

**Mortality Associated with Extreme Heat in Washington State:
The historical and projected public health burden**

Logan Arnold

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2021

Committee

Tania Busch Isaksen

Mark Scheuerell

Program Authorized to Offer Degree:

Quantitative Ecology and Resource Management

College of the Environment

©Copyright 2021

Logan Arnold

Abstract

Mortality Associated with Extreme Heat in Washington State:

The historical and projected public health burden

Logan Arnold

Chair of the Supervisory Committee:

Tania Busch Isaksen

Department of Environmental and Occupational Health Sciences

Extreme heat is one of the most important pathways illustrating the connection between climate and human health, including in temperate areas such as the Pacific Northwest. Moreover, climate change is expected to exacerbate this important public health issue. This research has two components. First, a time-stratified case-crossover analysis is used to characterize the historical (1980 – 2018) association between summertime (May - September) heat and non-traumatic mortality in Washington state. A separate analysis is conducted for each of the state's climate divisions to produce ten distinct exposure-response curves expressing odds of mortality as a function of humidex. Stratified analyses are used to assess the impact of age, sex, race/ethnicity, and select causes of death, and the reported results are pooled across all climate divisions using fixed-effects meta-analysis. Second, the historical heat-mortality relationship is combined with climate change projections to estimate the impact of a changing climate on heat-related deaths in 2030, 2050, and 2080 under two warming scenarios. The odds ratio and 95% confidence intervals of mortality at the 99th percentile of humidex compared to the 50th percentile did not include the null value in four of the ten climate divisions, and the point estimates in all ten divisions indicated increased risk of mortality at high values of humidex. Across Washington, the odds of mortality are 8% higher (6%, 10%) on 99th percentile days compared to 50th percentile days. Risk is higher for women than men and for Blacks than Whites. Risk also increases with age and for diabetic, circulatory, cardiovascular, ischemic, cerebrovascular, and respiratory deaths. 95% confidence intervals of projected heat-attributable mortality did not overlap with zero in three

of the ten climate divisions. In these three divisions, the average percent increase in heat-attributable deaths across both warming scenarios is 35%, 135%, and 603% in 2030, 2050, and 2080, respectively, over the historical period. This research is the most extensive study of heat-related mortality in Washington to date and can help inform public health initiatives which aim to improve both present and future health outcomes in the state.

Contents

Abstract	3
List of Appendices	6
List of Tables	7
List of Figures	8
Acknowledgments	9
1 Literature Review	10
1.1 Extreme Heat and Human Mortality	10
1.2 Quantifying Risk: Epidemiological and Synoptic Climatological Approaches	14
1.2.1 The Epidemiological Approach	14
1.2.2 The Synoptic Climatological Approach	19
1.3 Applications of the Heat-Mortality Exposure-Response Curve	20
1.3.1 Public Health Initiatives	20
1.3.2 Climate Change Forecasts	22
1.4 Climate Zones	24
1.5 Heat-Related Mortality in Washington State	25
1.6 References	27
2 Mortality Associated with Extreme Heat in Washington State: The historical and projected public health burden	38
2.1 Introduction	38
2.2 Methods	40
2.2.1 Data Sources	40
2.2.2 Exposure Assessment	42
2.2.3 Case-Crossover Analysis for the Historical Period	43
2.2.4 Projected Heat-Attributable Deaths due to Climate Change	44
2.3 Results	45
2.3.1 Historical Public Health Burden	45
2.3.2 Projected Public Health Burden	50
2.4 Discussion	52
2.5 Conclusion	59
2.6 References	60

List of Appendices

Appendix A. Wunderground vs. gridMET Humidex Values	68
Appendix B. Distribution of Humidex (1980-2018)	71
Appendix C. Full Effect Modification Results	73
Age/Cause of Death and Climate Division	74
Age/Cause of Death by Sex	77
Race/Ethnicity	79
Appendix D. Sensitivity Analysis: Case-Crossover vs. Time Series	81
Appendix E. Climate Classifications in Washington	83
Appendix F. Raw Code	85
gridMET Data Acquisition	86
CMIP5 GCMs Data Acquisition	95
Calibration Parameters	109
Mortality Data Pre-Processing	114
Model Selection	134
Case-Crossover Analysis and Historical Heat-Attributable Mortality	139
Meta-Analysis	143
Effect Modification	144
Sensitivity Analysis	148
Projected Heat-Attributable Deaths	154

List of Tables

2.1	Non-traumatic mortality characteristics.	45
2.2	Historical humidex descriptive statistics.	46
2.3	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, by age group, sex, and race/ethnicity.	49
2.4	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, by age group and cause of death.	49
2.5	Number of deaths attributable to heat above the 99 th percentile of humidex in the historical period and projected percent increase due to climate change.	51
A1	Humidex Comparison 1: 1980; SeaTac	69
A2	Humidex Comparison 2: 2001; SeaTac	69
A3	Humidex Comparison 3: 2018; SeaTac	69
A4	Humidex Comparison 4: 2018; Spokane	70
C1	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 0-4 and for each climate division.	74
C2	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 5-14 and for each climate division.	74
C3	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 15-44 and for each climate division.	75
C4	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 45-64 and for each climate division.	75
C5	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 65-84 and for each climate division.	76
C6	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for ages 85+ and for each climate division.	76
C7	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for females for each climate division.	77
C8	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for males for each climate division.	78
C9	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, by race/ethnicity for each climate division.	79
C10	Odds ratio (95% CI) of non-traumatic mortality comparing the 99 th and 50 th percentiles of humidex, for each racial/ethnic category by age, sex, and cause of death.	80
F1	Historical and projected mean humidex values for each climate division.	113
F2	Historical and projected 99 th percentile humidex values for each climate division.	113

List of Figures

2.1	Map of Washington’s climate divisions.	41
2.2	Humidex-mortality exposure-response curves for each of Washington’s climate divisions.	47
2.3	Odds ratio and 95% CI of non-traumatic mortality, comparing the 99 th and 50 th percentiles of humidex.	48
2.4	Spatial distribution of projected humidex anomalies.	50
B1	Spatial distribution of humidex in the historical period.	71
B2	Trend of yearly average and yearly 99 th percentile of humidex in the historical period.	71
D1	Humidex-mortality exposure-response curves for King County, produced using both a case-crossover analysis and a time series analysis.	82
E1	Köppen-Geiger climate classifications in Washington.	84

Acknowledgments

First and foremost, I must thank my research advisor and committee chair, Tania Busch Isaksen. As a student in the Quantitative Ecology and Resource Management (QERM) program, I am not the typical graduate student that Tania advises. Nevertheless, she supported me throughout my research while providing feedback, answering questions, and pushing me to be my best. This project would look very different if not for the time and energy that Tania poured into it. Moreover, I was welcomed also into her lab, the Collaborative on Extreme Event Resilience (CEER), an environment filled with intelligent and mindful students who also helped shape this project. Finally, I would like to thank Mark Scheuerell, the other half of my committee. Mark provided technical feedback for this project and helped improve my ability to think as a statistician and data scientist. I am quite grateful to have had the support and expertise of Tania, the CEER lab, and Mark during the last year.

I would like to thank the QERM Program, the Center for Quantitative Sciences, the College of the Environment, and Tania for providing funding that allowed me to conduct this research.

My QERM cohort (John Henry, Brielle Kwarta Thompson, and Kelly Mistry), my friends, and my family also deserve special recognition. Due to the COVID-19 pandemic, my experience as a graduate student was quite different from what I anticipated - and far more difficult emotionally. I am not sure that this project would have been completed without the support of my classmates in QERM, my friends, and, especially, my family. All of these people supported me from the beginning to the end and believed in my ability to succeed and excel, even during the points in time in which I did not believe that myself. I will never be able to articulate the true extent of my gratitude toward these people.

1 Literature Review

1.1 Extreme Heat and Human Mortality

Environmental health and epidemiological research increasingly produces evidence that illustrates the connection between climate and human health. One of the most important pathways driving this connection is the impact that abnormally hot temperatures have on morbidity and mortality. Kilbourne (1997) notes that the human body responds to extreme heat by increasing cardiovascular demand in an attempt to provide physiologic cooling and maintain thermal homeostasis. A failure to meet this demand leads to a rise in core body temperature that can manifest in several illnesses, including heatstroke, heat exhaustion, heat syncope, and heat cramps; of these possible outcomes, heatstroke is the one most likely to result in death (Kilbourne 1997). In the United States, annual heat-related mortality is higher than mortality due to hurricanes, lightning, tornadoes, floods, and earthquakes combined (Luber and McGeehin 2008).

The relationship between heat and mortality is often nonlinear and characterized by a threshold temperature frequently called the “minimum mortality threshold (MMT) temperature” (Gosling et al. 2009). Specifically, a “U-,” “V-,” or “J-shaped curve,” in which mortality risk decreases as temperature increases from the coldest values before rising again after temperature passes some warm temperature threshold, has been observed in many areas (Curriero et al. 2002; Stafoggia et al. 2006; Busch Isaksen et al. 2015). Heat waves that result in temperatures above the MMT can result in a sharp increase in crude mortality rates (Kilbourne 1997).

Although the United States Center for Disease Control and Prevention (CDC) attributed an average of 658 annual deaths to heat in the United States during 1999-2009 (Kochanek et al. 2011), deaths attributed specifically to heat often underestimate the true public health impact of heat (Kilbourne 1997). *Excess mortality*, i.e. the difference between observed and expected (“baseline”) mortality, is often used to obtain a more accurate description of this health outcome (Kilbourne 1997; Kinney et al. 2008; Gosling et al. 2009). For example, Kalkstein et al. (2011) analyzed 40 U.S. cities and identified approximately 1,300 excess deaths per summer related to extreme heat events between 1975 and 2004. Weinberger and colleagues (2020), in a study area that included 62% of the United States’ population, estimated that an average of 5,609 deaths between 1997-2006 were attributable to heat annually, while Shindell et al. (2020) estimated that, across

the contiguous United States, there have been 12,000 premature deaths associated with high temperatures annually during the past decade.

Typically, diagnosis of heatstroke is based on postmortem body temperature, and the CDC count of heat-related deaths given in Kochanek et al. (2011) is based mostly on such a diagnosis. A lag time between death and medical examination causes some deaths, especially those occurring outside of a hospital, to be misclassified (Kilbourne 1997). Moreover, extreme heat also harms individuals who are already sick, leading to a death that is classified as being caused by the pre-existing condition, since these more common causes of death can appear to be similar (Kilbourne 1997; Basagaña et al. 2011). Indeed, Kalkstein and Davis (1989) observed that extreme summer weather had a stronger relationship with all-cause mortality than with weather-related causes of death. A broader case definition can be used to account for this discrepancy; the Cook County Medical Examiner's Office, for example, included not only body temperature but also deaths not attributable to other causes during its assessment of mortality caused by the 1995 heat wave in Chicago (Disease Control and Prevention 1995).

In addition to the increased risk of death from heatstroke, deaths from cardiovascular diseases (Semenza et al. 1996; Medina-Ramón and Schwartz 2007; Basagaña et al. 2011; Busch Isaksen et al. 2015; Åström et al. 2020), respiratory disease (Basagaña et al. 2011; Fu et al. 2018), and diabetes (Schwartz 2005; Basagaña et al. 2011; Busch Isaksen et al. 2015) have also been shown to increase during a heat wave. The elderly (Kalkstein and Davis 1989; Stafoggia et al. 2006; Bell et al. 2008; Busch Isaksen et al. 2015; Fu et al. 2018), those living without air conditioning (Semenza et al. 1996; Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2005; Medina-Ramón and Schwartz 2007; Madrigano et al. 2018), Blacks specifically (O'Neill, Zanobetti, and Schwartz 2003, 2005; Madrigano et al. 2018) and non-Whites generally (Schwartz 2005), individuals without education beyond high school (Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2003), and outdoor agricultural workers (Spector et al. 2016) have been identified as being most at risk of experiencing injury, illness, and/or death during heat events. Downtown areas within cities are also more vulnerable to heat than suburbs (Reid et al. 2009), though rural populations are not necessarily at less risk than urban populations due to differences in baseline health status and infrastructure, including access to medical facilities and air conditioning (Odame et al. 2018). Note that there is no evidence for a biologic mechanism to explain disparities in heat-health outcomes across racial groups; instead, socioeconomic or sociopolitical status, which highly correlates with race (Jones 2001), is the true risk factor (Kilbourne

1997; O'Neill, Zanobetti, and Schwartz 2003, 2005). For instance, non-Hispanic blacks in New York City are less likely to possess air conditioning (Madrigano et al. 2018), and historical urban housing policies (i.e., redlining) have resulted in communities of color disproportionately inhabiting areas most susceptible to the urban heat island effect (Hoffman, Shandas, and Pendleton 2020). In other words, (systemic) racism produces different health outcomes for racially defined groups (Gravlee 2009; Boyd et al. 2020), including in the context of heat.

Humidity also plays a critical role in the health response to extreme heat, since elevated humidity can reduce the efficiency of thermoregulation and further compound the effects of high temperatures (Kilbourne 1997; Davis, McGregor, and Enfield 2016). As a result, previous studies have used a variety of metrics to quantify the heat-mortality relationship. Some have opted to ignore humidity by analyzing average daily temperature (Curriero et al. 2002) or minimum and maximum temperature (Schwartz 2005; Guo et al. 2011), while others have included humidity in their models through either apparent temperature (Stafoggia et al. 2006; Bell et al. 2008; Zanobetti and Schwartz 2008) or humidex (Jackson et al. 2010; Busch Isaksen et al. 2014, 2015; Spector et al. 2016). Barnett, Tong, and Clements (2010) analyzed the heat-mortality relationship in 107 U.S. cities using seven temperature measures – four that incorporated humidity and three that did not. They determined that there was no clear superior metric, and they suggest using practical considerations (for example, the metric with the least amount of missing data) to guide the variable choice, since these (often) highly correlated measures tend to provide similar predictive ability. In their systematic review, Xu and colleagues (2016) similarly found that no superior temperature indicator exists to assess mortality associated with heat events.

Notably, the comparison of exposure metrics by Barnett, Tong, and Clements (2010) did not include the spatial synoptic classification procedure, a more holistic meteorological assessment that allows for simultaneous and synergistic evaluation of numerous weather elements as they appear within air masses (Kalkstein 1991). For instance, wind speed and solar radiant heat energy, which also impact thermal homeostatic mechanisms (Kilbourne 1997), can be considered alongside the more common variables of temperature and humidity when using this approach. Applications of this methodology are often used to compare weather-related health outcomes across a large region such as the United States (Kalkstein and Greene 1997; Sheridan and Kalkstein 2010; Greene et al. 2011). However, synoptic classification has also been explored in studies of a smaller spatial scale (Busch Isaksen et al. 2015; Fonseca-Rodríguez et al. 2020).

Crucially, even if there is no clear superior metric to use when quantifying the heat-mortality relationship (Barnett, Tong, and Clements 2010), comparison studies (Hajat et al. 2010; Zhang et al. 2012) have shown that different exposure metrics lead to different definitions of “hot days” or “extreme heat events” and, consequently, a (potentially) different number of dangerous weather events and/or days classified as dangerous by one metric but not by another. In practice, many different definitions for “heat event” have been used to quantify the heat-mortality relationship. Examples include a comparison of mortality at an extreme value that could cause physiologic stress to moderate weather (O’Neill, Zanobetti, and Schwartz 2003; Stafoggia et al. 2006), a binary definition that utilizes a threshold value such as the 95th or 99th percentile of the exposure metric (Jackson et al. 2010; Wilson et al. 2013; Busch Isaksen et al. 2015), and a continuous variable that attempts to mimic the nonlinear exposure-response curve and describe the impacts associated with single-degree changes (Medina-Ramón and Schwartz 2007; Busch Isaksen et al. 2014).

Xu et al. (2016) argue that different heat wave definitions cause variation in the magnitude of effect estimates, though the direction of the effect remains the same; importantly, results from this review are potentially impacted by heterogeneity in study design and region. Wilson et al. (2013) conducted sensitivity analyses in which they applied the same statistical methodology (conditional logistic regression) to the same geographic region (Sydney, Australia) and obtained results that were broadly similar across heat wave definition. However, regardless of what definition is used, the mortality increase caused by exposure to high temperatures is likely not limited to so-called “heatwaves” (Zanobetti and Schwartz 2008).

Kalkstein and Davis (1989) provide evidence that the timing of high temperatures is important, as intra-seasonal acclimatization can reduce the impact of extreme weather that occurs late in the season, and that heat events longer in duration may be more dangerous than ones that are shorter but more intense. Likewise, vulnerability to late summer and autumn heat events has been shown to be lower than similar heat events in the spring (Sheridan and Kalkstein 2010). On the other hand, a systematic review of heatwave definitions (Xu et al. 2016) suggests that heatwave intensity has a more important impact on heat-related mortality than heatwave duration. Reviews conducted by Basu and Samet (2002) and Gosling et al. (2009) suggest that mortality related to heat events tends to have a maximum “lag” time of 3 days.

1.2 Quantifying Risk: Epidemiological and Synoptic Climatological Approaches

Gosling et al. (2009) identify two main methodologies to study the temperature-mortality relationship: the epidemiological approach and the synoptic climatological approach. However, the epidemiological approach is employed much more frequently, especially in the last decade.

1.2.1 The Epidemiological Approach

Broadly, the epidemiological approach attempts to explain an outcome variable (e.g. mortality) using predictor(s) (e.g. temperature) and, potentially, confounding variables (Basu and Samet 2002; Bhaskaran et al. 2013). The two most common methods to perform this task when studying heat-related health outcomes are regression models applied to time series data and the case-crossover study design (Gosling et al. 2009).

A time series analysis is commonly employed in epidemiological applications when the research goal is quantifying short-term associations of environmental exposures with health outcomes (Bhaskaran et al. 2013), including investigations of the heat-mortality relationship (Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2003, 2005; Shaposhnikov et al. 2014; Busch Isaksen et al. 2015; Heo et al. 2016). The unit of analysis (i.e. row of data) in time series regression studies is time itself (e.g. day, week, month, or year), while the outcome of interest is a count (e.g. of deaths, cases, or hospitalizations). This data structure necessitates that a form of count regression be used. Typically, Poisson regression is utilized, perhaps in the form of a quasi-Poisson model that allows for “over-dispersion” in the response variable (e.g. Zanobetti and Schwartz 2008). These models can also be used for rate data by including population size as an “offset” variable (e.g. Busch Isaksen et al. 2015).

A general additive model (GAM) framework is often applied to account for the non-linear relationship between heat and mortality. In GAMs, time and the environmental exposure variables of interest are typically included as predictors in the models using splines (Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2003, 2005; Heo et al. 2016; Wellenius et al. 2017), though some authors have chosen to approximate the spline with a piecewise linear model (i.e. a time- or exposure-stratified model with indicator variables for different strata) to improve interpretation and usefulness for public health officials and policymakers (Busch Isaksen et al. 2015).

GAMs are popular for their relative simplicity, but there are other possible modeling frameworks, including distributed lag nonlinear models [DLMNs; [Gasparrini, Armstrong, and Kenward \(2010\)](#)] and transitional regression models [TRMs; [Brumback et al. \(2000\)](#)]. The former can be implemented to study the impact of lagged variables on the nonlinear exposure-response relationship ([Guo et al. 2011](#)), while the latter has been proposed as a methodology to extend GLMs in a manner allowing for autocorrelation.

The use of population-level data in time series regression mean that many “standard” confounders in epidemiology, such as age and sex, do not apply, since these are unlikely to change during the unit of time used in analysis; potential confounders, therefore, are variables that do change rapidly and are likely related to the outcome, the predictors, or both ([Bhaskaran et al. 2013](#)). Humidity could confound results if only air temperature is included as a predictor ([Barnett, Tong, and Clements 2010](#)), and some authors have explored potential confounding by ambient air pollution ([Zanobetti and Schwartz 2008](#)) and wildfire-related air pollution ([Shaposhnikov et al. 2014](#)). To analyze effect modification by factors such as age, sex, race, and education, separate regression models must be fit to each category of interest ([O’Neill, Zanobetti, and Schwartz 2003](#); [Busch Isaksen et al. 2015](#)).

In recent years, the case-crossover study design has become a popular alternative to time series regression in environmental epidemiology. [Maclure \(1991\)](#) first introduced the study design to assess the impact of a brief exposure (“trigger”) on the change in risk of a rare acute-onset disease, such as myocardial infarction. The methodology has since been used in studies of not only heat-related mortality ([Stafoggia et al. 2006](#); [Medina-Ramón and Schwartz 2007](#); [Bell et al. 2008](#); [Basagaña et al. 2011](#); [Guo et al. 2011](#); [Wilson et al. 2013](#); [Fu et al. 2018](#); [Åström et al. 2020](#); [Fonseca-Rodríguez et al. 2020](#)), but also health effects related to particulate matter exposure during wildfires ([Alman et al. 2016](#); [Doubleday et al. 2020](#); [Yao et al. 2020](#)).

A case-crossover study design is akin to the traditional matched case-control design since selection is based on the outcome; however, like a crossover study, each case serves as its own control. Investigators can infer the risk of a health outcome by comparing a subject’s exposure during a time period relevant to the event (“hazard period”) to one or more control periods (“referent periods”) lacking the event ([Jaakkola 2003](#)). Matching in a case-crossover design, therefore, is an extreme form of stratified sampling in which there is only one individual per stratum; the choice to compare exposure distributions for the same case rather than comparing a case to one or more controls eliminates control-selection bias that can occur in case-

control studies ([Maclure 1991](#)). Conditional logistic regression is the typical quantitative technique used to compute the measure of effect in case-crossover designs ([Jaakkola 2003](#)), but conditional Poisson regression ([Armstrong, Gasparini, and Tobias 2014](#)) can also be used.

Choosing the duration and timing of the hazard and referent periods is a challenge when implementing the case-crossover design, and proper selection has important consequences with respect to potential methodological biases ([Janes, Sheppard, and Lumley 2005](#)). There must be variation in exposure between the hazard period and the control period(s) for the strata to contribute information to the analysis ([Mittleman and Mostofsky 2014](#)). When studying an environmental exposure such as heat, a benefit of case-crossover design is that exposure levels after an event are unlikely to be related to the event itself, so post-event control periods can be used ([Bateson and Schwartz 1999](#)). In studies of heat-related mortality, many authors take advantage of this characteristic by employing time-stratification in which the hazard period (day of death) is matched to the same day of the week of the same month and year ([Stafoggia et al. 2006](#); [Bell et al. 2008](#); [Basagaña et al. 2011](#)). Simulations by both Bateson and Schwartz (1999) and Levy et al. (2001) illustrate that this methodology controls for temporal confounding (day of week, seasonal, and/or long-term trends) by design (i.e. via matching), rather than by statistical modeling; the latter study further observes that increasing the number of control periods is an effective way to increase statistical efficiency.

In addition to the time-stratification technique described above, other referent selection strategies have also been employed. Heat-mortality studies have utilized a unidirectional and ambidirectional analysis ([Basu, Dominici, and Samet 2005](#)) as well as control days that occur every third day in the same month and year as the event ([Medina-Ramón and Schwartz 2007](#); [Zanobetti and Schwartz 2008](#)). Other environmental health applications have chosen to match by day over the entirety of a 32-day study period ([Alman et al. 2016](#)) or to use hourly (rather than daily) case and control windows ([Yao et al. 2020](#)). However, Levy et al. (2001) and Janes, Sheppard, and Lumley (2005) demonstrate that time-stratification, when compared to other referent strategies, leads to unbiased conditional logistic regression estimates, particularly by eliminating so-called “overlap bias” via the use of disjoint strata. As a result, time-stratification tends to be the preferred referent strategy. Regardless of which strategy is implemented, it is important to choose referent periods close to the event, since this helps eliminate possible selection bias ([Bateson and Schwartz 2001](#)). Furthermore, it is important to consider potential exposure autocorrelation between control period(s); although this autocorrelation does not introduce bias, it does reduce statistical efficiency ([Mittleman and Mostofsky 2014](#)).

Several studies have compared the use of time series regression to case-crossover designs in both in environmental health applications specifically and epidemiology generally. Maclure and Mittleman (2008) observe that both study designs aim to answer the question “Why now?” and not the question “Why them?” that is the subject of case-control studies. Lu and Zeger (2007) prove that the case-crossover approach is not an alternative to log-linear time series analysis but rather is the special case in which there is common exposure across subjects in each time period (e.g. in air pollution or heat studies). As a special case, Levy et al. (2001) show that using time-stratification in the case-crossover design and Poisson regression with dummy variables for the strata produce the same estimates. Subsequent work (Lu et al. 2008) uses this equivalence to propose a series of diagnostic tools to verify the modeling assumptions made in case-crossover analyses. Due to the differences in the statistical models used in each approach, it is common for the exposure-outcome relationship to be interpreted in terms of a relative risk/risk ratio when using time series regression, while an odds ratio typically describes the exposure-outcome relationship in case-crossover studies. In theory, this discrepancy can complicate comparison between the two methods, but the two estimates are often directly comparable in heat-mortality studies due to the rare disease assumption, since mortality is a rare event (Basu, Dominici, and Samet 2005).

The equivalence of time-series and case-crossover procedures indicates that the two methods should, in theory, produce similar results (Lu and Zeger 2007), with the caveat that both procedures require accurate decision-making on the part of the researcher to produce unbiased results (Fung et al. 2003). Carracedo-Martínez et al. (2010) argue that case-crossover designs require fewer arbitrary decisions: when using GAMs in time series regression, one must choose the type of smoother and the number of degrees of freedom to model time, while the case-crossover approach only requires one to select a referent strategy. (Note, however, that if the exposure metric is incorporated into a case-crossover analysis as a continuous variable, then the researcher must make similar decisions regarding the type of spline or polynomial to use.) Studies that have applied both methods verify that the two techniques generate results that are comparable quantitatively and identical qualitatively (Zanobetti and Schwartz 2008; Tong, Wang, and Guo 2012). Basu, Dominici, and Samet (2005) further demonstrate that time-stratification yields results that are less biased and more consistent.

Notably, in contrast to time series regression, a case-crossover analysis uses individuals (not days) as the unit of observation, and each individual is their own control; this eliminates confounding by constant

or slow-varying individual characteristics such as age, sex, and some behavioral factors (Maclure 1991). As a result, an advantage of the case-crossover approach is its suitability for exploring effect modification and individual susceptibility both in theory (Jaakkola 2003) and in practice (Stafoggia et al. 2006; Medina-Ramón and Schwartz 2007; Bell et al. 2008). This also makes the case-crossover design preferable when “denominator data” (i.e. population size) is small or unknown (Spector et al. 2016). Moreover, it is possible to include variables in conditional logistic regression models to control for potential confounding caused by features that are changing rapidly, including air pollution (Stafoggia et al. 2006; Bell et al. 2008; Basagaña et al. 2011).

Another desirable quality of the case-crossover methodology in the context of heat-mortality studies is that it permits the exploration of a nonlinear relationship between the environmental exposure and the health outcome. A penalized cubic spline (Stafoggia et al. 2006), a natural cubic spline (Bell et al. 2008; Guo et al. 2011), and a quadratic B-spline (Fu et al. 2018) have been used to model the exposure within the GAM framework. Furthermore, multiple studies have assessed both nonlinear and lagged effects of temperature on mortality with DLNMs (Guo et al. 2011; Fu et al. 2018; Fonseca-Rodríguez et al. 2020). However, in their supplemental material, Fu et al. (2018) demonstrate that frequentist spline-based methods to model the nonlinear relationship between temperature and mortality are sensitive to the number and placement of knots. Stringer, Brown, and Stafford (2020) have proposed a Bayesian approach to address this sensitivity and provide model-based estimates and uncertainty quantification for the smoothness of effects, but the methodology has yet to be applied in a novel setting. Notably, as in time series regression studies, the nonlinear heat-mortality relationship is sometimes reduced to a piecewise linear model (Stafoggia et al. 2006; Medina-Ramón and Schwartz 2007).

When compared to time series regression, the case-crossover design has some drawbacks. First, the price to pay for controlling for temporal confounding by design is reduced statistical efficiency (Bateson and Schwartz 1999). Using a longitudinal approach instead of conditional logistic regression has been proposed to improve statistical efficiency in the case-crossover design, though this methodology could lead to biased estimates if autocorrelation is present (Figueiras et al. 2005). Second, the favored quantitative technique of conditional logistic regression offers no way to allow for response data that is over-dispersed, and this can lead to biased values for standard errors and confidence intervals if the Poisson assumption is violated (Lu and Zeger 2007; Lu et al. 2008). Armstrong, Gasparri, and Tobias (2014) propose the use of conditional

Poisson models as an alternative to conditional logistic regression that can allow for over-dispersion and address autocorrelation using the framework presented by Brumback et al. (2000). Although conditional Poisson models yield results that are equivalent to those obtained using conditional logistic regression, provided that there is a common exposure across individuals (Armstrong, Gasparrini, and Tobias 2014), extant application of conditional Poisson regression is limited (Åström et al. 2020; Fonseca-Rodríguez et al. 2020).

1.2.2 The Synoptic Climatological Approach

While both time series regression and the case-crossover study design allow for the inclusion and modeling of confounding variables, the synoptic climatological approach instead focuses purely on meteorological and mortality data (Gosling et al. 2009). First proposed by Kalkstein (1991), the primary goal of the synoptic climatological approach is to identify “oppressive” or “high-risk” air masses that are associated with excess mortality by using a method such as principal component analysis (PCA) or cluster analysis (CA) to evaluate numerous weather elements simultaneously and realistically. Simultaneous evaluation of multiple variables avoids the choice of a single weather element that may not be representative of the total effect that weather has on mortality (Kalkstein 1991) and it allows for a synergistic evaluation in which the combined impact of multiple factors is greater than the sum of their individual impacts (Kalkstein and Greene 1997). Additionally, this approach may allow for better distinction between mortality caused by heat and mortality caused by air pollution (Kalkstein 1991).

While it is relatively straightforward to identify oppressive air masses, it is important to note that days experiencing these weather patterns tend to be characterized by a large standard deviation in daily mortality (Kalkstein et al. 1996). In other words, while there are clearly days with excess mortality associated with certain types of air masses, there are also many days in which the same air masses lead to mortality levels that are at or below the baseline. Stepwise multiple regression analysis can be used to identify the specific factors (both meteorological and non-meteorological) within these air masses that contribute to excess mortality, including the number of consecutive days the air mass has been present, the maximum temperature, and the time of season (Kalkstein et al. 1996; Kalkstein and Greene 1997).

In its initial formulation (Kalkstein 1991), the synoptic climatological approach was limited in that it was location-specific; an air mass identified as oppressive in one area may be benign in another. Kalkstein

and Greene (1997) updated the methodology by assessing major air masses that traverse a particular region. This new approach, called spatial synoptic classification, permits comparison between locations, including on the scale of the United States (Sheridan and Kalkstein 2010; Greene et al. 2011).

1.3 Applications of the Heat-Mortality Exposure-Response Curve

Anthropogenic activities are causing average global temperatures to rise (IPCC 2013) as well as the frequency, length, and intensity of future extreme heat events to increase (Meehl and Tebaldi 2004). Climate change may also impact the frequency of high-risk air masses during the summer (Kalkstein and Greene 1997). Furthermore, the areas within the contiguous United States that have the highest social vulnerability also appear to be warming faster than less vulnerable areas (Spangler and Wellenius 2020). As a result, knowledge about the relationship between high temperatures and human health is becoming increasingly important as this information is critical to both the immediate public health response and inquiries into future health outcomes.

1.3.1 Public Health Initiatives

By quantifying the impact of exposure to extreme temperatures, policymakers can develop preventative, life-saving measures, warning systems, and action plans that target susceptible groups, support health systems, and engage the community at large. For example, a heat-mortality relationship obtained using a synoptic climatological approach guided the development and implementation of the Philadelphia Hot Weather-Health Watch/Warning System, and the system generally predicted heat-related deaths at times when they actually occurred (Kalkstein et al. 1996). Ebi et al. (2004) estimate that this system saved 117 lives during a three-summer period. In addition to the synoptic methodology employed in Philadelphia, epidemiological approaches have been used to guide heat-health warning systems in Europe, though this has led to a discrepancy across systems in the choice of the exposure variable used to define heat waves (Lowe, Ebi, and Forsberg 2011; Casanueva et al. 2019). Days identified by an epidemiologic assessment of the temperature-mortality relationship have been shown to have the best association with excess mortality when compared to other techniques (Hajat et al. 2010). Independent of how they are developed and operated, all early warning systems must be able to forecast potential heat events, predict possible health consequences associated with

extreme heat, activate response plans in a timely manner, and effectively communicate pertinent information prior to the onset of temperatures about the MMT (Hajat et al. 2010; Lowe, Ebi, and Forsberg 2011).

Both Poisson regression (Petitti et al. 2016; Wellenius et al. 2017; Vaidyanathan et al. 2019) and case-crossover analysis (Åström et al. 2020) have been used to suggest changes to existing heat-warning systems. The use of local data rather than national or regional data in public health initiatives may further improve health outcomes during extreme heat events (Wellenius et al. 2017; Vaidyanathan et al. 2019; Åström et al. 2020), while multiple “trigger points” at which heat-health interventions are activated may help to target end-users of specific interventions more effectively (Petitti et al. 2016). Poisson regression has also been used to evaluate the effectiveness of a heat warning system (Nitschke et al. 2016).

A variety of quantitative methods have been employed to create a “heat vulnerability index” that illustrates how vulnerability changes due to variation in place and people (Bao, Li, and Yu 2015). Reid et al. (2009), for instance, used PCA to map heat vulnerability in major urban centers throughout the United States; this investigation identified downtown urban centers as particularly vulnerable to heat when compared to the rest of the metropolitan area, an important consideration when developing municipal-level public health initiatives.

Despite a warming climate, Kalkstein et al. (2011) observed a general decline across 40 U.S. cities in mortality rates attributable to extreme heat events since 1996 when compared to rates from 1975-1995. A review of research published between 2012 and 2015 (Hondula et al. 2015) also found a downward trend of negative health outcomes associated with heat events. Kalkstein et al. (2011) argue that that the reduced mortality rates can be attributed, at least in part, to resources related to education, notification, and response measures associated with heat-related health outcomes in many of the 40 cities, since Philadelphia implemented its heat warning system in 1995, and numerous other cities soon developed similar programs. Similarly, Heo et al. (2016) also observed lower mortality risk among the elderly in South Korea in 2008-2012 compared to 1996-2000 and suggest that the reduction can be attributed to the introduction of a heatwave forecast system in 2008 that was targeted at elderly individuals.

908 survey respondents aged 65+ in four North American cities provide further evidence for the importance of local public health efforts, as 90% were aware of their city’s heat warning system (Sheridan 2006). Notably, this survey also elucidated that most individuals tend to decide “on their own” if it is hot, so a major

area for improvement is education regarding early season vulnerability that occurs when temperatures are “warm” and not “hot” (Sheridan and Kalkstein 2010).

In contrast to the argument made by Kalkstein et al. (2011), Weinberger et al. (2018), in a study of 20 U.S. cities, argue that heat alerts were not associated with lower mortality rates in 19 of the cities (Philadelphia being the exception); however, it is important to acknowledge that there were only a small number of cities included in both studies, and the prevalence of heat response plans in the cities exclusive to the Weinberger et al. (2018) analysis was estimated to be low. Nitschke et al. (2016) also failed to observe a reduction in heat-related mortality following the implementation of a preventative program in South Australia, though morbidity outcomes were reduced significantly when comparing heat waves that occurred pre- and post-implementation.

Systematic reviews of the public health response to heat events (Bassil and Cole 2010; Boeckmann and Rohn 2014; Vu, Rutherford, and Phung 2019) corroborate individual findings that actions such as providing educational resources and implementing heat-health warning systems are associated with certain improved heat-health outcomes, but they also provide three important caveats. First, although the public is often nearly universally aware of heat events, few people understand that such warnings are connected to negative health consequences and therefore should be associated with behavioral changes; there may be a particular failure to target and educate individuals in vulnerable groups, including the elderly. The reviews also describe the difficulties that emerge when attempting to evaluate the effectiveness of public health responses in the context of heat-health outcomes. These difficulties, they argue, complicate a direct causal association between improved public health outcomes and individual adaptation measures. Lastly, extant work in developing countries is lacking.

1.3.2 Climate Change Forecasts

Enumerating the heat-mortality relationship also helps to inform projections of future heat-related deaths in light of climate change. Though there is existing work on the broad implications of climate change on human health (Patz, Grabow, and Limaye 2014), there are limited studies projecting heat-related health impacts attributable to climate change.

In its simplest form, projections of future excess mortality in a changing climate apply a historical

heat-mortality relationship to forecasts of future meteorological data (Hondula et al. 2015). Complexity quickly increases when trying to produce more accurate calculations due to spatio-temporal variation in heat-mortality relationships, uncertainty in demographic and climate projections, and an unknown rate of adaptation and acclimation to the increased frequency, duration, and intensity of heat events (Kinney et al. 2008; Gosling et al. 2009). However, the consideration of such details is necessary because increased urbanization and an aging population could amplify heat-related mortality (Luber and McGeehin 2008), while public health responses (Greene et al. 2011) as well as continued acclimatization and adaptation to heat (Medina-Ramón and Schwartz 2007; Knowlton et al. 2007) could reduce heat-related mortality [though results from Evangelopoulos et al. (2021) imply that populations may not readily adapt to climate change].

Despite different study areas and methodological choices, projections of heat-related health outcomes in future climate scenarios tend to converge to the same qualitative result. A systematic review (Huang et al. 2011) identified only 14 projections of future heat-related deaths due to climate change; nearly all studies predicted that climate change would significantly increase mortality due to heat events. A multi-country modeling study (Guo, Gasparrini, and Shanshan 2018) that considered a variety of global circulation models, greenhouse gas trajectories, population projections, and potential adaptations in 20 countries/regions also predicted heat-related excess mortality to increase under almost every combination of criteria. Shindell et al. (2020) forecast nearly 100,000 annual premature heat-related deaths in the contiguous United States by the end of the century under RCP8.5, though this number is much lower under moderate warming scenarios; independent of RCP, the incorporation of adaptation may reduce impacts by approximately 40-45%. While areas with a large number of cold-related deaths may see some public health benefits from a warmer climate, the net effect of climate change is expected to cause an increase in weather-related mortality (Kalkstein and Greene 1997; Medina-Ramón and Schwartz 2007; Weinberger et al. 2017).

Recently, a tutorial has been developed to facilitate projections of health outcomes under climate change scenarios (Vicedo-Cabrera, Sera, and Gasparrini 2019), and a separate comparative analysis has considered the specific problem of modeling adaptation to climate change (Gosling et al. 2017). Nevertheless, there remains limited application of these methods to support local and regional management of heat exposure and climate change adaptation.

