standardization	Mexico Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Age	0.9591	Low	1.0104	High
Gender	0.9993	-	1.0002	-
Cause of Death	0.9993	-	1.0002	-
Age + Gender	0.9575	Low	1.0108	High
Age + Cause of Death	0.9591	Low	1.0104	High
Gender + Cause of Death	0.9993	-	1.0002	-
Age + Gender + Cause of Death	0.9576	Low	1.0108	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

4.2 U.S.-Canada Border Analysis

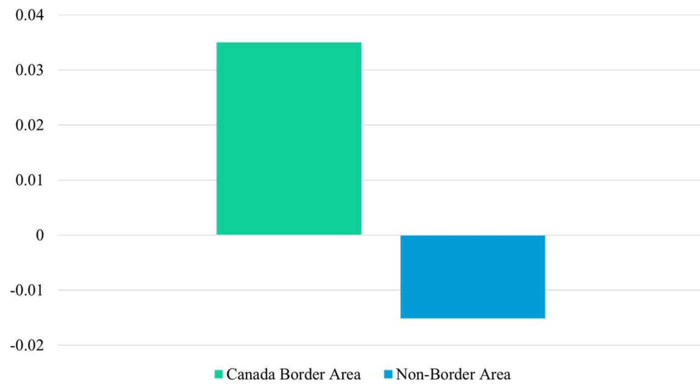
Just as with the U.S.-Mexico border, the first measure of comparison was the standardized mortality ratios for the U.S.-Canada border area versus the non-border area, where the twelve border states are combined to make up the standard population. The results in Table XII show that, in contrast to the U.S.-Mexico border area, the U.S.-Canada border area had an SMR that was significantly higher than the standard population and the SMR of the non-border area was significantly lower.

TABLE XII – SMR BY U.S.-CANADA BORDER AREA

Border Area	Observed Deaths	Expected Deaths	SMR	95% Confidence Interval	Significant*
Canada Border Area	4,132,324	3,992,665	1.0350	[1.0340, 1.0360]	High
Non-Border Area	9,099,417	9,239,076	0.9849	[0.9842, 0.9855]	Low

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 7 – GRAPH OF SMR BY U.S.-CANADA BORDER AREA



Source: CDC WONDER database (author's calculations)

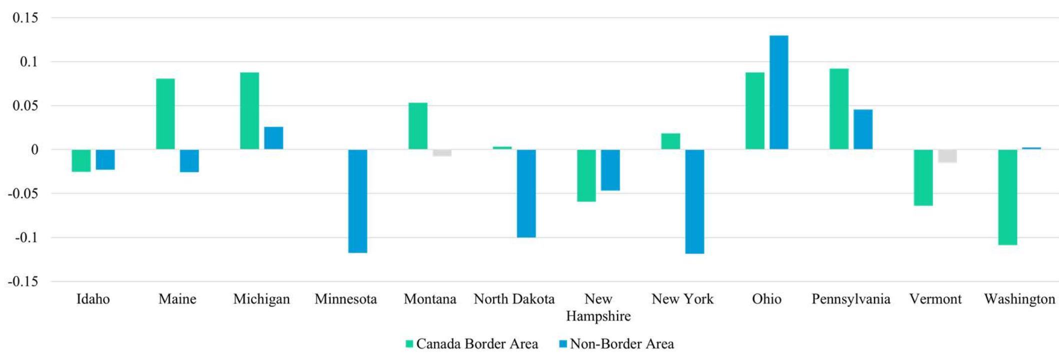
We then calculated the SMRs by state and border area with the same standard population as the first portion. There are six areas with SMRs over 1.05 and six areas with SMRs lower than 0.95 which provide further insight. The six areas with SMRs higher than 1.05 are the border areas of Montana (1.053), Maine (1.081), Michigan (1.088), Ohio (1.088), and Pennsylvania (1.092), and the non-border area of Ohio (1.130). Since five of the six areas with the highest SMRs are from the border region, these states mainly contributed to the overall high mortality seen at the border. The six areas with SMRs under 0.95 are the border areas of New Hampshire (0.941), Vermont (0.936), and Washington (0.891), and along with the non-border areas of North Dakota (0.900), Minnesota (0.883), and New York (0.882). Figure 3 shows that these areas are not clustered in one section of the border, but that the low SMRs and high SMRs are dispersed along the length of the border.

TABLE XIII – SMR BY STATE AND U.S.-CANADA BORDER AREA

State	Canada Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Idaho	0.9748	Low	0.9772	Low
Maine	1.0807	High	0.9744	Low
Michigan	1.0878	High	1.0259	High
Minnesota	1.000	-	0.8826	Low
Montana	1.0534	High	0.9922	-
North Dakota	1.0034	-	0.9001	Low
New Hampshire	0.9407	Low	0.9534	Low
New York	1.0183	High	0.8816	Low
Ohio	1.0877	High	1.1297	-
Pennsylvania	1.0921	High	1.0457	High
Vermont	0.9362	Low	0.9849	-
Washington	0.8913	Low	1.0025	High

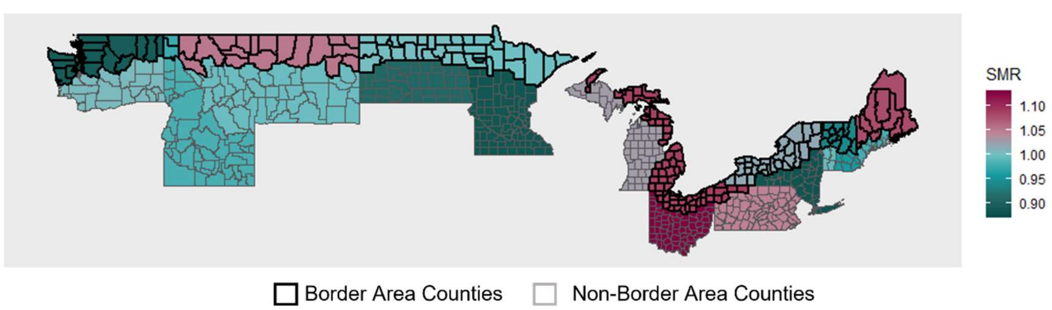
*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 8 – GRAPH OF SMR BY STATE AND U.S.-CANADA BORDER AREA



