Geographic Dimensions of Heat-Related Mortality in Seven United States Cities

David Michael Hondula
Charlottesville, Virginia

B.A., University of Virginia, 2006
M.S., University of Virginia, 2009

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Environmental Sciences

University of Virginia
August, 2013

Abstract

Extreme heat is the leading cause of weather-related death and illness in the United States. Adaptation and mitigation strategies designed to protect the public would greatly benefit from knowing specifically where heat-related deaths are most likely. This dissertation represents a comprehensive examination of intra-city heat-related mortality.

Over two million records of mortality spanning multiple decades and cities comprise the primary data set. Time series models were employed to relate mortality to temperature after accounting for seasonality and long-term time trends. Threshold temperatures were established for each city associated with significant increases in mortality above typical summertime conditions. The mortality rate within each postal code on threshold-exceeding days was then calculated, quantifying spatial variability in heat-related risk. These intra-city mortality patterns were compared to demographic and environmental variables using multivariate regression. Temporal variability in the spatial mortality patterns was evaluated in an iterative model building and testing framework.

Mortality on hot summer days significantly varied within the study cities, and intra-city variability was far greater than inter-city variability. Characteristics of high-risk zones included more elderly residents, more isolated residents, and more intensely developed land, although the specific set risk factors varied from one city to another. There was evidence that the location of high-risk zones are predictable. A city-specific, health data-driven approach offers an improved strategy for municipalities to identify, understand, and reduce the risks associated with the most deadly natural hazard in the United States.

Acknowledgements

Ten years ago this month I signed off my final broadcast as co-host and meteorologist of BRTV's "Weather Center Live" weekly high school television program in Bridgewater, New Jersey. It is remarkable and humbling to consider the many people who have helped me in the time since turn a childhood interest in the impact of weather and climate on people and society into the challenging and rewarding intellectual pursuit this dissertation represents.

The staff of the Department of Environmental Sciences must be among the country's finest within an academic unit. They have made the administrative aspects of my graduate school pursuits as streamlined as possible, always with smile. Thank you Cindy Allen, Chonna Gammon, Pam Hoover, Grace Lipscomb, Ashley McFadden, Debra Shifflet, and Henry White, for making the complicated seem easy and the last minute seem like a long one. Thank you also to the many wonderful members of the UVa Facilities Management team who help us every day by maintaining excellent workspaces and a friendly atmosphere in Clark Hall.

The disciplinary backgrounds and research interests of Environmental Sciences faculty span a wide range, but there is tremendous consistency in the welcoming and encouraging nature of those whom I have come to know in my time in the department. It is inspiring to spend every day with such a talented group of educators and researchers, and I am thankful to many who have always had their door open to me. Tom Biggs, Bob Dolan, and Manuel Lerdau are among those whose time has seemed endlessly generous.

The faculty members who serve on my dissertation committee have made this work better from its inception. Thank you Bruce Hayden, John Porter, and Matthew Trowbridge. Bruce, it was your Atmosphere and Weather class that hooked me on the Environmental Sciences program in fall 2003. You have been one of the instructors from whom I have learned the most, including how to attract the attention of the local police while completing science homework. John and Matt, your many great ideas and perspectives have challenged me to think in different ways and positioned me for a successful future.

Many friends and colleagues in the department have made Clark Hall a terrific and fun place to call home during graduate school. David Knight, Temple Lee, Luke Sitka, our many years together as officemates have created lifelong friendships. I have enjoyed thoroughly your company and been inspired by you to work hard, pursue interesting questions, and stay aware of the big picture. Jeff Atkins, Luke Cole, Amy Grady, Rosemary Malfi, Laura Reynolds, Drew Robison, Michael Saha, David Seekell, Sara Taube, thank you for your service to our department, the surprise appearance of baked goods at my desk, setting a good example, trips to the donut shop and grocery store, laughter, motivation, and making my experience as an EVSC graduate student a source of many fond memories in the years to come.

The undergraduate students with whom I have been priveledged to work over the past few years have added a new dimension to my enjoyment of continued study and research in the sciences. Matt Leisten, Michael Saha, Carleigh Wegner, and Lindsay Veazey, thank you for your patience, positive attitude, and outstanding effort as you helped me

find my way in my first semester of mentoring. Sydney Delmar, Rachael Dinega, Lyndsay Dowell, Courtney Good, Anjali Patel, Andrew Revelle, and Jennifer Sealey have also made contributions to this dissertation and related ongoing work, and are a continual source of energy and inspiration.