1.4 Climate Zones

An important limitation of most existing heat-mortality work, independent of the quantitative method used to enumerate the exposure-response curve, is the use of geopolitical units for analysis (e.g. counties) that may not capture the true spatial variation in health outcomes that occur during heat events and thus bias the results. Instead, spatial aggregation based on climatic patterns (e.g. temperature, rainfall) may be more appropriate, since individuals tend to adapt to their local weather conditions ([Medina-Ramón and Schwartz 2007](#)). Possible frameworks to use for this approach include local climate zones ([Stewart and Oke 2012](#)), NOAA's Climate Divisions ([Guttman and Quayle 1996](#)), and the Köppen-Geiger climate classification ([Kottek et al. 2006](#)); the spatial extent of the analysis should guide the selection of a specific climatological aggregation definition.

Knowledge about the heat-mortality relationship using climate zones or climate regions rather than geopolitical units has the potential to provide better guidance for policymakers attempting to mitigate heat-related deaths, including through the development of heat warning systems ([Vaidyanathan et al. 2019](#); [McElroy et al. 2020](#)). Spatial aggregation defined not by geopolitics but rather climatic conditions could be especially important since there may be profound differences in the increased frequency of future summer heat events across local climate zones within cities ([Geletič, Lehnert, and Dobrovolný 2019](#)). If necessary, for either scientific or political purposes, it is possible to pool region-specific estimates via meta-analysis ([Gasparri, Armstrong, and Kenward 2012](#); [Chang and Hoaglin 2017](#)) to obtain a single effect estimate for the entire study area ([Basagaña et al. 2011](#); [Fu et al. 2018](#)).

While the use of climate zones to aggregate medical and meteorological data in heat-mortality studies has the potential to yield more accurate and useful results, the approach is not widespread. [Wilson et al. \(2013\)](#) controlled for climate zones in the Sydney Greater Metropolitan Region via a dummy variable in a conditional logistic regression model and still found an association between high temperatures and mortality. Using climate zones to define exposure, [Guirguis et al. \(2018\)](#) detected a significant increase in hospitalizations during hot weather in comparison to hospitalizations observed over all days, independent of any temperature threshold, in the coastal region of San Diego County, but not in communities further inland that are characterized by higher average temperatures and greater air conditioning prevalence. Regional GAM analyses in South Korea ([Heo et al. 2016](#)) as well as regional case-crossover analyses in India ([Fu et al. 2018](#))

also yielded distinct temperature-mortality relationships for different climate regions. Evangelopoulos et al. (2021) applied different DLNMs to London residents born in different climate zones and observed different exposure-response associations for each group of inhabitants. However, Basagaña et al. (2011) failed to detect heterogeneity in the heat-mortality relationship across Catalonia's 14 climatic zones, and chose instead to focus on a single meta-analysis.

1.5 Heat-Related Mortality in Washington State

Washington state is an important research area for studies that both quantify the current heat-mortality relationship and forecast heat-related deaths in light of climate change. Temperate regions (for example, in the United States, the Pacific Coast, the Midwest, and the Northeast) are more sensitive to heat events (i.e., they are characterized by a lower MMT temperature), perhaps due to less appreciation of risk, lower prevalence of air conditioning, or weaker physiologic acclimatization (Kalkstein and Davis 1989; Kalkstein and Greene 1997; Curriero et al. 2002; Medina-Ramón and Schwartz 2007; Reid et al. 2009; Sheridan and Kalkstein 2010; Vaidyanathan et al. 2019; Shindell et al. 2020). The Pacific Northwest may also experience greater summertime warming due to climate change than most of the rest of the continental United States (Šeparović et al. 2013). Moreover, while this section of the country historically has experienced few severe heatwaves, climate models suggest that it will become increasingly susceptible to extreme heat events (Meehl and Tebaldi 2004). For example, a regional climate model projection for Washington state forecasts that south-Central Washington could experience up to 3 additional heatwaves in future summers compared to the control period (Salathé Jr. et al. 2010).

Jackson et al. (2010) examined four areas in Washington state and found that projected non-traumatic deaths due to heat days in the Seattle region were greatest for people in the 65+ age group in all years and warming scenarios analyzed, though the results for the rest of the state were not statistically significant. Busch Isaksen et al. (2014) built upon this work and computed a 1.83% increase in mortality for all ages, all non-traumatic mortality in King County for each one degree change in humidex above the “optimal alert” threshold. As a result of climate change, this exposure-response curve leads to a significant increase in projected mortality for the 85+ age group. A subsequent analysis (Busch Isaksen et al. 2015) that aimed only to quantify the historical heat-mortality relationship in King County extended the historical data to 2010

and found a 10% increase in the risk of death on a heat day versus a non-heat day for all ages, all causes. All three studies, however, establish a historical baseline using now-outdated data, only consider a subset of the state, do not utilize the novel case-crossover technique, and use geopolitical rather than climatological aggregation.

1.6 References

- Alman, B. L., G. Pfister, H. Hao, J. Stowell, X. Hu, Y. Liu, and M. J. Strickland. 2016. “The Association of Wildfire Smoke with Respiratory and Cardiovascular Emergency Department Visits in Colorado in 2012: A Case Crossover Study.” *Environ. Health* 54: 64.
- Armstrong, B. G., A. Gasparini, and A. Tobias. 2014. “Conditional Poisson Models: A Flexible Alternative to Conditional Logistic Case Cross-over Analysis.” *BMC Med. Res. Methodol.* 14: 122.
- Åström, D. O., C. Åström, B. Forsberg, A. M. Vicedo-Cabrera, A. Gasparini, A. Oudin, and K. Sundquis. 2020. “Heat Wave–Related Mortality in Sweden: A Case-Crossover Study Investigating Effect Modification by Neighbourhood Deprivation.” *Scandinavian Journal of Public Health* 48 (22): 6.
- Bao, J., X. Li, and C. Yu. 2015. “The Construction and Validation of the Heat Vulnerability Index, a Review.” *Int. J. Environ. Res. Public Health* 12: 7220–34.
- Barnett, A. G., S. Tong, and A. C. A. Clements. 2010. “What Measure of Temperature Is the Best Predictor of Mortality?” *Environ. Res.* 110 (6): 604–11.
- Basagaña, X., C Sartini, J. Barrera-Gómez, P. Dadvand, J. Cunillera, B. Ostro, J. Sunyer, and Medina-Ramón M. 2011. “Heat Waves and Cause-Specific Mortality at All Ages.” *Epidemiol.* 22 (6): 765–72.
- Bassil, K. L., and D. C. Cole. 2010. “Effectiveness of Public Health Interventions in Reducing Morbidity and Mortality During Heat Episodes: A Structured Review.” *Int. J. Environ. Res. Public Health* 22 (6): 765–72.
- Basu, R., F. Dominici, and J. M. Samet. 2005. “Temperature and Mortality Among the Elderly in the United States: A Comparison of Epidemiologic Methods.” *Epidemiol.* 16 (1): 58–66.
- Basu, R., and J. M. Samet. 2002. “Relation Between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence.” *Epidemiol. Rev.* 24: 190–202.

- Bateson, T. F., and J. Schwartz. 1999. "Control for Seasonal Variation and Time Trend in Case-Crossover Studies of Acute Effects of Environmental Exposures." *Epidemiol.* 10 (5): 539–44.
- . 2001. "Selection Bias and Confounding in Case-Crossover Analyses of Environmental Time-Series Data." *Epidemiol.* 12 (6): 654–61.
- Bell, M. L., M. S. O'Neill, N. Ranjit, V. H. Borja-Aburto, L. A. Cifuentes, and N. C. Gouveia. 2008. "Vulnerability to Heat-Related Mortality in Latin America: A Case-Crossover Study in São Paulo, Brazil, Santiago, Chile and Mexico City, Mexico." *Int. J. Of Epidemiol.* 37 (4): 796–804.
- Bhaskaran, K., A. Gasparrini, S. Hajat, L. Smeeth, and B. Armstrong. 2013. "Time Series Regression Studies in Environmental Epidemiology." *Int. J. Epidemiol.* 42 (4): 1187–95.
- Boeckmann, M., and I. Rohn. 2014. "Is Planned Adaptation to Heat Reducing Heat-Related Mortality and Illness? A Systematic Review." *BMC Public Health* 14: 1112.
- Boyd, R. W., E. G. Lindo, L. D. Weeks, and M. R. McLemore. 2020. "On Racism: A New Standard for Publishing on Racial Health Inequities." <https://www.healthaffairs.org/doi/10.1377/hblog20200630.939347/full/>.
- Brumback, B. A., L. M. Ryan, J. D. Schwartz, L. M. Neas, P. C. Stark, and H. A. Burge. 2000. "Transitional Regression Models, with Application to Environmental Time Series." *J. Am. Stat. Ass.* 95 (449): 16–27.
- Busch Isaksen, T., Fenskem R. A., E. K. Hom, Y. Ren, H. Lyons, and M. G. Yost. 2015. "Increased Mortality Associated with Extreme-Heat Exposure in King County, Washington, 1980-2010." *Int. J. Biometeorol.* 60 (1): 85–98.
- Busch Isaksen, T., M. Yost, E Hom, and R. Fenske. 2014. "Projected Health Impacts of Heat Events in Washington State Associated with Climate Change." *Rev. Environ. Health.* 29 (1-2): 119–23.
- Carracedo-Martínez, E., M. Taracido, A. Tobias, M. Saez, and A. Figueiras. 2010. "Case-Crossover Analysis of Air Pollution Health Effects: A Systematic Review of Methodology and Application." *Environ. Health Perspect.* 118 (8): 1173–82.

- Casanueva, A., A. Burgstall, S. Kotlarski, A. Messeri, M. Morabito, A. D. Flouris, L. Nybo, C. Spirig, and C. Schwierz. 2019. "Overview of Existing Heat-Health Warning Systems in Europe." *Int. J. Environ. Res. Public Health* 16: 2657.
- Chang, B. H., and D. C. Hoaglin. 2017. "Meta-Analysis of Odds Ratios: Current Good Practices." *Med. Care*. 55 (4): 328–35.
- Curriero, F. C., K. S. Heiner, J. M. Samet, S. L. Zeger, L. Strug, and J. A. Patz. 2002. "Temperature and Mortality in 11 Cities of the Eastern United States." *Am. J. Epidemiol.* 155 (1): 80–87.
- Davis, R. E., G. R. McGregor, and K. B. Enfield. 2016. "Humidity: A Review and Primer on Atmospheric Moisture and Human Health." *Environ. Res.* 144 (A): 106–16.
- Disease Control, Centers for, and Prevention. 1995. "Heat-Related Mortality-Chicago, July 1995." *MMWR* 44: 577–79.
- Doubleday, A., J. Schulte, L. Sheppard, M. Kadlec, R. Dhammapala, J. Fox, and T. Busch Isaksen. 2020. "Mortality Associated with Wildfire Smoke Exposure in Washington State, 2006–2017: A Case-Crossover Study." *Environ. Health* 19 (4).
- Ebi, K. L., T. J. Teisberg, L. S Kalkstein, L. Robinson, and R. F. Weiher. 2004. "Heat Watch/Warning Systems Save Lives: Estimated Costs and Benefits for Philadelphia 1995–98." *Bull. Amer. Meteor. Soc.* 85: 1067–74.
- Evangelopoulos, D., A. Analitis, C. Giannakopoulos, and K. Katsouyanni. 2021. "Does Climatic Zone of Birth Modify the Temperature-Mortality Association of London Inhabitants During the Warm Season? A Time-Series Analysis for 2004-2013." *Environ. Res.* 193: 110357.
- Figueras, A., E. Carracedo-Martínez, M. Saez, and M. Taracido. 2005. "Analysis of Case-Crossover Designs Using Longitudinal Approaches: A Simulation Study." *Epidemiol.* 16 (2): 239–346.
- Fonseca-Rodríguez, O., S. C. Sheridan, E. H. Lundevaller, and B. Schumann. 2020. "Hot and Cold Weather Based on the Spatial Synoptic Classification and Cause-Specific Mortality in Sweden: A Time-Stratified Case-Crossover Study." *Int. J. Biometeorol.* 64: 1435–49.

- Fu, S. H., A. Gasparrini, P. S. Rodriguez, and P. Jha. 2018. “Mortality Attributable to Hot and Cold Ambient Temperatures in India: A Nationally Representative Case-Crossover Study.” *PLoS Med.* 15 (7): e1002619.
- Fung, K. Y., D. Krewski, Y. Chen, R. Burnett, and S. Cakmak. 2003. “Comparison of Time Series and Case-Crossover Analyses of Air Pollution and Hospital Admission Data.” *Int. J. Epidemiol.* 32 (6): 1064–70.
- Gasparrini, A., B. Armstrong, and M. G. Kenward. 2010. “Distributed Lag Non-Linear Models.” *Stat Med* 29 (21): 2224–34.
- . 2012. “Multivariate Meta-Analysis for Non-Linear and Other Multi-Parameter Associations.” *Stat. Med.* 31 (29): 3821–39.
- Geletič, J., M. Lehnert, and M. Dobrovolný P. and Žžuvěla-Aloise. 2019. “Spatial Modelling of Summer Climate Indices Based on Local Climate Zones: Expected Changes in the Future Climate of Brno, Czech Republic.” *Climatic Change* 152: 487–502.
- Gosling, S. N., D. M. Hondula, A. Bunker, D. Ibarreta, J. Liu, X. Zhang, and R. Sauerborn. 2017. “Adaptation to Climate Change: A Comparative Analysis of Modeling Methods for Heat-Related Mortality.” *Environ. Health Perspect.* 125 (8): 087008.
- Gosling, S. N., J. A. Lowe, G. R. McGregor, M. Pelling, and B. D. Malamud. 2009. “Associations Between Elevated Atmospheric Temperature and Human Mortality: A Critical Review of the Literature.” *Climatic Change* 92: 299–341.
- Gravlee, C. C. 2009. “How Race Becomes Biology: Embodiment of Social Inequality.” *Am. J. Phys. Anthro.* 129: 47–57.
- Greene, S., L. S. Kalkstein, D. N. Mills, and J. Samenow. 2011. “An Examination of Climate Change on Extreme Heat Events and Climate–Mortality Relationships in Large u.s. Cities.” *Wea. Climate Soc.* 3: 281–92.
- Guirguis, K., R. Basu, W. K. Al-Delaimy, T. Benmarhnia, R. E. S. Clemesha, I. Corcos, J. Guzman-Morales, et al. 2018. “Heat, Disparities, and Health Outcomes in San Diego County’s Diverse

- Climate Zones.” *Geohealth* 2: 212–23.
- Guo, Y., A. G. Barnett, X. Pan, W. Yu, and Tong S. 2011. “The Impact of Temperature on Mortality in Tianjin, China: A Case-Crossover Design with a Distributed Lag Nonlinear Model.” *Environ. Health Perspect.* 119 (12): 1719–25.
- Guo, Y., A. Gasparri, and Li. et al. Shanshan. 2018. “Quantifying Excess Deaths Related to Heatwaves Under Climate Change Scenarios: A Multicountry Time Series Modelling Study.” *PLoS Med.* 15.
- Guttman, N. B., and R. G. Quayle. 1996. “A Historical Perspective of u.s. Climate Divisions.” *Bull. Amer. Meteor. Soc.* 77: 293–303.
- Hajat, S., S. C. Sheridan, M. J. Allen, M. Pascal, K. Laaidi, A. Yagouti, U. Bickis, et al. 2010. “Heat-Health Warning Systems: A Comparison of the Predictive Capacity of Different Approaches to Identifying Dangerously Hot Days.” *Am. J. Public Health* 100: 1137–44.
- Heo, S., E. Lee, B. Y. Kwon, S. Lee, K. H. Jo, and J. Kim. 2016. “Long-Term Changes in the Heat-Mortality Relationship According to Heterogeneous Regional Climate: A Time-Series Study in South Korea.” *BMJ Open* 6: e011786.
- Hoffman, J. S., V. Shandas, and N Pendleton. 2020. “The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas.” *Climate* 8 (12).
- Hondula, D. M., R. C. Balling, J. K. Vanos, and M. Georgescu. 2015. “Rising Temperatures, Human Health, and the Role of Adaptation.” *Curr. Clim. Change Rep.* 1: 144–54.
- Huang, C., A. G. Barnett, X. Wang, P. Vaneckova, G. FitzGerald, and S. Tong. 2011. “Projecting Future Heat-Related Mortality Under Climate Change Scenarios: A Systematic Review.” *Environ. Health Perspect.* 119 (12): 1681–90.
- IPCC. 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley. Cambridge, United Kingdom; New York, NY, USA: Cambridge University

Press.

- Jaakkola, J. J. 2003. "Case-Crossover Design in Air Pollution Epidemiology." *Eur. Respir. J. Suppl.* 40: 81s–85s.
- Jackson, J. E., M. G. Yost, C. Karr, C. Fitzpatrick, B. K. Lamb, Chung S. H., J. Avise, R. A. Rosenblatt, and R. A. Fenske. 2010. "Public Health Impacts of Climate Change in Washington State: Projected Mortality Risks Due to Heat Events and Air Pollution." *Climatic Change* 102: 159–86.
- Janes, H., L. Sheppard, and T. Lumley. 2005. "Overlap Bias in the Case-Crossover Design, with Application to Air Pollution Exposures." *Stat. Med.* 24 (2): 285–300.
- Jones, C. P. 2001. "Invited Commentary: "Race," Racism, and the Practice of Epidemiology." *Am. J. Epidemiol.* 154 (4): 299–306.
- Kalkstein, L. S. 1991. "A New Approach to Evaluate the Impact of Climate on Human Mortality." *Environ. Health Perspect.* 77: 1519–28.
- Kalkstein, L. S., and R. E. Davis. 1989. "Weather and Human Mortality: An Evaluation of Demographic and Interregional Responses in the United States." *Ann. Assoc. Am. Geogr.* 79: 44–64.
- Kalkstein, L. S., and J. S. Greene. 1997. "An Evaluation of Climate/Mortality Relationships in Large u.s. Cities and the Possible Impacts of a Climate Change." *Environ. Health Perspect.* 105: 84–93.
- Kalkstein, L. S., S. Greene, D. M. Mills, and J. Samenow. 2011. "An Evaluation of the Progress in Reducing Heat-Related Human Mortality in Major u.s. Cities." *Nat. Hazards* 56: 113–29.
- Kalkstein, L. S., P. F. Jamason, J. S. Greene, J. Libby, and L. Robinson. 1996. "The Philadelphia Hot Weather– Health Watch/Warning System: Development and Application, Summer 1995." *Bull. Amer. Meteor. Soc.* 77: 1519–28.
- Kilbourne, E. M. 1997. "The Public Health Consequences of Disasters." In, edited by E. K. Noji, 245–69. Oxford University Press, New York, Nw York.
- Kinney, P. L., M. S. O'Neill, M. L. Bell, and J. Schwartz. 2008. "Approaches for Estimating Effects of Climate Change on Heat-Related Deaths: Challenges and Opportunities." *Environ. Sci. Policy* 11 (1): 87–96.

- Knowlton, K., B. Lynn, R. A. Goldberg, C. Rosenzweig, C. Hogrefe, J. K. Rosenthal, and P. L. Kinney. 2007. "Projecting Heat-Related Mortality Impacts Under a Changing Climate in the New York City Region." *Am. J. Public Health* 97 (11): 2028–34.
- Kochanek, K., J. Xu, S. Murphy, A. Minino, and H. Kung. 2011. "Deaths: Final Data for 2009." *Natl. Vital Stat. Rep.* 60 (3).
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. "World Map of the köppen-Geiger Climate Classification Updated." *Meteorol. Z.* 15: 259–63.
- Levy, D., T. Lumley, L. Sheppard, J. Kaufman, and H. Checkoway. 2001. "Referent Selection in Case-Crossover Analyses of Acute Health Effects of Air Pollution." *Epidemiol.* 12 (2): 186–92.
- Lowe, D., K. L. Ebi, and B. Forsberg. 2011. "Heatwave Early Warning Systems and Adaptation Advice to Reduce Human Health Consequences of Heatwaves." *Int. J. Environ. Res. Public Health* 8 (12): 4623–48.
- Lu, Y., J. M. Symons, A. S. Geyh, and S. L. Zeger. 2008. "An Approach to Checking Case-Crossover Analyses Based on Equivalence with Time-Series Methods." *Epidemiol.* 19 (2): 169–75.
- Lu, Y., and S. L. Zeger. 2007. "On the Equivalence of Case-Crossover and Time Series Methods in Environmental Epidemiology." *Biostat.* 8 (2): 337–44.
- Luber, G., and M. McGeehin. 2008. "Climate Change and Extreme Heat Events." *Am. J. Prev. Med.* 35 (5): 429–35.
- Maclure, M. 1991. "The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events." *Am. J. Epidemiol.* 133: 144–53.
- Maclure, M., and M. A. Mittleman. 2008. "Case-Crossover Designs Compared with Dynamic Follow-up Designs." *Epidemiol.* 19 (2): 176–78.
- Madrigano, J., K. Lane, N. Petrovic, M. Ahmed, M. Blum, and T. Matte. 2018. "Awareness, Risk Perception, and Protective Behaviors for Extreme Heat and Climate Change in New York City." *Int. J. Environ. Res. Public Health* 15 (7): 1433.

- McElroy, S., L. Schwarz, H. Green, I. Corcos, K. Guirguis, A. Gershunov, and T. Benmarhnia. 2020. “Defining Heat Waves and Extreme Heat Events Using Sub-Regional Meteorological Data to Maximize Benefits of Early Warning Systems to Population Health.” *Science of The Total Environment* 721: 137678.
- Medina-Ramón, M., and J. Schwartz. 2007. “Temperature, Temperature Extremes, and Mortality: A Study of Acclimatisation and Effect Modification in 50 u.s. Cities.” *Occup. Environ. Med.* 64: 827–33.
- Meehl, G. A., and C. Tebaldi. 2004. “More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century.” *Science* 13: 994–97.
- Mittleman, M. A., and E. Mostofsky. 2014. “Exchangability in the Case-Crossover Design.” *Int. J. Of Epidemiol.* 43 (5): 1645–55.
- Nitschke, M., G. Tucker, A. Hansen, S. Williams, Y. Zhang, and P. Bi. 2016. “Evaluation of a Heat Warning System in Adelaide, South Australia, Using Case-Series Analysis.” *BMJ Open* 6: e012125.
- O’Neill, M. S., A. Zanobetti, and J. Schwartz. 2003. “Modifiers of the Temperature and Mortality Association in Seven US Cities.” *Am. J. Epidemiol.* 157 (12): 1074–82.
- . 2005. “Disparities by Race in Heat-Related Mortality in Four US Cities: The Role of Air Conditioning Prevalence.” *J. Urban Health* 82 (2): 191–97.
- Odame, E. A., Y. Li, S. Zheng, A. Vaidyanathan, and K. Silver. 2018. “Assessing Heat-Related Mortality Risks Among Rural Populations: A Systematic Review and Meta-Analysis of Epidemiological Evidence.” *Int. J. Environ. Res. Public Health* 15 (8): 1597.
- Patz, J. A., M. L. Grabow, and V. S. Limaye. 2014. “When It Rains, It Pours: Future Climate Extremes and Health.” *Ann. Glob. Health* 80 (4): 176–83.
- Petitti, D. B., D. M. Hondula, S. Yang, S. L. Harlan, and G. Chowell. 2016. “Multiple Trigger Points for Quantifying Heat-Health Impacts: New Evidence from a Hot Climate.” *Environ. Health Perspect.* 124 (2): 176–83.

- Reid, C. E., M. S. O'Neill, C. J. Gronlund, S. J. Brines, D. G. Brown, A. V. Diez-Roux, and J. Schwartz. 2009. "Mapping Community Determinants of Heat Vulnerability." *Environ. Health Perspect.* 117 (11): 1730–36.
- Salathé Jr., E. P., L. R. Leung, Y. Qian, and Y. Zhang. 2010. "Regional Climate Model Projections for the State of Washington." *Climatic Change* 102: 51–75.
- Schwartz, J. 2005. "Who Is Sensitive to Extremes of Temperature? A Case-Only Analysis." *Epidemiol.* 16 (1): 67–72.
- Semenza, J. C., C. H. Rubin, K. H. Falter, J. D. Selanikio, W. D. Flanders, H. L. Howe, and J. L. Wilhelm. 1996. "Heat-Related Deaths During the July 1995 Heat Wave in Chicago." *N Engl. J. Med.* 335: 84–90.
- Shaposhnikov, D., B. Revich, T. Bellander, G. B. Bedada, M. Bottai, T. Kharkova, E. Kvasha, et al. 2014. "Mortality Related to Air Pollution with the Moscow Heat Wave and Wildfire of 2010." *Epidemiol.* 25 (3): 359–64.
- Sheridan, S. C. 2006. "A Survey of Public Perception and Response to Heat Warnings Across Four North American Cities: An Evaluation of Municipal Effectiveness." *Int. J. Biometeorol.* 52: 3–15.
- Sheridan, S. C., and A. J. Kalkstein. 2010. "Seasonal Variability in Heat-Related Mortality Across the United States." *Nat. Hazards* 55: 291–305.
- Shindell, D., Y. Zhang, M. Scott, M. Ru, K. Stark, and K. L. Ebi. 2020. "The Effects of Heat Exposure on Human Mortality Throughout the United States." *Geohealth* 4 (4): e2019GH000234.
- Spangler, K. R., and G. A. Wellenius. 2020. "Spatial Patterns of Recent US Summertime Heat Trends: Implications for Heat Sensitivity and Health Adaptations." *Environ. Res. Comm.* 2 (3): 035002.
- Spector, J. T., D. K. Bonauto, L. Sheppard, T. Busch-Isaksen, M. Calkins, D. Adams, M. Lieblich, and R. A. Fenske. 2016. "A Case-Crossover Study of Heat Exposure and Injury Risk in Outdoor Agricultural Workers." *PLoS One* 11 (10): e0164498.
- Stafoggia, M., F. Forastiere, D. Agostini, A. Biggeri, L. Bisanti, E. Cadum, N. Caranci, et al. 2006. "Vulnerability to Heat-Related Mortality: A Multicity, Population-Based, Case-Crossover Analy-

- sis.” *Epidemiol.* 17 (3): 315–23.
- Stewart, I. D., and T. R. Oke. 2012. “Local Climate Zones for Urban Temperature Studies.” *Bull. Amer. Meteor. Soc.* 93: 1879–1900.
- Stringer, A., P. Brown, and J. Stafford. 2020. “Approximate Bayesian Inference for Case-Crossover Models.” *Biometrics*.
- Šeparović, L., A. Alexandru, R. Laprise, A. Martynov, L. Sushama, K. Winger, K. Tete, and M. Valin. 2013. “Present Climate and Climate Change over North America as Simulated by the Fifth-Generation Canadian Regional Climate Model.” *Clim. Dyn.* 41: 3167–3201.
- Tong, S., X. Y. Wang, and Y. Guo. 2012. “Assessing the Short-Term Effects of Heatwaves on Mortality and Morbidity in Brisbane, Australia: Comparison of Case-Crossover and Time Series Analyses.” *PLoS One* 7 (5): e37500.
- Vaidyanathan, A., S. Saha, A. M. Vicedo-Cabrera, A. Gasparrini, N. Abdurehman, R. Jordan, M. Hawkins, J. Hess, and A. Elixhauser. 2019. “Assessment of Extreme Heat and Hospitalizations to Inform Early Warning Systems.” *PNAS* 116 (12): 5420–27.
- Vicedo-Cabrera, A. M., F. Sera, and A. Gasparrini. 2019. “Hands-on Tutorial on a Modeling Framework for Projections of Climate Change Impacts on Health.” *Epidemiol.* 30 (3): 321–29.
- Vũ, A., S. Rutherford, and D. Phung. 2019. “Heat Health Prevention Measures and Adaptation in Older Populations - a Systematic Review.” *Int. J. Environ. Res. Public Health* 16 (22): 4370.
- Weinberger, K. R., D. Harris, K. R. Spangler, A. Zanobetti, and G. A. Wellenius. 2020. “Estimating the Number of Excess Deaths Attributable to Heat in 297 United States Counties.” *Environ. Epidemiol.* 4 (3): e096.
- Weinberger, K. R., L. Haykin, M. N. Eliot, J. D. Schwartz, A. Gasparrini, and G. A. Wellenius. 2017. “Projected Temperature-Related Deaths in Ten Large u.s. Metropolitan Areas Under Different Climate Change Scenarios.” *Environ. Int.* 107: 196–204.
- Weinberger, K. R., A. Zanobetti, J. Schwartz, and G. A. Wellenius. 2018. “Effectiveness of National Weather Service Heat Alerts in Preventing Mortality in 20 US Cities.” *Environ. Int.* 116: 30–38.

- Wellenius, G. A., M. N. Eliot, K. F. Bush, D. Holt, Lincoln R. A., Smith A. E., and J. Gold. 2017. "Heat-Related Morbidity and Mortality in New England: Evidence for Local Policy." *Environ. Res.* 156: 845–53.
- Wilson, L. A., G. G. Morgan, I. C. Hanigan, F. H. Johnston, H. Abu-Rayya, R. Broome, C. Gaskin, and B. Jalaludin. 2013. "The Impact of Heat on Mortality and Morbidity in the Greater Metropolitan Sydney Region: A Case Crossover Analysis." *Environ. Health* 12: 98.
- Xu, Z., G. FitzGerald, Y. Guo, B. Jalaludin, and S. Tong. 2016. "Impact of Heatwave on Mortality Under Different Heatwave Definitions: A Systematic Review and Meta-Analysis." *Environ. Int.* 89-90: 193–203.
- Yao, J., M. Brauer, J. Wei, K. M. McCraill, F. H. Johnston, and S. B. Henderson. 2020. "Sub-Daily Exposure to Fine Particulate Matter and Ambulance Dispatches During Wildfire Seasons: A Case-Crossover Study in British Columbia, Canada." *Environ. Health Perspect.* 128 (6): 67006.
- Zanobetti, A., and J. Schwartz. 2008. "Temperature and Mortality in Nine US Cities." *Epidemiol.* 19 (4): 563–70.
- Zhang, K., R. B. Rood, G. Michailidis, E. M. Oswald, J. D. Schwartz, A. Zanobetti, K. L. Ebi, and M. S. O'Neill. 2012. "Comparing Exposure Metrics for Classifying 'Dangerous Heat' in Heat Wave and Health Warning Systems." *Environ. Int.* 46: 23–39.

2 Mortality Associated with Extreme Heat in Washington State: The historical and projected public health burden

2.1 Introduction

The human body responds to extreme heat by increasing cardiovascular demand in an attempt to provide physiologic cooling and maintain thermal homeostasis, and a failure to meet this demand can lead to a variety of illnesses or death (Kilbourne 1997). Epidemiological studies of the association between heat and mortality primarily conclude that abnormally high temperatures are associated with excess mortality and that the exposure-response function is nonlinear, typically described by a V-, U-, or J-shaped curve (Basu and Samet 2002; Gosling et al. 2009). This relationship has been observed in locations with diverse climates and cultures, including the United States (Kalkstein and Davis 1989; Curriero et al. 2002; Medina-Ramón and Schwartz 2007; Busch Isaksen et al. 2014, 2015; Shindell et al. 2020; Weinberger et al. 2020), Europe (Stafoggia et al. 2006; Basagaña et al. 2011; Åström et al. 2020), Latin America (Bell et al. 2008), India (Fu et al. 2018), China (Guo et al. 2011), Russia (Shaposhnikov et al. 2014) and Australia (Wilson et al. 2013).

Deaths attributed specifically to heat often underestimate the true public health burden of high temperatures (Kilbourne 1997). Other commonly identified risk factors include cardiovascular diseases (Semenza et al. 1996; Medina-Ramón and Schwartz 2007; Basagaña et al. 2011; Busch Isaksen et al. 2015; Åström et al. 2020), respiratory diseases (Basagaña et al. 2011; Fu et al. 2018), diabetes (Schwartz 2005; Basagaña et al. 2011; Busch Isaksen et al. 2015), age (Kalkstein and Davis 1989; Stafoggia et al. 2006; Bell et al. 2008; Busch Isaksen et al. 2015; Fu et al. 2018), and lack of access to air conditioning (Semenza et al. 1996; Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2005; Medina-Ramón and Schwartz 2007; Madrigano et al. 2018). (Systemic) racism produces different health outcomes for racially defined groups (Gravlee 2009; Boyd et al. 2020), including in the context of heat-related mortality (O'Neill, Zanobetti, and Schwartz 2003, 2005; Schwartz 2005), since communities of color often have reduced access to air conditioning (Madrigano et al. 2018) and live disproportionately in areas most susceptible to the urban heat island effect (Hoffman, Shandas, and Pendleton 2020). It is also well-documented that temperate regions are more sensitive to heat events, perhaps due to less appreciation of risk, lower prevalence of air conditioning, or weaker physiologic acclimatization (Kalkstein and Davis 1989; Kalkstein and Greene 1997; Curriero et al.

2002; Medina-Ramón and Schwartz 2007; Reid et al. 2009; Sheridan and Kalkstein 2010; Shindell et al. 2020). Moreover, anthropogenic activities are causing average global temperatures to rise (IPCC 2013) as well as the frequency, length, and intensity of future extreme heat events to increase (Meehl and Tebaldi 2004). Therefore, knowledge about the relationship between high temperatures and human health is becoming increasingly important, as this information is critical for both the immediate public health response and inquiries into future health outcomes.

Washington state, an area largely characterized by a temperate climate, is an important location for research into heat-related health outcomes. Jackson et al. (2010) examined four areas in Washington state and obtained relative risk values in the greater Seattle area that indicated a significant relationship between heat event duration and increased daily mortality rates for non-traumatic deaths in persons aged 45 and above. While results in the other 3 areas were not statistically significant, the patterns of elevated relative risk estimates suggest that there is also a real difference in mortality rates during heat events in these areas, however sample size issues prevented statistically significant estimates. Busch Isaksen et al. (2014) computed a 1.83% (95% CI: 0.77%, 2.91%) increase in mortality for all ages, all non-traumatic mortality in King County for each one degree change in humidex above the optimal alert threshold of 35.7 humidex. A subsequent analysis (Busch Isaksen et al. 2015) also found a 10% increase (95% CI: 6%, 14%) in the risk of death on a heat day versus a non-heat day for all ages, all causes.

The Pacific Northwest historically has experienced few severe heatwaves, but climate models suggest that this area will become increasingly susceptible to extreme heat events (Meehl and Tebaldi 2004) and that it may experience greater summertime warming due to climate change than most of the rest of the continental United States (Šeparović et al. 2013). Thus, a region that is already vulnerable to heat may soon face a climate to which it is not adapted nor acclimated. Jackson et al. (2010) found that projected non-traumatic deaths due to heat days in the Seattle region were greatest for people in the 65+ age group in all years and warming scenarios analyzed, though the results for the rest of the state were not statistically significant. In King County, Busch Isaksen et al. (2014) identified the 85+ age group as being most at-risk of heat-related mortality in a changing climate.

An important limitation of existing heat-related mortality research in Washington (Jackson et al. 2010; Busch Isaksen et al. 2014, 2015) is the use of geopolitical units for analysis (cities or counties) that may

not capture the true spatial variation in exposure, and subsequent health outcomes, that occurs during heat events. Instead, spatial aggregation based on climatic patterns may be more appropriate, since individuals tend to adapt to their local weather conditions (Medina-Ramón and Schwartz 2007). Furthermore, there is also a need for a revised analysis utilizing up-to-date data that spans the entire state to support public health decision-making. For instance, the Washington Tracking Network, produced by the Washington State Department of Health, is a program that aims to produce publicly-available maps and datasets containing environmental and public health data. This initiative includes a section on heat stress and notes that future extreme heat events are predicted to have increased frequency and intensity, yet it fails to recognize the reality that hospitalizations and deaths explicitly attributable to heat under-represent the true public health burden of heat-related health outcomes in the state or provide any indication as to how this public health problem varies across the state.

This study uses a case-crossover approach to investigate the heat-mortality relationship in Washington state and explores possible heterogeneity in the dose-response curve across the state's ten climate divisions, as defined by the National Oceanic and Atmospheric Administration [NOAA; Guttman and Quayle (1996)]. The impact of heat on non-traumatic mortality is assessed for all-ages, in addition to possible effect modification by age, sex, race/ethnicity and cause of death. Pooled odds ratio (OR) estimates and 95% confidence intervals (CIs) are reported, comparing the odds of mortality at the 99th percentile of summertime humidex to the 50th percentile. The historical results are also combined with climate projections to investigate the impact that climate change will have on excess mortality attributable to extreme heat.

2.2 Methods

2.2.1 Data Sources

2.2.1.1 Mortality Data Mortality data from the Washington State Department of Health covers the entire state and the years 1980-2018. Only deaths during the months of May through September were included in the analysis, resulting in 153 days/calendar year and 5,967 days for the entire study period. Non-traumatic mortality across all ages was the primary outcome of interest (ICD-9: 0-799; ICD-10: A00-R99). Investigations of specific subsets of non-traumatic mortality were determined *a priori* based on existing literature (Jackson et al. 2010; Busch Isaksen et al. 2014, 2015) and included diabetes (ICD-9: 250; ICD-10: E08-

E13), circulatory (ICD-9: 390–459; ICD-10: I00–I99, G45, G46), cardiovascular (ICD-9: 393–429; ICD-10: I05–I52), ischemic (ICD-9: 410–414; ICD-10: I20–I25), cerebrovascular (ICD-9: 430–438; ICD-10: I67), respiratory (ICD-9: 460–519; ICD-10: J00–J99), nephritis and nephrotic syndromes (ICD-9: 580–589; ICD-10: N17–N19), acute renal failure (ICD-9: 584; ICD-10: N17), and mental disorders (ICD-9: 290–316; ICD-10: F01–F69). Other individual-level characteristics include age (divided into six age groups: 0–4, 5–14, 15–44, 45–64, 65–84, and 85+), sex, and race/ethnicity. Each death had an associated latitude and longitude value, enabling aggregation of cases into one of the ten climate divisions in Washington (Guttman and Quayle 1996), depicted in Figure 2.1.

According to the Washington State Institutional Review Board, the use of Washington State mortality data is exempt from review and that an application for official exempt status review is not needed.