Source: CDC WONDER database (author's calculations)

FIGURE 9 – MAP OF SMR BY STATE AND U.S.-CANADA BORDER AREA



Source: CDC WONDER database (author's calculations)

When the population is divided by cause of death in Table XIV, the Canada border area is shown to have a significantly high SMR for 8 out of 10 causes of death. The maps for the SMRs by county for each cause of death is included in Appendix B. The causes of death that the border region is significantly low for are 'Certain infectious and parasitic diseases' and 'Other cause'. The two counties with the highest rates of 'Certain infectious and parasitic diseases' are Philadelphia County, PA (SMR = 2.463) and Bronx County, NY (SMR = 2.371) which are both metropolitan counties in the non-border area.

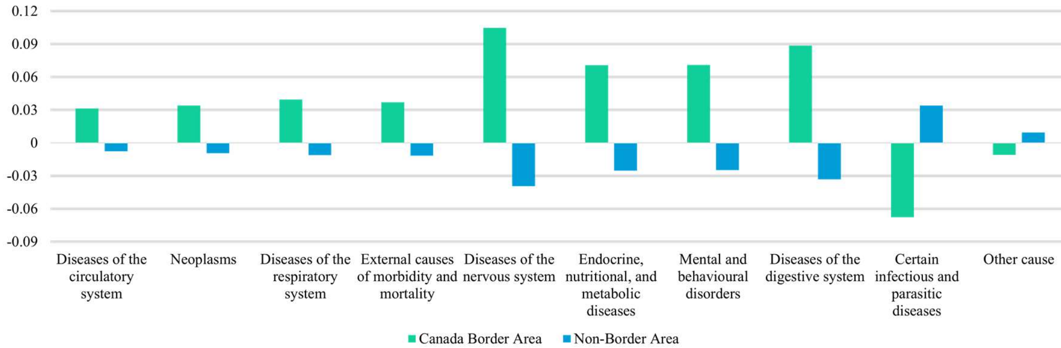
Due to the introduction of BAQS and the concerns of heightened rates of respiratory illnesses near the high-trafficked border crossings, we will take a closer look at the ICD chapter, 'Diseases of the respiratory system'. There are counties with high SMRs for this cause of death along the border (Figure B.3), however, the clustering of poor SMRs is in southern Ohio, a non-border area. This region has 5 of the top 6, and 10 of the top 20 worst county SMRs. Since the research for BAQS is focused on the small area immediately next to the border crossings, county-level data may be too wide of an area to show any difference.

TABLE XIV – SMR BY CAUSE OF DEATH AND U.S.-CANADA BORDER AREA

Cause of Death	Canada Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Diseases of the circulatory system	1.0312	High	0.9924	Low
Neoplasms	1.0340	High	0.9907	Low
Diseases of the respiratory system	1.0395	High	0.9888	Low
External causes of morbidity and mortality	1.0371	High	0.9885	Low
Diseases of the nervous system	1.1047	High	0.9606	Low
Endocrine, nutritional, and metabolic diseases	1.0707	High	0.9748	Low
Mental and behavioural disorders	1.0710	High	0.9752	Low
Diseases of the digestive system	1.0885	High	0.9668	Low
Certain infectious and parasitic diseases	0.9323	Low	1.0340	High
Other cause	0.9892	Low	1.0094	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 10 – GRAPH OF SMR BY CAUSE OF DEATH AND U.S.-CANADA BORDER AREA



Source: CDC WONDER database (author's calculations)

We also carried out the different standardization methods for this border and our results were consistent with the U.S.-Mexico border in that the gender and cause of death variables did not have much effect on the SMRs.

TABLE XV – SMR BY U.S.-CANADA BORDER AREA AND STANDARDIZATION METHOD

Confounders in Standardization	Canada Border Area		Non-Border Area	
	SMR	Significant*	SMR	Significant*
Age	1.0350	High	0.9849	Low
Gender	1.0460	High	0.9804	Low
Cause of Death	1.0460	High	0.9804	Low
Age + Gender	1.0341	High	0.9853	Low
Age + Cause of Death	1.0350	High	0.9849	Low
Gender + Cause of Death	1.0460	High	0.9804	Low
Age + Gender + Cause of Death	1.0343	High	0.9855	Low

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

4.3 Border Comparison

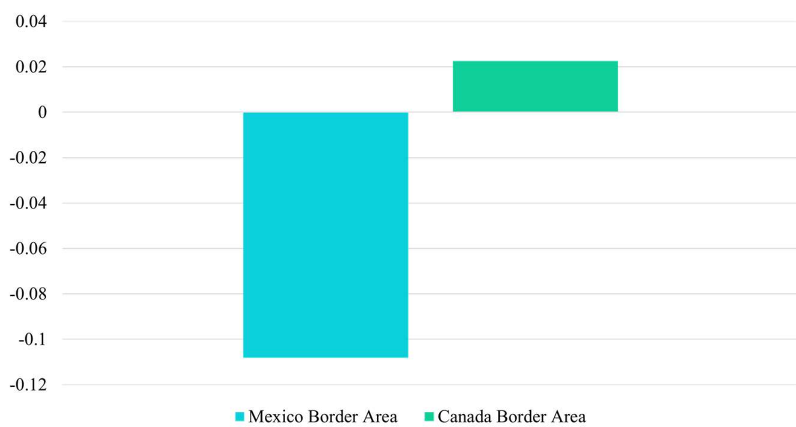
For the previous two sections, we have been using the border states collectively to make up the standard population. Now, to compare the two borders to the same population, we will use the entire United States as the standard. In Table XVI, the 95% confidence intervals of the SMRs are shown, standardized by age. The U.S.-Mexico border area is still significantly lower, and the U.S.-Canada border area is still significantly higher. This shows that the border areas not only have significantly different mortality than the non-border areas but than the overall United States population.