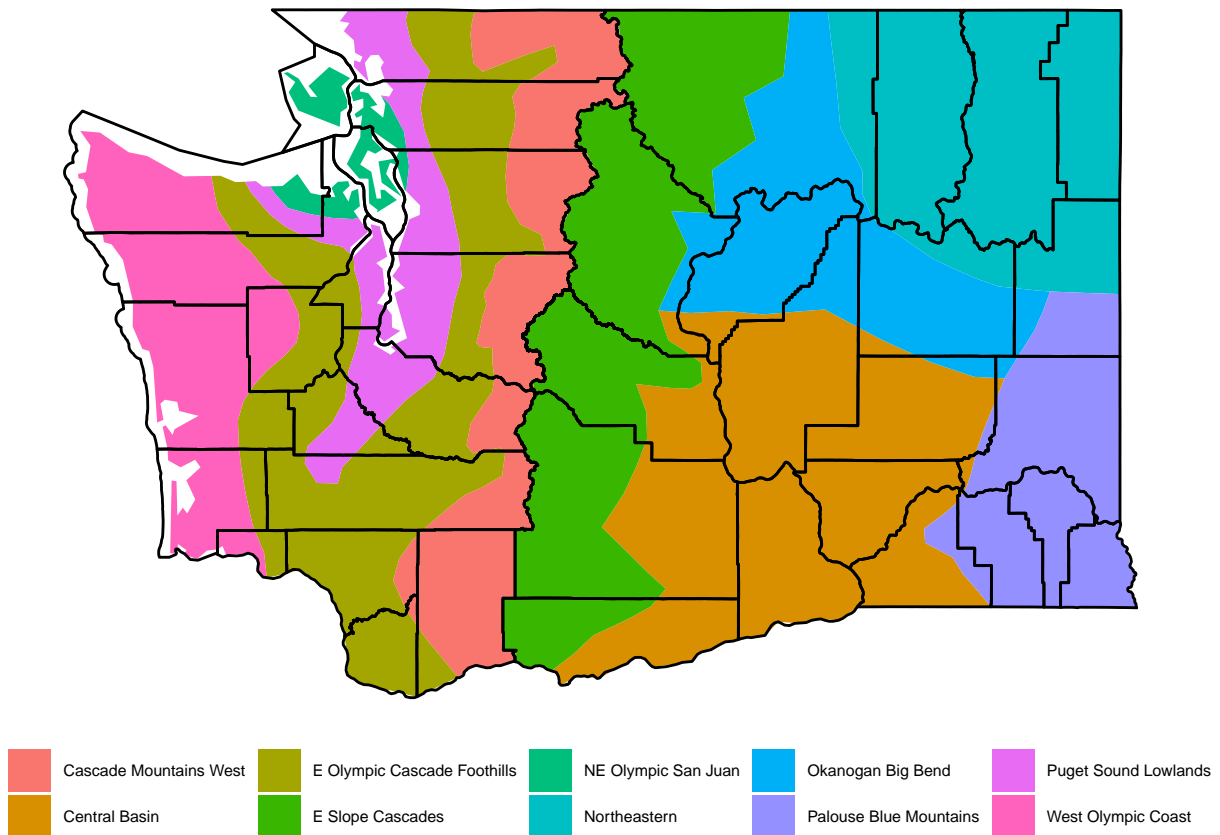


Figure 2.1: Map of Washington’s climate divisions.

2.2.1.2 Historical Meteorological Data The gridMET data (Abatzoglou 2013) from the University of California Merced Climatology Lab was used to obtain historical meteorological data (daily temperature

and relative humidity values) to characterize exposure to heat. This data, which is intended to be used in ecological applications and modeling, was produced using a hybrid method that combines temporally rich data from the North American Land Data Assimilation System Phase 2 [NLDAS; [Mitchell et al. \(2004\)](#)] with spatially rich data from the Parameter-elevation Regressions on Independent Slopes Model [PRISM; [Daly et al. \(2008\)](#)]. The gridMET data has observations at a spatial resolution of 1/24 degree (~4km × 4km) and was validated using weather stations in the western United States ([Abatzoglou 2013](#)).

2.2.1.3 Climate Projections The Multivariate Adaptive Constructed Analogs [MACA; [Abatzoglou and Brown \(2012\)](#)] statistical downscaling of 20 global circulation models (GCMs) of the Coupled Model Inter-Comparison Project 5 (CMIP5) from the [University of California Merced Climatology Lab](#) was used to obtain projected meteorological data for the years 2030, 2050, and 2080. This data has been used previously to forecast heat-related deaths ([Weinberger et al. 2017](#)). To account for uncertainty in future greenhouse gas emissions, two representative concentration pathways (RCP) scenarios were used to represent intermediate (RCP4.5) and worst-case (RCP8.5) trajectories. For each combination of year and RCP, the results from all 20 models were averaged to obtain a single projected temperature and relative humidity value for every grid point and every day from May 1 through September 31.

2.2.2 Exposure Assessment

For both the historical and projected analyses, temperature and relative humidity values were combined to construct the exposure metric, daily maximum humidex. Humidex is a unitless apparent temperature measurement that incorporates both air temperature and humidity ([Masterton and Richardson 1979](#)):

$$\text{Humidex} = T + \frac{5}{9}(\nu - 10) \quad (1)$$

where T is the air temperature (°C), H is the relative humidity (%), and $\nu = (6.112 \times 10^{\frac{7.5T}{237.7+T}}) \frac{H}{100}$ is the vapor pressure (kPA). Humidex was chosen because it incorporates humidity, which is known to influence the physiological response to extreme heat ([Kilbourne 1997](#); [Davis, McGregor, and Enfield 2016](#)), and for comparison to previous studies, particularly those in the Pacific Northwest ([Jackson et al. 2010](#); [Busch Isaksen et al. 2014, 2015](#)), that also utilize humidex as the exposure metric. A nearest neighbors calculation

was used to assign individual-level humidex exposure by comparing the latitude and longitude of death to the gridded meteorological data.

Humidex calculations using the maximum temperature and maximum relative humidity are biased upward if these two metrics are not maximized at the same time of day. To account for this bias, humidex was calculated using daily maximum temperature and daily minimum relative humidity. Minimum relative humidity was chosen from daily maximum, daily minimum, and daily average (mean of maximum and minimum) relative humidity because it produced humidex values that most closely aligned with observations from weather stations in Seattle and Spokane at various points during the study's 39-year time frame (see [Appendix A](#)). For consistency, daily maximum temperature and daily minimum relative humidity values were used to calculate both historical and future humidex values. The projected humidex values were also calibrated to the historical data to correct for climate model biases ([Ho et al. 2012](#)).

2.2.3 Case-Crossover Analysis for the Historical Period

The case-crossover study design ([Maclure 1991](#)) was used to quantify the historical (1980-2018) relationship between humidex and non-traumatic mortality. A case-crossover study design is akin to the traditional matched case-control design since selection is based on the outcome; however, like a crossover study, each case serves as its own control. Risk of mortality is inferred by comparing a person's exposure to humidex on the date of death to multiple control or "referent" periods ([Jaakkola 2003](#)), and matching each case to itself eliminates confounding by constant or slow-varying individual characteristics such as age and sex ([Maclure 1991](#)). Control days were selected by matching the date of death to all other days that share the same day of week, month, and year of the death. This approach has been used extensively in previous heat-related mortality studies ([Stafoggia et al. 2006](#); [Bell et al. 2008](#); [Basagaña et al. 2011](#); [Wilson et al. 2013](#); [Åström et al. 2020](#)) and is ideal because it controls for seasonality, long-term trends, and any potential effects related to the day of the week by design ([Sheppard and Lumley 2005](#)).

A conditional logistic regression model was created for each climate division in Washington. To model the nonlinear relationship between heat and mortality, both polynomials and natural cubic splines were used to incorporate individual-level humidex as a continuous variable. AIC was used for model selection, and ultimately a quadratic polynomial was chosen to produce an exposure-response curve that describes the odds

of non-traumatic mortality at all humidex values. The odds of non-traumatic mortality at the 99th percentile of humidex were compared to the odds at the 50th percentile to create a single odds ratio (OR), by climate division. Stratified analyses were conducted to explore effect modification by both age and cause of death. For both the full and stratified analyses, the ORs in each climate division were pooled using fixed-effects meta-analysis to obtain a single, state-level estimate.

All analyses were conducted using **R** 3.6.2 (R Core Team 2019). Conditional logistic regression was implemented with the `survival` package (Therneau 2020), while the `meta` package (Balduzzi, Rücker, and Schwarzer 2019) was used for meta-analysis.

2.2.4 Projected Heat-Attributable Deaths due to Climate Change

Calculations for projected heat-related were guided by Vicedo-Cabrera, Sera, and Gasparrini (2019). Specifically, it was assumed that populations and outcome rates will remain constant in the future. While unrealistic, this approach isolates the effect of climate from other important trends such as age structure to answer the question, “How would the current population respond if exposed to warmer temperatures projected in the future?”

First, the number of deaths attributable to heat above the 99th percentile in each climate division were calculated for the historical period using the following formula:

$$D_{\text{attr}} = D \times (1 - e^{-\beta \mathbf{H}}) \quad (2)$$

where D is the total number of deaths occurring on days above the 99th percentile of humidex and $(1 - e^{-\beta \mathbf{H}})$ is the fraction of deaths attributable to (a quadratic function of) humidex, \mathbf{H} , restricted to humidex values above the 99th percentile. β corresponds to the coefficients of the exposure-response curve derived in the case-crossover analysis. (Note that this formula reduces to $\frac{RR-1}{RR}$ in the case of a linear or binary relationship.)

After obtaining the number of deaths attributable to extreme heat in the historical period, projected, individual-level humidex values were assigned to each death in the years 2030, 2050, and 2080 and under the emissions scenarios RCP4.5 and RCP8.5 based on the grid point in which the death is located. Then,

Equation 2 was applied using the historical 99th percentile of humidex in each climate division as a cutoff to calculate the projected public health burden. The results of the projected heat-attributable deaths calculations are reported as percent increases over the historical period.

2.3 Results

2.3.1 Historical Public Health Burden

Table 2.1: Non-traumatic mortality characteristics.

Characteristic	Climate Division										
	All	Cascade Mountains West	Central Basin	E Olympic Cascade Foothills	E Slope Cascades	NE Olympic San Juan	Northeastern	Okanogan Big Bend	Palouse Blue Mountains	Puget Sound Lowlands	West Olympic Coast
Total	563,365	1,020	62,794	85,498	4,329	14,711	52,685	6,479	7,155	316,999	11,695
Age											
0 - 4	7,133	13	1,033	810	42	88	636	52	59	4,311	89
5 - 14	970	-	118	141	-	-	91	12	-	569	12
15 - 44	18,948	35	2,006	2,599	94	233	1,575	158	176	11,800	272
45 - 64	98,702	233	10,393	15,996	775	1,838	8,914	1,109	1,059	56,145	2,240
65 - 84	270,513	514	30,422	42,247	2,174	7,240	25,592	3,279	3,449	149,498	6,098
85+	167,090	222	18,822	23,704	1,236	5,302	15,876	1,868	2,406	94,670	2,984
Sex											
Female	285,395	452	31,588	42,359	1,982	7,366	26,836	3,148	3,596	162,397	5,671
Male	277,955	568	31,204	43,137	2,347	7,345	25,848	3,331	3,559	154,592	6,024
Race/Ethnicity											
White	518,873	997	57,769	82,544	4,241	14,321	50,758	6,078	6,902	284,114	11,149
Black	14,429	-	701	592	-	37	572	23	21	12,433	40
Asian	12,341	-	326	700	-	106	324	17	37	10,780	42
Native American	6,058	11	923	669	26	122	663	288	30	2,983	343
Hispanic	5,980	-	2,842	400	36	47	74	56	17	2,454	49
Native Hawaiian or Other Pacific Islander	4,296	-	106	392	-	33	139	-	-	3,571	43
Cause of Death											
Diabetes	10,004	20	1,276	1,576	64	226	919	140	132	5,427	224
Circulatory	213,937	335	25,109	31,318	1,607	5,380	20,184	2,434	2,760	120,153	4,657
Cardiovascular	160,118	261	18,970	23,972	1,215	3,970	14,822	1,813	2,005	89,483	3,607
Ischemic	102,827	148	13,129	15,296	839	2,528	9,731	1,236	1,290	56,218	2,412
Cerebrovascular	19,905	16	2,367	2,480	121	373	1,919	194	273	11,802	360
Respiratory	54,393	92	5,894	8,375	395	1,398	5,716	657	758	29,845	1,263
Nephritis and nephrotic	4,160	-	513	552	22	78	417	46	72	2,371	82
Acute Renal Failure	684	-	81	96	-	-	69	-	16	394	-
Mental Disorders	12,108	23	1,146	1,938	86	269	1,013	143	165	7,099	226

- denotes less than 10 observations

Non-traumatic mortality counts broken down by age, sex, race/ethnicity, and cause of death are given in Table 2.1 for Washington and each climate division. In total, 563,365 deaths were included in the analysis. 56% of the deaths occurred in the Puget Sound Lowlands, the division in which the majority of Seattle, Washington's most populous city, is located. The range of deaths in each climate division is very large, with only 0.18% of deaths occurring in the least populous division, the Cascade Mountains West. The two largest age groups are 65 - 84 (48%) and 85+ (30%), respectively. The distribution across sexes is fairly equal, although more women than men are included in the data (50.7% vs. 49.3%). White individuals are by far the largest racial/ethnic category (92.1%), and the vast majority of non-White deaths are in the Puget Sound

Lowlands.

Table 2.2 shows the distribution of humidex during the historical period in each of the climate divisions. The distribution of humidex, like the distribution of mortality, is also quite variable across climate divisions. There is a 26% difference between the climate division with the lowest median humidex value (West Olympic Coast) to the division with the highest median humidex value (Central Basin). Similarly, there is a 20% difference in the 99th percentiles of humidex between these two climate divisions. Multiple pairwise t-tests using the Bonferroni correction were conducted to assess humidex heterogeneity across all climate divisions at the $\alpha = 0.05$ level, and the only two that were identified as having the same mean were East Slope Cascades and Puget Sound Lowlands. However, these two divisions are not adjacent and consequently could not be combined. Figures showing the spatial and temporal distribution of humidex are given in [Appendix B](#).

Table 2.2: Historical humidex descriptive statistics.

Climate Division	Humidex (1980-2018)								
	Mean	SD	Min	Max	25 th	50 th	75 th	95 th	99 th
Cascade Mountains West	22.74	6.82	1.23	41.36	17.80	22.91	27.78	33.51	36.77
Central Basin	27.24	6.55	5.66	48.40	22.53	27.50	32.11	37.56	40.56
E Olympic Cascade Foothills	24.97	6.39	-1.29	48.14	20.40	24.97	29.44	35.55	39.57
E Slope Cascades	23.43	6.69	1.43	42.24	18.65	23.70	28.38	33.98	37.16
NE Olympic San Juan	21.77	4.70	6.11	39.30	18.53	21.71	24.93	29.72	32.87
Northeastern	24.70	6.71	2.85	44.92	19.79	24.99	29.73	35.25	38.01
Okanogan Big Bend	25.49	6.65	3.48	46.08	20.66	25.68	30.45	36.03	39.14
Palouse Blue Mountains	25.19	6.85	3.91	44.78	20.25	25.56	30.28	35.83	38.84
Puget Sound Lowlands	23.50	5.91	5.57	46.44	19.26	23.46	27.64	33.42	36.98
West Olympic Coast	21.29	4.61	5.88	44.83	18.31	21.20	24.03	29.21	33.77

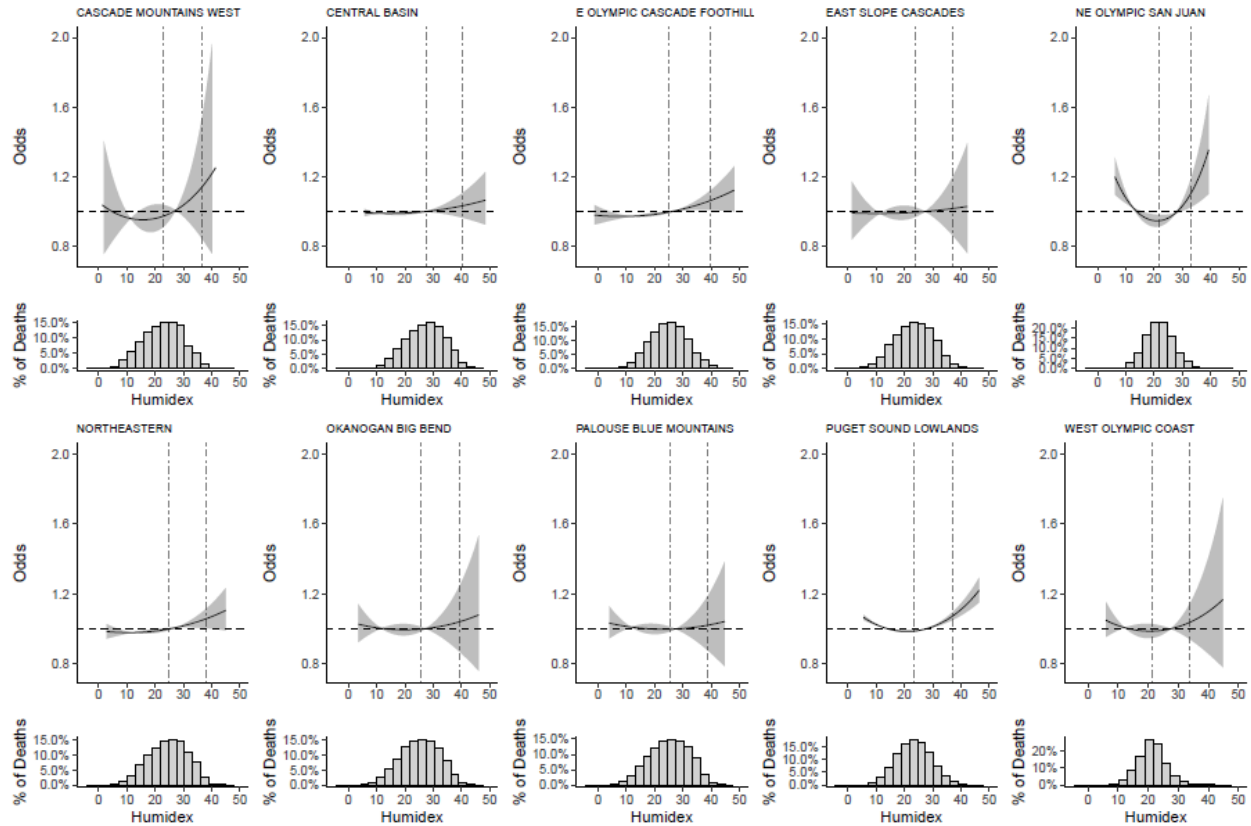


Figure 2.2: Humidex-mortality exposure-response curves for each of Washington’s climate divisions.

Figure 2.2 shows the humidex-mortality exposure-response function for each climate division. The function describes the odds of non-traumatic mortality (and 95% CI) at each humidex value, with the horizontal dashed line indicating the null value of 1.0. The vertical dashed lines denote the 50th and 99th percentiles of humidex for that climate division. Despite humidex distributions that vary from one climate division to another, the dose-response curves in Figure 2.2 are quite similar, and all generally resemble a J- or U-shape. Excluding the Cascade Mountains West division, which had the fewest number of deaths during the study period (1,020), the value for the estimated odds of non-traumatic mortality at the 99th percentile of humidex is quite similar across all climate divisions. The most notable difference is in the associated 95% CI: it is quite tight and mostly non-overlapping with the null value in the East Olympic Cascade Foothills, NE Olympic San Juan, Northeastern, and Puget Sound Lowlands regions but frequently overlaps with the null value in the other six climate divisions. In all regions, the estimated function becomes less precise moving away from the 50th and, especially, 99th percentiles.

Figure 2.3 is a forest plot depicting ORs (and 95% CIs) of non-traumatic mortality, comparing the

50th and 99th percentiles of humidex, for each climate division as well as a single, statewide value produced via meta-analysis. All climate divisions had point estimates above 1.00, but only four (E Olympic Cascade Foothills, NE Olympic San Juan, Northeastern, and Puget Sound Lowlands) had 95% CIs that did not overlap with the null value. The pooled OR was 1.08 (1.06, 1.10), driven primarily by the OR for Puget Sound Lowlands (1.09 [1.06, 1.11]).

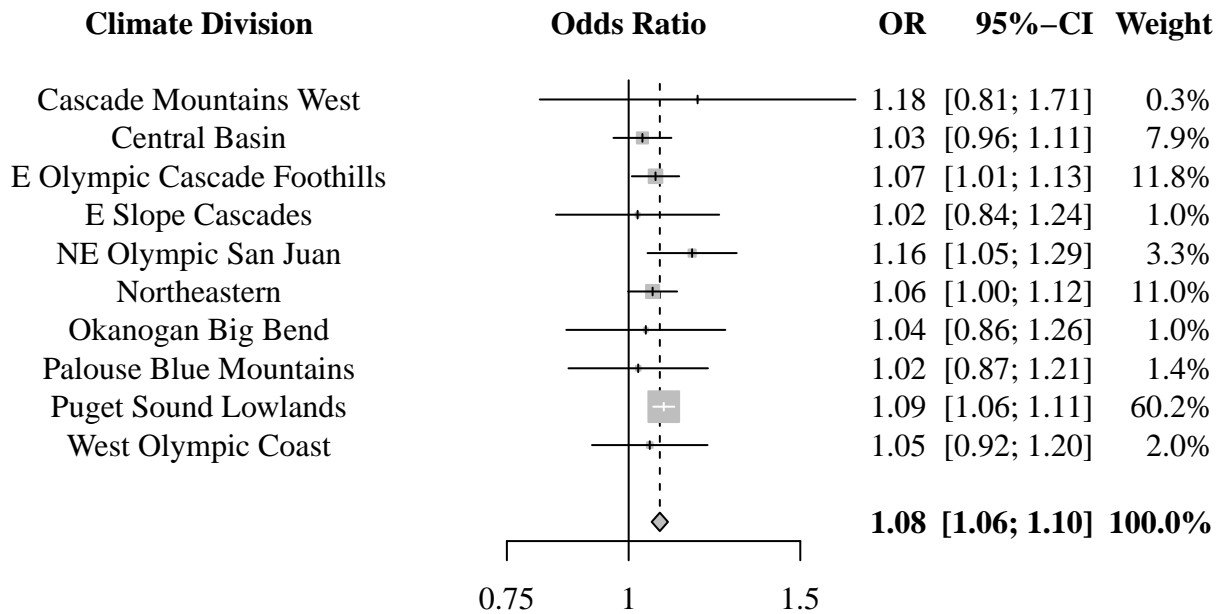


Figure 2.3: Odds ratio and 95% CI of non-traumatic mortality, comparing the 99th and 50th percentiles of humidex.

The results for effect modification by age, sex, and race/ethnicity are given in Table 2.3 and by both age and cause of death in Table 2.4. All values represent ORs pooled across all ten climate divisions. The OR for non-traumatic mortality increases with age, with the 85+ age group at greatest risk of non-traumatic mortality during a heat event. The point estimate is higher for females than males as well as for Blacks than Whites, though the confidence intervals overlap somewhat. The pooled OR was not significant for other racial/ethnic categories. Diabetic, circulatory, cardiovascular, ischemic, cerebrovascular, and respiratory deaths all have 95% CIs that are non-overlapping with 1. With the exception of diabetic and cerebrovascular deaths, the OR for cause-specific mortality also increases with age; both diabetic deaths and cerebrovascular deaths only have a significant OR for the 65-84 age group. More extensive effect modification results are given in Appendix C.

Table 2.3: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, by age group, sex, and race/ethnicity.

Group	Observations (%)	OR (95% CI)
Total	563,365 (100)	1.08 (1.06, 1.10)
Age		
0 – 4	7,133 (1.3)	.90 (.75, 1.08)
5 – 14	970 (0.2)	.90 (.56, 1.45)
15 – 44	18,948 (3.4)	1.05 (.94, 1.17)
45 – 64	98,702 (17.5)	1.06 (1.02, 1.12)
65 – 84	270,513 (48.0)	1.07 (1.04, 1.10)
85+	167,090 (29.7)	1.09 (1.06, 1.13)
Sex		
Female	285,395 (50.7)	1.08 (1.05, 1.11)
Male	277,955 (49.3)	1.06 (1.03, 1.09)
Race/Ethnicity		
White	518,873 (92.1)	1.07 (1.05, 1.09)
Black	14,429 (2.6)	1.15 (1.02, 1.30)
Asian	12,341 (2.2)	1.06 (.93, 1.21)
Native American	6,058 (1.1)	1.08 (.89, 1.32)
Hispanic	5,980 (1.1)	1.01 (.84, 1.21)
Native Hawaiian or Other Pacific Islander	4,296 (0.8)	.97 (.77, 1.20)

Results are pooled across all climate divisions using fixed-effects meta-analysis.
 Bolded rows indicate 95% CIs that did not include 1.

Table 2.4: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, by age group and cause of death.

Cause of Death (Total Cases)	All Ages	0-4	5-14	15-44	45-64	65-84	85+
All Non-Traumatic Causes (563,365)	1.08 (1.06, 1.10)	0.90 (0.75, 1.08)	0.90 (0.56, 1.45)	1.05 (0.94, 1.17)	1.06 (1.02, 1.12)	1.07 (1.04, 1.10)	1.10 (1.06, 1.13)
Diabetes (10,004)	1.20 (1.03, 1.39)	NA	NA	1.25 (0.60, 2.63)	1.12 (0.81, 1.56)	1.25 (1.03, 1.52)	1.12 (0.80, 1.56)
Circulatory (213,937)	1.10 (1.07, 1.14)	1.10 (0.40, 3.01)	0.68 (0.10, 4.71)	1.08 (0.86, 1.35)	1.09 (1.00, 1.19)	1.09 (1.04, 1.15)	1.11 (1.05, 1.17)
Cardiovascular (160,118)	1.10 (1.06, 1.14)	1.36 (0.45, 4.14)	0.30 (0.02, 3.56)	1.12 (0.87, 1.44)	1.09 (0.99, 1.20)	1.09 (1.04, 1.16)	1.11 (1.04, 1.18)
Ischemic (102,827)	1.09 (1.04, 1.14)	NA	NA	1.12 (0.77, 1.64)	1.03 (0.92, 1.16)	1.10 (1.03, 1.18)	1.10 (1.02, 1.20)
Cerebrovascular (19,905)	1.16 (1.05, 1.29)	0.12 (0.00, 27.94)	NA	1.12 (0.77, 1.64)	1.27 (0.87, 1.86)	1.21 (1.04, 1.41)	1.10 (0.93, 1.30)
Respiratory (54,494)	1.08 (1.02, 1.15)	0.36 (0.10, 1.36)	1.63 (0.15, 18.28)	0.81 (0.47, 1.37)	1.17 (0.97, 1.40)	1.05 (0.96, 1.14)	1.14 (1.02, 1.28)
Nephritis and nephrotic (4,160)	1.03 (0.82, 1.28)	1.83 (0.15, 22.31)	NA	1.60 (0.34, 7.85)	0.75 (0.40, 1.42)	1.04 (0.75, 1.45)	1.09 (0.76, 1.57)
Acute Renal Failure (684)	1.14 (0.66, 2.00)	NA	NA	NA	0.75 (0.40, 1.42)	0.87 (0.39, 1.91)	1.09 (0.76, 1.57)
Mental Disorders (12,108)	1.02 (0.90, 1.17)	NA	NA	0.80 (0.43, 1.49)	0.89 (0.60, 1.32)	1.08 (0.86, 1.37)	1.05 (0.87, 1.27)

Results are pooled across all climate divisions using fixed-effects meta-analysis. Bolded values indicate 95% CIs that did not include 1.

2.3.2 Projected Public Health Burden

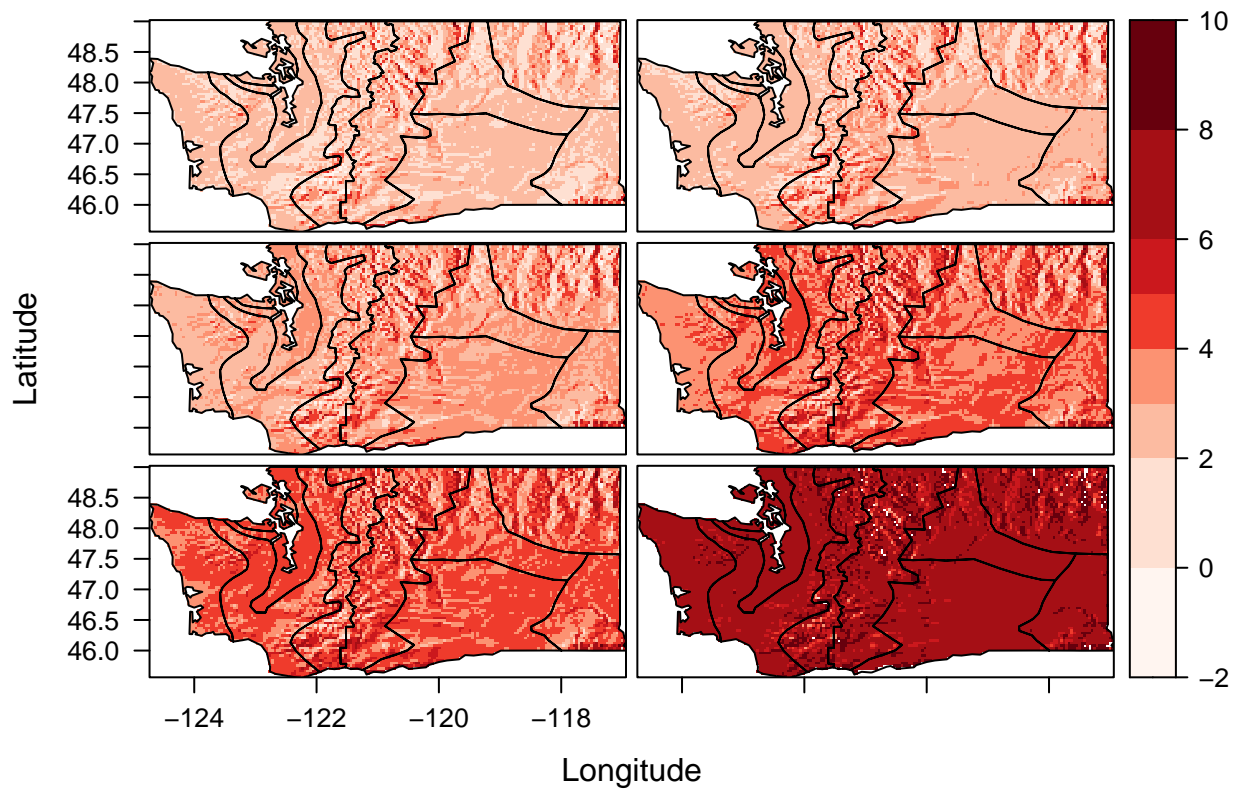


Figure 2.4: Spatial distribution of projected humidex anomalies.

Notes: Columns denote RCPs: left is RCP4.5, right is RCP8.5. Rows indicate years: top is 2030, middle is 2050, and bottom is 2080.

Figure 2.4 shows maps of projected humidex anomalies (projected values minus historical values) in Washington for 2030, 2050, and 2080 under both RCP4.5 and RCP8.5. In 2030, the difference between projected and historical humidex values is less than 2 for most of the state under both RCPs. By 2050, the two RCPs start to produce divergent results. Warming remains somewhat modest under RCP4.5, and very few parts of the state have projected humidex anomalies exceeding 4. Conversely, most of the state exceeds this threshold in 2050 under RCP8.5. This discrepancy continues in 2080; most projected humidex anomalies are between 4 and 6 under RCP4.5, while the humidex anomalies are between 8 and 10 for the entire state under RCP8.5. For reference, the average daily maximum temperatures in the Puget Sound Lowlands (where Seattle is located) and Northeastern (where Spokane is located) under RCP4.5 and RCP8.5 in 2080 is similar to current daily average July temperatures in Boise and Nashville (for the Puget Sound

Lowlands) or Charlotte and Houston (for Northeastern), respectively.

The total number of deaths attributable to heat during the historical period (1980-2018) and the percent increase in 2030, 2050, and 2080 under RCP4.5 and RCP8.5 are given in Table 2.5. Only three climate divisions - E Olympic Cascade Foothills, NE Olympic San Juan, and Puget Sound Lowlands - had 95% CIs for the percent increase over the historical period that did not overlap with 0; in these three zones, the CI did not include 0 in any year or under either RCP.

Table 2.5: Number of deaths attributable to heat above the 99th percentile of humidex in the historical period and projected percent increase due to climate change.

Climate Division	Historical (N)	Percent Increase					
		2030		2050		2080	
		RCP4.5	RCP8.5	RCP4.5	RCP8.5	RCP4.5	RCP8.5
Cascade Mountains West	2	0	100	100	200	200	900
Central Basin	25	32	68	112	224	252	1156
E Olympic Cascade Foothills	61	11	28	57	133	152	556
E Slope Cascades	1	100	100	100	200	200	1100
NE Olympic San Juan	21	43	62	119	238	252	1652
Northeastern	33	42	76	124	245	276	1306
Okanogan Big Bend	3	33	33	100	200	233	1067
Palouse Blue Mountains	1	100	200	300	500	700	2900
Puget Sound Lowlands	283	23	42	87	178	196	812
West Olympic Coast	6	0	0	33	100	133	617

Bolded rows indicate climate divisions where 95% CIs did not include 0.

The percent increases in Table 2.5 reflect the warming patterns given in Figure 2.4. In each year, the percent increase of attributable heat deaths is higher under RCP8.5 than RCP4.5, and the discrepancy between the two greenhouse gas emissions scenarios increases with time. Of the three climate divisions with statistically significant results, NE Olympic San Juan is projected to have the greatest increase in deaths attributable to heat (values range from 43% in 2030 under RCP 4.5 to 1652% in 2080 under RCP8.5), while E Olympic Cascade Foothills has the smallest projected percent increase (values range from 13% to 559%). Across these three climate divisions, the average percent increase in heat-attributable deaths is 35%, 135%, and 603% in 2030, 2050, and 2080, respectively.

2.4 Discussion

This study used a case-crossover analysis to describe the historical relationship between heat and non-traumatic mortality in Washington's ten climate divisions. Fixed-effects meta-analysis was used to simplify the exposure-response curves to a single, state-wide value, the OR of non-traumatic mortality at the 99th percentile of summertime humidex compared to the 50th percentile. Subgroup analyses were also conducted to investigate effect modification by age, sex, race/ethnicity, and cause of death. Finally, the historical exposure-response curves were combined with climate projections to investigate the impact of climate change on an existing public health issue. The results provide evidence that high values of humidex are associated with increased non-traumatic mortality in Washington and indicate that climate change could lead to an increased number of deaths attributable to extreme heat.

The state-wide OR value produced in this study was 1.08 (1.06, 1.10). This value for risk of mortality associated with same-day heat events is quite similar to results from Seattle (O'Neill, Zanobetti, and Schwartz 2003; Jackson et al. 2010) and King County (Busch Isaksen et al. 2015) as well as Sydney, Australia (Wilson et al. 2013), South Korea (Heo et al. 2016), and Sweden (Åström et al. 2020). However, it is less than risk estimates produced in southern Europe (Stafoggia et al. 2006; Basagaña et al. 2011) and greater than values from other parts of the United States (O'Neill, Zanobetti, and Schwartz 2005; Medina-Ramón and Schwartz 2007); these differences are likely driven by varied access to air conditioning or physiological acclimatization.

The J- and U-shaped dose-response curves in Figure 2.2 are characteristic of the relationship between heat and mortality (Basu and Samet 2002; Gosling et al. 2009) and indicate elevated risks of mortality at high values of humidex (for J-shaped curves) or both low and high values of humidex (for U-shaped curves). Busch Isaksen and colleagues (2014, 2015) have previously described a J-shaped curve for King County, and identified 30 humidex as the value at which risk of mortality increases. The East Olympic Cascade Foothills, NE Olympic San Juan, Northeastern, and Puget Sound Lowlands climate divisions are similarly described by an elevated risk threshold of approximately 30 humidex.

Knowledge about the heat-mortality relationship using climate zones or climate regions rather than geopolitical units has the potential to provide better guidance for policymakers attempting to mitigate heat-related deaths, including through the development of heat warning systems (Vaidyanathan et al. 2019; McEl-

roy et al. 2020). Nevertheless, such an approach is not widespread in studies of heat-related mortality. Guirguis et al. (2018) detected a significant increase in hospitalizations during hot weather in comparison to hospitalizations observed over all days, independent of any temperature threshold, in the coastal region of San Diego County, but not in communities further inland that are characterized by higher average temperatures and greater air conditioning prevalence. Regional time series analyses in South Korea (Heo et al. 2016) as well as regional case-crossover analyses in India (Fu et al. 2018) also yielded distinct temperature-mortality relationships for different climate regions. However, Basagaña et al. (2011) failed to detect heterogeneity in the heat-mortality relationship across Catalonia's 14 climatic zones, and chose instead to focus on a single meta-analysis. Overall, the spatial results from this study are somewhat mixed. While the exposure-response curves suggest qualitative differences in the heat-mortality relationship between zones in Washington (Figure 2.2), the similar ORs comparing the 99th and 50th percentiles of humidex (Figure 2.3) may suggest that individuals in the state respond to their differing local temperatures in a similar manner. It is unclear whether the differences in results are due to methodological choices, the region of study, or some other factor, and the use of climate variables to aggregate health outcomes in epidemiological studies remains an emerging area of research with unanswered questions.

Stratified analyses for individual-level characteristics indicate that risk of non-traumatic mortality on 99th percentile days increases with age, particularly in the 85+ age group, and that risk is higher among women than men and among Blacks than Whites (Table 2.3). Age is well-established risk factor for heat-related mortality (Kalkstein and Davis 1989; Stafoggia et al. 2006; Bell et al. 2008; Busch Isaksen et al. 2015; Fu et al. 2018). Greater risk among women has been identified previously in Italy (Stafoggia et al. 2006) and Spain (Basagaña et al. 2011), though conflicting results have been observed in King County, Washington (Busch Isaksen et al. 2015) and elsewhere in the U.S. (O'Neill, Zanobetti, and Schwartz 2003; Wellenius et al. 2017) as well as in Latin America (Bell et al. 2008), Russia (Shaposhnikov et al. 2014), and South Korea (Heo et al. 2016) where either men were at greater risk, or there was no difference across sexes. An area for future research is to elucidate the cause of these conflicting results, e.g. differences in study demographics or methodology. The disparities across racial groups are consistent with previous studies (O'Neill, Zanobetti, and Schwartz 2003, 2005; Schwartz 2005; Madrigano et al. 2018) and are likely due to socioeconomic differences across racial groups perpetuated by institutionalized racism (Jones 2001; Madrigano et al. 2018; Hoffman, Shandas, and Pendleton 2020). To reiterate, in the context of heat-related

mortality, race is best thought of as a *risk marker* rather than a *risk factor* (Kaplan and Bennett 2003), or a *social construct* rather than a *biological determinant* (Jones 2001). Nevertheless, the results stratified by race/ethnicity are useful because they can help to inform the allocation of medical/public health resources (Kaufman and Cooper 2001; Bhopal 2006).

Diabetic, circulatory, cardiovascular, ischemic, cerebrovascular, and respiratory deaths are also significantly associated with heat (Table 2.4). The results for all causes of death except for ischemic are largely consistent with previous results (Semenza et al. 1996; Schwartz 2005; Medina-Ramón and Schwartz 2007; Basagaña et al. 2011; Busch Isaksen et al. 2015; Åström et al. 2020), though Busch Isaksen and colleagues (2015) did not obtain a significant association for respiratory deaths. A significant association for ischemic deaths is consistent with Wilson et al. (2013) but contradicts other existing work (Busch Isaksen et al. 2015; Fu et al. 2018). Finally, null results for nephritis and nephrotic, acute renal failure and mental disorder deaths mostly agree with prior research (Wilson et al. 2013; Busch Isaksen et al. 2015), though Basagaña et al. (2011) found significant results for mental disorders. Explaining these contrasting results for certain causes of death is a potential area for future research.

The projected heat-attributable deaths is higher than the historical value in almost every climate division under both RCP4.5 and RCP8.5 and in 2030, 2050, and 2080 (Table 2.5). The direction of this result agrees with most existing studies, despite different study areas and methodological choices (Knowlton et al. 2007; Huang et al. 2011; Guo, Gasparrini, and Shanshan 2018; Wellenius et al. 2017; Shindell et al. 2020). In the Puget Sound Lowlands, there were an average of 7.3 annual heat-related deaths in the historical period, almost identical to the value of 7.1 obtained by Busch Isaksen et al. (2014) for King County. However, it is difficult to compare projections from this study directly to results from Busch Isaksen et al. (2014), since different methodologies were used, and, consequently, the magnitude of the projected values represent different quantities. The approach used by Jackson et al. (2010) is more similar to this study, and the results are identical qualitatively (increase in heat-related mortality over the historical period, a gap between warming scenarios/RCPs that grows with time, and results that are statistically significant only in the western part of the state) and similar quantitatively (large increase under moderate warming/RCP4.5 and drastic increase under high warming/RCP8.5).

Projected increases in heat-attributable mortality in the New York City region in 2050 (Knowlton et al.

2007) as well as select American metropolitan areas in both 2050 and 2090 (Wellenius et al. 2017) are lower than results from this study, possibly indicating that Washington is more vulnerable to heat-related mortality in a changing climate than other parts of the country. The projected heat-attributable mortality values are similar in 2030 under RCP4.5 and RCP8.5, but the difference between these scenarios grows with time and is quite large by 2080; the values for 2050 under RCP8.5 are also quite similar to the values for 2080 under RCP4.5. This temporal pattern is consistent with results from Wellenius et al. (2017), and Shindell et al. (2020) also observed a large discrepancy between these RCPs in the 2090s. In contrast, Guo et al. (2018) observed percent increases that were more consistent from one scenario to the next. Overall, the somewhat contrasting results in projected heat-attributable mortality across studies likely stem from different data and methodology, such as the temperature threshold used to calculate the number of heat-attributable deaths.

A case-crossover analysis conducted in Sweden illustrated the importance of incorporating regional information about the heat-mortality relationship when developing public health interventions (Åström et al. 2020). Kalkstein et al. (2011) argue that a general decline across 40 U.S. cities in mortality rates attributable to extreme heat events since 1996 when compared to rates from 1975-1995 – a decline that occurred despite a warming climate – can be attributed, at least in part, to resources related to education, notification, and response measures associated with heat-related health outcomes in many of the 40 cities. Ebi et al. (2004) estimate that the Philadelphia Hot Weather-Health Watch/Warning System saved 117 lives during a three-summer period, and Heo et al. (2016) attribute lower mortality risk during heat events among the elderly in 2008-2012 compared to 1996-2000 to the introduction of a heatwave forecast system in 2008 that targeted elderly individuals. Therefore, there is precedent for using results similar to the historical analysis in this study to improve heat-related health outcomes in a changing climate. In contrast to these results from other locations, however, it is possible that heat-related mortality has increased in Seattle over time (Davis et al. 2003). Consequently, there is a clear need to incorporate this knowledge into public health decision-making in Washington.

The research design expands upon existing work in Washington (Jackson et al. 2010; Busch Isaksen et al. 2014, 2015) in a number of ways. First, the case-crossover approach is novel in the state. Using this technique was one of the study goals, and it was a beneficial choice because it does not require “denominator” data (i.e., population size), which could be difficult to obtain for Washington’s climate divisions. Previous research in King County, Washington (Busch Isaksen et al. 2014, 2015) has used a Poisson time

series model to describe the relationship between humidex and mortality. Notably, studies that have applied both a time series and case-crossover approach assert that the two techniques generate results which are comparable quantitatively and identical qualitatively (Zanobetti and Schwartz 2008; Tong, Wang, and Guo 2012). This was confirmed by using King County to conduct a sensitivity analysis (see Appendix D). The major difference between the two approaches is that a case-crossover analysis produces odds ratios, while a Poisson time series model produces relative risks. However, Basu, Dominici, and Samet (2005) observe that estimates produced with either method are often directly comparable, since mortality is a rare event. Thus, despite the utilization of a different technique, it is possible to compare results from this study to results in Busch Isaksen et al. (2014, 2015).

In addition to the use of a novel analytic technique, this study also incorporated data from the entire state of Washington. Previous work has been restricted to King County (Busch Isaksen et al. 2014, 2015) or selected populous areas in the state (Jackson et al. 2010). The use of data through the year 2018 also resulted in an analysis with a larger historical timeframe than any existing study. Since this study used NOAA's climate divisions (Guttman and Quayle 1996) to group deaths, it is the first in Washington to use a form of spatial aggregation not based on geopolitical boundaries. Finally, regarding the projections component of this study, the use of 20 different climate models implicitly accounts for uncertainty in climate model physics, which may be more important in projections of heat-related mortality than uncertainty associated with greenhouse gas emissions (Gosling, McGregor, and Lowe 2012). This is a notable improvement over the two existing projections studies in the state (Jackson et al. 2010; Busch Isaksen et al. 2014), each of which use only two climate models.

There are a number of limitations to this study. Although humidex exposure was assigned on an individual-level using a nearest neighbors calculation, there is still likely to be some exposure misclassification; for example, individuals may be far from their residence on the control days preceding death or even on the date of death itself, so the assigned humidex exposure may not represent conditions that these individuals actually experienced. The price to pay for controlling for temporal confounding using in the case-crossover design is reduced statistical efficiency (Bateson and Schwartz 1999). This was likely demonstrated by only four of the ten climate divisions having statistically significant ORs due to the small sample sizes in some parts of the state. The lack of air pollution data, a potential confounding variable, could overestimate the magnitude of heat-related mortality. This study also did not incorporate socioeconomic variables, which

may be important since those variables are likely to be highly correlated with access to air conditioning, a documented modifier of the heat-health relationship (Semenza et al. 1996; Curriero et al. 2002; O'Neill, Zanobetti, and Schwartz 2005; Guirguis et al. 2018). A synoptic climatological approach was not employed, so the exposure-response relationship may fail to model all weather elements that contribute to mortality (Kalkstein 1991). Similarly, this study may also fail to produce the true heat-mortality curve because it did not include the impact of lagged and/or multi-day exposure.

Using areas defined by climatic variables has notable benefits, but the lack of previous work utilizing this method of spatial aggregation, particularly in the Pacific Northwest, complicated direct comparison of effect estimates. Climate is also not the only factor that influences the heat-mortality relationship, so the use of climatological rather than geopolitical aggregation could mask some other important and locally-varying factors, such as public health messaging or socioeconomic variables, that influence the heat-mortality relationship. Furthermore, the specific use of climate divisions identified by NOAA may be problematic because the boundaries are not necessarily intended to create areas of climatological homogeneity (instead, the divisions were often based on crop considerations and/or drainage basins), and some divisional boundaries are chosen to coincide with county boundaries (Guttman and Quayle 1996). Indeed, on a single day, meteorological variables can vary greatly across a single climate division.

An alternative to NOAA's climate divisions is the Köppen-Geiger climate classification (Kottek et al. 2006), which was used directly by Fu et al. (2018) in India and indirectly by Guirguis et al. (2018) in San Diego County. While this scheme is also based primarily on five distinct vegetation types, each of these major classifications are divided further based on explicitly defined temperature and precipitation ranges. Although there is some microclimate heterogeneity, most of Washington falls into one of three main Köppen-Geiger climate types: areas west of the Cascades are mostly warm temperate climate (C), while east of the Cascades is a mix of snowy climates (D) in the northeast and both arid climates (B) and warm temperate climate (C) in the southeast (see Appendix E and, specifically, Figure E1). Interestingly, three of the four climate divisions with significant humidex-mortality exposure-response curves (East Olympic Cascade, NE Olympic San Juan, and Puget Sound Lowlands) lie west of the Cascades and are characterized primarily by warm, temperate climates with dry summers. This region of the state is divided into two subtypes based on precipitation, but it is possible that this division is not important for heat-related mortality and that, for the purposes of public health applications, these three climate divisions should be grouped together.

Independent of the specific boundaries chosen, a problem that will persist is climatological heterogeneity across a single region. This heterogeneity is likely to increase as the size of the region increases, and may favor the use of cities or counties for aggregation, since these smaller-scale regions could have less heterogeneity, even if their boundaries are not influenced by climate variables. When appropriate, the optimal approach may be to divide a city or county into climate zones, as in Guirguis et al. (2018); this approach requires both true meteorological heterogeneity across the city/county as well as a spatial distribution of population that allows for the necessary statistical power to ascertain differences from one climate zone to the next. An important area for further research is the use of climatological variables for spatial aggregation when studying heat-related health outcomes as well as the importance of the choice of a specific climate classification.

The second portion of the study, the projected public health burden, incorporated uncertainty in greenhouse gas trajectories by using two different RCPs, but it did not include an aggressive decarbonization scenario (i.e. RCP2.6). This is a notable disadvantage when compared to existing projections in the state (Jackson et al. 2010; Busch Isaksen et al. 2014), which include low, moderate, and high warming scenarios. However, more recent work in major U.S. cities (Weinberger et al. 2017) and London (Vicedo-Cabrera, Sera, and Gasparrini 2019) also include only RCP4.5 and RCP8.5 in their projections, so there is precedent for excluding a low-emission trajectory. There are also limitations associated with the use of bias-corrected climate projections; for example, the downscaled data will inherit any imperfections in the training data (Abatzoglou and Brown 2012), and the calibration procedure in Ho et al. (2012) specifically may yield an increase in the 99th percentile that is greater than the mean. This study does not incorporate uncertainty in adaptation scenarios, which may be even greater than the uncertainty associated with climate modeling (Gosling et al. 2017). Adaptation was not incorporated because it requires arbitrary modeling choices related to the extent and timing of adaptation mechanisms (Vicedo-Cabrera, Sera, and Gasparrini 2019). The assumptions of no adaptation and a constant population, while simplistic, yield projections that have isolated the effect of climate. Finally, the reported heat-attributable deaths are only for values of humidex above the historical 99th percentile and may not represent the full extent of mortality associated with extreme heat.

2.5 Conclusion

By quantifying the impact of exposure to extreme temperatures, policymakers can develop preventative, life-saving measures, warning systems, and action plans that target susceptible groups, support health systems, and engage the community at large (Lowe, Ebi, and Forsberg 2011; Casanueva et al. 2019). The application of a novel quantitative method and spatial unit of aggregation combined with contemporary data that spans the entire state has provided further evidence that heat-related mortality is an extant public health problem in Washington. The combination of these results with climate change projections suggest that this existing issue could be exacerbated within the coming decades as the duration, intensity, and frequency of heat events increases. By following the successes observed in other locations, the results of this study can help inform initiatives and adaptations within Washington that aim to improve heat-related health outcomes and ameliorate the challenges imposed by a changing climate.

2.6 References

- Abatzoglou, J. T. 2013. “Development of Gridded Surface Meteorological Data for Ecological Applications and Modelling.” *Int. J. Climatol.* 33: 121–31.
- Abatzoglou, J. T., and T. J. Brown. 2012. “A Comparison of Statistical Downscaling Methods Suited for Wildfire Applications.” *Int. J. Climatol.* 32: 772–80.
- Åström, D. O., C. Åström, B. Forsberg, A. M. Vicedo-Cabrera, A. Gasparrini, A. Oudin, and K. Sundquis. 2020. “Heat Wave–Related Mortality in Sweden: A Case-Crossover Study Investigating Effect Modification by Neighbourhood Deprivation.” *Scandinavian Journal of Public Health* 48 (22): 6.
- Balduzzi, S., G. Rücker, and G. Schwarzer. 2019. “How to Perform a Meta-Analysis with R: A Practical Tutorial.” *Evidence-Based Mental Health*, no. 22: 153–60.
- Basagaña, X., C Sartini, J. Barrera-Gómez, P. Dadvand, J. Cunillera, B. Ostro, J. Sunyer, and Medina-Ramón M. 2011. “Heat Waves and Cause-Specific Mortality at All Ages.” *Epidemiol.* 22 (6): 765–72.
- Basu, R., F. Dominici, and J. M. Samet. 2005. “Temperature and Mortality Among the Elderly in the United States: A Comparison of Epidemiologic Methods.” *Epidemiol.* 16 (1): 58–66.
- Basu, R., and J. M. Samet. 2002. “Relation Between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence.” *Epidemiol. Rev.* 24: 190–202.
- Bateson, T. F., and J. Schwartz. 1999. “Control for Seasonal Variation and Time Trend in Case-Crossover Studies of Acute Effects of Environmental Exposures.” *Epidemiol.* 10 (5): 539–44.
- Bell, M. L., M. S. O’Neill, N. Ranjit, V. H. Borja-Aburto, L. A. Cifuentes, and N. C. Gouveia. 2008. “Vulnerability to Heat-Related Mortality in Latin America: A Case-Crossover Study in são Paulo, Brazil, Santiago, Chile and Mexico City, Mexico.” *Int. J. Of Epidemiol.* 37 (4): 796–804.
- Bhopal, R. 2006. “Race and Ethnicity: Responsible Use from Epidemiological and Public Health Perspectives.” *J. Law Med. Ethics* 34 (3): 500–507.

- Boyd, R. W., E. G. Lindo, L. D. Weeks, and M. R. McLemore. 2020. "On Racism: A New Standard for Publishing on Racial Health Inequities." <https://www.healthaffairs.org/doi/10.1377/hblog20200630.939347/full/>.
- Busch Isaksen, T., Fenskem R. A., E. K. Hom, Y. Ren, H. Lyons, and M. G. Yost. 2015. "Increased Mortality Associated with Extreme-Heat Exposure in King County, Washington, 1980-2010." *Int. J. Biometeorol.* 60 (1): 85–98.
- Busch Isaksen, T., M. Yost, E Hom, and R. Fenske. 2014. "Projected Health Impacts of Heat Events in Washington State Associated with Climate Change." *Rev. Environ. Health.* 29 (1-2): 119–23.
- Casanueva, A., A. Burgstall, S. Kotlarski, A. Messeri, M. Morabito, A. D. Flouris, L. Nybo, C. Spirig, and C. Schwierz. 2019. "Overview of Existing Heat-Health Warning Systems in Europe." *Int. J. Environ. Res. Public Health* 16: 2657.
- Curriero, F. C., K. S. Heiner, J. M. Samet, S. L. Zeger, L. Strug, and J. A. Patz. 2002. "Temperature and Mortality in 11 Cities of the Eastern United States." *Am. J. Epidemiol.* 155 (1): 80–87.
- Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. A. Pasteris. 2008. "Physiographically-Sensitive Mapping of Temperature and Precipitation Across the Conterminous United States." *Int. J. Climatol.* 28: 2031–64.
- Davis, R. E., P. C. Knappenberger, P. J. Michales, and W. M. Novicoff. 2003. "Changing Heat-Related Mortality in the United States." *Environ. Health Perspect.* 111 (14): 1712–18.
- Davis, R. E., G. R. McGregor, and K. B. Enfield. 2016. "Humidity: A Review and Primer on Atmospheric Moisture and Human Health." *Environ. Res.* 144 (A): 106–16.
- Ebi, K. L., T. J. Teisberg, L. S Kalkstein, L. Robinson, and R. F. Weiher. 2004. "Heat Watch/Warning Systems Save Lives: Estimated Costs and Benefits for Philadelphia 1995–98." *Bull. Amer. Meteor. Soc.* 85: 1067–74.
- Fu, S. H., A. Gasparrini, P. S. Rodriguez, and P. Jha. 2018. "Mortality Attributable to Hot and Cold Ambient Temperatures in India: A Nationally Representative Case-Crossover Study." *PLoS Med.* 15 (7): e1002619.

- Gosling, S. N., D. M. Hondula, A. Bunker, D. Ibarreta, J. Liu, X. Zhang, and R. Sauerborn. 2017. “Adaptation to Climate Change: A Comparative Analysis of Modeling Methods for Heat-Related Mortality.” *Environ. Health Perspect.* 125 (8): 087008.
- Gosling, S. N., J. A. Lowe, G. R. McGregor, M. Pelling, and B. D. Malamud. 2009. “Associations Between Elevated Atmospheric Temperature and Human Mortality: A Critical Review of the Literature.” *Climatic Change* 92: 299–341.
- Gosling, S. N., G. R. McGregor, and J. A. Lowe. 2012. “The Benefits of Quantifying Climate Model Uncertainty in Climate Change Impacts Assessment: An Example with Heat-Related Mortality Change Estimates.” *Climatic Change* 112: 217–31.
- Gravlee, C. C. 2009. “How Race Becomes Biology: Embodiment of Social Inequality.” *Am. J. Phys. Anthro.* 129: 47–57.
- Guirguis, K., R. Basu, W. K. Al-Delaimy, T. Benmarhnia, R. E. S. Clemesha, I. Corcos, J. Guzman-Morales, et al. 2018. “Heat, Disparities, and Health Outcomes in San Diego County’s Diverse Climate Zones.” *Geohealth* 2: 212–23.
- Guo, Y., A. G. Barnett, X. Pan, W. Yu, and Tong S. 2011. “The Impact of Temperature on Mortality in Tianjin, China: A Case-Crossover Design with a Distributed Lag Nonlinear Model.” *Environ. Health Perspect.* 119 (12): 1719–25.
- Guo, Y., A. Gasparrini, and Li. et al. Shanshan. 2018. “Quantifying Excess Deaths Related to Heatwaves Under Climate Change Scenarios: A Multicountry Time Series Modelling Study.” *PLoS Med.* 15.
- Guttman, N. B., and R. G. Quayle. 1996. “A Historical Perspective of US Climate Divisions.” *Bull. Amer. Meteor. Soc.* 77: 293–303.
- Heo, S., E. Lee, B. Y. Kwon, S. Lee, K. H. Jo, and J. Kim. 2016. “Long-Term Changes in the Heat–Mortality Relationship According to Heterogeneous Regional Climate: A Time- Series Study in South Korea.” *BMJ Open* 6: e011786.

- Ho, C. K., D. B. Stephenson, M. Collins, C. A. T. Ferro, and S. J. Brown. 2012. "Calibration Strategies: A Source of Additional Uncertainty in Climate Change Projections." *Bull. Amer. Meteor. Soc.* 93 (1): 21–26.
- Hoffman, J. S., V. Shandas, and N Pendleton. 2020. "The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas." *Climate* 8 (12).
- Huang, C., A. G. Barnett, X. Wang, P. Vaneckova, G. FitzGerald, and S. Tong. 2011. "Projecting Future Heat-Related Mortality Under Climate Change Scenarios: A Systematic Review." *Environ. Health Perspect.* 119 (12): 1681–90.
- IPCC. 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by T. F Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley. Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press.
- Jaakkola, J. J. 2003. "Case-Crossover Design in Air Pollution Epidemiology." *Eur. Respir. J. Suppl.* 40: 81s–85s.
- Jackson, J. E., M. G. Yost, C. Karr, C. Fitzpatrick, B. K. Lamb, Chung S. H., J. Avise, R. A. Rosenblatt, and R. A. Fenske. 2010. "Public Health Impacts of Climate Change in Washington State: Projected Mortality Risks Due to Heat Events and Air Pollution." *Climatic Change* 102: 159–86.
- Jones, C. P. 2001. "Invited Commentary: "Race," Racism, and the Practice of Epidemiology." *Am. J. Epidemiol.* 154 (4): 299–306.
- Kalkstein, L. S. 1991. "A New Approach to Evaluate the Impact of Climate on Human Mortality." *Environ. Health Perspect.* 77: 1519–28.
- Kalkstein, L. S., and R. E. Davis. 1989. "Weather and Human Mortality: An Evaluation of Demographic and Interregional Responses in the United States." *Ann. Assoc. Am. Geogr.* 79: 44–64.
- Kalkstein, L. S., and J. S. Greene. 1997. "An Evaluation of Climate/Mortality Relationships in Large u.s. Cities and the Possible Impacts of a Climate Change." *Environ. Health Perspect.* 105: 84–93.

- Kalkstein, L. S., S. Greene, D. M. Mills, and J. Samenow. 2011. "An Evaluation of the Progress in Reducing Heat-Related Human Mortality in Major u.s. Cities." *Nat. Hazards* 56: 113–29.
- Kaplan, J. B., and T. Bennett. 2003. "Use of Race and Ethnicity in Biomedical Publication." *JAMA* 289 (20): 2709–16.
- Kaufman, J. S., and R. S. Cooper. 2001. "Commentary: Considerations for Use of Racial/Ethnic Classification in Etiologic Research." *Am. J. Epidemiol.* 154 (4): 291–98.
- Kilbourne, E. M. 1997. "The Public Health Consequences of Disasters." In, edited by E. K. Noji, 245–69. Oxford University Press, New York, Nw York.
- Knowlton, K., B. Lynn, R. A. Goldberg, C. Rosenzweig, C. Hogrefe, J. K. Rosenthal, and P. L. Kinney. 2007. "Projecting Heat-Related Mortality Impacts Under a Changing Climate in the New York City Region." *Am. J. Public Health* 97 (11): 2028–34.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. "World Map of the köppen-Geiger Climate Classification Updated." *Meteorol. Z.* 15: 259–63.
- Lowe, D., K. L. Ebi, and B. Forsberg. 2011. "Heatwave Early Warning Systems and Adaptation Advice to Reduce Human Health Consequences of Heatwaves." *Int. J. Environ. Res. Public Health* 8 (12): 4623–48.
- Maclure, M. 1991. "The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events." *Am. J. Epidemiol.* 133: 144–53.
- Madrigano, J., K. Lane, N. Petrovic, M. Ahmed, M. Blum, and T. Matte. 2018. "Awareness, Risk Perception, and Protective Behaviors for Extreme Heat and Climate Change in New York City." *Int. J. Environ. Res. Public Health* 15 (7): 1433.
- Masterton, J. M., and F. A. Richardson. 1979. *Humidex ; a Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity*. Downsview, Ont.: Environment Canada, Atmospheric Environment,.
- McElroy, S., L. Schwarz, H. Green, I. Corcos, K. Guirguis, A. Gershunov, and T. Benmarhnia. 2020. "Defining Heat Waves and Extreme Heat Events Using Sub-Regional Meteorological Data to Max-

- imize Benefits of Early Warning Systems to Population Health.” *Science of The Total Environment* 721: 137678.
- Medina-Ramón, M., and J. Schwartz. 2007. “Temperature, Temperature Extremes, and Mortality: A Study of Acclimatisation and Effect Modification in 50 u.s. Cities.” *Occup. Environ. Med.* 64: 827–33.
- Meehl, G. A., and C. Tebaldi. 2004. “More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century.” *Science* 13: 994–97.
- Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, et al. 2004. “The Multi-Institution North American Land Data Assimilation System (NLDAS): Utilizing Multiple GCIP Products and Partners in a Continental Distributed Hydrological Modeling System.” *J. Of Geo. Res* 109: D07S90.
- O’Neill, M. S., A. Zanobetti, and J. Schwartz. 2003. “Modifiers of the Temperature and Mortality Association in Seven US Cities.” *Am. J. Epidemiol.* 157 (12): 1074–82.
- . 2005. “Disparities by Race in Heat-Related Mortality in Four US Cities: The Role of Air Conditioning Prevalence.” *J. Urban Health* 82 (2): 191–97.
- R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reid, C. E., M. S. O’Neill, C. J. Gronlund, S. J. Brines, D. G. Brown, A. V. Diez-Roux, and J. Schwartz. 2009. “Mapping Community Determinants of Heat Vulnerability.” *Environ. Health Perspect.* 117 (11): 1730–36.
- Schwartz, J. 2005. “Who Is Sensitive to Extremes of Temperature? A Case-Only Analysis.” *Epidemiol.* 16 (1): 67–72.
- Semenza, J. C., C. H. Rubin, K. H. Falter, J. D. Selanikio, W. D. Flanders, H. L. Howe, and J. L. Wilhelm. 1996. “Heat-Related Deaths During the July 1995 Heat Wave in Chicago.” *N Engl. J. Med.* 335: 84–90.

- Shaposhnikov, D., B. Revich, T. Bellander, G. B. Bedada, M. Bottai, T. Kharkova, E. Kvasha, et al. 2014. “Mortality Related to Air Pollution with the Moscow Heat Wave and Wildfire of 2010.” *Epidemiol.* 25 (3): 359–64.
- Sheppard, J. H., and T. Lumley. 2005. “Overlap Bias in the Case-Crossover Design, with Application to Air Pollution Exposures.” *Stat. Med.* 24: 285–300.
- Sheridan, S. C., and A. J. Kalkstein. 2010. “Seasonal Variability in Heat- Related Mortality Across the United States.” *Nat. Hazards* 55: 291–305.
- Shindell, D., Y. Zhang, M. Scott, M. Ru, K. Stark, and K. L. Ebi. 2020. “The Effects of Heat Exposure on Human Mortality Throughout the United States.” *Geohealth* 4 (4): e2019GH000234.
- Stafoggia, M., F. Forastiere, D. Agostini, A. Biggeri, L. Bisanti, E. Cadum, N. Caranci, et al. 2006. “Vulnerability to Heat-Related Mortality: A Multicity, Population-Based, Case-Crossover Analysis.” *Epidemiol.* 17 (3): 315–23.
- Šeparović, L., R. Alexandru A. Laprise, A. Martynov, L. Sushama, K. Winger, K. Tete, and M. Valin. 2013. “Present Climate and Climate Change over North America as Simulated by the Fifth-Generation Canadian Regional Climate Model.” *Clim. Dyn.* 41: 3167–3201.
- Therneau, T. M. 2020. *A Package for Survival Analysis in r*. <https://CRAN.R-project.org/package=survival>.
- Tong, S., X. Y. Wang, and Y. Guo. 2012. “Assessing the Short-Term Effects of Heatwaves on Mortality and Morbidity in Brisbane, Australia: Comparison of Case-Crossover and Time Series Analyses.” *PLoS One* 7 (5): e37500.
- Vaidyanathan, A., S. Saha, A. M. Vicedo-Cabrera, A. Gasparrini, N. Abdurehman, R. Jordan, M. Hawkins, J. Hess, and A. Elixhauser. 2019. “Assessment of Extreme Heat and Hospitalizations to Inform Early Warning Systems.” *PNAS* 116 (12): 5420–27.
- Vicedo-Cabrera, A. M., F. Sera, and A. Gasparrini. 2019. “Hands-on Tutorial on a Modeling Framework for Projections of Climate Change Impacts on Health.” *Epidemiol.* 30 (3): 321–29.

- Weinberger, K. R., D. Harris, K. R. Spangler, A. Zanobetti, and G. A. Wellenius. 2020. “Estimating the Number of Excess Deaths Attributable to Heat in 297 United States Counties.” *Environ. Epidemiol.* 4 (3): e096.
- Weinberger, K. R., L. Haykin, M. N. Eliot, J. D. Schwartz, A. Gasparri, and G. A. Wellenius. 2017. “Projected Temperature-Related Deaths in Ten Large u.s. Metropolitan Areas Under Different Climate Change Scenarios.” *Environ. Int.* 107: 196–204.
- Wellenius, G. A., M. N. Eliot, K. F. Bush, D. Holt, Lincoln R. A., Smith A. E., and J. Gold. 2017. “Heat-Related Morbidity and Mortality in New England: Evidence for Local Policy.” *Environ. Res.* 156: 845–53.
- Wilson, L. A., G. G. Morgan, I. C. Hanigan, F. H. Johnston, H. Abu-Rayya, R. Broome, C. Gaskin, and B. Jalaludin. 2013. “The Impact of Heat on Mortality and Morbidity in the Greater Metropolitan Sydney Region: A Case Crossover Analysis.” *Environ. Health* 12: 98.
- Zanobetti, A., and J. Schwartz. 2008. “Temperature and Mortality in Nine US Cities.” *Epidemiol.* 19 (4): 563–70.

Appendix A. Wunderground vs. gridMET Humidex Values

The [gridMET](#) data ([Abatzoglou 2013](#)) includes daily maximum and minimum values for temperature and relative humidity. The desired heat exposure metric was daily maximum humidex, which is calculated using temperature and relative humidity. However, it is possible that a humidex calculation that uses both daily maximum temperature and daily maximum relative humidity would be biased upwards if these two meteorological variables are not maximized at the same time of day. To ascertain how best to calculate humidex from the gridMET data, a comparison was made to weather station observations obtained from [Weather Underground](#) (Wunderground).

- Wunderground humidex values are calculated using daily maximum temperature and daily maximum dew point
- gridMET humidex values are calculated using daily maximum temperature and various relative humidity values
 - Max RH: maximum relative humidity
 - Min RH: minimum relative humidity
 - Avg RH: average of maximum and minimum relative humidity

Four total comparisons were made: observations from SeaTac (Seattle) were obtained for 1980, 2001, and 2018, and observations from Spokane were obtained for 2018. The years were chosen to represent the beginning, middle, and end of the study time period, while the locations were chosen based on Wunderground data availability. In each year, only values from May - September were considered, as these were the months included in the analysis. The distributions of observed and modeled humidex are given in the tables below, along with RMSE values for the three different calculations.

Table A1: Humidex Comparison 1: 1980; SeaTac

	gridMET Relative Humidity			
	SeaTac	Min RH ¹	Avg RH ²	Max RH ³
Min	10.23	10.57	12.02	13.31
25 th	17.98	17.26	19.10	20.68
Median	20.98	21.04	23.09	24.92
Mean	21.27	21.24	23.22	25.20
75 th	24.60	25.08	27.09	29.28
Max	34.52	35.85	38.93	41.01

RMSE values ¹ 2.21 ² 3.11 ³ 4.8

Table A2: Humidex Comparison 2: 2001; SeaTac

	gridMET Relative Humidity			
	SeaTac	Min RH ¹	Avg RH ²	Max RH ³
Min	10.96	10.68	11.74	12.80
25 th	18.25	17.96	19.57	21.26
Median	22.36	21.78	23.75	25.81
Mean	22.25	21.92	23.96	26.00
75 th	26.19	26.41	28.32	30.76
Max	35.14	34.74	38.55	42.35

RMSE values ¹ 2.55 ² 3.29 ³ 4.93

Table A3: Humidex Comparison 3: 2018; SeaTac

	gridMET Relative Humidity			
	SeaTac	Min RH ¹	Avg RH ²	Max RH ³
Min	16.43	15.73	16.54	17.18
25 th	21.27	20.57	21.85	23.37
Median	24.67	24.25	26.40	28.42
Mean	25.71	25.19	27.90	30.61
75 th	29.65	29.09	32.52	36.61
Max	38.21	37.87	44.93	52.03

RMSE values ¹ 1.12 ² 3.09 ³ 6.17

Table A4: Humidex Comparison 4: 2018; Spokane

	gridMET Relative Humidity			
	SeaTac	Min RH ¹	Avg RH ²	Max RH ³
Min	10.65	9.80	10.79	11.78
25 th	20.52	19.96	22.77	25.94
Median	24.87	24.40	27.86	31.74
Mean	25.60	24.55	28.44	32.32
75 th	30.24	28.81	33.60	38.62
Max	40.17	37.70	44.15	50.69

RMSE values ¹ 1.44 ² 3.2 ³ 7.2

In all comparisons, the RMSE value was lowest when minimum relative humidity was used to calculate humidex from the gridMET data. Therefore, minimum relative humidity was chosen to calculate the exposure variable used in both the historical case-crossover analysis and in the projected heat-attributable mortality calculations.

Appendix B. Distribution of Humidex (1980-2018)

Figure B1 maps the spatial distribution of average daily maximum humidex values during the historical period (1980-2018). Humidex generally increases from northern to southern Washington. Moving inland from the coast, humidex increases until dropping significantly at the Cascades. On the eastern side of the Cascades, humidex increases sharply before declining once again in the eastern third of the state; however, all values east of the Cascades exceed the values west of the Cascades.

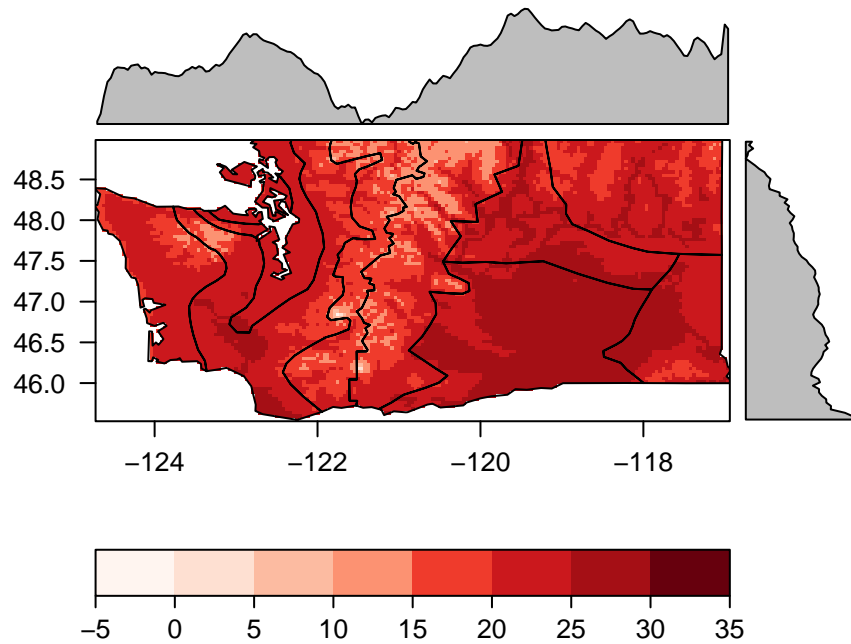


Figure B1: Spatial distribution of humidex in the historical period.

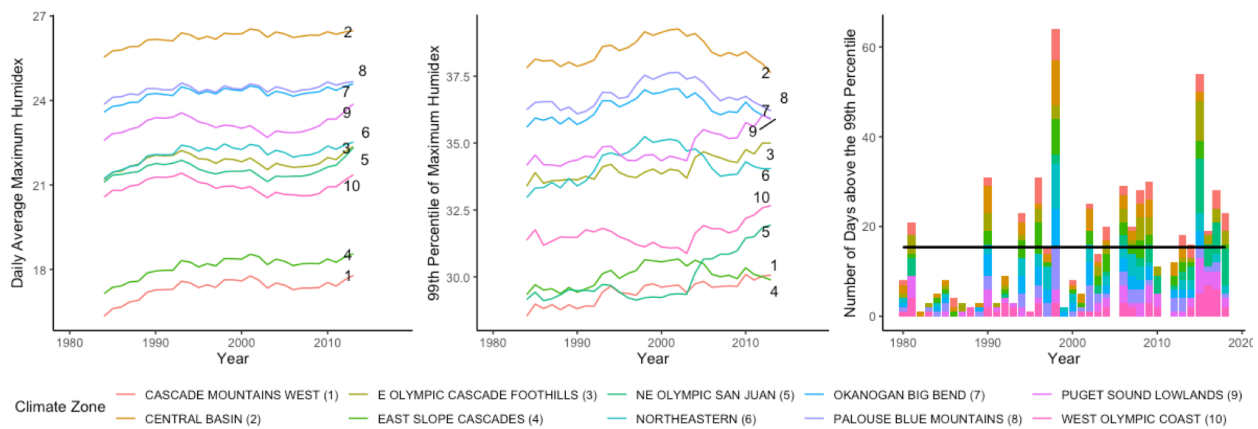


Figure B2: Trend of yearly average and yearly 99th percentile of humidex in the historical period.

Figure B2 displays the temporal trend of the yearly average (left) and yearly 99th percentile (middle) values of humidex from 1980 to 2018. Although the trends vary slightly across climate divisions, the average humidex value has increased from the beginning to the end of the historical period. The pattern is not as straightforward for the 99th percentile. Some divisions see a spike between 1990 to 2010, while this metric remains relatively constant in other divisions until around 2005, after which it increases sharply.

The right panel of Figure B2 shows the number of days in each year that exceed the 99th percentile of humidex for the entire historical period, separated by climate division. The horizontal black line indicates the expected height of the bar for each year, if the distribution was uniform (1.53 days per year per division, or 15.3 total days per year). By this metric, 1998 and 2015 are the hottest and second-hottest, respectively, years on record, and years in the second half of the historical period are much more likely to have a day exceeding the 99th percentile of humidex than years in the first half.

Appendix C. Full Effect Modification Results

The effect modification results given in Tables [2.3](#) and [2.4](#) are pooled across all climate zones. This section gives the results for each climate division.