TABLE XVI – SMR BY BORDER AREA

Border Area	Observed Deaths	Expected Deaths	SMR	95% Confidence Interval	Significant*
Mexico Border Area	1,950,101	2,186,455	0.8919	[0.8906, 0.8932]	Low
Canada Border Area	4,132,324	4,040,869	1.0226	[1.0216, 1.0236]	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

FIGURE 11 – GRAPH OF SMR BY BORDER AREA



Source: CDC WONDER database (author's calculations)

The same variations of confounders in standardization in the previous sections are applied in Table XVII. All variations in the Mexico border region produced a significantly low SMR and the Canada border was always significantly high. Unlike in the previous sections, adding gender or cause of death to the standardization method did affect the SMR, albeit slightly. In this case, adding either gender or cause of death (or both) caused both borders' SMRs to move further from 1, that is, the U.S.-Mexico border's SMR was reduced further and the U.S.-Canada border area's increased, making them both more significant.

TABLE XVII – SMR BY BORDER AREA AND STANDARDIZATION METHOD

Confounders in Standardization	Mexico Border Area		Canada Border Area	
	SMR	Significant*	SMR	Significant*
Age	0.8919	Low	1.0226	High
Gender	0.8205	Low	1.1026	High
Cause of Death	0.8207	Low	1.1026	High
Age + Gender	0.8884	Low	1.0226	High
Age + Cause of Death	0.8919	Low	1.0226	High
Gender + Cause of Death	0.8205	Low	1.1026	High
Age + Gender + Cause of Death	0.8883	Low	1.0228	High

*Significantly High or Low at the 5% level compared to the standard population
Source: CDC WONDER database (author's calculations)

CHAPTER 5 - CONCLUSION

Despite the heavy interest in mortality at the U.S.-Mexico border, our analysis shows that the overall mortality in the border area is significantly lower than mortality in the border states and lower than the United States as a whole. We also found that the causes of death that the border population has been shown to be at an increased risk for do not apply to all counties within the 100 km of the border. Instead, these are issues only in certain counties or states. For instance, diabetes mellitus has been shown to have high rates in the border region, but we found that this was mainly in New Mexico and Texas and that when the Arizona and California border regions are included, the area actually has low levels of death from 'Endocrine, nutritional and metabolic diseases'.

Given the depth of research on the Hispanic paradox in the border region, we analysed the area by Hispanic origin. Our results validated the paradox since both the border and non-border Hispanic population had significantly low SMRs. In addition, the SMR for the non-Hispanic border population was also significantly low confirming that it is not only the ethnicity of the border region that is causing the low SMR.

The mortality at the U.S.-Canada border has not been a common topic of research, but we found that there is a significantly higher level of mortality in the area, and that these elevated rates are shown for 8 out of the 10 cause of death categories.

With regards to the different methods of standardization used, we confirmed that age is the most important confounder in standardization as it had the greatest effect on the SMR. Through all the analysis, standardization by cause of death had the smallest effect on the SMR, only changing the SMR in the ten-thousandths place, or not at all.

As discussed in Section 3.2, the main limitation of this study was the suppressed death counts in the WONDER database. We overcame this by removing the smallest counties from the analysis, but having the unsuppressed death data would have been ideal. Access to this information would also make it possible to use direct standardization instead of indirect standardization which would have also been preferable. We would also be able to analyse the mortality for a shorter time period which would allow us to compare the mortality across time and analyse trends. There is potential for further research to investigate time trends and determine if the differences are diverging or converging with time.

Another limitation was the lack of literature available on the mortality in the U.S.-Canada border region. Since there was only a small body of research, there was less direction and precedent, compared to the U.S.-Mexico border. As a recommendation for more research, there is a need to further explore the potential causes of the higher mortality that we found at the U.S.-Canada border.

Through this research and analysis, we learned that geography and proximity to the border is a substantial variable in mortality. It is well known that age and gender are two variables that have a strong correlation with mortality, but when accounting for these confounders, there were still significant differences at both borders. In addition, the county level analysis for cause of death showed that rates can vary drastically just over county lines. Thus, the county of residence could be used to inform mortality projections.

REFERENCES

Abraído-Lanza, A. F., Dohrenwend, B.P., Ng-Mak, D.S. and Turner, J.B. (1999). 'The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses.', *American Journal of Public Health*, 89(10), pp. 1543–1548. doi: 10.2105/AJPH.89.10.1543.

Adler, N. E. *et al.* (1994). 'Socioeconomic status and health: The challenge of the gradient', *American Psychologist*, 49(1), pp. 15–24. doi: 10.1037/0003-066X.49.1.15.

Band, P., Zielinski, J., Jiang, H. and Liu, L. (2006). 'Canada-US border air quality strategy (BAQS): preliminary results of mortality and cancer incidence in Windsor, Canada'. doi: 10.2495/AIR060761.

Banegas, M. P. *et al.* (2011). 'Breast Cancer Knowledge, Attitudes, and Early Detection Practices in United States-Mexico Border Latinas', *Journal of Women's Health*, 21(1), pp. 101–107. pp. 101–107. doi: 10.1089/jwh.2010.2638.

Becker, T. M., Wiggins, C., Key, C. R., and Samet, J. M. (1988). 'Ischemic heart disease mortality in Hispanics, American Indians, and non-Hispanic whites in New Mexico, 1958-1982.', *Circulation*, 78(2), pp. 302–309. doi: 10.1161/01.CIR.78.2.302.

Breslow, N. E. and Day, N. E. (1987) 'Statistical methods in cancer research. Volume II--The design and analysis of cohort studies', *IARC scientific publications*, (82), pp. 1–406.