Age/Cause of Death by Climate Division

Table C1: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 0-4 and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	9.5 (0.1, 878.77)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Central Basin	0.97 (0.64, 1.47)	NA	0.95 (0.1, 8.68)	2.19 (0.19, 25.18)	NA	NA	0.62 (0.03, 14.16)	NA	NA	NA
E Olympic Cascade Foothills	0.52 (0.28, 0.95)	NA	11.47 (0.13, 1001.26)	79.09 (0.27, 23374.69)	NA	0 (0, 54737051990.11)	0.13 (0, 7.54)	NA	NA	NA
E Slope Cascades	0.22 (0.02, 2.88)	NA	NA	NA	NA	NA	NA	NA	NA	NA
NE Olympic San Juan	1.24 (0.14, 11.07)	NA	1.64 (0, 1087.91)	NA	NA	NA	NA	NA	NA	NA
Northeastern	1.12 (0.64, 1.95)	NA	2.24 (0.07, 71.11)	NA	NA	0.16 (0, 62788514587.29)	0.04 (0, 3.7)	1.41 (0.01, 329.36)	NA	NA
Okanogan Big Bend	0.97 (0.09, 10.71)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Palouse Blue Mountains	0.55 (0.06, 4.81)	NA	NA	2.08 (0, 2176.48)	NA	NA	NA	NA	NA	NA
Puget Sound Lowlands	0.93 (0.73, 1.18)	NA	0.94 (0.25, 3.47)	0.85 (0.2, 3.56)	NA	0.16 (0, 45.34)	0.5 (0.09, 2.74)	1.75 (0.1, 32.07)	NA	NA
West Olympic Coast	1 (0.1, 10.15)	NA	0.02 (0, 252.94)	0.22 (0, 23665.07)	NA	NA	NA	NA	NA	NA
Fixed Effects Pooled OR	0.9 (0.75, 1.08)	NA	1.12 (0.41, 3.08)	1.36 (0.44, 4.16)	NA	0.12 (0, 27.95)	0.36 (0.1, 1.36)	1.83 (0.15, 22.31)	NA	NA

74

Table C2: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 5-14 and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Central Basin	0.98 (0.26, 3.59)	NA	1 (0.02, 48.61)	1.07 (0.02, 60.41)	NA	NA	0.56 (0.01, 59.17)	NA	NA	NA
E Olympic Cascade Foothills	0.62 (0.14, 2.79)	NA	0.11 (0, 327.91)	0.27 (0, 772)	NA	NA	2.9 (0, 3931.34)	NA	NA	NA
E Slope Cascades	0.87 (0, 854.19)	NA	NA	NA	NA	NA	NA	NA	NA	NA
NE Olympic San Juan	25.4 (0.09, 7553.66)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Northeastern	0.64 (0.14, 2.8)	NA	0 (0, 2142.95)	0 (0, 2142.95)	NA	NA	NA	NA	NA	NA
Okanogan Big Bend	0.79 (0.02, 28.1)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Palouse Blue Mountains	0.49 (0, 113.59)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Puget Sound Lowlands	1 (0.54, 1.86)	NA	0.83 (0.08, 9.01)	0.15 (0, 5.68)	NA	NA	2.45 (0.1, 57.89)	NA	NA	NA
West Olympic Coast	0.56 (0.02, 15.12)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Fixed Effects Pooled OR	0.9 (0.56, 1.45)	NA	0.68 (0.1, 4.71)	0.3 (0.02, 3.56)	NA	NA	1.64 (0.15, 18.52)	NA	NA	NA

Table C3: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 15-44 and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	0.47 (0.02, 10.9)	NA	0.05 (0, 24.24)	0.21 (0, 76.52)	1.7 (0, 1053.41)	NA	NA	NA	NA	NA
Central Basin	1.12 (0.82, 1.52)	1.32 (0.22, 7.92)	1.15 (0.63, 2.08)	1.14 (0.58, 2.24)	1.19 (0.44, 3.17)	0.77 (0.06, 10.45)	1.18 (0.22, 6.33)	0.74 (0.01, 48.85)	NA	0.71 (0.09, 5.4)
E Olympic Cascade Foothills	1.02 (0.74, 1.4)	1.88 (0.23, 15.54)	0.81 (0.44, 1.51)	0.93 (0.48, 1.82)	1.05 (0.4, 2.75)	0.8 (0.06, 11.11)	1.04 (0.26, 4.08)	0.18 (0, 44.79)	NA	0.14 (0.02, 0.88)
E Slope Cascades	1.22 (0.23, 6.37)	61.57 (0, 7531181.75)	0.58 (0.05, 6.64)	0.46 (0.02, 13.7)	0.1 (0, 19.35)	NA	NA	NA	NA	NA
NE Olympic San Juan	1.35 (0.51, 3.59)	NA	2.89 (0.27, 31.22)	4.44 (0.32, 61.96)	NA	0.08 (0, 1495.07)	0.02 (0, 9.9)	NA	NA	NA
Northeastern	1.03 (0.73, 1.47)	1.44 (0.12, 17.68)	1.52 (0.72, 3.2)	1.37 (0.58, 3.26)	1.71 (0.51, 5.71)	2.88 (0.19, 43.23)	0.43 (0.1, 1.93)	2.94 (0.04, 220.7)	NA	1.41 (0.28, 7.02)
Okanogan Big Bend	1.17 (0.38, 3.65)	60.93 (0.19, 19047.03)	2.94 (0.36, 23.77)	4.01 (0.37, 43.97)	102.43 (0.31, 33914.08)	1.85 (0, 724.48)	0.15 (0, 27.42)	NA	NA	1.17 (0, 336.79)
Palouse Blue Mountains	1.41 (0.49, 4.1)	0.64 (0.01, 65.62)	1.57 (0.23, 10.68)	0.72 (0.08, 6.61)	0.29 (0.01, 5.93)	NA	0.76 (0, 141.89)	NA	NA	0.01 (0, 247.69)
Puget Sound Lowlands	1.03 (0.9, 1.18)	0.94 (0.34, 2.56)	1.04 (0.78, 1.39)	1.11 (0.79, 1.54)	1.03 (0.62, 1.71)	0.64 (0.22, 1.82)	0.9 (0.44, 1.82)	2.5 (0.31, 20.41)	NA	1.06 (0.48, 2.32)
West Olympic Coast	0.88 (0.23, 3.4)	NA	1.11 (0.1, 12.09)	1.19 (0.08, 18.54)	1.34 (0.11, 17.11)	NA	0.19 (0, 286.92)	0.95 (0, 273.76)	NA	0.09 (0, 40.44)
Fixed Effects Pooled OR	1.05 (0.94, 1.17)	1.25 (0.6, 2.62)	1.07 (0.86, 1.34)	1.12 (0.87, 1.44)	1.12 (0.77, 1.63)	1.12 (0.77, 1.63)	0.81 (0.48, 1.38)	1.61 (0.34, 7.67)	NA	0.81 (0.44, 1.48)

Table C4: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 45-64 and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	1.22 (0.47, 3.15)	6.54 (0.39, 109.34)	1.01 (0.47, 2.19)	1.29 (0.54, 3.09)	1.27 (0.39, 4.11)	NA	2.05 (0.43, 9.69)	NA	NA	0.67 (0.03, 16.18)
Central Basin	1.1 (0.96, 1.26)	1.17 (0.79, 1.72)	1.07 (0.98, 1.16)	1.09 (0.98, 1.2)	1.06 (0.94, 1.19)	1.05 (0.79, 1.39)	0.99 (0.82, 1.19)	0.93 (0.5, 1.73)	0.73 (0.15, 3.53)	0.99 (0.65, 1.51)
E Olympic Cascade Foothills	1 (0.89, 1.14)	1.1 (0.75, 1.63)	1.1 (1.01, 1.2)	1.09 (0.98, 1.2)	1.09 (0.96, 1.24)	1.26 (0.92, 1.73)	1.07 (0.9, 1.27)	1.27 (0.68, 2.4)	2.2 (0.62, 7.77)	0.98 (0.69, 1.37)
E Slope Cascades	1.06 (0.64, 1.77)	0.65 (0.09, 4.69)	0.99 (0.69, 1.42)	1 (0.67, 1.51)	1.09 (0.67, 1.78)	1.41 (0.32, 6.15)	1 (0.47, 2.1)	0.82 (0.04, 16.47)	NA	1.78 (0.41, 7.69)
NE Olympic San Juan	1.04 (0.71, 1.54)	2.63 (0.98, 7.06)	1.27 (1.01, 1.58)	1.27 (0.98, 1.65)	1.17 (0.84, 1.63)	1.36 (0.6, 3.1)	0.8 (0.51, 1.26)	NA	NA	1.32 (0.53, 3.24)
Northeastern	1.05 (0.91, 1.21)	1.47 (0.93, 2.3)	1.09 (0.99, 1.2)	1.08 (0.96, 1.21)	1.08 (0.94, 1.24)	1.14 (0.84, 1.54)	1.14 (0.95, 1.36)	0.66 (0.32, 1.33)	0.63 (0.1, 4.18)	1.26 (0.84, 1.9)
Okanogan Big Bend	0.91 (0.58, 1.4)	1.62 (0.44, 5.9)	1.21 (0.9, 1.61)	1.13 (0.8, 1.58)	1.36 (0.9, 2.05)	1.22 (0.46, 3.27)	0.86 (0.49, 1.51)	0.42 (0.05, 3.56)	NA	1.02 (0.29, 3.63)
Palouse Blue Mountains	0.74 (0.48, 1.14)	1.56 (0.48, 5.04)	1.25 (0.96, 1.61)	1.36 (1, 1.83)	1.13 (0.77, 1.65)	1.26 (0.52, 3.04)	0.99 (0.59, 1.68)	1.53 (0.35, 6.76)	NA	0.5 (0.15, 1.72)
Puget Sound Lowlands	1.09 (1.02, 1.16)	1.13 (0.92, 1.38)	1.1 (1.05, 1.15)	1.09 (1.04, 1.15)	1.09 (1.02, 1.16)	1.17 (1.01, 1.35)	1.12 (1.03, 1.22)	1.07 (0.79, 1.45)	1.03 (0.47, 2.26)	1.01 (0.85, 1.21)
West Olympic Coast	1.01 (0.67, 1.52)	1.13 (0.33, 3.81)	1.18 (0.89, 1.58)	1.18 (0.85, 1.62)	1.27 (0.84, 1.9)	1.31 (0.39, 4.36)	0.96 (0.55, 1.68)	2.19 (0.25, 19.27)	NA	0.68 (0.2, 2.25)
Fixed Effects Pooled OR	1.06 (1.02, 1.12)	1.2 (1.03, 1.39)	1.1 (1.07, 1.14)	1.1 (1.06, 1.14)	1.09 (1.04, 1.14)	1.16 (1.05, 1.29)	1.08 (1.02, 1.15)	1.03 (0.82, 1.28)	1.15 (0.66, 2)	1.02 (0.9, 1.17)

Table C5: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 65-84 and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	0.92 (0.5, 1.71)	0.76 (0.01, 106.59)	0.45 (0.14, 1.42)	0.44 (0.11, 1.73)	0.53 (0.09, 3.06)	1.2 (0.01, 279.48)	1.19 (0.14, 10.1)	NA	NA	5.38 (0.03, 961.13)
Central Basin	1.04 (0.96, 1.13)	1.37 (0.82, 2.28)	1.07 (0.94, 1.21)	1.07 (0.92, 1.24)	1.02 (0.86, 1.21)	1.16 (0.78, 1.72)	1 (0.78, 1.29)	1.15 (0.46, 2.86)	0.67 (0.06, 7.08)	1.16 (0.55, 2.43)
E Olympic Cascade Foothills	1.08 (1, 1.17)	1.19 (0.72, 1.97)	1.09 (0.96, 1.23)	1.08 (0.93, 1.25)	1.08 (0.9, 1.29)	1.41 (0.91, 2.17)	1.05 (0.84, 1.31)	1.36 (0.55, 3.33)	2.2 (0.4, 11.97)	1.03 (0.57, 1.86)
E Slope Cascades	1.02 (0.75, 1.39)	0.51 (0.04, 6.67)	1.25 (0.75, 2.1)	1.37 (0.75, 2.48)	1.57 (0.78, 3.18)	2.05 (0.33, 12.73)	1.08 (0.43, 2.68)	0.16 (0, 54.78)	0.31 (0, 244.52)	0.98 (0.04, 26.41)
NE Olympic San Juan	1.12 (0.93, 1.36)	2.41 (0.64, 9.07)	1.16 (0.84, 1.61)	1.27 (0.87, 1.86)	1.28 (0.8, 2.04)	1.58 (0.52, 4.86)	0.62 (0.33, 1.17)	3.22 (0.19, 55.3)	0.4 (0, 2019.88)	1.6 (0.31, 8.39)
Northeastern	1.09 (1, 1.18)	1.4 (0.76, 2.56)	1.12 (0.97, 1.29)	1.09 (0.93, 1.29)	1.13 (0.93, 1.37)	1.38 (0.91, 2.12)	1.22 (0.96, 1.54)	0.93 (0.33, 2.63)	0.57 (0.03, 10.26)	1.1 (0.53, 2.25)
Okanogan Big Bend	1.07 (0.83, 1.37)	2.16 (0.41, 11.45)	1.42 (0.93, 2.15)	1.38 (0.85, 2.23)	1.74 (0.98, 3.06)	1.75 (0.35, 8.8)	0.87 (0.43, 1.78)	0.07 (0, 3.09)	0.01 (0, 23955.1)	3.45 (0.44, 26.89)
Palouse Blue Mountains	0.98 (0.77, 1.24)	2.09 (0.38, 11.41)	1.24 (0.84, 1.82)	1.47 (0.93, 2.31)	1.22 (0.7, 2.13)	0.8 (0.21, 3.1)	0.87 (0.44, 1.75)	0.38 (0.02, 8.14)	0.21 (0, 15.96)	0.55 (0.06, 5.08)
Puget Sound Lowlands	1.07 (1.03, 1.11)	1.16 (0.88, 1.53)	1.08 (1.02, 1.16)	1.08 (1, 1.16)	1.1 (1, 1.2)	1.13 (0.92, 1.39)	1.05 (0.93, 1.18)	1.01 (0.65, 1.58)	0.77 (0.25, 2.3)	1.04 (0.76, 1.42)
West Olympic Coast	1.08 (0.85, 1.38)	0.88 (0.12, 6.33)	1.22 (0.81, 1.85)	1.21 (0.76, 1.93)	1.27 (0.71, 2.25)	1.61 (0.31, 8.43)	1.02 (0.5, 2.08)	1.36 (0.05, 37.67)	NA	1.31 (0.2, 8.48)
Fixed Effects Pooled OR	1.07 (1.04, 1.1)	1.25 (1.03, 1.52)	1.09 (1.04, 1.15)	1.09 (1.04, 1.16)	1.1 (1.03, 1.18)	1.21 (1.04, 1.41)	1.05 (0.96, 1.14)	1.04 (0.75, 1.45)	0.87 (0.39, 1.91)	1.08 (0.86, 1.37)

Table C6: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for ages 85+ and for each climate division.

Climate Division	All Non-Traumatic Causes	Diabetes	Circulatory	Cardiovascular	Ischemic	Cerebrovascular	Respiratory	Nephritis and nephrotic	Acute Renal Failure	Mental Disorders
Cascade Mountains West	1.94 (0.76, 4.92)	2.3 (0.06, 83.25)	4.55 (1.16, 17.88)	6.78 (1.39, 33.01)	3.35 (0.32, 35.28)	0 (0, 82204927.92)	3.97 (0.23, 69.37)	NA	NA	0.22 (0, 14.28)
Central Basin	0.98 (0.89, 1.09)	0.99 (0.44, 2.24)	1.02 (0.89, 1.18)	1.08 (0.91, 1.28)	1.11 (0.9, 1.37)	0.82 (0.53, 1.27)	0.96 (0.69, 1.34)	0.77 (0.29, 2.03)	0.81 (0.09, 7.16)	0.82 (0.45, 1.48)
E Olympic Cascade Foothills	1.11 (1, 1.22)	0.54 (0.2, 1.52)	1.15 (0.99, 1.34)	1.12 (0.94, 1.33)	1.2 (0.96, 1.51)	1.28 (0.76, 2.15)	1.11 (0.79, 1.55)	1.03 (0.35, 3.05)	2.69 (0.24, 30.02)	1.04 (0.64, 1.69)
E Slope Cascades	1.01 (0.68, 1.5)	0.84 (0.02, 46.32)	0.77 (0.41, 1.45)	0.72 (0.35, 1.45)	0.8 (0.32, 2)	0.29 (0.01, 7.95)	1.51 (0.3, 7.45)	1.02 (0.02, 52.8)	NA	4.46 (0.58, 34.52)
NE Olympic San Juan	1.24 (1, 1.55)	2.36 (0.28, 19.97)	1.35 (0.96, 1.89)	1.21 (0.8, 1.81)	1.03 (0.6, 1.78)	1.13 (0.33, 3.93)	1.15 (0.58, 2.31)	0.39 (0.01, 11.06)	NA	1.37 (0.41, 4.54)
Northeastern	1.02 (0.92, 1.14)	2.29 (0.89, 5.93)	1.02 (0.87, 1.19)	1.02 (0.85, 1.23)	0.98 (0.78, 1.25)	0.77 (0.47, 1.26)	1.02 (0.72, 1.43)	0.63 (0.21, 1.9)	0.91 (0.07, 11.43)	1.34 (0.74, 2.4)
Okanogan Big Bend	1.06 (0.76, 1.48)	0.01 (0, 0.87)	0.99 (0.62, 1.59)	0.84 (0.48, 1.47)	0.87 (0.42, 1.79)	0.67 (0.16, 2.82)	0.88 (0.3, 2.59)	2.43 (0.1, 56.66)	NA	0.74 (0.1, 5.36)
Palouse Blue Mountains	1.24 (0.94, 1.64)	2.74 (0.37, 20.04)	1.43 (0.97, 2.12)	1.63 (1.02, 2.61)	1.63 (0.88, 3.03)	1.25 (0.38, 4.1)	1.37 (0.56, 3.36)	2.1 (0.41, 10.79)	NA	0.58 (0.1, 3.28)
Puget Sound Lowlands	1.12 (1.07, 1.18)	1.05 (0.65, 1.68)	1.13 (1.05, 1.22)	1.12 (1.03, 1.22)	1.1 (0.98, 1.23)	1.25 (1.01, 1.56)	1.23 (1.05, 1.43)	1.25 (0.77, 2.05)	1.25 (0.39, 4.02)	1.06 (0.83, 1.35)
West Olympic Coast	1.03 (0.72, 1.47)	3.56 (0.22, 57.11)	1.07 (0.65, 1.78)	1.11 (0.63, 1.98)	1.34 (0.63, 2.87)	0.94 (0.08, 10.62)	0.76 (0.22, 2.58)	10.75 (0.06, 2065.06)	NA	0.49 (0.07, 3.53)
Fixed Effects Pooled OR	1.09 (1.06, 1.13)	1.11 (0.8, 1.55)	1.11 (1.05, 1.17)	1.11 (1.04, 1.18)	1.1 (1.02, 1.2)	1.1 (0.93, 1.3)	1.14 (1.02, 1.28)	1.09 (0.76, 1.57)	1.09 (0.76, 1.57)	1.05 (0.87, 1.27)

Age/Cause of Death by Sex

Table C7: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for females for each climate division.

Cause of Death	All Ages	0-4	5-14	15-44	45-64	64-84	85+
All Non-Traumatic Causes	1.08 (1.05, 1.11)	0.89 (0.68, 1.17)	0.95 (0.46, 1.92)	1.06 (0.90, 1.25)	1.10 (1.02, 1.18)	1.09 (1.04, 1.14)	1.03 (1.04, 1.13)
Diabetes	1.16 (0.95, 1.43)	NA	NA	1.18 (0.31, 4.53)	1.16 (0.71, 1.89)	1.29 (0.99, 1.70)	0.91 (0.58, 1.41)
Circulatory	1.11 (1.06, 1.16)	1.82 (0.38, 8.77)	0.74 (0.03, 16.76)	1.04 (0.69, 1.57)	1.14 (0.98, 1.34)	1.13 (1.05, 1.21)	1.10 (1.03, 1.17)
Cardiovascular	1.11 (1.05, 1.17)	1.62 (0.26, 10.23)	0.39 (0.01, 13.56)	1.03 (0.63, 1.70)	1.16 (0.97, 1.39)	1.13 (1.04, 1.22)	1.09 (1.01, 1.18)
Ischemic	1.10 (1.02, 1.18)	NA	NA	1.39 (0.57, 3.40)	1.11 (0.88, 1.41)	1.10 (1.00, 1.23)	1.08 (0.97, 1.20)
Cerebrovascular	1.20 (1.04, 1.37)	NA	NA	1.39 (0.57, 4.00)	1.32 (0.75, 2.33)	1.27 (1.03, 1.55)	1.13 (0.94, 1.37)
Respiratory	1.11 (1.02, 1.22)	0.61 (0.10, 3.79)	1.18 (0.04, 31.58)	0.88 (0.40, 1.94)	1.28 (0.98, 1.68)	1.04 (0.92, 1.18)	1.18 (1.02, 1.38)
Nephritis and nephrotic	1.12 (0.82, 1.54)	NA	NA	2.62 (0.30, 22.72)	0.65 (0.25, 1.69)	1.18 (0.73, 1.91)	1.18 (0.73, 1.91)
Acute Renal Failure	1.38 (0.65, 2.91)	NA	NA	2.62 (0.30, 22.72)	0.55 (0.03, 11.03)	1.18 (0.39, 3.55)	1.59 (0.50, 5.04)
Mental Disorders	0.99 (0.83, 1.18)	NA	NA	1.18 (0.31, 4.53)	0.81 (0.40, 1.66)	0.98 (0.70, 1.36)	1.01 (0.81, 1.26)

Table C8: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for males for each climate division.

Cause of Death	All Ages	0-4	5-14	15-44	45-64	64-84	85+
All Non-Traumatic Causes	1.06 (1.03, 1.09)	0.91 (0.71, 1.15)	0.90 (0.47, 1.72)	1.03 (0.90, 1.19)	1.04 (0.98, 1.11)	1.05 (1.01, 1.09)	1.12 (1.05, 1.19)
Diabetes	1.23 (1.00, 1.52)	NA	NA	1.29 (0.48, 3.47)	1.08 (0.70, 1.68)	1.20 (0.91, 1.60)	1.43 (0.84, 2.43)
Circulatory	1.09 (1.04, 1.14)	0.77 (0.18, 3.29)	0.68 (0.05, 9.20)	1.11 (0.84, 1.45)	1.07 (0.97, 1.19)	1.07 (1.00, 1.14)	1.13 (1.04, 1.24)
Cardiovascular	1.09 (1.03, 1.14)	1.01 (0.21, 4.97)	0.41 (0.02, 6.98)	1.17 (0.87, 1.58)	1.07 (0.95, 1.19)	1.07 (1.00, 1.15)	1.13 (1.02, 1.26)
Ischemic	1.09 (1.02, 1.16)	NA	NA	1.05 (0.69, 1.62)	1.01 (0.88, 1.16)	1.10 (1.01, 1.20)	1.14 (1.01, 1.30)
Cerebrovascular	1.11 (0.94, 1.32)	0.24 (0.00, 54.18)	NA	0.85 (0.27, 2.72)	1.23 (0.73, 2.07)	1.15 (0.92, 1.44)	1.14 (1.01, 1.30)
Respiratory	1.05 (0.96, 1.15)	0.27 (0.04, 1.70)	3.70 (0.18, 74.16)	0.82 (0.39, 1.71)	1.06 (0.83, 1.37)	1.05 (0.94, 1.18)	1.09 (0.92, 1.31)
Nephritis and nephrotic	0.92 (0.66, 1.27)	2.70 (0.08, 92.42)	NA	0.86 (0.09, 7.75)	0.74 (0.30, 1.84)	0.92 (0.57, 1.47)	0.96 (0.55, 1.68)
Acute Renal Failure	0.86 (0.38, 1.96)	NA	NA	NA	0.74 (0.30, 1.84)	0.47 (0.14, 1.62)	0.96 (0.55, 1.68)
Mental Disorders	1.06 (0.87, 1.31)	NA	NA	0.69 (0.31, 1.54)	0.93 (0.58, 1.48)	1.20 (0.86, 1.67)	1.14 (0.80, 1.61)

Race/Ethnicity

Table C9: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, by race/ethnicity for each climate division.

Climate Division	White	Black	Asian	Hispanic	Native American	Native Hawaiian or Other Pacific Islander
Cascade Mountains West	1.17 (0.75, 1.83)	NA	NA	NA	NA	NA
Central Basin	1.04 (0.98, 1.1)	0.99 (0.6, 1.66)	0.96 (0.43, 2.13)	1.05 (0.82, 1.35)	1.07 (0.68, 1.69)	0.84 (0.24, 2.99)
E Olympic Cascade Foothills	1.07 (1.01, 1.13)	1 (0.55, 1.83)	1 (0.58, 1.73)	1.12 (0.51, 2.47)	0.77 (0.41, 1.42)	0.96 (0.46, 1.98)
E Slope Cascades	1.01 (0.81, 1.25)	0.01 (0, 20.58)	NA	2.73 (0.16, 45.18)	1.66 (0.08, 34.55)	NA
NE Olympic San Juan	1.16 (1.02, 1.34)	NA	0.85 (0.12, 6.04)	0.74 (0.05, 11.63)	1.29 (0.3, 5.5)	NA
Northeastern	1.06 (1, 1.12)	1.08 (0.6, 1.93)	1.01 (0.47, 2.15)	0.75 (0.14, 4)	1.04 (0.59, 1.83)	0.85 (0.24, 3.01)
Okanogan Big Bend	1.04 (0.86, 1.25)	1.13 (0.09, 14.73)	1.41 (0.05, 37.85)	0.45 (0.06, 3.16)	1.18 (0.5, 2.76)	NA
Palouse Blue Mountains	1.02 (0.86, 1.21)	1.74 (0.11, 26.51)	0.91 (0.08, 10.69)	2.84 (0.11, 70.94)	2.1 (0.09, 48.85)	NA
Puget Sound Lowlands	1.08 (1.05, 1.12)	1.17 (1.03, 1.34)	1.07 (0.93, 1.23)	0.95 (0.7, 1.29)	1.2 (0.91, 1.59)	0.98 (0.77, 1.25)
West Olympic Coast	1.06 (0.88, 1.27)	NA	0.84 (0.01, 103.98)	1.37 (0.03, 73.3)	0.47 (0.14, 1.52)	1.07 (0.03, 33.56)
Fixed Effects Pooled OR	1.07 (1.05, 1.09)	1.15 (1.02, 1.3)	1.06 (0.93, 1.21)	1.01 (0.84, 1.21)	1.08 (0.89, 1.32)	0.97 (0.77, 1.2)

Table C10: Odds ratio (95% CI) of non-traumatic mortality comparing the 99th and 50th percentiles of humidex, for each racial/ethnic category by age, sex, and cause of death.

Category	White	Black	Asian	Native American	Hispanic	Native Hawaiian or Other Pacific Islander
Total	1.07 (1.05, 1.09)	1.15 (1.02, 1.3)	1.06 (0.93, 1.21)	1.01 (0.84, 1.21)	1.08 (0.89, 1.32)	0.97 (0.77, 1.2)
Age						
0 – 4	0.88 (0.71, 1.08)	0.82 (0.45, 1.49)	0.96 (0.26, 3.49)	1.39 (0.51, 3.8)	0.92 (0.5, 1.7)	1.66 (0.59, 4.69)
5 – 14	1 (0.58, 1.72)	0.31 (0.06, 1.6)	1.66 (0.18, 15.46)	0.52 (0.03, 8.4)	1.1 (0.11, 11.32)	1.46 (0.09, 22.83)
15 – 44	1.02 (0.91, 1.15)	1.12 (0.74, 1.69)	1.23 (0.63, 2.41)	1.17 (0.64, 2.14)	1.26 (0.74, 2.15)	0.89 (0.46, 1.72)
45 – 64	1.05 (1, 1.11)	1.14 (0.92, 1.43)	1.26 (0.92, 1.74)	1.3 (0.94, 1.79)	1.09 (0.76, 1.55)	0.95 (0.62, 1.44)
65 – 84	1.07 (1.04, 1.1)	1.27 (1.05, 1.54)	0.98 (0.81, 1.2)	0.95 (0.71, 1.28)	0.96 (0.71, 1.31)	0.94 (0.66, 1.35)
85+	1.1 (1.06, 1.14)	1.14 (0.81, 1.6)	1.06 (0.84, 1.35)	0.86 (0.47, 1.58)	0.84 (0.5, 1.42)	1.04 (0.52, 2.05)
Sex						
Female	1.08 (1.05, 1.11)	1.26 (1.05, 1.52)	1.12 (0.93, 1.34)	1.18 (0.9, 1.54)	0.93 (0.7, 1.24)	0.96 (0.69, 1.34)
Male	1.06 (1.03, 1.09)	1.09 (0.92, 1.28)	1 (0.83, 1.21)	1 (0.77, 1.3)	1.08 (0.84, 1.39)	0.98 (0.72, 1.34)
Cause of Death						
Diabetes	1.19 (1.02, 1.39)	1.41 (0.71, 2.81)	1.37 (0.5, 3.79)	1.4 (0.47, 4.17)	1.08 (0.38, 3.03)	1.09 (0.28, 4.18)
Circulatory	1.1 (1.06, 1.13)	1.19 (0.97, 1.47)	1.07 (0.86, 1.34)	1.25 (0.89, 1.76)	1.09 (0.78, 1.52)	1.25 (0.85, 1.85)
Cardiovascular	1.09 (1.05, 1.14)	1.18 (0.93, 1.5)	1.1 (0.83, 1.46)	1.24 (0.84, 1.83)	1.09 (0.73, 1.62)	1.24 (0.78, 1.97)
Ischemic	1.09 (1.04, 1.14)	1.18 (0.86, 1.63)	1.25 (0.87, 1.78)	1.29 (0.78, 2.14)	1.14 (0.68, 1.92)	1.14 (0.61, 2.11)
Cerebrovascular	1.16 (1.04, 1.29)	1.26 (0.61, 2.58)	1.2 (0.53, 2.73)	1.08 (0.29, 4)	0.91 (0.25, 3.38)	2.92 (0.73, 11.6)
Respiratory	1.08 (1.02, 1.16)	1.25 (0.77, 2.02)	1.08 (0.66, 1.77)	0.73 (0.39, 1.38)	0.82 (0.38, 1.76)	1.38 (0.56, 3.37)
Nephritis and nephrotic	1.02 (0.8, 1.29)	1.06 (0.39, 2.88)	1.44 (0.46, 4.54)	0.92 (0.04, 22.86)	0.56 (0.09, 3.56)	0.55 (0.06, 5.21)
Acute Renal Failure	1.21 (0.69, 2.13)	0.77 (0.02, 34.53)	0.26 (0, 20.77)			
Mental Disorders	1.03 (0.9, 1.19)	0.78 (0.3, 2.04)	0.7 (0.26, 1.88)	1.23 (0.41, 3.72)	0.59 (0.12, 2.78)	1.53 (0.15, 15.54)

Results are pooled across all climate divisions using fixed-effects meta-analysis.

Appendix D. Sensitivity Analysis: Case-Crossover vs. Time Series

Studies that have applied both a time series and case-crossover approach assert that the two techniques generate results which are comparable quantitatively and identical qualitatively ([Zanobetti and Schwartz 2008](#); [Tong, Wang, and Guo 2012](#)). This result was verified using King County ($n = 149,362$) to conduct a sensitivity analysis.

- The case-crossover approach, implemented using a conditional logistic regression model, incorporates individual-level humidex using a quadratic polynomial. Time is controlled for by-design.
- The time series approach, implemented using a generalized additive quasi-Poisson model, incorporates daily averages of individual-level humidex values using a quadratic polynomial. Time is controlled for by using a spline for year, indicator variables for month and day-of-week (to mimic the case-crossover format), and post-censal population data from the [Office of Financial Management](#).

Notably, the case-crossover approach produces results in terms of odds/odds ratios, while the time series approach produces results in terms of relative risk. However, since the outcome of interest, mortality, is a rare event, the RR and OR should be approximately equal, allowing direct comparison ([Basu, Dominici, and Samet 2005](#)). The results of these analyses are:

- Case-Crossover Analysis
 - OR (99th vs. 50th): 1.10 (1.06, 1.14)
 - Minimum Mortality Temperature: 20.09
- Time Series Analysis
 - RR (99th vs. 50th): 1.10 (1.08, 1.12)
 - Minimum Mortality Temperature: 20.78

As expected, the results are identical qualitatively almost identical quantitatively, with the time series approach producing a more precise estimate. The dose-response curves from each approach are given in [D1](#).

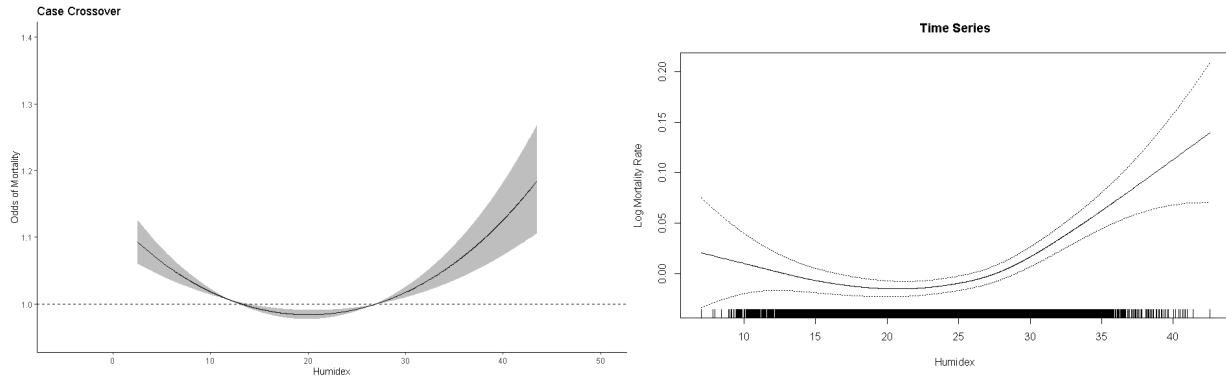


Figure D1: Humidex-mortality exposure-response curves for King County, produced using both a case-crossover analysis and a time series analysis.

Appendix E. Climate Classifications in Washington

Guttman and Quayle (1996) outline the history of climatic divisions in the contiguous United States. Initially (in 1912), Washington was simply divided into two climatology sections - east and west of the Cascades - with this separation chosen for the purpose of summarizing and publishing data for economical distribution; in other words, they were based on mailing practicality and not climate homogeneity. These sections were soon refined based on local agricultural or water needs and/or communication considerations. In 1949, a more systematic definition was created by adopting the U.S. Department of Agriculture Bureau of Agricultural Economics Crop Reporting Districts as divisions; this decision was justified due to the relationship between crop type and climate classification. Some minor revisions followed, but in most cases the definitions have not changed since 1949. In Washington, this has resulted in the ten climate divisions now identified by NOAA, mapped in Figure 2.1. However, it is possible that the divisional boundaries do not create regions of climatological homogeneity, and instead are based on drainage basins and/or major crops.

An alternative to NOAA's climate divisions is the [Köppen-Geiger climate classification](#), a scheme based primarily on five distinct vegetation types and secondarily on temperature and precipitation (Kottek et al. 2006). This approach may be preferable because its consideration of climatological variables is explicit. Data from [ESRI](#) was used to map the Köppen-Geiger classification for Washington in Figure E1, with the boundaries of NOAA's climate divisions overlaid for comparison (left); a finer resolution map for the Köppen-Geiger classification, obtained from [WikiMedia](#), is also included (right).

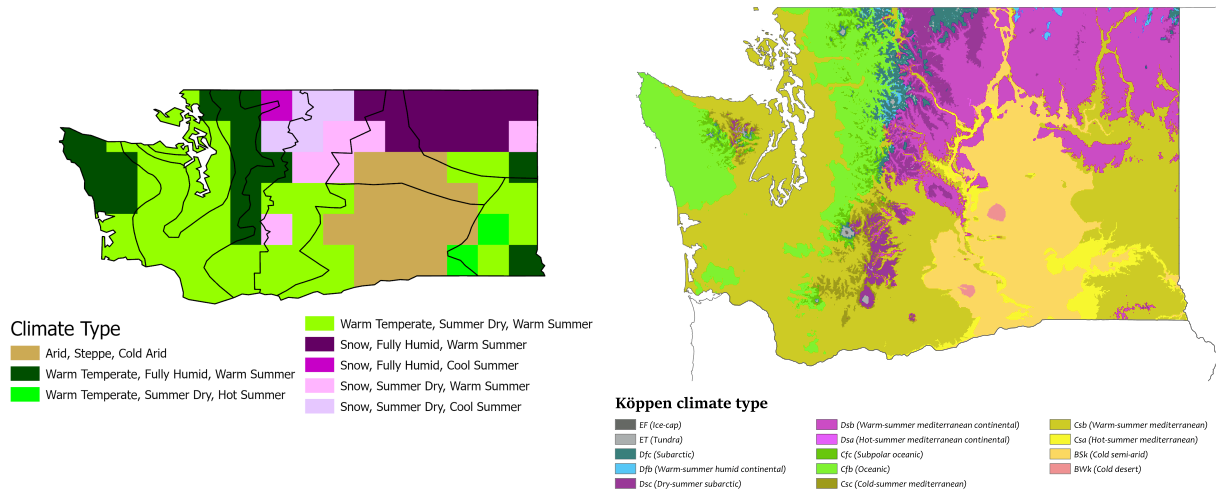


Figure E1: Köppen-Geiger climate classifications in Washington.

Figure E1 suggests that Washington may be mostly described by three Köppen-Geiger types of climates - Arid, Warm Temperate, and Snowy - with some further sub-division based on precipitation patterns. This number is much lower than the ten climate divisions outlined by NOAA. It is unclear which classification scheme is preferable to aggregate heat-related health outcomes, but it is possible that some of the uniformity across climate divisions as described in Figures 2.2 and 2.3 is due to climatological homogeneity across NOAA's climate divisions.

Appendix F. Raw Code

The sections below include all of the code used in this research. The code should be run sequentially to reproduce all results. Data paths and file names should be changed as needed to reflect local paths/names.

gridMET Data Acquisition

Using the [climateR package](#) to obtain temperature (maximum daily temperature, `tmax`) and relative humidity (minimum daily relative humidity, `rhmin`) data from the [gridMET](#) data set ([Abatzoglou 2013](#)).

This code:

- Obtains data for each grid point in the months and years of interest and creates a `.csv` file for each year and variable with the raw data (Part 1).
- Combines the individual `.csv` files into a single `.csv` file with all meteorological variables, the computed humidex value (using maximum temperature and minimum relative humidity), and the climate division for each grid point (Part 2).