Bureau of Transportation Statistics (2021). *Border Crossing Entry Data | US-Canada Ports of Entry*. U.S. Department of Transportation. https://data.bts.gov/Research-and-Statistics/US-Canada-Ports-of-Entry/s8z8-5vi3

Centers for Disease Control and Prevention, National Center for Health Statistics (2020). *Underlying Cause of Death 1999-2019*. CDC WONDER Online Database. < http://wonder.cdc.gov/ucd-border.html>

Cossman, J., James, W., Cosby, A., and Cossman, R. (2010). 'Underlying causes of the emerging nonmetropolitan mortality penalty', *American Journal of Public Health* 100(8): 1417–1419. doi:10.2105/AJPH.2009.174185.

Crimmins, E. M., Kim, J. K. and Seeman, T. E. (2009). 'Poverty and Biological Risk: The Earlier "Aging" of the Poor', *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 64A(2), pp. 286–292. doi: 10.1093/gerona/gln010.

Díaz-Apodaca, B. A., Ebrahim, S., McCormack, V., de Cosío, F. G. and Ruiz-Holguín, R. (2010). 'Prevalence of type 2 diabetes and impaired fasting glucose: cross-sectional study of multiethnic adult population at the United States-Mexico border', *Revista Panamericana de Salud Pública*, 28, pp. 174–181. doi: 10.1590/S1020-49892010000900007.

Dwyer-Lindgren, L. et al. (2016). 'US County-Level Trends in Mortality Rates for Major Causes of Death, 1980-2014', *JAMA*, 316(22), pp. 2385–2401. doi:10.1001/jama.2016.13645.

Eschbach, K., Ostir, G. V., Patel, K. V., Markides, K. S. and Goodwin, J. S. (2004). 'Neighborhood Context and Mortality Among Older Mexican Americans: Is There a Barrio Advantage?', *American Journal of Public Health*, 94(10), pp. 1807–1812. doi: 10.2105/AJPH.94.10.1807.

Feeny, D., Kaplan, M. S., Huguet, N. and McFarland, B. H. (2010). 'Comparing population health in the United States and Canada', *Population Health Metrics*, 8(1), p. 8. doi: 10.1186/1478-7954-8-8.

Herrera, D. G. *et al.* (2012). 'Cervical Cancer Screening in the US–Mexico Border Region: A Binational Analysis', *Maternal and child health journal*, 16(0 2), pp. 298–306. doi: 10.1007/s10995-012-1130-8.

Hummer, R. A. Powers, D. A., Pullum, S.G., Gossman, G. L. and Frisbie, W. P. (2007). 'Paradox Found (Again): Infant Mortality Among the Mexican-Origin Population in the United States', *Demography*, 44(3), pp. 441–457.

Hummer, R. A., Rogers, R. G., Amir, S. H., Forbes, D. and Frisbie, W. P. (2000). 'Adult Mortality Differentials among Hispanic Subgroups and Non-Hispanic Whites', *Social Science Quarterly*, 81(1), pp. 459–476.

Israel, E. and Batalova, J. (2020). 'Mexican Immigrants in the United States', *migrationpolicy.org.* Available at: https://www.migrationpolicy.org/article/mexican-immigrants-united-states-2019.

Israel, E. and Batalova, J. (2021). 'Canadian Immigrants in the United States', *migrationpolicy.org.* Available at: https://www.migrationpolicy.org/article/canadian-immigrants-united-states-2021#distribution.

James, W., Cossman, J. and Wolf, J.K. (2018) 'Persistence of death in the United States: The remarkably different mortality patterns between America's Heartland and Dixieland', *Demographic Research*, 39, pp. 897–910.

Kim, I., Lim, H. K., Kang, H. Y. and Khang, Y. H. (2020) 'Comparison of three small-area mortality metrics according to urbanity in Korea: the standardized mortality ratio, comparative mortality figure, and life expectancy', *Population Health Metrics*, 18(1), p. 3. doi:10.1186/s12963-020-00210-7.

Krueger, P. M., Bhaloo, T. and Rosenau, P. V. (2009). 'Health Lifestyles in the United States and Canada: Are We Really So Different?*', *Social Science Quarterly*, 90(5), pp. 1380–1402. doi: https://doi.org/10.1111/j.1540-6237.2009.00660.x.

Kumar, D. and Klefsjö, B. (1994) 'Proportional hazards model: a review', *Reliability Engineering & System Safety*, 44(2), pp. 177–188. doi: 10.1016/0951-8320(94)90010-8.

Liao, Y. *et al.* (1998). 'Mortality patterns among adult Hispanics: findings from the NHIS, 1986 to 1990.', *American Journal of Public Health*, 88(2), pp. 227–232. doi: 10.2105/AJPH.88.2.227.

Markides, K. S. and Coreil, J. (1986). 'The health of Hispanics in the southwestern United States: an epidemiologic paradox.', *Public Health Reports*, 101(3), pp. 253–265.

McDonald, J. and Paulozzi, L. (2019). 'Parsing the Paradox: Hispanic Mortality in the US by Detailed Cause of Death', *Journal of Immigrant and Minority Health*, 21. doi: 10.1007/s10903-018-0737-2.

Miller, C. E. and Vasan, R. S. (2021) 'The southern rural health and mortality penalty: A review of regional health inequities in the United States', *Social Science & Medicine*, 268, p. 113443. doi:10.1016/j.socscimed.2020.113443.

Moya, E. M., Chávez-Baray, S. M., Wood, W. W. and Martinez, O. (2016). 'Nuestra Casa: An advocacy initiative to reduce inequalities and tuberculosis along the US-Mexico border', *International public health journal*, 8(2), pp. 107–119.

Oyana, T. J., Rogerson, P. and Lwebuga-Mukasa, J. S. (2004). 'Geographic Clustering of Adult Asthma Hospitalization and Residential Exposure to Pollution at a United States–Canada Border Crossing', *American Journal of Public Health*, 94(7), pp. 1250–1257. doi: 10.2105/AJPH.94.7.1250.