Note that this code requires the `WA_geo_data.csv` file to assign points into each climate zone. This file was created with a nearest neighbors analysis; it is *much* faster to use this file than to use the nearest neighbor analysis to assign each point into a climate zone.

```
#### Setup ####

#Libraries

#install.packages("devtools")
#devtools::install_github("mikejohnson51/climateR")
#devtools::install_github("ropensci/USAboundariesData")
library(climateR)
library(raster)

#Kelvin to Celsius function
KtoC <- function(x) {
  x - 273.15
}

#Humidex function
getHumidex <- function(T, RH){
  v = 6.112*10^(7.5*T/(237.7 + T))*RH/100
  T + 5/9*(v-10)
}

#set directory to hold the generated data
DataPath_met <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State"
```

```

#read WA_geo_data.csv file to assign points into climate zones
DataPath_geo <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State
WA_geo <- read.csv(paste(DataPath_geo, "WA_geo_data.csv", sep = ""))

#define area of interest
#WA data isn't working on its own
#solution: take entire CONUS, then subset it to include only WA
CONUS = aoi_get(state = "conus")
WA = aoi_get(state = "Washington")
WA_xmin <- -124.7258
WA_xmax <- -116.916
WA_ymin <- 45.54432
WA_ymax <- 49.00249

#define time frame of interest
year_start <- 1980
year_end <- 2018
years <- c(year_start:year_end)

#####
# Part 1: Read-in Met Data
# uses climateR package to read-in gridMET data and save to .csv files
#####

#### tmax ####
for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 30)

  #get met data for year
  met.data = getGridMET(CONUS, param = c('tmax'),
                        startDate = sD, endDate = eD)

  #convert to data frame
  met.data.raster = raster::stack(met.data$tmax)
  met.data.df = as.data.frame(met.data.raster, xy = TRUE)

  #remove points not in Washington
  met.data.df <- met.data.df[met.data.df$x >= WA_xmin,]
  met.data.df <- met.data.df[met.data.df$x <= WA_xmax,]
  met.data.df <- met.data.df[met.data.df$y >= WA_ymin,]
  met.data.df <- met.data.df[met.data.df$y <= WA_ymax,]

```

```

#remove points with NA
met.data.df <- met.data.df[complete.cases(met.data.df), ]

#convert Kelvin to degrees Celsius
met.data.df <- cbind(met.data.df[1:2],
                    apply(met.data.df[3:ncol(met.data.df)],2, KtoC) )

#### assign each point to climate zone and county ####
print("Merging with geo data and saving")

met.data.df <- merge(WA_geo, met.data.df, by = c("x", "y"))

#save data
write.csv(met.data.df, paste(DataPath_met,
                             "tmax/tmax_", year, "_total.csv",
                             sep = ""), row.names=FALSE)

print(" ")
}

#### tmin ####
for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 30)

  #get met data for year
  met.data = getGridMET(CONUS, param = c('tmin'),
                       startDate = sD, endDate = eD)

  #convert to data frame
  met.data.raster = raster::stack(met.data$tmin)
  met.data.df = as.data.frame(met.data.raster, xy = TRUE)

  #remove points not in Washington
  met.data.df <- met.data.df[met.data.df$x >= WA_xmin,]
  met.data.df <- met.data.df[met.data.df$x <= WA_xmax,]
  met.data.df <- met.data.df[met.data.df$y >= WA_ymin,]
  met.data.df <- met.data.df[met.data.df$y <= WA_ymax,]

  #remove points with NA
  met.data.df <- met.data.df[complete.cases(met.data.df), ]

```

```

#convert Kelvin to degrees Celsius
met.data.df <- cbind(met.data.df[1:2],
                    apply(met.data.df[3:ncol(met.data.df)],2, KtoC) )

#### assign each point to climate zone and county ####
print("Merging with geo data and saving")

met.data.df <- merge(WA_geo, met.data.df, by = c("x", "y"))

#save data
write.csv(met.data.df, paste(DataPath_met,
                             "tmin/tmin_", year, "_total.csv",
                             sep = ""), row.names=FALSE)

print(" ")
}

#### rhmin ####
for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 1)

  #get met data for year
  met.data = getGridMET(CONUS, param = c('rhmin'),
                       startDate = sD, endDate = eD)

  #convert to data frame
  met.data.raster = raster::stack(met.data$rhmin)
  met.data.df = as.data.frame(met.data.raster, xy = TRUE)

  #remove points not in Washington
  met.data.df <- met.data.df[met.data.df$x >= WA_xmin,]
  met.data.df <- met.data.df[met.data.df$x <= WA_xmax,]
  met.data.df <- met.data.df[met.data.df$y >= WA_ymin,]
  met.data.df <- met.data.df[met.data.df$y <= WA_ymax,]

  #remove points with NA
  met.data.df <- met.data.df[complete.cases(met.data.df), ]

  #### assign each point to climate zone and county ####
  print("Merging with geo data and saving")
}

```

```

met.data.df <- merge(WA_geo, met.data.df, by = c("x", "y"))

#save data
write.csv(met.data.df, paste(DataPath_met,
                             "rhmin/rhmin_", year, "_total.csv",
                             sep = ""), row.names=FALSE)

print(" ")
}

#### rhmax ####
for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 30)

  #get met data for year
  met.data = getGridMET(CONUS, param = c('rhmax'),
                        startDate = sD, endDate = eD)

  #convert to data frame
  met.data.raster = raster::stack(met.data$rhmax)
  met.data.df = as.data.frame(met.data.raster, xy = TRUE)

  #remove points not in Washington
  met.data.df <- met.data.df[met.data.df$x >= WA_xmin,]
  met.data.df <- met.data.df[met.data.df$x <= WA_xmax,]
  met.data.df <- met.data.df[met.data.df$y >= WA_ymin,]
  met.data.df <- met.data.df[met.data.df$y <= WA_ymax,]

  #remove points with NA
  met.data.df <- met.data.df[complete.cases(met.data.df), ]

  #### assign each point to climate zone and county ####
  print("Merging with geo data and saving")

  met.data.df <- merge(WA_geo, met.data.df, by = c("x", "y"))

  #save data
  write.csv(met.data.df, paste(DataPath_met,
                               "rhmax/rhmax_", year, "_total.csv",
                               sep = ""), row.names=FALSE)
}

```

```

print(" ")
}

#####
# Part 2: Calculate Humidex
# Reads the data files created in part 1 to calculate humidex values.
# Humidex is calculated using daily maximum temperature
# and daily minimum relative humidity.
#####

#### calculate humidex ####
for(year in years){

  print(paste("Calculating humidex for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 30)

  #read-in data
  temp.min.dat.total <-
    read.csv(paste(DataPath_met, "tmin/tmin_", year, "_total.csv", sep = ""))
  temp.max.dat.total <-
    read.csv(paste(DataPath_met, "tmax/tmax_", year, "_total.csv", sep = ""))
  rh.max.dat.total <-
    read.csv(paste(DataPath_met, "rhmax/rhmax_", year, "_total.csv", sep = ""))
  rh.min.dat.total <-
    read.csv(paste(DataPath_met, "rhmin/rhmin_", year, "_total.csv", sep = ""))

  n_obs_temp.min <- nrow(temp.min.dat.total)
  n_obs_temp.max <- nrow(temp.max.dat.total)
  n_obs_rh.max <- nrow(rh.max.dat.total)
  n_obs_rh.min <- nrow(rh.min.dat.total)

  #convert from wide to long format
  temp.min.dat.total <-
    cbind(temp.min.dat.total[1:4],
          unlist(temp.min.dat.total[5:ncol(temp.min.dat.total)]))
  temp.max.dat.total <-
    cbind(temp.max.dat.total[1:4],
          unlist(temp.max.dat.total[5:ncol(temp.max.dat.total)]))
  rh.max.dat.total <-
    cbind(rh.max.dat.total[1:4],
          unlist(rh.max.dat.total[5:ncol(rh.max.dat.total)]))
  rh.min.dat.total <-

```

```

cbind(rh.min.dat.total[1:4],
      unlist(rh.min.dat.total[5:ncol(rh.min.dat.total)]))

#get date data for all observations
temp.min.dat.total$date <- NA
temp.max.dat.total$date <- NA
rh.max.dat.total$date <- NA
rh.min.dat.total$date <- NA
j_temp.min = 1
j_temp.max = 1
j_rh.max = 1
j_rh.min = 1

start <- as.Date(sD)
end <- as.Date(eD)
theDate <- start

while (theDate <= end){

  #put date into data frame
  temp.min.dat.total$date[j_temp.min:(j_temp.min + n_obs_temp.min - 1)] <-
    as.character(theDate)
  temp.max.dat.total$date[j_temp.max:(j_temp.max + n_obs_temp.max - 1)] <-
    as.character(theDate)
  rh.max.dat.total$date[j_rh.max:(j_rh.max + n_obs_rh.max - 1)] <-
    as.character(theDate)
  rh.min.dat.total$date[j_rh.min:(j_rh.min + n_obs_rh.min - 1)] <-
    as.character(theDate)

  #update counters
  theDate <- theDate + 1
  j_temp.min = j_temp.min + n_obs_temp.min
  j_temp.max = j_temp.max + n_obs_temp.max
  j_rh.max = j_rh.max + n_obs_rh.max
  j_rh.min = j_rh.min + n_obs_rh.min

}

#merge temp
temp.dat.total <- merge(temp.min.dat.total, temp.max.dat.total,
                       by = c("x", "y", "date"))
temp.dat.total <- temp.dat.total[,-c(7:8)] #remove second zonenumber variable
colnames(temp.dat.total) <- c("x", "y", "date", "zonenumber", "county", "tmin", "tmax")

#merge RH and get average RH from max and min
rh.total <- merge(rh.max.dat.total, rh.min.dat.total,
                 by = c("x", "y", "date"))

```

```

rh.total <- rh.total[,-c(7:8)] #remove second zonename variable
colnames(rh.total) <- c("x", "y", "date", "zonename", "county", "rhmax", "rhmin")
rh.total$rhavg <- (rh.total$rhmax + rh.total$rhmin)/2

#merge temp data with RH data by lat and long
dat.total <- merge(temp.dat.total, rh.total, by = c("x", "y", "date"))
dat.total <- dat.total[,-c(8:9)] #remove second zonename variable
colnames(dat.total) <- c("x", "y", "date", "zonename", "county",
                        "tmin", "tmax", "rhmax", "rhmin", "rhavg")

#calculate humidex (using maximum temperature and minimum relative humidity)
dat.total$humidex<- mapply(getHumidex, dat.total$tmax, dat.total$rhmin)

#save data for the year
write.csv(dat.total, paste(DataPath_met,
                           "met_data_", year, ".csv",
                           sep = ""), row.names=FALSE)
}

#####
# Alternative: Nearest neighbor calculation to assign points into climate divisions
#####

library(sf) #for spatial operations

#read shapefiles
DataPath_sh <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State/"
cz_shp <- st_read(paste(DataPath_sh,
                        "climateZones/Climate_Zones_WA.shp", sep = ""))
WA_shp <- st_read(paste(DataPath_sh,
                        "WAcountries/WA_County_Bndys.shp", sep = ""))

# create points
p_matrix = matrix(c(met.data.df$x, met.data.df$y), ncol = 2)

#see where points intersect with climate zones
p_matrix_cz <- apply(p_matrix, 1, function(row) {

  # transformation to planar is required,
  # since sf library assumes planar projection
  cz_pl <- st_transform(cz_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,
  #
                                  dimnames = list("", c("lat", "long"))))
  pnt_sf <- st_transform(st_sfc(st_point(row),crs = 4326), 2163)
  # st_intersects with sparse = FALSE returns a logical matrix
  # with rows corresponds to argument 1 (points) and

```

```
# columns to argument 2 (polygons)  
cz_pl[which(st_intersects(pnt_sf, cz_pl, sparse = FALSE)), ]$NAME  
})
```

CMIP5 GCMs Data Acquisition

Using the [climateR package](#) to obtain projected temperature (maximum daily temperature, `tmax`) and relative humidity (minimum daily relative humidity, `rhmin`) from the [University of California Merced Climatology Lab](#) for the years 2030, 2050, and 2080 under RCP4.5 and RCP8.5. Specifically, MACA downscaling ([Abatzoglou and Brown 2012](#)) of 20 GCMs of the Coupled Model Inter-Comparison Project 5 (CMIP5) were used to produce the climate projections. More information, including a further details about MACA and a list of GCMs included in CMIP5, can be found [here](#). This code:

- Obtains data for each grid point in the months, years, and RCPs of interest and creates a `.csv` file for each year, variable, climate model, and RCP with the raw data (Part 1).
- Combines the individual `.csv` files for a given year, variable, climate model, and RCP into a single `.csv` file that contains the average values of either `tmax` or `rhmin` for each grid point across all models in that year and RCP (Part 2).
- Combines the files containing `tmax` and `rhmin` data for a particular year and RCP to create a final `.csv` file with the calculated humidex value at all grid points as well as the county and climate division of each grid point (Part 3).

Note that in order to calibrate ([Ho et al. 2012](#)) the projected (CMIP5) data ([Abatzoglou and Brown 2012](#)) to the historical (gridMET) data ([Abatzoglou 2013](#)), this script should be modified so that it also obtains data for the historical period (1980-2018).

```
#### Setup ####

#Libraries

#install.packages("devtools")
#devtools::install_github("mikejohnson51/climateR")
#devtools::install_github("ropensci/USAboundariesData")
library(climateR)
library(raster)

#Kelvin to Celsius function
KtoC <- function(x) {
  x - 273.15
}
```

```

#Humidex function
getHumidex <- function(T, RH){
  v = 6.112*10^(7.5*T/(237.7 + T))*RH/100
  T + 5/9*(v-10)
}

#set directory to hold the generated data
DataPath_met <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State,

#read WA_geo_data.csv file to assign points into climate zones
DataPath_geo <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State,
WA_geo <- read.csv(paste(DataPath_geo, "WA_geo_data.csv", sep = ""))

#read shapefiles
DataPath_sh <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State/
cz_shp <- st_read(paste(DataPath_sh, "climateZones/Climate_Zones_WA.shp", sep = ""))
WA_shp <- st_read(paste(DataPath_sh, "WAcountries/WA_County_Bndys.shp", sep = ""))

#define area of interest
#WA data isn't working on its own
#solution: take entire CONUS, then subset it to include only WA
CONUS = aoi_get(state = "conus")
WA = aoi_get(state = "Washington")
WA_xmin <- -124.7258
WA_xmax <- -116.916
WA_ymin <- 45.54432
WA_ymax <- 49.00249

#set years, variables, and RCPs of interest
#also, get list of models
years <- c(2030, 2050, 2080)
vars <- c("tmax", "rhmin")
RCPs = c("rcp45", "rcp85")
models <- model_meta$maca[,1]

#####
# Part 1: Read-in Projected Data
# uses climateR package to read-in MACA CMIP5 data and save to .csv files
#####

#### get data ####

#loop through years
for(year in years){

  #loop through RCPs

```

```

for(RCP in RCPs){

  #loop through meteorological variables
  for(var in vars){

    #loop through climate models
    for(mod in models){

      skip_to_next <- FALSE #for error checking

      print(paste("Reading", var, "data for year", year,
                  "model", mod,
                  "and scenario", RCP,
                  sep = " "))

      #define start and end dates
      sD <- paste(year, "-05-01", sep = "") #start date (May 1)
      eD <- paste(year, "-09-30", sep = "") #end date (Sept 1)

      #get projections for year
      tryCatch({
        proj.data = getMACA(WA, param = c(var),
                            model = mod,
                            scenario = RCP,
                            startDate = sD, endDate = eD)
      }, error=function(e){
        cat("ERROR: could not get data for ", mod, "\n")
        skip_to_next <<- TRUE
      })
      if(skip_to_next){next}

      print("Data is read in")
      print(".....")

      #### Cleaning up and Saving ####
      print("Cleaning and saving")

      #convert to data frame
      proj.data.raster = raster::stack(proj.data)
      proj.data.df = as.data.frame(proj.data.raster, xy = TRUE)

      #delete data to free up memory
      rm(proj.data)
      rm(proj.data.raster)

      #remove points with NA
      proj.data.df <- proj.data.df[complete.cases(proj.data.df), ]
    }
  }
}

```

```

    #convert Kelvin to degrees Celsius
    if(var == "tmax"){
      proj.data.df <-
        cbind(proj.data.df[1:2],
              apply(proj.data.df[3:ncol(proj.data.df)],2, KtoC) )
    }

    #save data
    write.csv(proj.data.df, paste(DataPath_proj,
                                  year, "/",
                                  RCP, "/",
                                  var, "/", var,
                                  "_", mod,
                                  "_", year, ".csv", sep = ""),
              row.names=FALSE)

    #delete data to free up memory
    rm(proj.data.df)

    print(".....")
  }
}
}
}

#####
# Part 2: Aggregate to Get Mean for each Grid Cell Location
# Creates a single .csv file for each year, meteorological variable, and RCP
#####

#### tmax ####

#create empty data frame
dat.tmax <- NULL

#loop through years
for(year in years){

  #loop through RCPs
  for(RCP in RCPs){

    #loop through climate models

```

```

for(mod in models){

  print(paste("Reading tmax data for year", year,
             "model", mod,
             "and scenario", RCP,
             sep = " "))

  #combine the data for this year, model, and RCP
  #to the rest of the data for this year and RCP
  dat.tmax <- rbind(dat.tmax, read.csv(paste(DataPath_proj,
                                           year, "/",
                                           RCP,
                                           "/tmax/tmax_",
                                           mod,
                                           "_", year,
                                           ".csv", sep = "")))

  print(".....")
}

#get average at each grid point
print(paste("Averaging tmax data for year", year,
           "and scenario", RCP,
           "across all models",
           sep = " "))
dat.tmax.avg <- aggregate(dat.tmax,
                        by = list(dat.tmax$x, dat.tmax$y),
                        mean)

#remove second x/y coordinate
dat.tmax.avg <- dat.tmax.avg[,-c(1:2)]

#save the single .csv file for tmax and year
print("Saving...")

write.csv(dat.tmax.avg, paste(DataPath_proj,
                             year, "/",
                             RCP, "/",
                             "tmax/tmax",
                             "_", year,
                             "_avg.csv", sep = ""),
         row.names=FALSE)

#delete data to start over
dat.tmax <- NULL
rm(dat.tmax.avg)
print(".....")
}

```

```

}

#### rhmin ####
#different because not every model has rhmin data
models <- models[-c(5, 20)] #remove models without rhmin data

#create empty data frame
dat.rhmin <- NULL

#loop through years
for(year in years){

  #loop through RCPs
  for(RCP in RCPs){

    #loop through climate models
    for(mod in models){

      print(paste("Reading rhmin data for year", year,
                  "model", mod,
                  "and scenario", RCP,
                  sep = " "))

      #combine the data for this year, model, and RCP
      #to the rest of the data for this year and RCP
      dat.rhmin <- rbind(dat.rhmin, read.csv(paste(DataPath_proj,
                                                  year, "/",
                                                  RCP,
                                                  "/rhmin/rhmin_",
                                                  mod,
                                                  "_", year,
                                                  ".csv", sep = "")))

      print(".....")
    }

    #get average at each grid point
    print(paste("Averaging rhmin data for year", year,
                "and scenario", RCP,
                "across all models",
                sep = " "))

    dat.rhmin.avg <- aggregate(dat.rhmin,
                              by = list(dat.rhmin$x, dat.rhmin$y),
                              mean)

    #remove second x/y coordinate data
    dat.rhmin.avg <- dat.rhmin.avg[,-c(1:2)]
  }
}

```

```

#save the single .csv file for rhmin and year
print("Saving...")
write.csv(dat.rhmin.avg, paste(DataPath_proj,
                              year, "/",
                              RCP, "/",
                              "rhmin/rhmin",
                              "_", year,
                              "_avg.csv", sep = ""),
          row.names=FALSE)

#delete data to start over
dat.rhmin <- NULL
rm(dat.rhmin.avg)
print(".....")

}
}

#####
# Part 3: Obtain Projected Humindex
# Creates a single .csv file for each year and RCP
# The final file contains tmax, rhmin, and humindex
# as well as county and climate division
#####

#create empty data frames
dat.rhmin <- NULL
dat.tmax <- NULL
dat.proj <- NULL

#loop over years
for(year in years){

  #loop over RCPs
  for(RCP in RCPs){

    #### read-in tmax data ####
    print(paste("Reading in tmax data for year", year,
               "and scenario", RCP,
               sep = " "))

    dat.tmax<- read.csv(paste(DataPath_proj,
                              year, "/",
                              RCP, "/",
                              "tmax/tmax",
                              "_", year,
                              "_avg.csv", sep = ""))
  }
}

```

```

#### assign tmax points into climate divisions ####
print("Assigning points into each CZ")

# create points
p_matrix = matrix(c(dat.tmax$x, dat.tmax$y), ncol = 2)

#see where points intersect with climate zones
idx <- 1 # counter for progress
p_matrix_cz <- apply(p_matrix, 1, function(row) {

  #update progress
  idx <<- idx + 1
  if(idx %% 1000 == 0){
    print(paste("Reached observation ",
                idx, " of ", nrow(p_matrix),
                sep = ""))
  }

  # transformation to planar is required,
  # since sf library assumes planar projection
  cz_pl <- st_transform(cz_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,
                                  dimnames = list("", c("lat", "long"))))
  pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
  # st_intersects with sparse = FALSE returns a logical matrix
  # with rows corresponds to argument 1 (points) and
  # columns to argument 2 (polygons)

  cz_pl[which(st_intersects(pnt_sf, cz_pl, sparse = FALSE)), ]$NAME
})

#### assign tmax points into counties ####
print("Assigning points into each county")

#see where points intersect with counties
idx <- 1 # counter for progress
p_matrix_WA <- apply(p_matrix, 1, function(row) {

  #update progress
  idx <<- idx + 1
  if(idx %% 1000 == 0){
    print(paste("Reached observation ",
                idx, " of ", nrow(p_matrix),
                sep = ""))
  }

  # transformation to palnar is required,

```

```

# since sf library assumes planar projection
WA_pl <- st_transform(WA_shp, 2163)
#coords <- as.data.frame(matrix(row, nrow = 1,
#                                dimnames = list("", c("lat", "long"))))
pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
# st_intersects with sparse = FALSE returns a logical matrix
# with rows corresponds to argument 1 (points) and
# columns to argument 2 (polygons)

WA_pl[which(st_intersects(pnt_sf, WA_pl, sparse = FALSE)), ]$JURNM
})

#### Removing Points That could not be assigned into a climate division ####
# not all points could be classified
# because the "WA" points are actually in a big rectangle around the state,
# i.e. not within the state boundaries

print("Cleaning up...")

#Part 1: climate divisions
n_obs <- 1:length(p_matrix_cz) #total number of observations

# these are the indices of observations with location data
l_obs <- NA
for(i in 1:10){
  l_obs <- c(l_obs, which(p_matrix_cz == i)) #matching to 1 of the 10 climate zones
}

#remove first NA
l_obs <- l_obs[-1]
# order from least to greatest
l_obs <- sort(l_obs)
#the observations to remove are any observation without location data
r_obs_cz <- n_obs[-l_obs]

#Part 2: Counties
r_obs_WA <- which(is.na(p_matrix_WA == 1))

#identify observations without location data
#points to remove are those which could not be classified
# into at least one of: CZ, County
r_obs <- unique(c(r_obs_cz, r_obs_WA))

#remove observations without location data
dat.tmax <- dat.tmax[-r_obs, ]
p_matrix_cz <- p_matrix_cz[-r_obs]
p_matrix_WA <- p_matrix_WA[-r_obs]

```

```

#combine met data with climate division data
dat.tmax <- cbind(unlist(p_matrix_cz), unlist(p_matrix_WA), dat.tmax)
colnames(dat.tmax)[1:2] <- c("zonename", "county")

#delete variables to free memory
rm(p_matrix)
rm(p_matrix_cz)
rm(p_matrix_WA)

print(".....")

#####

#### read-in rhmin data ####
print(paste("Reading in rhmin data for year", year,
           "and scenario", RCP,
           sep = " "))

dat.rhmin <- read.csv(paste(DataPath_proj,
                           year, "/",
                           RCP, "/",
                           "rhmin/rhmin",
                           "_", year,
                           "_avg.csv", sep = ""))

#### assign rhmin points into climate divisions ####
print("Assigning points into each CZ")

# create points
p_matrix = matrix(c(dat.rhmin$x, dat.rhmin$y), ncol = 2)

#see where points intersect with climate zones
idx <- 1 # counter for progress
p_matrix_cz <- apply(p_matrix, 1, function(row) {

  #update progress
  idx <<- idx + 1
  if(idx %% 1000 == 0){
    print(paste("Reached observation ",
               idx, " of ", nrow(p_matrix),
               sep = ""))
  }

  # transformation to planar is required,
  # since sf library assumes planar projection
  cz_pl <- st_transform(cz_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,

```

```

#                                     dimnames = list("", c("lat", "long"))))
pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
# st_intersects with sparse = FALSE returns a logical matrix
# with rows corresponds to argument 1 (points) and
# columns to argument 2 (polygons)

cz_pl[which(st_intersects(pnt_sf, cz_pl, sparse = FALSE)), ]$NAME
})

#### assign rhmin points into counties ####
print("Assigning points into each county")

#see where points intersect with counties
idx <- 1 # counter for progress
p_matrix_WA <- apply(p_matrix, 1, function(row) {

  #update progress
  idx <<- idx + 1
  if(idx %% 1000 == 0){
    print(paste("Reached observation ",
                idx, " of ", nrow(p_matrix),
                sep = ""))
  }

  # transformation to planar is required,
  # since sf library assumes planar projection
  WA_pl <- st_transform(WA_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,
  #                               dimnames = list("", c("lat", "long"))))
  pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
  # st_intersects with sparse = FALSE returns a logical matrix
  # with rows corresponds to argument 1 (points) and
  # columns to argument 2 (polygons)

  WA_pl[which(st_intersects(pnt_sf, WA_pl, sparse = FALSE)), ]$JURNM
})

#### Removing Points That could not be assigned into a climate division ####
# not all points could be classified
# because the "WA" points are actually in a big rectangle around the state,
# i.e. not within the state boundaries

print("Cleaning up...")

#Part 1: climate divisions
n_obs <- 1:length(p_matrix_cz) #total number of observations

```

```

# these are the indices of observations with location data
l_obs <- NA
for(i in 1:10){
  l_obs <- c(l_obs, which(p_matrix_cz == i)) #matching to 1 of the 10 climate zones
}

#remove first NA
l_obs <- l_obs[-1]
# order from least to greatest
l_obs <- sort(l_obs)
#the observations to remove are any observation without location data
r_obs_cz <- n_obs[-l_obs]

#Part 2: Counties
r_obs_WA <- which(is.na(p_matrix_WA == 1))

#identify observations without location data
#points to remove are those which could not be classified
# into at least one of: CZ, County
r_obs <- unique(c(r_obs_cz, r_obs_WA))

# remove observations without location data
dat.rhmin <- dat.rhmin[-r_obs, ]
p_matrix_cz <- p_matrix_cz[-r_obs]
p_matrix_WA <- p_matrix_WA[-r_obs]

#combine met data with cz data
dat.rhmin <- cbind(unlist(p_matrix_cz), unlist(p_matrix_WA), dat.rhmin)
colnames(dat.rhmin)[1:2] <- c("zonename", "county")

#delete variables to free memory
rm(p_matrix)
rm(p_matrix_cz)
rm(p_matrix_WA)

print(".....")

#####

#### convert from wide to long format ####
print("Converting from wide to long format...")

#define start and end dates
sD <- paste(year, "-05-01", sep = "") #start date (May 1)
eD <- paste(year, "-09-30", sep = "") #end date (Sept 1)

n_obs_temp <- nrow(dat.tmax)

```

```

n_obs_rh <- nrow(dat.rhmin)

#convert from wide to long
dat.tmax <- cbind(dat.tmax[1:4],
                 unlist(dat.tmax[5:ncol(dat.tmax)]))
dat.rhmin <- cbind(dat.rhmin[1:4],
                 unlist(dat.rhmin[5:ncol(dat.rhmin)]))

#get date data for all observations
dat.tmax$date <- NA
dat.rhmin$date <- NA
j_temp = 1
j_rh = 1

start <- as.Date(sD)
end <- as.Date(eD)
theDate <- start

#loop through all dates
while (theDate <= end){

  #put date into data frame
  dat.tmax$date[j_temp:(j_temp + n_obs_temp - 1)] <-
    as.character(theDate)
  dat.rhmin$date[j_rh:(j_rh + n_obs_rh - 1)] <-
    as.character(theDate)

  #update counters
  theDate <- theDate + 1
  j_temp = j_temp + n_obs_temp
  j_rh = j_rh + n_obs_rh

}

#### merge data by lat and long ####
print("Merging tmax and rhmin data...")
dat.total <- merge(dat.tmax, dat.rhmin, by = c("x", "y", "date"))
dat.total <- dat.total[, -c(7,8)] #remove second zonename and county variable
colnames(dat.total) <- c("x", "y", "date",
                       "zonename", "county", "tmax", "rhmin")

#### calculate humidex ####
print("Calculating humidex...")

#calculate humidex
dat.total$humidex<- mapply(getHumidex,
                          dat.total$tmax,

```

```
dat.total$rhmin)

#### Save Data ####
print("Saving...")

write.csv(dat.total, paste(DataPath_proj,
                           year, "/",
                           RCP, "/",
                           "proj_data_",
                           year, "_avg_total.csv",
                           sep = ""), row.names=FALSE)

#delete variable to free memory
rm(dat.total)

print(".....")
}
}
```

Calibration Parameters

Typically, when using climate projections to assess climate impacts (e.g., to model heat-related mortality), it is necessary to adjust or *calibrate* or *bias-correct* the projected data to correct for any bias in the climate models rather than simply using the raw projections from the climate model output. Specifically, this process should correct for any biases in the mean and variation in the modeled (CMIP5) data (Abatzoglou and Brown 2012) when compared to observed (gridMET) data (Abatzoglou 2013). Therefore, the mean and standard deviation of historical (1980-2018) data are needed for both the observed and modeled data in addition to the mean and standard deviation of the modeled data in all years (2030, 2050, and 2080) and greenhouse gas emission scenarios (RCP4.5 and RCP8.5) used for projections. More information on the two specific calibration processes used (bias correction and change factor) can be found in Ho et al. (Ho et al. 2012). Note that one downside of this calibration process is that the 99th percentile tends to increase more than the mean, relative to the uncorrected projections. (See Tables F1 and F2 for the uncalibrated and calibrated humidex values for each year and RCP.) The code below obtains the parameters necessary for calibration.

```
#### Data paths ####
#historical values of gridMET data
DataPath_met <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State

#historical values of CIMIP5 models
DataPath_hist <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State

#projected values of CIMIP5 models
DataPath_proj <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State

#### Data Frame and List to Hold Parameters ####
param.list <- vector(mode = "list", length = 10)

param.df <- data.frame(Type = character(8),
                      Year = character(8),
                      RCP = character(8),
                      Mean = double(8), SD = double(8))
param.df$Type <- c("gridMET Observations",
                  "Historical Model Values",
                  "Future Model Values",
                  "Future Model Values",
                  "Future Model Values")
```

```

        "Future Model Values",
        "Future Model Values",
        "Future Model Values")
param.df$Year <- c("1980-2018", "1980-2018",
                  "2030", "2030",
                  "2050", "2050",
                  "2080", "2080")
param.df$RCP <- c("NA", "NA",
                  "rcp45", "rcp85",
                  "rcp45", "rcp85",
                  "rcp45", "rcp85")

for(i in 1:10){
  param.list[[i]] <- param.df
}

#### Historical Observed (gridMET) Data ####

#timeframe
year_start <- 1980
year_end <- 2018
years <- c(year_start:year_end)

humidex.obs <- NA
cz.obs <- NA #climate zones are saved numerically, but 1-10 correspond to alphabetical order

#read-in data
for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #read-in data for give year and combine w/ previous year
  met.data.year <- read.csv(paste(DataPath_met,
                                "met_data_", year, ".csv", sep = ""))
  humidex.obs <- c(humidex.obs, met.data.year$humidex)
  cz.obs <- c(cz.obs, met.data.year$zonename)

  rm(met.data.year)

}

humidex.obs <- humidex.obs[-1] #remove first row of NAs
cz.obs <- cz.obs[-1] #remove first row of NAs

obs <- data.frame(cz.obs, humidex.obs)

#obs.mean <- mean(humidex.obs)

```

```

#obs.sd <- sd(humidex.obs)
#param.df$Mean[1] <- obs.mean
#param.df$SD[1] <- obs.sd

obs.means <- aggregate(obs$humidex.obs, list(obs$cz.obs), mean)
obs.sds <- aggregate(obs$humidex.obs, list(obs$cz.obs), sd)

for(i in 1:10){
  param.list[[i]]$Mean[1] <- obs.means[i,2]
  param.list[[i]]$SD[1] <- obs.sds[i,2]
}

#remove variables to free-up memory
rm(humidex.obs)
rm(cz.obs)
rm(obs)

#### Historical Modeled Data ####
#hist.df <- read.csv(paste(DataPath_hist, "humidex_avg.csv", sep = ""))

#hist.mean <- mean(as.matrix(hist.df[,-c(1:4)]))
#hist.sd <- sd(as.matrix(hist.df[,-c(1:4)]))
#rm(hist.df)
#param.df$Mean[2] <- hist.mean
#param.df$SD[2] <- hist.sd

#values below are from the humidex_avg.csv file above
#values are for each climate zone
hist.means <- c(17.58174, 26.73511, 22.14051, 18.79228, 22.00259,
               22.50464, 24.66285, 24.81683, 23.52716, 21.41202)

hist.sds <- c(5.237409, 4.733775, 5.110079, 5.910122, 3.177318,
              5.152231, 4.874128, 5.117695, 3.849598, 4.216839)

for(i in 1:10){
  param.list[[i]]$Mean[2] <- hist.means[i]
  param.list[[i]]$SD[2] <- hist.sds[i]
}

#total
#param.df$Mean[2] <- 22.69708
#param.df$SD[2] <- 5.716126

#### Projections ####
years <- c(2030, 2050, 2080)

```

```

RCPs = c("rcp45", "rcp85")
counter = 3 #corresponds to row of parameter matrix

for(year in years){

  for(RCP in RCPs){

    #create data path for the year and RCP
    data.path <- paste(DataPath_proj,
                      year, "/",
                      RCP, "/", sep = "")

    #define empty data frame
    proj.data.df <- matrix(NA, nrow = 1, ncol = 8)
    colnames(proj.data.df) <- c("x", "y", "date",
                               "zonename", "county",
                               "tmax", "rhmin", "humidex")

    #read-in data
    print(paste("Reading projection data for year", year,
               "and scenario", RCP,
               sep = " "))
    proj.data.df <- read.csv(paste(data.path,
                                  "proj_data_",
                                  year, "_avg_total.csv", sep = ""))
    proj.data.df <- proj.data.df[-1,] #remove first row of NAs

    #analyze data
    print("Analyzing data...")

    mod.means <- aggregate(proj.data.df$humidex,
                          list(proj.data.df$zonename), mean)
    mod.sds <- aggregate(proj.data.df$humidex,
                        list(proj.data.df$zonename), sd)

    for(i in 1:10){
      param.list[[i]]$Mean[counter] <- mod.means[i,2]
      param.list[[i]]$SD[counter] <- mod.sds[i,2]
    }

    #param.df$Mean[counter] <- mean(proj.data.df$humidex)
    #param.df$SD[counter] <- sd(proj.data.df$humidex)

    counter <- counter + 1

    rm(proj.data.df)
  }
}

```

```

print(".....")

}
}

#### Save Final Matrix ####
DataPath <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State/Data
save(param.list, file = paste(DataPath, "BiasCorrectionParameters.Rda", sep = ""))

```

Table F1: Historical and projected mean humidex values for each climate division.

Climate Division	1980-2018	2030						2050						2080					
	Historical	RCP4.5			RCP8.5			RCP4.5			RCP8.5			RCP4.5			RCP8.5		
		U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF
Cascade Mountains West	22.75	25.17	24.88	24.66	25.71	25.41	25.02	26.21	25.91	25.50	27.29	26.99	26.53	27.46	27.16	26.66	30.36	30.06	29.51
Central Basin	27.25	29.23	28.67	28.95	29.66	29.10	29.39	30.15	29.59	29.86	31.10	30.54	30.81	31.25	30.69	30.95	34.13	33.57	33.84
E Olympic Cascade Foothills	24.98	26.98	26.67	26.79	27.37	27.06	27.10	27.84	27.53	27.62	28.92	28.62	28.65	29.10	28.80	28.85	31.93	31.64	31.71
E Slope Cascades	23.48	26.17	25.44	25.18	26.56	25.83	25.54	27.03	26.30	25.98	28.03	27.30	26.98	28.13	27.40	27.08	31.01	30.28	29.91
NE Olympic San Juan	21.78	23.74	23.35	23.57	23.97	23.58	23.80	24.49	24.20	24.41	25.57	25.18	25.40	25.68	25.28	25.46	28.71	28.32	28.51
Northeastern	24.72	26.70	26.26	26.46	27.17	26.74	26.82	27.56	27.12	27.24	28.48	28.04	28.14	28.70	28.27	28.34	31.76	31.33	31.37
Okanogan Big Bend	25.50	27.45	27.00	27.22	27.72	27.53	28.26	27.81	28.05	29.16	28.71	28.95	29.31	28.86	29.10	32.27	31.82	32.05	
Palouse Blue Mountains	25.20	27.68	27.22	26.87	28.28	27.82	27.47	28.57	28.17	27.77	29.56	29.11	28.78	29.78	29.32	28.95	32.80	32.35	31.99
Puget Sound Lowlands	23.52	25.54	25.19	25.31	25.87	25.53	25.61	26.46	26.12	26.21	27.49	27.15	27.24	27.64	27.30	27.40	30.71	30.37	30.47
West Olympic Coast	21.31	23.26	22.82	23.04	23.43	22.90	23.10	23.86	23.42	23.71	24.83	24.39	24.61	24.99	24.55	24.98	27.56	27.12	27.63

U = Uncalibrated; BC = Bias Correction; CF = Change Factor

Table F2: Historical and projected 99th percentile humidex values for each climate division.

Climate Division	1980-2018	2030						2050						2080					
	Historical	RCP4.5			RCP8.5			RCP4.5			RCP8.5			RCP4.5			RCP8.5		
		U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF	U	BC	CF
Cascade Mountains West	36.84	33.41	33.10	38.79	33.55	33.24	39.15	34.79	34.48	39.63	35.45	35.15	40.66	35.97	35.68	40.79	39.09	38.78	43.65
Central Basin	40.55	37.45	36.89	42.30	37.45	36.89	42.74	39.21	38.65	43.21	39.29	38.73	44.17	40.02	39.46	44.30	43.00	42.43	47.20
E Olympic Cascade Foothills	39.61	34.60	34.30	41.34	35.10	34.80	41.65	35.82	35.52	42.17	36.51	36.21	43.19	37.34	37.04	43.40	39.98	39.68	46.26
E Slope Cascades	37.19	34.57	33.83	39.21	34.63	33.90	39.57	36.31	35.58	40.01	36.56	35.83	41.01	37.39	36.66	41.11	40.21	39.48	43.94
NE Olympic San Juan	32.93	29.64	29.25	34.66	30.18	29.79	34.90	31.03	30.64	35.50	31.15	31.76	36.49	32.37	21.97	36.66	35.74	35.35	39.60
Northeastern	38.01	34.43	33.99	39.71	34.62	34.19	40.07	36.56	36.12	40.50	35.92	35.49	41.39	37.63	37.19	41.59	40.36	39.92	44.62
Okanogan Big Bend	39.12	35.79	35.35	40.86	35.69	35.24	41.17	37.31	36.86	41.69	37.62	37.17	42.59	38.35	37.90	42.74	41.28	40.83	45.69
Palouse Blue Mountains	38.84	37.21	36.76	40.23	37.40	36.95	40.83	39.03	38.58	41.13	38.86	38.41	42.13	39.91	39.46	42.30	42.52	42.07	45.34
Puget Sound Lowlands	37.01	32.31	31.96	38.84	32.51	32.17	39.14	33.59	33.25	29.74	34.41	34.07	40.77	35.59	35.25	40.93	37.85	37.51	44.00
West Olympic Coast	33.81	30.95	30.51	35.36	31.37	30.93	35.42	31.59	31.15	36.03	33.01	32.57	36.93	33.63	33.19	37.30	36.26	35.82	39.95

U = Uncalibrated; BC = Bias Correction; CF = Change Factor

Mortality Data Pre-Processing

Before any analysis could be conducted, the individual-level mortality data needed to be pre-processed.

Considerations include:

- Combining data from different years into a single file
- Assigning deaths into climate divisions
- Assigning humidex from the gridMET data to each death on the day of death as well as on the control days
- Obtain information for covariates (cause of death, age group, race category)

The first block of R code below pre-processes mortality data from 2007-2018. The second block of R code combines this cleaned data with pre-existing mortality data for 1980-2006 and converts the data into the case-crossover format by assigning individual-level humidex exposure on all of the appropriate days. Note that the data from 1980-2006 included non-traumatic deaths only (ICD-9: 0-799; ICD-10: A00-R99), while the data from 2007-2018 included all causes of death. Therefore, the traumatic deaths (ICD-10: S01-Z99) had to be removed from the 2007-2018 data.

The final result is a single `csv` file with 563,365 deaths and 2,493,073 total rows called `case_crossover_data.csv`. (There are more rows than deaths since each death has multiple days - the day of death and one or more control days - associated with it.) The variables included in this file are, listed from the first to last column, are:

- `certno` - The ID associated with each death
- `index.month` - Month of referent days. In the final `.csv` file, this is the same as the month of death (`month`) variable, since the referent days are selected to be in the same month that the death occurred.
- `day` - The day of death (`date`) or a label for the referent day
- `Date.CC` - The date of the observation, in YYYY-MM-DD format
- `month` - Month of death

- `case` - A binary vector that is 1 on the day of death (`day == date`) and 0 otherwise
- `Date.Death` - The date of death associated with each observation
- `Year` - the year of death
- `lat` and `long` - The latitude and longitude, respectively, associated with each death
- `count` - A vector of 1's to indicate that each observation corresponds to an individual
- `resp`, `circ`, `cardio`, `ischem`, `cerebro`, `diabetes`, `nephritis`, `renal`, `mental`, `accident`, `suicide`, and `homicide` - Vectors where a value of 1 indicates that the individual died of the corresponding cause of death
- `sex` - The sex of the individual
- `age` - The age of the individual
- `race` and `hisp` - The race and Hispanic ethnicity of the individual
- `smoking` - The smoking status of the individual
- `tmin`, `tmax`, `rhmax`, `rhmin`, `rhavg`, and `humidex` - Individual-level meteorological variables on the date of the observation
- `age.grp` - The age group of the individual, based on DOH categories: 0-4 (1), 5-14 (2), 15-44 (3), 45-64 (4), 65-84 (5), and 85+ (6)
- `race.cat` and `race.cat.num` - The racial category of the individual, based on DOH categories: White (1), Black (2), Native American (3), Hispanic (4), NHOPI (5), Asian (6)
- `Traumatic` - A vector where a value of 1 indicates that a death is traumatic. Since all of the traumatic deaths have been removed, all values of this variable are NA.

More information on some of these variables is provided by the [Washington State Department of Health \(DOH\)](#).

```
library(openxlsx)
```

```
#####
```

```

# Create File for 2007-2018
# Data file already existed for 1980-2006
#####

#### Read-in raw Mortality Data ####

DataPath_raw = "//pacific/bilab/WildfireSmoke_Data/Mortality\ data/Annie\ Death\ Data/death.fi
DataPath_raw_2016 = "//pacific/bilab/WildfireSmoke_Data/Mortality\ data/Death\ Data\ WA/Origin

dat_raw_2007 <- read.csv(paste(DataPath_raw,
                              "death.2007.csv", sep = "" ))
dat_raw_2008 <- read.csv(paste(DataPath_raw,
                              "death.2008.csv", sep = "" ))
dat_raw_2009 <- read.csv(paste(DataPath_raw,
                              "death.2009.csv", sep = "" ))
dat_raw_2010 <- read.csv(paste(DataPath_raw,
                              "death.2010.csv", sep = "" ))
dat_raw_2011 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2011.csv", sep = "" ))
dat_raw_2012 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2012.csv", sep = "" ))
dat_raw_2013 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2013.csv", sep = "" ))
dat_raw_2014 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2014.csv", sep = "" ))
dat_raw_2015 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2015.csv", sep = "" ))
dat_raw_2016 <- read.xlsx(paste(DataPath_raw_2016,
                              "DeathStatF2016.xlsx", sep = "" ))
dat_raw_2017 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2017.csv", sep = "" ))
dat_raw_2018 <- read.csv(paste(DataPath_raw,
                              "DeathStatF2018.csv", sep = "" ))

#### Remove variables and dates that aren't of interest ####
# Different years have different variables that need to be removed

#function for 2007-2010
get_data1 <- function(df){

  #date of death
  df$Date <- as.Date(paste(substr(df$dth_date, 1, 4),
                          substr(df$dth_date, 5, 6),
                          substr(df$dth_date, 7, 8),
                          sep = "-"))

```

```

#variables of interest
vars <- c("certno", "sex", "Date", "race",
          "age", "hispanic", "smoking", "underly", "mltcse2")
df.new <- df[,vars]

#month and Year
df.new$Date = as.Date(as.character(df.new$Date),format="%Y-%m-%d")
df.new$Year = format(df.new$Date,'%Y')
df.new$Month = months(df.new$Date)
df.new <- subset(df.new, Month == "May" |
                 Month == "June" |
                 Month == "July" |
                 Month == "August" |
                 Month == "September") #May - Sept

#re-order by date
ord = order(df.new$Date)
df.new = df.new[ord,]

return(df.new)
}

#function for 2011-2015
get_data2 <- function(df){

  #date of death
df$Date <- paste(df$dth_yr, df$dth_mo, df$dth_da, sep = "-")

  #variables of interest
vars <- c("certno", "sex", "Date", "race",
          "age", "hispanic", "smoking", "underly", "mltcse2")
df.new <- df[,vars]

  #month and Year
df.new$Date = as.Date(as.character(df.new$Date),format="%Y-%m-%d")
df.new$Year = format(df.new$Date,'%Y')
df.new$Month = months(df.new$Date)
df.new <- subset(df.new, Month == "May" |
                 Month == "June" |
                 Month == "July" |
                 Month == "August" |
                 Month == "September") #May - Sept

  #re-order by date
ord = order(df.new$Date)
df.new = df.new[ord,]

```

```

return(df.new)

}

#function for 2016
get_data3 <- function(df){

  #date of death
  df$Date <- paste(df$`Date.of.Death.-.Year`,
                  df$`Date.of.Death.-.Month`,
                  df$`Date.of.Death.-.Day`,
                  sep = "-")

  df$hispanic = 0
  df$hispanic[df$Hispanic.No == "N"] = 1

  #variables of interest
  vars <- c("State.File.Number", "Sex", "Date", "Race.Calculation",
           "Age", "hispanic", "Tobacco", "Underlying.COD.Code",
           "Record.Axis.Code.2")

  df.new <- df[,vars]
  colnames(df.new) <- c("certno", "sex", "Date", "race",
                      "age", "hispanic", "smoking", "underly", "mltsc2")

  #month and Year
  df.new$Date = as.Date(as.character(df.new$Date),format="%Y-%m-%d")
  df.new$Year = format(df.new$Date,'%Y')
  df.new$Month = months(df.new$Date)
  df.new <- subset(df.new, Month == "May" |
                  Month == "June" |
                  Month == "July" |
                  Month == "August" |
                  Month == "September") #May - Sept

  #re-order by date
  ord = order(df.new$Date)
  df.new = df.new[ord,]

  return(df.new)

}

#function for 2017-2018
get_data4 <- function(df){

  #date of death

```

```

df$Date <- paste(df$`Date.of.Death...Year`,
                 df$`Date.of.Death...Month`,
                 df$`Date.of.Death...Day`,
                 sep = "-")

df$hispanic = 0
df$hispanic[df$Hispanic.No == "N"] = 1

#variables of interest
vars <- c("State.File.Number", "Sex", "Date", "Race.Calculation",
          "Age", "hispanic", "Tobacco", "Underlying.COD.Code",
          "Record.Axis.Code.2")

df.new <- df[,vars]
colnames(df.new) <- c("certno", "sex", "Date", "race",
                    "age", "hispanic", "smoking", "underly", "mltscse2")

#month and Year
df.new$Date = as.Date(as.character(df.new$Date),format="%Y-%m-%d")
df.new$Year = format(df.new$Date,'%Y')
df.new$Month = months(df.new$Date)
df.new <- subset(df.new, Month == "May" |
                Month == "June" |
                Month == "July" |
                Month == "August" |
                Month == "September") #May - Sept

#re-order by date
ord = order(df.new$Date)
df.new = df.new[ord,]

return(df.new)
}

dat_raw_2007 <- get_data1(dat_raw_2007)
dat_raw_2008 <- get_data1(dat_raw_2008)
dat_raw_2009 <- get_data1(dat_raw_2009)
dat_raw_2010 <- get_data1(dat_raw_2010)
dat_raw_2011 <- get_data2(dat_raw_2011)
dat_raw_2012 <- get_data2(dat_raw_2012)
dat_raw_2013 <- get_data2(dat_raw_2013)
dat_raw_2014 <- get_data2(dat_raw_2014)
dat_raw_2015 <- get_data2(dat_raw_2015)
dat_raw_2016 <- get_data3(dat_raw_2016)
dat_raw_2017 <- get_data4(dat_raw_2017)
dat_raw_2018 <- get_data4(dat_raw_2018)

```

```

#### Bring Everything Together ####
dat_raw_2007_2018 <- rbind(dat_raw_2007,
                          dat_raw_2008,
                          dat_raw_2009,
                          dat_raw_2010,
                          dat_raw_2011,
                          dat_raw_2012,
                          dat_raw_2013,
                          dat_raw_2014,
                          dat_raw_2015,
                          dat_raw_2016,
                          dat_raw_2017,
                          dat_raw_2018)

#### cause of death ####
dat_raw_2007_2018$count = 1

#Assume a death isn't from a specific cause, then change count to 1 if assumption is wrong
dat_raw_2007_2018$resp = 0
dat_raw_2007_2018$circ = 0
dat_raw_2007_2018$cardio = 0
dat_raw_2007_2018$ischem = 0

# resp counts
# use the following ICD-10 codes from Tania's 2015 paper: J00-J99
resp_codes <- NA
for(i in 1:100){
  if(i <= 10){
    resp_codes[i] <- paste("J0", i - 1, sep = "")
  }
  else{
    resp_codes[i] <- paste("J", i - 1, sep = "")
  }
}

for(i in 1:length(resp_codes)){
  code <- resp_codes[i]
  deaths_from_code <- grep(code, dat_raw_2007_2018$underly) #primary cause
  dat_raw_2007_2018$resp[deaths_from_code] <- 1
}

# circ counts
# use the following ICD-10 codes from Tania's 2015 paper: I00-I99, G45, G46
circ_codes <- NA
for(i in 1:100){
  if(i <= 10){

```

```

    circ_codes[i] <- paste("I0", i - 1, sep = "")
  }
  else{
    circ_codes[i] <- paste("I", i - 1, sep = "")
  }
}
circ_codes[101:102] <- c("G45, G46")

for(i in 1:length(circ_codes)){
  code <- circ_codes[i]
  deaths_from_code <- grep(code, dat_raw_2007_2018$underly) #primary cause
  dat_raw_2007_2018$circ[deaths_from_code] <- 1
}

# cardio counts
# use the following ICD-10 codes from Tania's 2015 paper: I05-I52
cardio_codes <- NA
j = 0
for(i in 5:52){
  if(i <= 9){
    cardio_codes[j] <- paste("I0", i, sep = "")
  }
  else{
    cardio_codes[j] <- paste("I", i, sep = "")
  }
  j = j + 1
}

for(i in 1:length(cardio_codes)){
  code <- cardio_codes[i]
  deaths_from_code <- grep(code, dat_raw_2007_2018$underly) #primary cause
  dat_raw_2007_2018$cardio[deaths_from_code] <- 1
}

# ischem counts
# use the following ICD-10 codes from Tania's 2015 paper
ischem_codes <- c("I20", "I21", "I22", "I23", "I24", "I25")

for(i in 1:length(ischem_codes)){
  code <- ischem_codes[i]
  deaths_from_code <- grep(code, dat_raw_2007_2018$underly) #primary cause
  dat_raw_2007_2018$ischem[deaths_from_code] <- 1
}

```

```

}

#### Read-in Geocoded Mortality Data ####

# Annie already has a file with certno and lat/long for 2006-2017
# (it lacked other variables of interest, though)
DataPath_geo = "//pacific/bilab/WildfireSmoke_Data/Mortality\data\Annie\ Death\ Data\death.fi
dat_full_2006_2017 <- read.csv(paste(DataPath_geo, "join_all_complete.csv", sep = ""))

#2018
DataPath_geo = "//pacific/bilab/WildfireSmoke_Data/Mortality\data\Annie\ Death\ Data\geocodes
dat_geo_2018 <- read.xlsx(paste(DataPath_geo, "DeathGeocodeF2018.xlsx", sep = ""))
colnames(dat_geo_2018)[1] = "certno"
colnames(dat_geo_2018)[7] = "lat"
colnames(dat_geo_2018)[8] = "long"

# extract lat/long data
dat_lat_long_2006_2017 <- dat_full_2006_2017[,c("certno", "lat", "long")]
dat_lat_long_2018 <- dat_geo_2018[,c("certno", "lat", "long")]

#combine
dat_lat_long <- rbind(dat_lat_long_2006_2017, dat_lat_long_2018)
dat_lat_long$lat <- as.numeric(dat_lat_long$lat)
dat_lat_long$long <- as.numeric(dat_lat_long$long)

# merge with raw data by certno
dat_2007_2018 <- merge(dat_raw_2007_2018, dat_lat_long, by = c("certno"), all.x = T)

# remove deaths w/out geo data
# BEFORE: 252780
# AFTER: 245996
# 6784/252780 = 2.68%
dat_2007_2018 <- dat_2007_2018[complete.cases(dat_2007_2018[,c("lat", "long")]),]

#### Assign to Climate Zones ####
library(sf) #for spatial operations

#read shapefile
DataPath_cz <- "//pacific/bilab/Code_data_FROM_MEMBRANE/CCHRISK/SpatialAnalysis/Data/climateZon
cz_shp <- st_read(paste(DataPath_cz, "Climate_Zones_WA.shp", sep = ""))

# create points
p_matrix = matrix(c(dat_2007_2018$long, dat_2007_2018$lat), ncol = 2)

#see where points intersect with climate zones

```

```

p_matrix_cz <- apply(p_matrix, 1, function(row) {

  # transformation to planar is required,
  # since sf library assumes planar projection
  cz_pl <- st_transform(cz_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,
  #                                dimnames = list("", c("lat", "long"))))
  pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
  # st_intersects with sparse = FALSE returns a logical matrix
  # with rows corresponds to argument 1 (points) and
  # columns to argument 2 (polygons)

  cz_pl[which(st_intersects(pnt_sf, cz_pl, sparse = FALSE)), ]$NAME
})

#### Removing points that could not be assigned into a climate zone ####
# not all points with lat/long data could be classified

n_obs <- 1:length(p_matrix_cz) #total number of observations

l_obs <- NA #observations with location data
for(i in 1:10){
  l_obs <- c(l_obs, which(p_matrix_cz == i)) #matching to 1 of the 10 climate zones
}

#remove first NA
l_obs <- l_obs[-1]
#order from least to greatest - these are the indices of observations that are kept
l_obs <- sort(l_obs)
#the observations to remove are any observation without location data
r_obs <- n_obs[-l_obs]

#remove observations without location data
# ORIGINAL: 252780
# LAT/LONG DATA: 245996
# POINTS IN CZ: 243140
# 243140/252780 = 96.2%
dat_2007_2018 <- dat_2007_2018[-r_obs, ]

#### Update Final Data Frame with Climate Zone Data ####

#order of climate zone names used in other files
cz_names <- c("WEST OLYMPIC COAST",
              "NE OLYMPIC SAN JUAN",
              "PUGET SOUNDS LOWLANDS",
              "E OLYMPICS CASCADE FOOTHILLS",

```

```

        "CASCADE MOUNTAINS WEST",
        "EAST SLOPE CASCADES",
        "OKANOGAN BIG BEND",
        "CENTRAL BASIN",
        "NORTHEASTERN",
        "PALOUSE BLUE MOUNTAINS")

#add cz data to observations with appropriate location data
dat_2007_2018$zonename <- unlist(p_matrix_cz)

dat_2007_2018$zonenum <- NA
for(i in 1:length(cz_names)){
  dat_2007_2018$zonenum[dat_2007_2018$zonename == cz_names[i]] = i
}

cz_names <- levels(unlist(p_matrix_cz))

#### Re-order by date ####
ord = order(dat_2007_2018$Date, dat_2007_2018$zonename)
dat_2007_2018 = dat_2007_2018[ord,]

#### Aggregate individual-level data for each day and climate zone ####

dat_2007_2018_total <- aggregate(cbind(count,resp,circ,cardio,ischem)
                                ~ zonename + Date,
                                data = dat_2007_2018, FUN=sum)

#### Save new data files ####
setwd("C:/Users/larnold3/Desktop")

write.csv(dat_2007_2018, "2007_2018_CZ_daily_mortality_individual.csv")

write.csv(dat_2007_2018_total, "2007_2018_CZ_daily_mortality_total.csv")

#####
# Merge 1980 - 2006 Data with 2007 - 2018 Data
#####

#### Packages ####

library(reshape)
library(RANN) #for nearest neighbor analysis to assign exposure

#climate zones
cz_names <- c("WEST OLYMPIC COAST",

```

```

    "NE OLYMPIC SAN JUAN",
    "PUGET SOUND LOWLANDS",
    "E OLYMPIC CASCADE FOOTHILLS",
    "CASCADE MOUNTAINS WEST",
    "EAST SLOPE CASCADES",
    "OKANOGAN BIG BEND",
    "CENTRAL BASIN",
    "NORTHEASTERN",
    "PALOUSE BLUE MOUNTAINS")

cz_names <- sort(cz_names) #put in alphabetical order

#### Read-in raw Mortality Data ####

DataPath = "C:/Users/larnold3/desktop/"

dat_1980_2006 <- read.csv(paste(DataPath, "MortalityData/1980_2006_CZ_daily_mortality_individuals.csv", sep = ""))
dat_2007_2018 <- read.csv(paste(DataPath, "MortalityData/2007_2018_CZ_daily_mortality_individuals.csv", sep = ""))

#keep only the needed data
dat_1980_2006 <- dat_1980_2006[,c(18, #certno
                                2, 4, #date
                                13, 14, 3, 5, #climate zone and lat/lon
                                25, 19, 26, 27, 28, 29, #cause of death
                                15, 16, 21, 22, 23)] #individual level characteristics
colnames(dat_1980_2006)[3] <- "Year"

dat_2007_2018 <- dat_2007_2018[,c(2, #certno
                                  4, 11, #date
                                  20, 21, 18, 19, #climate zone and lat/lon
                                  9, 13, 14, 15, 16, 17, #cause of death
                                  3, 6, 5, 7, 8)] #individual level characteristics

#combine the separate mortality data files
mort.dat <- rbind(dat_1980_2006, dat_2007_2018)
mort.dat$Date <- as.Date(mort.dat$Date)

#####
# Get counties
# This section can be skipped if county data is not needed
#####
library(sf)

WA_shp <- st_read(paste(DataPath, "Shapefiles/WAcounties/WA_County_Bndys.shp", sep = ""))

```

```

p_matrix = matrix(c(mort.dat$long, mort.dat$lat), ncol = 2) # create points

print("Assigning points into each county")

#see where points intersect with counties
p_matrix_WA <- apply(p_matrix, 1, function(row) {

  # transformation to palnar is required, since sf library assumes planar projection
  WA_pl <- st_transform(WA_shp, 2163)
  #coords <- as.data.frame(matrix(row, nrow = 1,
  #                                dimnames = list("", c("lat", "long"))))
  pnt_sf <- st_transform(st_sfc(st_point(row), crs = 4326), 2163)
  # st_intersects with sparse = FALSE returns a logical matrix
  # with rows corresponds to argument 1 (points) and
  # columns to argument 2 (polygons)

  WA_pl[which(st_intersects(pnt_sf, WA_pl, sparse = FALSE)), ]$JURNM
})

r_obs_WA <- which(is.na(p_matrix_WA == 1)) #points that could not be classified

mort.dat <- mort.dat[-r_obs_WA, ]
p_matrix_WA <- p_matrix_WA[-r_obs_WA]

mort.dat <- cbind(unlist(p_matrix_WA), mort.dat)
colnames(mort.dat)[1] <- "county"
write.csv(mort.dat, "GeoCodedMortalityData_County.csv")

#####
# Convert Data to Case Crossover Format
#####

#for each case, identify the set of referent days
#want to match by day of week during the month
mort.dat$ref.day1 <- as.factor(mort.dat$Date - 28)
mort.dat$ref.day2 <- as.factor(mort.dat$Date - 21)
mort.dat$ref.day3 <- as.factor(mort.dat$Date - 14)
mort.dat$ref.day4 <- as.factor(mort.dat$Date - 7)
mort.dat$ref.day5 <- as.factor(mort.dat$Date + 7)
mort.dat$ref.day6 <- as.factor(mort.dat$Date + 14)
mort.dat$ref.day7 <- as.factor(mort.dat$Date + 21)
mort.dat$ref.day8 <- as.factor(mort.dat$Date + 28)
mort.dat$index.month <- as.numeric(format(mort.dat$Date, format="%m"))

#convert wide to long
df.cases.long <- melt(mort.dat,
                      id.vars=c("certno", "index.month"),

```

```

        measure.vars=c("Date",
                      "ref.day1", "ref.day2",
                      "ref.day3", "ref.day4",
                      "ref.day5", "ref.day6",
                      "ref.day7", "ref.day8"))
colnames(df.cases.long)[3:4] <- c("day", "Date")

#get rid of referent days in April and October, since they are outside the study period
df.cases.long$month <- as.numeric(format(df.cases.long$Date, format="%m"))
df.referents <- subset(df.cases.long, df.cases.long$month > 4 & df.cases.long$month <10)

#assign case days and control days
df.referents$case <- ifelse(df.referents$day == "Date", 1, 0)

#keep only referent days within the same month as the case day
df.ref.select <- subset(df.referents,
                      df.referents$month == df.referents$index.month)

#merge with mortality data by certno to get individual characteristics
df.cc <- merge(df.ref.select, mort.dat[,1:17], by=c("certno"))
#Date for case-crossover analysis, i.e. date of death OR date of reference
colnames(df.cc)[4] <- "Date.CC"
colnames(df.cc)[7] <- "Date.Death" #Date of Death

# # Sample code to use this dataset to fit a case-crossover model for one climate zone
# df.subset <- df.cc[df.cc$zonenum == 1,]
# mod.cc = clogit(case ~ exposure + strata(certno),
#                weights = count,
#                method = 'breslow',
#                data = df.subset)
# summary(mod.cc)

#####
# Assign Exposure to Each Case and Referent Day
#####

# function that takes in:
# lat, and long of death
# and dataframe of meteorological data
# then, returns met values for the given date that are closest to lat/long death data
# NOTE: assumes that the date of death (or for referent day)
# matches the date of met data
# and that the the met data has been subsetted to a climate zone
# to match the climate zone of death
closest_met <- function(LAT, LONG, met.df){

  #nearest neighbor calculation

```

```

#want to find the point in the first argument that is closest to each point in the second ar
closest <- nn2(cbind(met.df$lat, met.df$long), cbind(LAT, LONG), 1)

met.data.closest <- unlist(met.df[closest$nn.idx, 6:11]) #return a vector of only the met va

}

#create empty columns in case-crossover dataframe to hold met data
df.cc$tmin <- NA
df.cc$tmax <- NA
df.cc$rhmax <- NA
df.cc$rhmin <- NA
df.cc$rhavg <- NA
df.cc$humidex <- NA

#### Read-in and Merge Met Data ####

#timeframe
year_start <- 1980
year_end <- 2018
years <- c(year_start:year_end)

for(year in years){

  print(paste("Reading met data for year", year, sep = " "))

  #read-in data for give year and combine w/ previous year
  met.data.year <- read.csv(paste(DataPath,
                                "MetData/met_data_", year,
                                ".csv", sep = ""))

  #convert date to character
  met.data.year$date <- as.character(met.data.year$date)
  #rename "x" and "y" to "long" and "lat"
  colnames(met.data.year)[1:2] <- c("long", "lat")

  print(paste("Joing met data for year", year, sep = " "))

  #define start and end dates
  sD <- paste(year, "-05-01", sep = "") #start date (May 1)
  eD <- paste(year, "-09-30", sep = "") #end date (Sept 1)

  start <- as.Date(sD)
  end <- as.Date(eD)
  theDate <- start

  #loop through dates between start and end
  while (theDate <= end){

```

```

print(paste("Today is ", theDate, sep = ""))

#subset met data for the date
met.data.day <- met.data.year[met.data.year$date == theDate, ]

#subset met data for the CZ
for(name in cz_names){
  print(paste("Assigning in ", name, sep = ""))

  met.data.day.cz <- met.data.day[met.data.day$zonename == name, ]

  #If the case-crossover date (date of death or referent day)
  #matches "theDate,"
  #and if the climate zone matches "name"
  #then get met data for theDate
  obs <- which(df.cc$date.CC == theDate & df.cc$zonename == name)
  for(i in obs){
    df.cc[i,23:28] <- closest_met(df.cc$lat[i],
                                df.cc$long[i],
                                met.data.day.cz)
  }
}

theDate <- theDate + 1

print(" ")

}

print("Saving")

#save data
write.csv(df.cc, paste(DataPath,
                       "case_crossover_data.csv", sep = ""),
          row.names=FALSE)

print(" ")

}

#####
# Individual-Level Covariates
#####

#### Age and Race ####

```

```

#create age group categories according to DOH categories
df.cc$age.grp <- with(df.cc,
  ifelse(age < 5, 1,
    ifelse(age >= 5 & age <15, 2,
      ifelse(age >= 15 & age < 45, 3,
        ifelse(age >= 45 & age < 65, 4,
          ifelse(age >= 65 & age < 85, 5,
            ifelse(age >= 85, 6, 0)))))))))

#create race category according to DOH categories
df.cc$race.cat <- with(df.cc, ifelse(race == "1", "White",
  ifelse(race == "2", "Black",
    ifelse(race == "3", "Native American",
      ifelse(race == "C", "Hispanic",
        ifelse(race %in% c("A", "B", "F", "H"), "NHOPI",
          ifelse(race %in% c("4", "5", "7", "D", "E", "G"), "Asian",
            ifelse(race == 6, "Other", "Not reported")))))))))))

#Create numerical race category
df.cc$race.cat.num <- with(df.cc, ifelse(race.cat == "White", 1,
  ifelse(race.cat == "Black", 2,
    ifelse(race.cat == "Native American", 3,
      ifelse(race.cat == "Hispanic", 4,
        ifelse(race.cat == "NHOPI", 5,
          ifelse(race.cat == "Asian", 6, 0)))))))))

#### cause of death ####
#cardiovascular, circulatory, respiratory, and ischemic have already been classified

#need to classify the following:
cerebro <- NA
diabetes <- NA
nephritis <- NA
renal <- NA
mental <- NA
accident <- NA
suicide <- NA
homicide <- NA

df.cc <- add_column(df.cc, cerebro, .after = 18)
df.cc <- add_column(df.cc, diabetes, .after = 19)
df.cc <- add_column(df.cc, nephritis, .after = 20)
df.cc <- add_column(df.cc, renal, .after = 21)
df.cc <- add_column(df.cc, mental, .after = 22)
df.cc <- add_column(df.cc, accident, .after = 23)
df.cc <- add_column(df.cc, suicide, .after = 24)
df.cc <- add_column(df.cc, homicide, .after = 25)

```

```

#Cerebro: 430-438 I67
cerebro_codes <- c("430", "431", "432", "434", "435", "436", "437", "438", "I67")

for(i in 1:length(cerebro_codes)){

  #exact match - e.g. just 430
  deaths_from_code <- which(df.cc$underly == cerebro_codes[i])
  df.cc$cerebro[deaths_from_code] <- 1

  #other matches - e.g. 4301, 4302, etc.
  code <- paste(cerebro_codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$cerebro[deaths_from_code] <- 1
}

#Diabetes: 250 E08-E13
diabetes_codes <- c("250", "E08", "E09", "E10", "E11", "E12", "E13")

for(i in 1:length(diabetes_codes)){

  #exact match - e.g. just 430
  deaths_from_code <- which(df.cc$underly == diabetes_codes[i])
  df.cc$diabetes[deaths_from_code] <- 1

  #other matches - e.g. 4301, 4302, etc.
  code <- paste(diabetes_codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$diabetes[deaths_from_code] <- 1
}

#nephritis: 580-589 N17-N19
nephritis_codes <- c("580", "581", "582", "583", "584", "585", "586", "587",
                    "588", "589", "N17", "N18", "N19")

for(i in 1:length(nephritis_codes)){

  #exact match - e.g. just 430
  deaths_from_code <- which(df.cc$underly == nephritis_codes[i])
  df.cc$nephritis[deaths_from_code] <- 1

  #other matches - e.g. 4301, 4302, etc.
  code <- paste(nephritis_codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$nephritis[deaths_from_code] <- 1
}

#acute renal failure: 584 N17
renal_codes <- c("584", "N17")

```

```

for(i in 1:length(renal_codes)){
  #exact match - e.g. just 430
  deaths_from_code <- which(df.cc$underly == renal_codes[i])
  df.cc$renal[deaths_from_code] <- 1

  #other matches - e.g. 4301, 4302, etc.
  code <- paste(renal_codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$renal[deaths_from_code] <- 1
}

#mental disorders: 290-316, F01-F69
mental_codes <- NA
j = 0
for(i in 1:69){
  if(i <= 9){
    mental_codes[j] <- paste("F0", i, sep = "")
  }
  else{
    mental_codes[j] <- paste("F", i, sep = "")
  }
  j = j + 1
}
for(i in 290:316){
  mental_codes <- c(mental_codes, i)
}

for(i in 1:length(mental_codes)){
  #exact match - e.g. just 430
  deaths_from_code <- which(df.cc$underly == mental_codes[i])
  df.cc$mental[deaths_from_code] <- 1

  #other matches - e.g. 4301, 4302, etc.
  code <- paste(mental_codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$mental[deaths_from_code] <- 1
}

#Non-Traumatic: A00 - R99
#Traumatic: S00 - Z99
letters <- c("S", "T", "U", "V", "W", "X", "Y", "Z")
Traumatic_Codes <- NA
j = 1
for(k in 1:length(letters)){
  for(i in :99){
    if(i < 10){

```

```

    Traumatic_Codes[j] <- paste(letters[k], "0", i, sep = "")
  }else{
    Traumatic_Codes[j] <- paste(letters[k], i, sep = "")
  }
  j = j + 1
}
}

df.cc$Traumatic <- NA
for(i in 1:length(Traumatic_Codes)){
  print(Traumatic_Codes[i])

  #exact match - e.g. just V01
  deaths_from_code <- which(df.cc$underly == Traumatic_Codes[i])
  df.cc$Traumatic[deaths_from_code] <- 1

  #other matches - e.g. V011, V012, etc.
  code <- paste(Traumatic_Codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.cc$underly) #primary cause
  df.cc$Traumatic[deaths_from_code] <- 1
}

#remove traumatic deaths!
df.cc <- subset(df.cc, is.na(Traumatic))

#####
# Save final data
#####
write.csv(df.cc, paste(DataPath,
                       "case_crossover_data.csv", sep = ""),
          row.names=FALSE)

```

Model Selection

Model selection was used to identify the best way to incorporate humidex using a nonlinear function in the case-crossover analysis (conditional logistic regression model). Both natural cubic splines (with various knot numbers and placements; Part 1) and polynomials (Part 2) were explored. Ultimately, based on AIC and ease of implementation and interpretation, a quadratic polynomial was chosen.

```
#### setup ####

#load libraries
library(gnm)
library(survival)
library(splines)

#read-in data
DataPath <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State/Data
df.cc <- read.csv(paste(DataPath, "case_crossover_data.csv", sep = ""))

#climate zones
cz_names <- c("WEST OLYMPIC COAST",
              "NE OLYMPIC SAN JUAN",
              "PUGET SOUND LOWLANDS",
              "E OLYMPIC CASCADE FOOTHILLS",
              "CASCADE MOUNTAINS WEST",
              "EAST SLOPE CASCADES",
              "OKANOGAN BIG BEND",
              "CENTRAL BASIN",
              "NORTHEASTERN",
              "PALOUSE BLUE MOUNTAINS")

cz_names <- sort(cz_names) #put in alphabetical order

#### clogit model ####
# Case is a 0/1 vector to indicate if the row is a
  # case (1) or
  # control (0)
# for a given stratum
# Weights are the counts of the observations (should be 1 for this data).
# Method = 'breslow' because default method can't use weight.
# Stratum is a unique identifier for each person (e.g. the cert number).

#####
```

```

# Part 1: Natural Cubic Splines
#####

#### Go through 2 to 10 df ####
#Default knot placement

#data frames to hold information
mod.AIC <- matrix(NA, nrow = 9, ncol = 11)
mod.AIC <- as.data.frame(mod.AIC)
colnames(mod.AIC) <- c("DF", cz_names)
mod.AIC$DF <- 2:10

#Loop through climate zones
for(i in 1:length(cz_names)){
  print(paste("Fitting models for ", cz_names[i], sep = ""))

  dat.cz <- df.cc[df.cc$zonename == cz_names[i],]

  #Loop through degrees of freedom
  for(j in 1:nrow(mod.AIC)){

    print(paste("using ", j + 1, " degrees of freedom"), sep = "")

    #fit model
    mod <- clogit(case ~ ns(humidex, df = j + 1) + strata(certno),
                  weights = count,
                  method = 'breslow',
                  data = dat.cz)

    #save AIC for this model
    mod.AIC[j,i + 1] <- AIC(mod)
  }

  print(" ")
}

#see which df is best
for(i in 1:length(cz_names)){
  print(which.min(mod.AIC[,i + 1]))
}

#### 2 df with a knot at the .99 percentile ####

#loop through climate zones
for(i in 1:length(cz_names)){
  print(paste("Fitting model for ", cz_names[i], sep = ""))
}

```

```

dat.cz <- df.cc[df.cc$zonename == cz_names[i],]
dat.cz <- dat.cz %>%
  mutate(humidex.percentile = percent_rank(humidex))

#get the humidex spline
fit.ns <- ns(dat.cz$humidex,
            df = 2,
            knots = quantile(dat.cz$humidex, .99))

#fit the model
mod <- clogit(case ~ fit.ns + strata(certno),
             weights = count,
             method = 'breslow',
             data = dat.cz)

#get AIC for the model
print(AIC(mod))
}

#### Two Knots: one at .50 and one at .75 through .99 ####

#get sequence of upper knots to test
knot.prcnt <- seq(.75, .99, by = .01)

#data frames to hold information
mod.AIC <- matrix(NA, nrow = length(knot.prcnt), ncol = 11)
mod.AIC <- as.data.frame(mod.AIC)
colnames(mod.AIC) <- c("Knot", cz_names)
mod.AIC$Knot <- knot.prcnt

#loop through climate zones
for(i in 1:length(cz_names)){

  print(paste("Fitting models for ", cz_names[i], sep = ""))

  dat.cz <- df.cc[df.cc$zonename == cz_names[i],]

  #loop through knots
  for(j in 1:nrow(mod.AIC)){

    print(paste("placing knot at", 100*knot.prcnt[j],
               "percentile"), sep = "")

    #fit model
    mod <- clogit(case ~ ns(humidex, df = 3,
                          knots = quantile(dat.cz$humidex,

```

```

c(.50, knot.prcnt[j])) )
+ strata(certno),
weights = count,
method = 'breslow',
data = dat.cz)

#get AIC for this model
mod.AIC[j,i + 1] <- AIC(mod)
}

print(" ")
}

#see which percentile is best
for(i in 1:length(cz_names)){
  print(mod.AIC[which.min(mod.AIC[,i + 1]),1])
  print(min(mod.AIC[,i + 1]))
}

#####
# Part 2: Polynomials
# Go through 2 to 10 degrees,
# Use AIC and BIC for model selection
#####

#data frames to hold information for AIC
mod.AIC <- matrix(NA, nrow = 9, ncol = 11)
mod.AIC <- as.data.frame(mod.AIC)
colnames(mod.AIC) <- c("DF", cz_names)
mod.AIC$DF <- 2:10

#data frames to hold information for BIC
mod.BIC <- matrix(NA, nrow = 9, ncol = 11)
mod.BIC <- as.data.frame(mod.BIC)
colnames(mod.BIC) <- c("DF", cz_names)
mod.BIC$DF <- 2:10

#loop through climate zones
for(i in 1:length(cz_names)){

  print(paste("Fitting models for ", cz_names[i], sep = ""))

  dat.cz <- df.cc[df.cc$zonename == cz_names[i],]

  #loop through degrees

```

```

for(j in 1:nrow(mod.AIC)){

  print(paste("using ", j + 1, " degrees of freedom"), sep = "")

  #fit model
  mod <- clogit(case ~ poly(humidex, j + 1) + strata(certno),
                weights = count,
                method = 'breslow',
                data = dat.cz)

  #get AIC and BIC
  mod.AIC[j,i + 1] <- AIC(mod)
  mod.BIC[j,i + 1] <- BIC(mod)
}
print(" ")
}

#see which degree is best
for(i in 1:length(cz_names)){
  print(cz_names[i])
  print(which.min(mod.AIC[,i + 1]) + 1)
  print(which.min(mod.BIC[,i + 1]) + 1)
}