Palloni, A. and Arias, E. (2004). 'Paradox lost: Explaining the hispanic adult mortality advantage', *Demography*, 41(3), pp. 385–415. doi: 10.1353/dem.2004.0024.

Payan, T. and Cruz, P. L. (2017). 'Managing the U.S.-Mexico Border First Requires Defining It', *Issue brief no. 04.20.17. Rice University's Baker Institute for Public Policy,* p. 8.

Pew Research Center (2020). 'Immigrants in America: Current Data and Demographics', *Pew Research Center's Hispanic Trends Project*. www.pewresearch.org/hispanic/2020/08/20/facts-on-u-s-immigrants-current-data/.

Salinas, J. J., Su, D. and Al Snih, S. (2013). 'Border Health in the Shadow of the Hispanic Paradox: Issues in the Conceptualization of Health Disparities in Older Mexican Americans Living in the Southwest', *Journal of cross-cultural gerontology*, 28(3), pp. 251–266. doi: 10.1007/s10823-013-9202-9.

Schoenbach, V.J. and Rosamond, Wayne D. (2000) 'Fundamentals of epidemiology - an evolving text'. *Department of Epidemiology School of Public Health University* *of North Carolina at Chapel Hill*. http://www.epidemiolog.net/evolving/FundamentalsOfEpidemiology.pdf.

Shai, D. and Rosenwaike, I. (1987). 'Mortality among Hispanics in metropolitan Chicago: An examination based on vital statistics data', *Journal of Chronic Diseases*, 40(5), pp. 445–451. doi: 10.1016/0021-9681(87)90178-0.

Siddiqi, A. and Hertzman, C. (2007) 'Towards an epidemiological understanding of the effects of long-term institutional changes on population health: a case study of Canada versus the USA', *Social Science & Medicine* (1982), 64(3), pp. 589–603. doi: 10.1016/j.socscimed.2006.09.034.

Sorlie, P. D., Backlund, E. and Keller, J. B. (1995). 'US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study', *American Journal of Public Health*, 85(7), pp. 949–956. doi: 10.2105/ajph.85.7.949.

Sorlie, P. D., Backlund, E., Johnson, N. J., and Rogot, E. (1993). 'Mortality by Hispanic Status in the United States', *JAMA*, 270(20), pp. 2464–2468. doi: 10.1001/jama.1993.03510200070034.

Stoddard, P., He, G., Vijayaraghavan, M. and Schillinger, D. (2010). 'Disparities in undiagnosed diabetes among United States-Mexico border populations', *Revista Panamericana de Salud Pública*, 28, pp. 198–206. doi: 10.1590/S1020-49892010000900010.

US–Mexico Border Health Commission (2020). *Healthy Border 2020: A Prevention & Health Promotion Initiative. US Department of Health and Human Services*. https://www.hhs.gov/sites/default/files/res_2805.pdf

APPENDIX A – U.S.-MEXICO BORDER REGION MAPS

Figure A.1 – Cause of Death: Diseases of the circulatory system

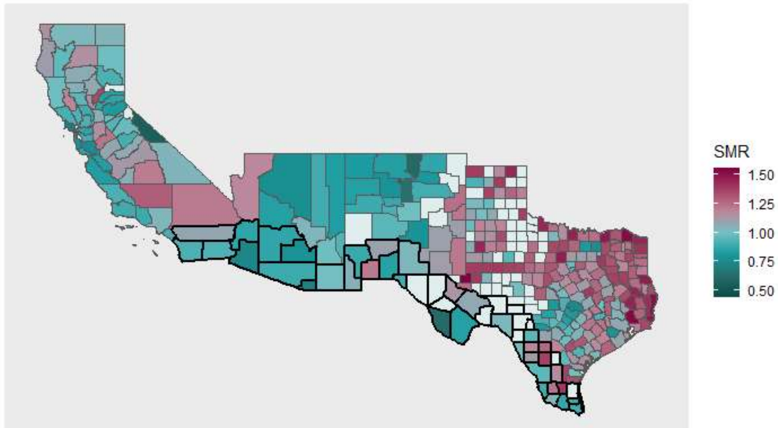


Figure A.2 – Cause of Death: Neoplasms

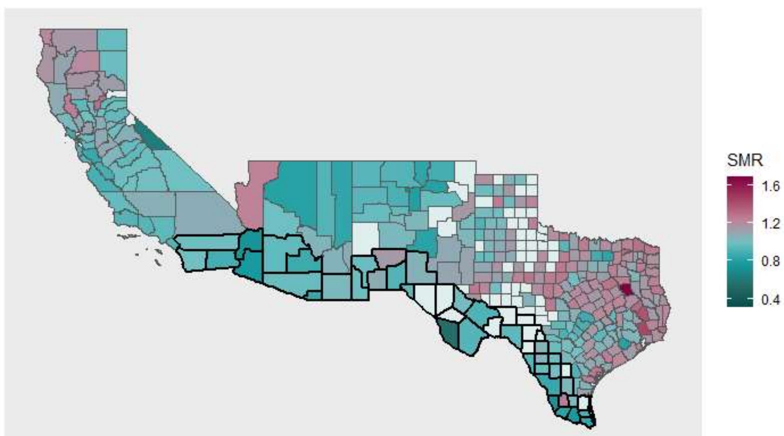


Figure A.3 – Cause of Death: Diseases of the respiratory system

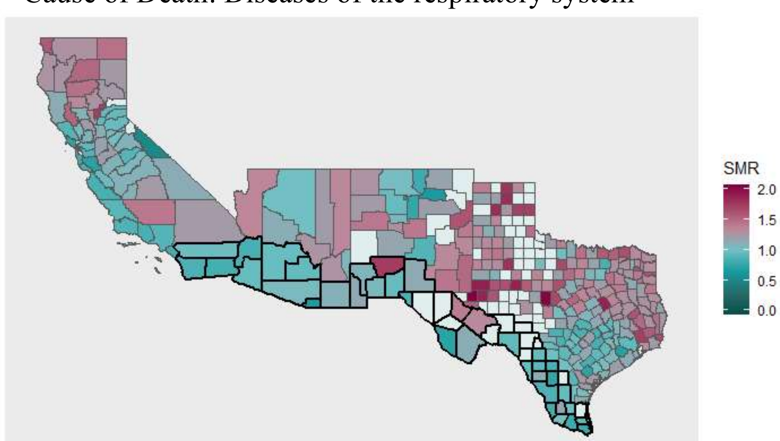


Figure A.4 – Cause of Death: External causes of morbidity and mortality

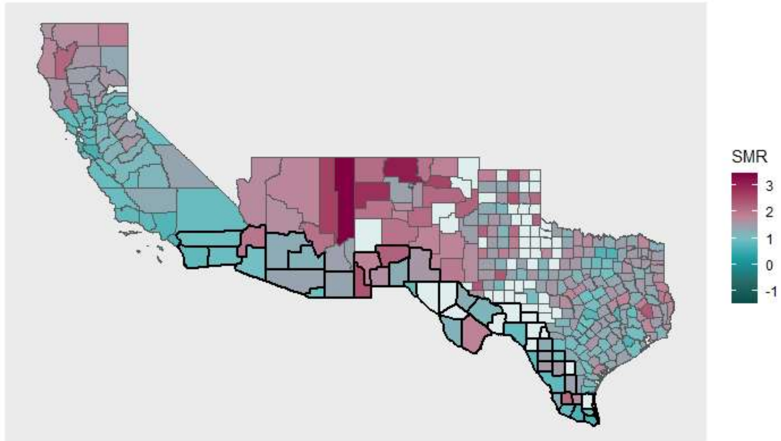


Figure A.5 – Cause of Death: Diseases of the nervous system