```

Case-Crossover Analysis and Historical Heat-Attributable Mortality

The `clogit()` function from the `survival` package (Therneau 2020) was used to conduct the case-crossover analysis for the historical period. The analysis was conducted using a conditional logistic regression model for each climate division with control days matched by day of week, month, and year. Humidex was included as a quadratic polynomial. Note that while looping through each climate division to conduct the case-crossover analysis, the code also calculates mortality attributable to heat above the 99th percentile and provides a 95% CI for this estimate.

```
#### clogit model ####
# Case is a 0/1 vector to indicate if the row is a
  # case (1) or
  # control (0)
# for a given stratum
# Weights are the counts of the observations (should be 1 for this data).
# Method = 'breslow' because default method can't use weight.
# Stratum is a unique identifier for each person (e.g. the cert number).

#####
## Fit quadratic model, save coefficients and odds ratios
#####

#read-in data
DataPath <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State/Data
df.cc <- read.csv(paste(DataPath, "case_crossover_data.csv", sep = ""))

#climate zones
cz_names <- c("WEST OLYMPIC COAST",
              "NE OLYMPIC SAN JUAN",
              "PUGET SOUND LOWLANDS",
              "E OLYMPIC CASCADE FOOTHILLS",
              "CASCADE MOUNTAINS WEST",
              "EAST SLOPE CASCADES",
              "OKANOGAN BIG BEND",
              "CENTRAL BASIN",
              "NORTHEASTERN",
              "PALOUSE BLUE MOUNTAINS")
cz_names <- sort(cz_names) #put in alphabetical order

#get quadratic polynomial for humidex
fit.poly <- poly(df.cc$humidex, 2)
df.cc$fit.poly1 <- fit.poly[,1]
df.cc$fit.poly2 <- fit.poly[,2]
```

```

#list to hold plots for each climate zone
myplots <- vector('list', length(cz_names))

#matrix to hold coefficients
#beta 1 is linear, beta 2 is quadratic
coefs <- matrix(NA, nrow = 10, ncol = 6)
coefs <- as.data.frame(coefs)
rownames(coefs) <- cz_names
colnames(coefs) <- c("Beta1", "Beta2",
                    "Beta1_lower", "Beta2_lower",
                    "Beta1_upper", "Beta2_upper")
coefs.se <- matrix(NA, nrow = 10, ncol = 2)
coefs.se <- as.data.frame(coefs.se)
rownames(coefs.se) <- cz_names
colnames(coefs.se) <- c("Beta1", "Beta2")
coefs.vcov <- vector('list', length(cz_names))

#list to hold odds ratios comparing mortality at various humidex percentiles
percentiles <- seq(.51, .99, by = .01)
OR.estimates <- vector(mode = "list", length = 10)
for(i in 1:10){
  OR.estimates[[i]] <- matrix(NA, nrow = 49, ncol = 3)
  colnames(OR.estimates[[i]]) <- c("OR", "Lower", "Upper")
  rownames(OR.estimates[[i]]) <- percentiles
}

#loop through climate zones and fit model for each one
for(i in 1:length(cz_names)){
  print(paste("Fitting model for ", cz_names[i], sep = ""))

  #subset to the climate zone i
  dat.cz <- df.cc[df.cc$zonename == cz_names[i],]
  dat.cz <- dat.cz %>%
    mutate(humidex.percentile = percent_rank(humidex))

  #fit model
  mod <- clogit(case ~ fit.poly1 + fit.poly2 + strata(certno),
                weights = count,
                method = 'breslow',
                data = dat.cz)

  #save coefficients
  coefs[i, 1:2] <- coefficients(mod)
  coefs[i, 3:4] <- confint(mod)[,1]
  coefs[i, 5:6] <- confint(mod)[,2]
  coefs.se[i,] <- summary(mod)$coefficients[,3]
  coefs.vcov[[i]] <- vcov(mod)
}

```

```

#get fitted dose-response curve
dat.fit <- cbind(dat.cz$humidex,
               dat.cz$humidex.percentile,
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 1:2])),
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 3:4])),
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 5:6])))
dat.fit <- as.data.frame(dat.fit)
colnames(dat.fit) <- c("Humidex", "Humidex.Percentile",
                      "Odds", "Lower", "Upper")
dat.fit <- data.frame(dat.fit,
                     OR = dat.fit$Odds/min(dat.fit$Odds))

#calculate odds ratios
fifty.percentile <- which.min(abs(dat.fit$Humidex.Percentile-.50))
fifty.percentile.fit <- dat.fit[fifty.percentile, ]
for(ii in 1:49){
  this.percentile <-
    which.min(abs(dat.fit$Humidex.Percentile - percentiles[ii]))
  this.percentile.fit <- dat.fit[this.percentile, ]
  OR.estimates[[i]][ii,1] <-
    this.percentile.fit$Odds/fifty.percentile.fit$Odds
  OR.estimates[[i]][ii,2] <-
    this.percentile.fit$Lower/fifty.percentile.fit$Lower
  OR.estimates[[i]][ii,3] <-
    this.percentile.fit$Upper/fifty.percentile.fit$Upper
}

#Number of deaths attributable to heat above 99th percentile
humidex.50 <- quantile(dat.cz$humidex, .50)
humidex.50.closest <- which.min(abs(dat.fit$Humidex - humidex.50))
odds.50 <- dat.fit$Odds[humidex.50.closest]
odds.50.lower <- dat.fit$Lower[humidex.50.closest]
odds.50.upper <- dat.fit$Upper[humidex.50.closest]

dat.cz.99 <- subset(dat.cz, humidex.percentile > .99)

print("The total number of deaths occurring on days
      above the 99th percentile are")
print(sum(dat.cz.99$case == 1))
print("The number of deaths attributable to heat above 99th percentile is")
print(sum((1 - exp(-cbind(dat.cz.99$fit.poly1,dat.cz.99$fit.poly2)
                       %*%t(coefs[i, 1:2]/odds.50))))*dat.cz.99$case))
print("and the confidence interval is")
print(paste("Lower:",

```

```
        sum((1 - exp(-cbind(dat.cz.99$fit.poly1,dat.cz.99$fit.poly2)
                           %*%t(coefs[i, 3:4]/odds.50.lower))))*dat.cz.99$case),
        sep = "")
print(paste("Upper:",
           sum((1 - exp(-cbind(dat.cz.99$fit.poly1,dat.cz.99$fit.poly2)
                           %*%t(coefs[i, 5:6]/odds.50.upper))))*dat.cz.99$case),
        sep = "")
}
```

Meta-Analysis

Meta analysis using the `metagen()` function from the `meta` package (Balduzzi, Rucker, and Schwarzer 2019) to obtain a single, statewide OR (and 95% CI) for mortality.

```
#### read-in data ####
# Data is log odds ratio of mortality,
# comparing 99th to 50th percentile of humidex.
#
# Data is obtained from case-crossover analysis,
# where humidex is expressed as quadratic polynomial.
#
# There is a value and standard error for each climate division

#results directory
DataPath_res <- "//Users/LoganArnold2/Desktop/Tania/Heat-Related Mortality in Washington State

logOdds_99vs50 <- read.csv(paste(DataPath_res, "logOdds_99vs50.csv", sep = ""),
                          header = F)
colnames(logOdds_99vs50) <- c("ES", "SE", "Climate Division")
logOdds_99vs50`Climate Division` <- cz_names

#### meta-analysis ####
m <- meta::metagen(ES,
                   SE,
                   data = logOdds_99vs50,
                   sm = "OR",
                   studlab=paste(`Climate Division`),
                   comb.fixed = TRUE,
                   comb.random = FALSE)

forest(m, leftcols = c("Climate Division"))
```

Effect Modification

Investigating effect modification is straightforward in conditional logistic regression models. Note that in a case-crossover analysis, each stratum in the conditional logistic regression model corresponds to an individual person; therefore, the only covariates that can be included in the model are ones that are time-varying on the level of an individual, e.g. humidex. Including, for example, a covariate for race would not yield an estimated coefficient because an individual's race does not change from the control periods to the hazard periods. Instead, effect modification is performed by subsetting the data based on the variable(s) of interest, and creating a new conditional logistic regression model on this subset of data. For example, if `df.cc` is the case-crossover data frame, then the following code performs effect modification for age group 6 (85+) and cardiovascular deaths, saving the dose-response curve coefficients and odds ratios (comparing the 51st through 99th percentiles to the 50th percentile) for each climate division. This code is easily modifiable for other age groups and/or causes of death, and can also be used to adjust for other individual-level variables (e.g. sex or race).

```
#### setup ####  
  
#load libraries  
library(survival)  
library(dplyr)  
library(metafor)  
  
#subset data to age group 6  
df.cc <- subset(df.cc, age.grp == 6)  
  
#climate zones  
cz_names <- c("WEST OLYMPIC COAST",  
             "NE OLYMPIC SAN JUAN",  
             "PUGET SOUND LOWLANDS",  
             "E OLYMPIC CASCADE FOOTHILLS",  
             "CASCADE MOUNTAINS WEST",  
             "EAST SLOPE CASCADES",  
             "OKANOGAN BIG BEND",  
             "CENTRAL BASIN",  
             "NORTHEASTERN",  
             "PALOUSE BLUE MOUNTAINS")  
  
cz_names <- sort(cz_names) #put in alphabetical order
```

```

#####
# Cause of Death: Cardio
#####

#matrix to hold dose-response coefficients
#beta 1 is linear, beta 2 is quadratic
coefs <- matrix(NA, nrow = 10, ncol = 6)
coefs <- as.data.frame(coefs)
rownames(coefs) <- cz_names
colnames(coefs) <- c("Beta1", "Beta2",
                    "Beta1_lower", "Beta2_lower",
                    "Beta1_upper", "Beta2_upper")
coefs.se <- matrix(NA, nrow = 10, ncol = 2)
coefs.se <- as.data.frame(coefs.se)
rownames(coefs.se) <- cz_names
colnames(coefs.se) <- c("Beta1", "Beta2")

#list to hold OR values for each climate division
percentiles <- seq(.51, .99, by = .01)
OR.estimates <- vector(mode = "list", length = 10)
for(i in 1:10){
  OR.estimates[[i]] <- matrix(NA, nrow = 49, ncol = 3)
  colnames(OR.estimates[[i]]) <- c("OR", "Lower", "Upper")
  rownames(OR.estimates[[i]]) <- percentiles
}

#loop through climate zones
for(i in 1:length(cz_names)){

  print(paste("Fitting model for ", cz_names[i], sep = ""))

  dat.cz <- df.cc[df.cc$zonename == cz_names[i],]
  dat.cz <- dat.cz %>%
    mutate(humidex.percentile = percent_rank(humidex))

  #subset the data for this climate zone to include only cardio deaths
  dat.cz <- subset(dat.cz, cardio == 1)

  #check to make sure that there is data for this subset
  #if no data, move to next climate zone
  if(nrow(dat.cz) == 0){
    print("no data")
    next
  }
}

```

```

#fit model
mod <- clogit(case ~ poly(humidex, 2) + strata(certno),
              weights = count,
              method = 'breslow',
              data = dat.cz)
coefs[i, 1:2] <- coefficients(mod)
coefs[i, 3:4] <- confint(mod)[,1]
coefs[i, 5:6] <- confint(mod)[,2]

coefs.se[i,] <- summary(mod)$coefficients[,3]

fit.poly <- poly(dat.cz$humidex, 2)

dat.fit <- cbind(dat.cz$humidex,
                dat.cz$humidex.percentile,
                exp(fit.poly%*%t(coefs[i, 1:2])),
                exp(fit.poly%*%t(coefs[i, 3:4])),
                exp(fit.poly%*%t(coefs[i, 5:6])))
dat.fit <- as.data.frame(dat.fit)
colnames(dat.fit) <- c("Humidex", "Humidex.Percentile",
                     "Odds", "Lower", "Upper")

print("Calculating odds ratios...")
fifty.percentile <- which.min(abs(dat.fit$Humidex.Percentile-.50))
fifty.percentile.fit <- dat.fit[fifty.percentile, ]
for(ii in 1:49){
  this.percentile <- which.min(abs(dat.fit$Humidex.Percentile -
                                percentiles[ii]))
  this.percentile.fit <- dat.fit[this.percentile, ]
  OR.estimates[[i]][ii,1] <-
    this.percentile.fit$Odds/fifty.percentile.fit$Odds
  OR.estimates[[i]][ii,2] <-
    this.percentile.fit$Lower/fifty.percentile.fit$Lower
  OR.estimates[[i]][ii,3] <-
    this.percentile.fit$Upper/fifty.percentile.fit$Upper
}
}

#####
# Meta-Analysis
#####

OR.effect.size <- matrix(NA, nrow = 10, ncol = 3)

#get .99 vs .50 value
for(i in 1:10){

```

```

  OR.effect.size[i,] <- OR.estimated[[i]][49,]
}

log.OR.effect.size <- log(OR.effect.size)
log.OR.SE <- (log.OR.effect.size[,3] - log.OR.effect.size[,2])/3.96
weights <- 1/log.OR.SE^2

log.OR.meta <- sum(weights*log.OR.effect.size[,1])/sum(weights)
log.OR.meta.SE <- sqrt(1/sum(weights))
log.OR.meta.lower <- log.OR.meta - 1.96*log.OR.meta.SE
log.OR.meta.upper <- log.OR.meta + 1.96*log.OR.meta.SE

dat.meta <- cbind(log.OR.effect.size[,1], log.OR.SE)
dat.meta <- as.data.frame(dat.meta)
dat.meta$zonename <- cz_names
colnames(dat.meta) <- c("ES", "SE", "Zone")
# Replace Inf in data by NA
# Some groups have small sample sizes, so estimates can't be produced
dat.meta <- do.call(data.frame,
                    lapply(dat.meta,
                            function(x) replace(x, is.infinite(x), NA)))

#Fixed Effect Model - same as above
model.FE <- rma(ES, sei = SE, data = dat.meta, method="FE")
model.FE
log.OR.FE <- model.FE$beta
log.OR.FE.lower <- model.FE$ci.lb
log.OR.FE.upper <- model.FE$ci.ub

#create dataframe
dat.forest <- exp(log.OR.effect.size)
dat.forest <- as.data.frame(dat.forest)
colnames(dat.forest) <- c("OR", "OR.Lower", "OR.Upper")
dat.forest$zonename <- cz_names
dat.forest <- rbind(dat.forest,
                    c(exp(log.OR.FE),
                      exp(log.OR.FE.lower),
                      exp(log.OR.FE.upper),
                      "FE Model for all Climate Zones"))

#save data
write.csv(dat.forest, "EffectModification_Age6_Cardio.csv")

```

Sensitivity Analysis

The code below conducts the sensitivity analysis comparing the case-crossover approach (conditional logistic regression) to the time series approach (generalized additive model) in King County, as described in see [Appendix D](#).

```
#### setup ####
countyName = 'King County'

#Packages
library(readxl) #read-in xlsx data
library(lubridate) #for date manipulation
library(dplyr)

library(survival)
#library(mgcv) #for time series model
library(gam) #for time series model

library(ggplot2)

## re-plot all the figures ? 0 = No, 1 = Yes ##
plotId = 0

## Specify directories ##

DataPath = "C:/Users/larnold3/desktop/"
OutPath = "C:/Users/larnold3/desktop/"

years <- 1980:2018

#### load mortality data ####

#aggregated into counties
df.county <- read.csv(paste(DataPath,
                           "case_crossover_data_county.csv", sep = ""))

#Remove traumatic deaths from 2007 - 2018
#Non-Traumatic: A00 - R99
#Traumatic: S, T, U, V, W, X, Y, Z
letters <- c("S", "T", "U", "V", "W", "X", "Y", "Z")
Traumatic_Codes <- NA

j = 1
for(k in 1:length(letters)){
  for(i in 0:99){
```

```

if(i < 10){
  Traumatic_Codes[j] <- paste(letters[k], "0", i, sep = "")
}else{
  Traumatic_Codes[j] <- paste(letters[k], i, sep = "")
}
j = j + 1
}
}

df.county$Traumatic <- NA
for(i in 1:length(Traumatic_Codes)){
  print(Traumatic_Codes[i])

  #exact match - e.g. just V01
  deaths_from_code <- which(df.county$underly == Traumatic_Codes[i])
  df.county$Traumatic[deaths_from_code] <- 1

  #other matches - e.g. V011, V012, etc.
  code <- paste(Traumatic_Codes[i], "[0-9]", sep = "")
  deaths_from_code <- grep(code, df.county$underly) #primary cause
  df.county$Traumatic[deaths_from_code] <- 1
}

df.county_NonTraumatic <- subset(df.county, is.na(Traumatic))

#subset to King County
countyMort = subset(df.county_NonTraumatic , county == countyName)

#####
## Case-Crossover Model
#####

fit.poly <- poly(countyMort$humidex, 2)
countyMort$fit.poly1 <- fit.poly[,1]
countyMort$fit.poly2 <- fit.poly[,2]

#fit model
mod.cc <- clogit(case ~ fit.poly1 + fit.poly2 + strata(certno),
  weights = count,
  method = 'breslow',
  data = countyMort)

#get coefficients
beta <- coefficients(mod.cc)
beta.lower <- confint(mod.cc)[,1]
beta.upper <- confint(mod.cc)[,2]

```

```

#gather data for plotting
dat.fit <- cbind(countyMort$humidex,
                exp(cbind(countyMort$fit.poly1,
                          countyMort$fit.poly2)%*%beta),
                exp(cbind(countyMort$fit.poly1,
                          countyMort$fit.poly2)%*%beta.lower),
                exp(cbind(countyMort$fit.poly1,
                          countyMort$fit.poly2)%*%beta.upper))
dat.fit <- as.data.frame(dat.fit)
colnames(dat.fit) <- c("Humidex", "Odds", "Lower", "Upper")

#graphing
ggplot(data=dat.fit, aes(x=Humidex, y=Odds)) +
  geom_line() +
  xlab("Humidex") +
  xlim(-5, 50) +
  geom_ribbon(data=dat.fit, aes(ymin=Lower, ymax=Upper), alpha=0.3) +
  ylab("Odds of Mortality") +
  ylim(.95, 1.4) +
  theme_classic() +
  theme(legend.position = "none") +
  geom_hline(yintercept = 1, linetype = 'dashed') +
  ggtitle("Case Crossover")

#calculate OR comparing 99th to 50th percentiles
humidex.50 <- quantile(countyMort$humidex, .50)
humidex.99 <- quantile(countyMort$humidex, .99)

humidex.50.closest <- which.min(abs(dat.fit$Humidex - humidex.50))
humidex.99.closest <- which.min(abs(dat.fit$Humidex - humidex.99))

odds.50 <- dat.fit$Odds[humidex.50.closest]
odds.50.lower <- dat.fit$Lower[humidex.50.closest]
odds.50.upper <- dat.fit$Upper[humidex.50.closest]

odds.99 <- dat.fit$Odds[humidex.99.closest]
odds.99.lower <- dat.fit$Lower[humidex.99.closest]
odds.99.upper <- dat.fit$Upper[humidex.99.closest]

OR <- cbind(odds.99/odds.50,
            odds.99.lower/odds.50.lower,
            odds.99.upper/odds.50.upper)
OR

#get MMT
MMT.CC <- dat.fit$Humidex[which.min(dat.fit$Odds)]
MMT.CC

```

```

#####
## Time Series Model
#####

# df.cc is for case-crossover analysis, so referent days must be removed
countyMort.TS <- subset(countyMort, case == 1)

#get total deaths for each day
countyMort.count = aggregate(count ~ Date.Death,
                             data = countyMort.TS ,
                             FUN=sum)

#get average humidex for each day
countyMort.humidex = aggregate(humidex ~ Date.Death,
                               data = countyMort.TS,
                               FUN = mean)

#merge count and humidex data
countyMort.date <- merge(countyMort.count,
                        countyMort.humidex,
                        by = "Date.Death")

# add time
nDays = 39*153
countyMort.date$time <- c(1: nDays)
countyMort.date$Date.Death <- as.Date(countyMort.date$Date.Death)
countyMort.date$DOW <- wday(countyMort.date$Date) #day of week (1-7, Sunday is 1)
countyMort.date$DOW.Factor <- as.factor(countyMort.date$DOW) #turn DOW into factor
countyMort.date$month <- month(countyMort.date$Date.Death)
countyMort.date$Month.Factor <- as.factor(countyMort.date$month)
countyMort.date$Year <- year(countyMort.date$Date.Death)

#include month and DOW using indicator variable w/ all levels included
months <- model.matrix( ~ Month.Factor - 1, data = countyMort.date)
DOWs <- model.matrix( ~ DOW.Factor - 1, data = countyMort.date)

countyMort.date[colnames(DOWs)] <- DOWs
countyMort.date[colnames(months)] <- months

#population
pop <- read_excel(paste(DataPath, "KingCountyPopulation.xlsx", sep = ""))
pop <- pop[22:64] #get 1980-2018
pop <- as.numeric(pop[-c(2,13, 24, 35)]) #remove duplicate counts from census years; convert to
pop <- cbind(pop, years)
colnames(pop) <- c("Population", "Year")
countyMort.date <- merge(countyMort.date, pop, by = "Year") #Merge Mortality and Population Data

```

```

#GAM time series model
mod.TS <- gam(count ~
              s(time, df = 4) +
              poly(humidex, 2) +
              Month.Factor5 + Month.Factor6 + Month.Factor7 +
              Month.Factor8 + Month.Factor9 +
              DOW.Factor1 + DOW.Factor2 + DOW.Factor3 + DOW.Factor4 +
              DOW.Factor5 + DOW.Factor6 + DOW.Factor7 +
              offset(log(Population)),
              family = quasi(link=log,variance=mu),
              data=countyMort.date)

## diagnostics
res <- residuals(mod.TS)
res <- res[!is.na(res)]
par(mfrow = c(1,3))
qqnorm(res, pch = 1, frame = FALSE)
qqline(res, col = "steelblue", lwd = 2)
acf(res, main = "ACF of Residuals")
pacf(res, main = "PACF of Residuals")

#plot(mod.TS, se = T, ask = T) #select 2 to get heat-mortality curve

#Get humidex values and plot them, fixing axis labels
preplot.obj <- preplot(mod.TS, se = T)
#preplot.obj.humidex <- preplot.obj$`poly(humidex, 2)`
preplot.obj.humidex <- preplot.obj$`s(humidex, df = 4)`
par(mfrow = c(1,1))
plot(preplot.obj.humidex, se = T,
      xlab = "Humidex", ylab = "Log Mortality Rate",
      main = "Time Series")

#calculate RR comparing 99th to 50th percentiles
humidex.50 <- quantile(countyMort.date$humidex, .50)
humidex.99 <- quantile(countyMort.date$humidex, .99)

humidex.50.closest <- which.min(abs(preplot.obj.humidex$x - humidex.50))
humidex.99.closest <- which.min(abs(preplot.obj.humidex$x - humidex.99))

logMortalityRate.50 <- preplot.obj.humidex$y[humidex.50.closest]
logMortalityRate.50.lower <-
  preplot.obj.humidex$y[humidex.50.closest] -
  1.96*preplot.obj.humidex$se.y[humidex.50.closest]
logMortalityRate.50.upper <-
  preplot.obj.humidex$y[humidex.50.closest] +
  1.96*preplot.obj.humidex$se.y[humidex.50.closest]

logMortalityRate.99 <- preplot.obj.humidex$y[humidex.99.closest]

```

```

logMortalityRate.99.lower <-
  preplot.obj.humidex$y[humidex.99.closest] -
  1.96*preplot.obj.humidex$se.y[humidex.99.closest]
logMortalityRate.99.upper <-
  preplot.obj.humidex$y[humidex.99.closest] +
  1.96*preplot.obj.humidex$se.y[humidex.99.closest]

RR <- cbind(exp(logMortalityRate.99)/exp(logMortalityRate.50)
            exp(logMortalityRate.99.lower)/exp(logMortalityRate.50.lower),
            exp(logMortalityRate.99.upper)/exp(logMortalityRate.50.upper))
RR

#get MMT
MMT.TS <- preplot.obj.humidex$x[which.min(preplot.obj.humidex$y)]
MMT.TS

```

Projected Heat-Attributable Deaths

Using projected humidex data and the historical humidex-mortality exposure-mortality curve to calculate future heat-attributable deaths under RCP4.5 and RCP8.5 in the years 2030, 2050, and 2080. This code first determines which grid point of the projected data is closest to the latitude/longitude coordinates associated with each death. Then, the code conducts the historical case-crossover analysis and saves the necessary variables to calculate heat-attributable deaths using Equation 2. Finally, the code calculates projected heat-attributable mortality in each climate division for each combination of year and RCP. Note that the calculations are done for both the bias correction and change factor methods (Ho et al. 2012) to calibrate the projected (CMIP5) meteorological data (Abatzoglou and Brown 2012) to the observed (gridMET) data (Abatzoglou 2013); when applying these calibration techniques, it is assumed that the standard deviation is the same for both the observed (historical) and modeled/projected data. The results reported in text (Table 2.5) are the average of these two methods.

The results are output to a .txt file that contains the following information for each combination of year, RCP, and climate division:

- Number of deaths (labeled “hot days”) with assigned humidex values above the historical 99th percentile; results given for both bias correction and change factor approaches
- The uncalibrated and calibrated mean and 99th percentile of humidex. The calibrated values are given for both the bias correction and change factor approaches.
- Number of projected heat-attributable deaths (including 95% CI) using the bias correction approach for a given climate division
- Number of projected heat-attributable deaths (including 95% CI) using the change factor approach for a given climate division

```
#####  
# Setup  
#####  
  
#### Packages ####  
  
library(reshape)
```

```

library(dplyr)
library(RANN) #for nearest neighbor analysis to assign exposure
library(lubridate)#for working with dates
library(survival) #for clogit

#climate zones
cz_names <- c("WEST OLYMPIC COAST",
              "NE OLYMPIC SAN JUAN",
              "PUGET SOUND LOWLANDS",
              "E OLYMPIC CASCADE FOOTHILLS",
              "CASCADE MOUNTAINS WEST",
              "EAST SLOPE CASCADES",
              "OKANOGAN BIG BEND",
              "CENTRAL BASIN",
              "NORTHEASTERN",
              "PALOUSE BLUE MOUNTAINS")
cz_names <- sort(cz_names)

#### Read-in Data ####

DataPath = "C:/Users/larnold3/desktop/"

#grid points of projected data
load(paste(DataPath, "proj.data.grid.cells.RData", sep = ""))

#historical mortality data (for climate zones), in case-crossover format
df.cc <- read.csv(paste(DataPath, "case_crossover_data.csv", sep = ""))
#subset to remove referent days
df.cc <- subset(df.cc, day == "Date")

#historical mortality data (for counties)
df.county <- read.csv(paste(DataPath, "GeoCodedMortalityData_County.csv", sep = ""))

#merge cz and county data by certno
df.cc <- merge(df.cc, df.county[,c("county", "certno")], by = "certno")

#####
# Match Death Location to Nearest Grid Point in Projected Data
#####

#also get day of year death occurred

#read-in one set of projected data
df.proj <- read.csv(paste(DataPath, "proj_data_2030_avg_total.csv", sep = ""))

#list: for all deaths in each climate division
# get lat/long of projected data closest to the death

```

```

# as well as the month and day of death
proj.data.grid.cells <- vector(mode = "list", length = length(cz_names))

for(i in 1:length(cz_names)){

  print(paste("Assigning closest point for ", cz_names[i], sep = ""))

  #subset by climate zone
  df.cc.sub <- subset(df.cc, zonename == cz_names[i])
  df.proj.sub <- subset(df.proj, zonename == cz_names[i])

  #get unique lat, long coordinates of projected met data
  #to use for nearest neighbor
  df <- unique(as.data.frame(cbind(df.proj.sub$y, df.proj.sub$x)))

  #get lat, long coordinates of historical mortality data
  #to use for nearest neighbor
  #also get day and month of death to assign future exposure
  p_matrix <- data.frame("LAT" = df.cc.sub$lat,
                        "LONG" = df.cc.sub$long,
                        "MonthDay" = paste(month(df.cc.sub$Date.Death),
                                           day(df.cc.sub$Date.Death),
                                           sep = "-"))

  #go through all lat,long locations of death
  #and match to nearest neighbor from projected data
  for(ii in 1:nrow(p_matrix)){
    if(ii %% 1000 == 0){
      print(paste("Reached observation ", ii, " of ",
                  nrow(p_matrix), sep = ""))
    }

    pts <- p_matrix[ii,1:2]
    closest <- nn2(data = df, query = pts, k = 1)

    proj.data.grid.cells[[i]] <- rbind(proj.data.grid.cells[[i]],
                                       data.frame(df[closest$nn.idx,],
                                                  p_matrix[ii,3]))
  }

  colnames(proj.data.grid.cells[[i]]) <- c("y", "x", "MonthDay")
}

#save file
save(proj.data.grid.cells,

```

```

file = paste(DataPath, "proj.data.grid.cells.RData", sep = "")

#####
# Fit Case-Crossover Models with Historical Data
#####
# Necessary for attributable fraction calculations
# Fit quadratic model, save coefficients

#matrix to hold coefficients
#beta 1 is linear, beta 2 is quadratic
coefs <- matrix(NA, nrow = 10, ncol = 6)
coefs <- as.data.frame(coefs)
rownames(coefs) <- cz_names
colnames(coefs) <- c("Beta1", "Beta2",
                    "Beta1_lower", "Beta2_lower",
                    "Beta1_upper", "Beta2_upper")
coefs.se <- matrix(NA, nrow = 10, ncol = 2)
coefs.se <- as.data.frame(coefs.se)
rownames(coefs.se) <- cz_names
colnames(coefs.se) <- c("Beta1", "Beta2")

#re-read-in data (it was subsetted above)
df.cc <- read.csv(paste(DataPath,
                       "case_crossover_data.csv", sep = ""))
df.county <- read.csv(paste(DataPath,
                            "GeoCodedMortalityData_County.csv", sep = ""))

#merge cz and county data by certno
df.cc <- merge(df.cc, df.county[,c("county", "certno")], by = "certno")

#create quadratic polynomial of humidex
fit.poly <- poly(df.cc$humidex, 2)
df.cc$fit.poly1 <- fit.poly[,1]
df.cc$fit.poly2 <- fit.poly[,2]

#list for historical humidex values on days of death
humidex.historical <- vector(mode = "list", length = length(cz_names))

#vector for historical 99th percentiles
humidex.99.historical <- rep(NA, length(cz_names))

#vector for historical odds at 50th percentile
odds.50.historical <- rep(NA, length(cz_names))

#fit the model for each climate division
for(i in 1:length(cz_names)){
  print(paste("Fitting model for ", cz_names[i], sep = ""))
}

```

```

dat.cz <- df.cc[df.cc$zonename == cz_names[i],]
dat.cz <- dat.cz %>%
  mutate(humidex.percentile = percent_rank(humidex))

mod <- clogit(case ~ fit.poly1 + fit.poly2 + strata(certno),
             weights = count,
             method = 'breslow',
             data = dat.cz)
coefs[i, 1:2] <- coefficients(mod)
coefs[i, 3:4] <- confint(mod)[,1]
coefs[i, 5:6] <- confint(mod)[,2]

coefs.se[i,] <- summary(mod)$coefficients[,3]

dat.fit <- cbind(dat.cz$humidex,
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 1:2])),
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 3:4])),
               exp(cbind(dat.cz$fit.poly1,
                         dat.cz$fit.poly2)%*%t(coefs[i, 5:6])))
dat.fit <- as.data.frame(dat.fit)
colnames(dat.fit) <- c("Humidex", "Odds", "Lower", "Upper")

humidex.50 <- quantile(dat.cz$humidex, .50)
humidex.50.closest <- which.min(abs(dat.fit$Humidex - humidex.50))
odds.50 <- dat.fit$Odds[humidex.50.closest]
odds.50.historical[[i]] <- odds.50

humidex.99.historical[i] <- quantile(dat.cz$humidex, .99)
dat.cz <- subset(dat.cz, day == "Date")
humidex.historical[[i]] <- dat.cz[,c("lat", "long", "Date.Death",
                                   "tmax", "rhmin", "humidex",
                                   "county")]
}

#####
# Projected Heat-Attributable Mortality
#####

#### Functions for Calibration ####

#Bias Correction Approach
BC <- function(mu_o, #mean of observation
              sd_o, #standard deviation of observations

```

```

        sd_m_historical, #standard deviation of historical model
        future_modeled_temp, #future modeled temperature
        mu_m_historical) #mean of historical model
    {
    mu_o + sd_o/sd_m_historical*(future_modeled_temp - mu_m_historical)
    }

#Change Factor Approach
CF <- function(mu_m_future, #mean of future modeled temp
              sd_m_future, #sd of future modeled temp
              sd_m_historical, #sd of historical modeled temp
              observed_temp, #observed temp
              mu_m_historical) #mean of historical modeled temp
    {
    mu_m_future + sd_m_future/sd_m_historical*(observed_temp - mu_m_historical)
    }

#### load data for calculations ####

#load bias correction parameters
#first is average, second has specific values for each CZ
#BiasCorrectionParameters <- read.csv(paste(DataPath,
#"BiasCorrectionParameters_Humidex.csv", sep = ""))
load(paste(DataPath, "BiasCorrectionParameters.Rda", sep = ""))

#load grid cells for projected data
load(paste(DataPath, "proj.data.grid.cells_county.RData", sep = ""))

#data path that contains projected data
DataPath_proj <- "C:/Users/larnold3/desktop/ProjectionData/"

#set years and RCPs
years <- c(2030, 2050, 2080)
RCPs = c("rcp45", "rcp85")
nyears <- length(1980:2018)
BC_counter <- 3 #counter for calibration; start in row 3

#### Projections ####

sink("ProjectionsResults.txt")

#loop through years
for(year in years){

    #loop through RCPs
    for(RCP in RCPs){

```

```

cat(paste("Reading in data for",
          year,
          "and scenario",
          RCP, sep = " "))

cat("\n")
cat(".....")
cat("\n")

#create data path for the year and RCP
data.path <- paste(DataPath_proj,
                  year, "/",
                  RCP, "/", sep = "")

#read-in data
proj.data.df <- read.csv(paste(data.path,
                              "proj_data_",
                              year,
                              "_avg_total.csv",
                              sep = ""))

#loop through climate divisions
for(i in 1:length(cz_names)){
  cat(paste("Doing calculations for ", cz_names[i], sep = ""))
  cat("\n")

  #Get Bias Correction Parameters for this CZ
  BiasCorrectionParameters <- param.list[[i]]

  #get grid points and historical humidex for the climate zone
  points.cz <- proj.data.grid.cells[[i]]
  humidex.cz <- humidex.historical[[i]]

  points.cz$HistoricalTmax <- humidex.cz$tmax
  points.cz$HistoricalRhmin <- humidex.cz$rhmin
  points.cz$HistoricalHumidex <- humidex.cz$humidex
  points.cz$MergeVar <- paste(points.cz$x, points.cz$y,
                              points.cz$MonthDay, sep = ":")

  #subset entire projected data set
  #to include only the grid points near deaths
  proj.data.df.cz <- proj.data.df[paste(proj.data.df$x,
                                         proj.data.df$y,
                                         sep=":")
                                   %in% paste(points.cz$x,
                                              points.cz$y,
                                              sep=":"),]
}

```

```

#merge projected data with historical deaths
#match on Day and Month as well as lat and long
proj.data.df.cz$date <- as.Date(proj.data.df.cz$date)
proj.data.df.cz$MonthDay <- paste(month(proj.data.df.cz$date),
                                day(proj.data.df.cz$date),
                                sep = "-")
proj.data.df.cz$MergeVar <- paste(proj.data.df.cz$x,
                                proj.data.df.cz$y,
                                proj.data.df.cz$MonthDay,
                                sep = ":")

merged.data.cz <- merge(points.cz, proj.data.df.cz, by = "MergeVar")

#Get values needed for calibration
mu_o <- BiasCorrectionParameters$Mean[1] #mean(merged.data.cz$HistoricalHumidex)
sd_o <- 1 #BiasCorrectionParameters$SD[1] #sd(merged.data.cz$HistoricalHumidex)
sd_m_historical <- 1 #BiasCorrectionParameters$SD[2]
future_modeled_temp <- merged.data.cz$humidex
mu_m_historical <- BiasCorrectionParameters$Mean[2]

mu_m_future <- BiasCorrectionParameters$Mean[BC_counter] #mean(merged.data.cz$humidex)
sd_m_future <- 1 #BiasCorrectionParameters$SD[BC_counter] #sd(merged.data.cz$humidex)
sd_m_historical <- 1 #BiasCorrectionParameters$SD[2]
observed_temp <- merged.data.cz$HistoricalHumidex
mu_m_historical <- BiasCorrectionParameters$Mean[2]

#do calibration
humidex.BC <- BC(mu_o,
                sd_o,
                sd_m_historical,
                future_modeled_temp,
                mu_m_historical)
humidex.CF <- CF(mu_m_future,
                sd_m_future,
                sd_m_historical,
                observed_temp,
                mu_m_historical)

#get historical and future humidex series
HUMIDEX <- data.frame(cbind(merged.data.cz$HistoricalHumidex,
                           merged.data.cz$humidex,
                           humidex.BC,
                           humidex.CF))
colnames(HUMIDEX) <- c("Historical",
                     "Projected",
                     "BiasCorrected",
                     "ChangeFactor")

```

```

#comparing to 99th percentile
cat(paste("historical hot days: ",
          sum(HUMIDEX$Historical > humidex.99.historical[i])))
cat("\n")
cat(paste("Projected hot days: ",
          sum(HUMIDEX$Projected > humidex.99.historical[i])))
cat("\n")
cat(paste("Bias corrected hot days: ",
          sum(HUMIDEX$BiasCorrected > humidex.99.historical[i])))
cat("\n")
cat(paste("Change factor hot days: ",
          sum(HUMIDEX$ChangeFactor > humidex.99.historical[i])))
cat("\n")
cat("\n")

cat(paste("Mean of projected data: ",
          mean(HUMIDEX$Projected)))
cat("\n")
cat(paste("Mean of BC data: ",
          mean(HUMIDEX$BiasCorrected)))
cat("\n")
cat(paste("Mean of CF data: ",
          mean(HUMIDEX$ChangeFactor)))
cat("\n")

cat(paste("99th Percentile of projected data: ",
          quantile(HUMIDEX$Projected, probs = c(.99))))
cat("\n")
cat(paste("99th Percentile of BC data: ",
          quantile(HUMIDEX$BiasCorrected, probs = c(.99))))
cat("\n")
cat(paste("99th Percentile of CF data: ",
          quantile(HUMIDEX$ChangeFactor, probs = c(.99))))
cat("\n")
cat("\n")

#apply polynomial transformation
#apply to bias corrected humidex values above historical 99th
#humidex was transformed into an orthogonal quadratic polynomial;
#apply those same coefficients to the future data
projected.humidex.poly.BC <- predict(fit.poly,
                                   subset(HUMIDEX,
                                           humidex.BC >
                                           humidex.99.historical[i])$BiasCorrected)
projected.humidex.poly.CF <- predict(fit.poly,
                                   subset(HUMIDEX,
                                           humidex.CF >

```

```

humidex.99.historical[i])$ChangeFactor)

#get this zone's odds of mortality at 50th percentile of humidex
odds.50 <- odds.50.historical[[i]]

#do calculations
cat("The number of total deaths attributable to heat above 99th percentile is")
cat("\n")

cat("PROJECTED (Bias Corrected)")
cat("\n")
cat(sum((1 - exp(-projected.humidex.poly.BC**%t(coefs[i, 1:2])/odds.50))))
cat("\n")
cat(paste("Lower:",
          sum((1 - exp(-projected.humidex.poly.BC**%t(coefs[i, 3:4])/odds.50)))),
      sep = " ")
cat("\n")
cat(paste("Upper:",
          sum((1 - exp(-projected.humidex.poly.BC**%t(coefs[i, 5:6])/odds.50))))),
      sep = " ")
cat("\n")

cat("PROJECTED (Change Factor)")
cat("\n")
cat(sum((1 - exp(-projected.humidex.poly.CF**%t(coefs[i, 1:2])/odds.50))))
cat("\n")
cat(paste("Lower:",
          sum((1 - exp(-projected.humidex.poly.CF**%t(coefs[i, 3:4])/odds.50))))),
      sep = " ")
cat("\n")
cat(paste("Upper:",
          sum((1 - exp(-projected.humidex.poly.CF**%t(coefs[i, 5:6])/odds.50))))),
      sep = " ")
cat("\n")

cat(".....")
cat("\n")
}

cat("updating bias correction counter")
cat("\n")

#update counter for bias correction parameters
BC_counter = BC_counter + 1

cat(".....")
cat("\n")

```

```
}  
  
cat(".....")  
cat("\n")  
  
}  
  
sink()
```