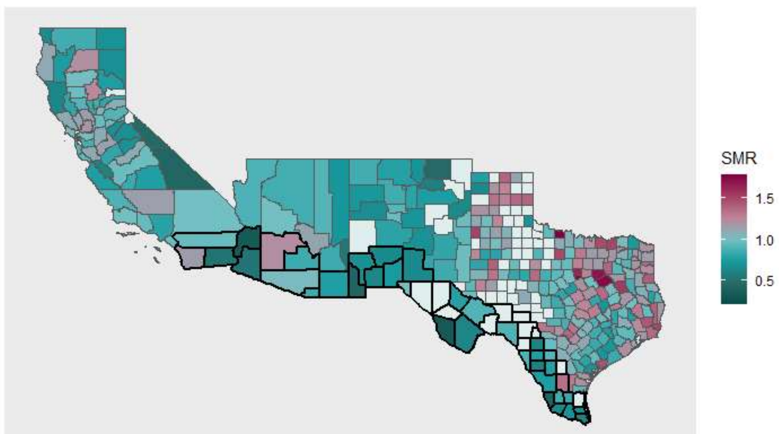


Figure A.6 – Cause of Death: Endocrine, nutritional, and metabolic diseases

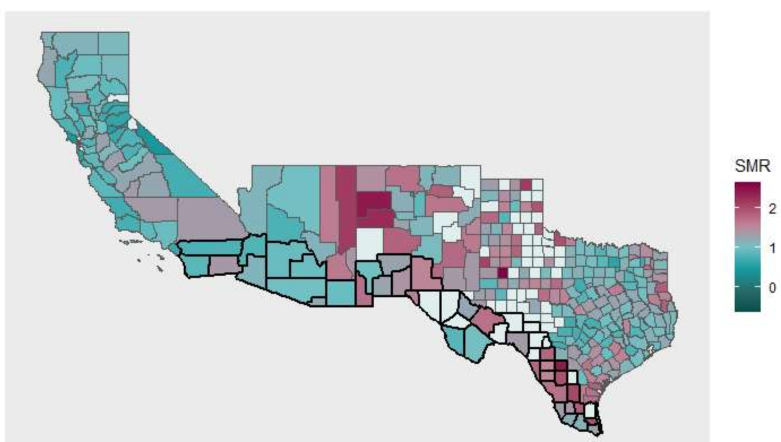


Figure A.7 – Cause of Death: Mental and behavioural disorders

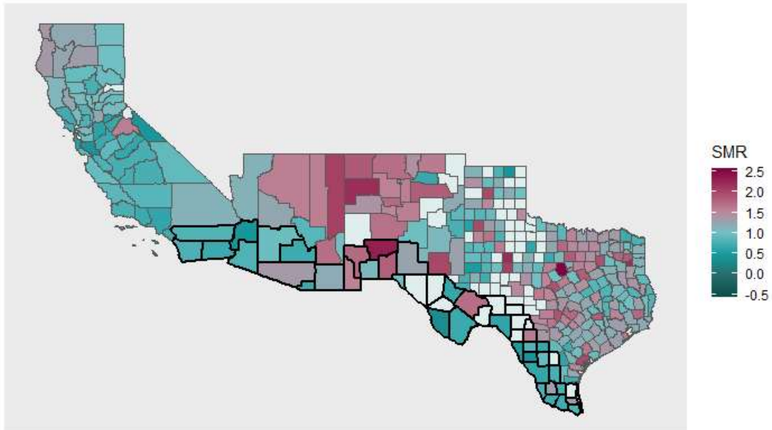


Figure A.8 – Cause of Death: Diseases of the digestive system

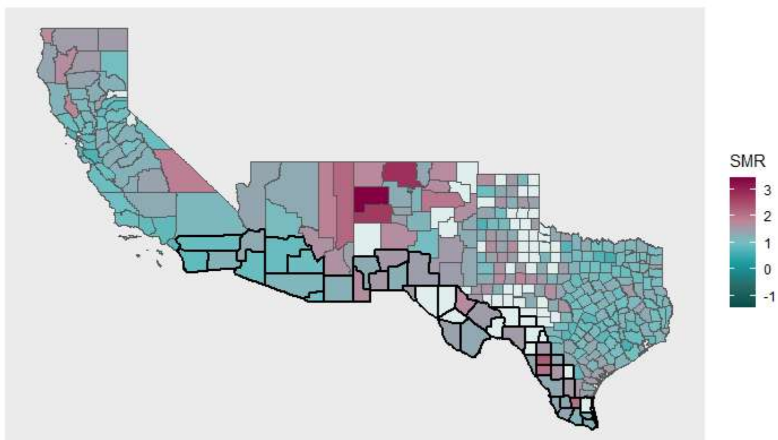


Figure A.9 – Cause of Death: Certain infectious and parasitic diseases

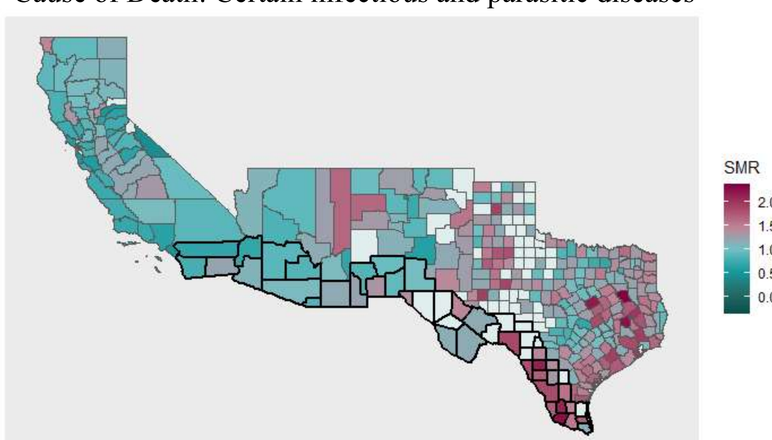
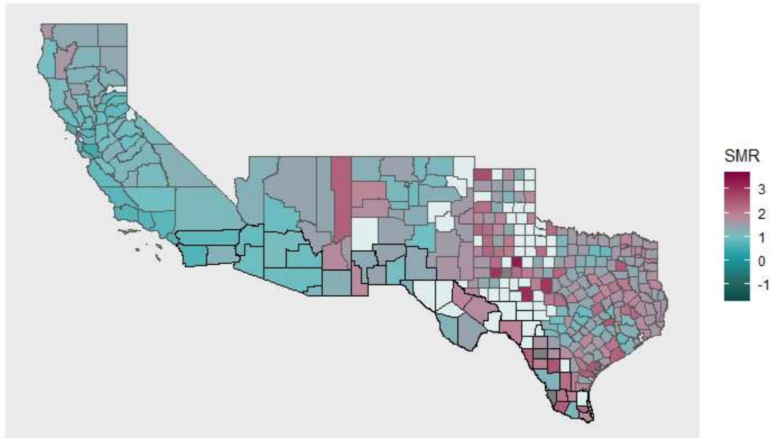


Figure A.10 – Cause of Death: Other cause



APPENDIX B – U.S.-CANADA BORDER REGION MAPS

Figure B.1– Cause of Death: Diseases of the circulatory system

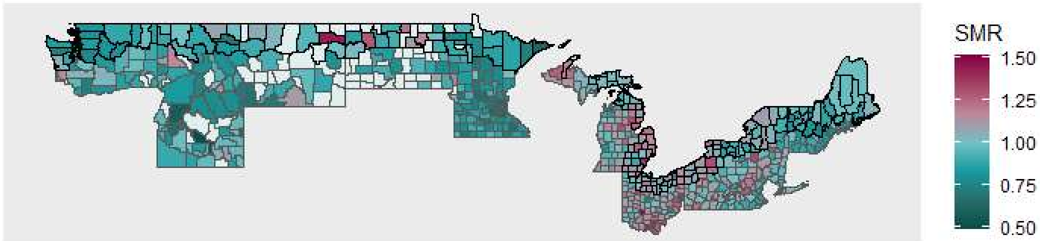


Figure B.2 – Cause of Death: Neoplasms

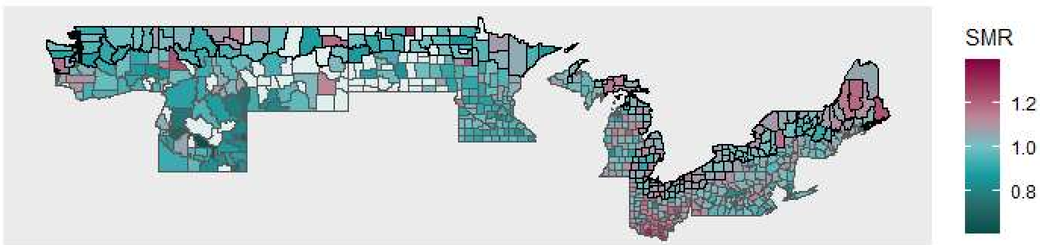


Figure B.3 – Cause of Death: Diseases of the respiratory system

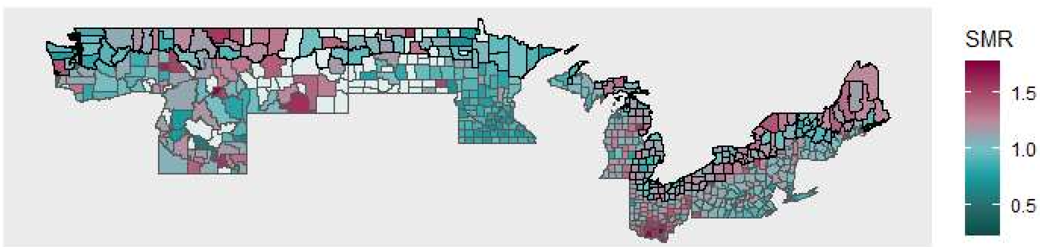


Figure B.4 – Cause of Death: External causes of morbidity and mortality

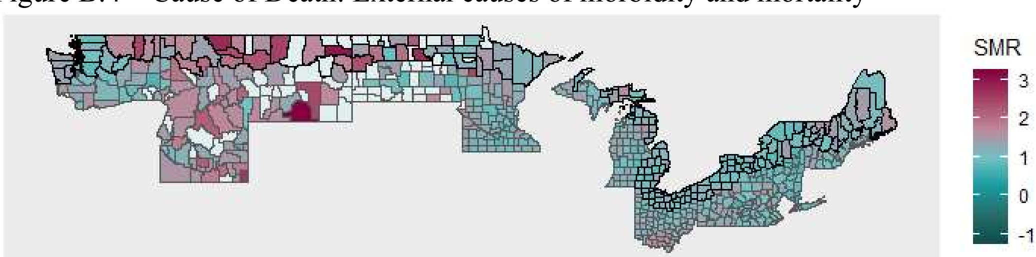


Figure B.5 – Cause of Death: Diseases of the nervous system

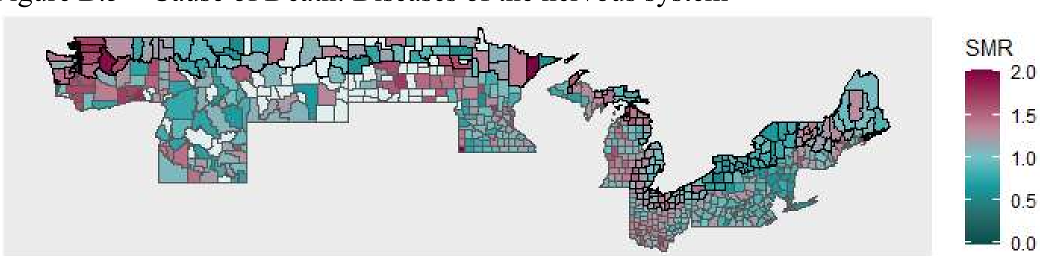


Figure B.6 – Cause of Death: Endocrine, nutritional, and metabolic diseases

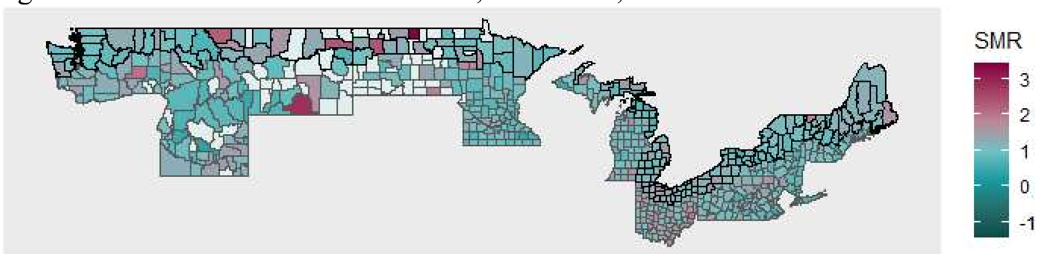


Figure B.7 – Cause of Death: Mental and behavioural disorders

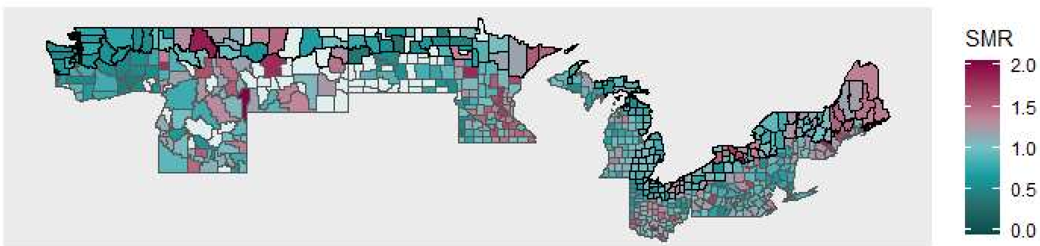


Figure B.8 – Cause of Death: Diseases of the digestive system

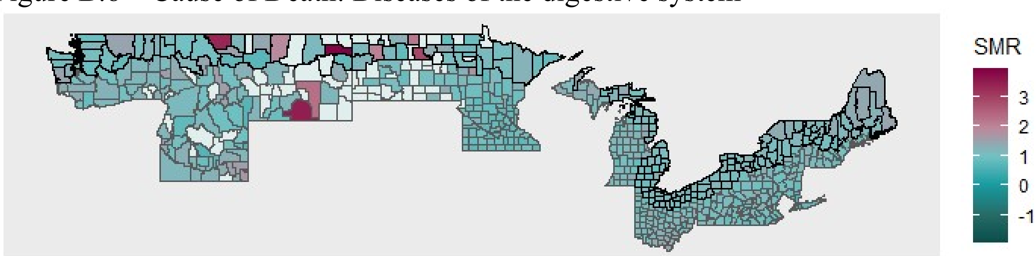


Figure B.9 – Cause of Death: Certain infectious and parasitic diseases

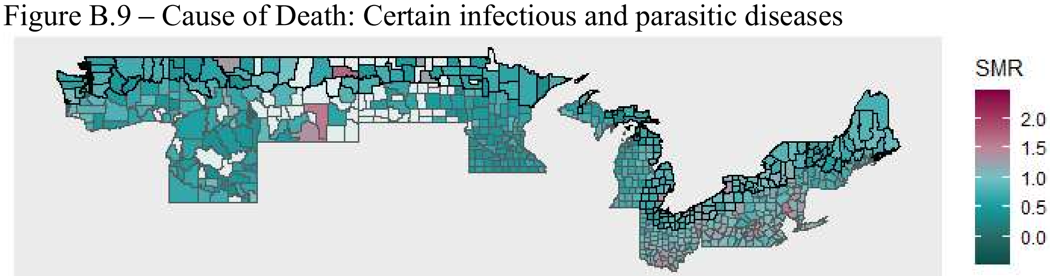


Figure B.10 – Cause of Death: Other Cause