Two outstanding young statisticians welcomed me to their home countries for extended stays during my dissertation research to contribute to my training, develop new research projects, and experience different cultures. Joacim Rocklöv at Umeå University in northern Sweden warmly replied to an e-mail from an American graduate student he had never heard of (and his more reptuable supervisor) in the fall of 2010, which started an unpredictable chain of events that has made a tremendous impact on me intectually and personally. The Nordic Research Opportunity program for U.S. National Science Foundation Graduate Research Fellowship recipients afforded me the chance to spend fall 2011 with him in Umeå. It was there that I met Adrian Barnett, who I would spend the next summer with at Queensland University of Technology in Brisbane, Australia as part of NSF's East-Asia Pacific Summer Research Institute. To both Joacim and Adrian I am deeply indebted for their exceedingly generous hospitality, introducing me to fika and kangaroos, and high-quality instruction and mentoring that have advanced my work in many ways. Other researchers who have contributed ideas and encouragement to this effort include Larry Kalkstein, Scott Sheridan, and Karen Smoyer-Tomic.

My family has been incredibly supportive of this pursuit. Mom and Dad, I will graduate soon, I promise. Thank you for affording me every opportunity to be surrounded by the best and brightest and follow my dreams. To my sister Kelly, a fellow double 'Hoo, thank

you for your thoughtful, creative, and smart perspectives on this work and for the reminders that science is fun. To Stephanie, much love. Sharing the joys of this work with you make them all the brighter.

I am thankful to the National Science Foundation and Environmental Protection Agency, whose graduate research fellowship programs supported me during graduate school. Other funding from the Mike Garstang Award, the Association of American Geographers, the Department of Environmental Sciences, the Graduate School of Arts and Sciences, the Virginia Space Grant Consortium, the Australian Academy of Sciences, and the Swedish Research Council has helped advance this research.

This dissertation is dedicated to Bob Davis, who for a decade has been my supervisor, teacher, role model, lunch companion, and golf partner. Bob, one of my biggest dreams for the future is that I can be a mentor to someone else like you have been to me. You have called me (publicly) unfailingly optimistic, which is certainly reflective of my outlook for my career in the sciences after spending ten years in your company. You have given me freedom to explore and discover my interests, made me part of the academic family, and grown within me the motivation and abilities to achieve the highest standards in teaching, presenting, writing, research, mentoring, and service to the university that you demonstrate every day. Your willingness to take me on as an independent study student as a second-year undergraduate will forever be one of the small moments that has had the biggest impact on my life.

Table of Contents

Chapter 1. Introduction……..…..……1

1. Heat and Human Health…… 1

2. Identifying Impacts……... 3

3. Spatial Variability……..… 9

4. Temporal Variability……... 21

5. Response Strategies……..… 27

6. Motivation for Study……... 31

7. Dissertation Structure…….. 32

References…… 35

Chapter 2. Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis…….....……51

Abstract……. 52

Background……..… 54

Methods…… 57

Results……..… 63

Discussion…… 66

Conclusions……..… 73

Figures…….…… 75

Tables…… 81

References…… 83

Chapter 3. A time series approach for evaluating intra-city heat-related mortality…….87

Abstract…… 88

Introduction…….. 89

Methods……. 92

Results……...……. 97

Discussion…… 100

Figures…….. 103

References……..…… 108

Supplementary Material…….. 113

Chapter 4. Geographic dimensions of heat-related mortality in seven U.S. Cities……...115

Abstract…….. 115

Introduction……. 117

Methods……..…… 121

Results……..…... 128

Discussion……..…... 134

Conclusion……..….. 155

Figures……..… 157

Tables……..… 166

References…….. 172

Appendix…… 180

Chapter 5. The predictability of high-risk zones for heat-related mortality in seven U.S. Cities…….……186

Abstract…….. 187

Introduction……... 188

Data…… 191

Methods…….. 193

Results……..… 201

Discussion…….. 206

Conclusion……. 212

Figures…… 213

Tables…… 218

References……..… 221

Chapter 6. Review…….....226

1. Summary…….. 226

2. Conclusions……..…… 229

3. Future Research Opportunities… 232

Chapter 1. Introduction

The main topic of this dissertation is the relationship between high temperatures and human health. This chapter introduces the topic and reviews relevant scientific literature. Overviews of the underlying physiology (Section 1), known impacts of heat on human health (Section 2), sources of spatial variability in risk (Section 3), sources of temporal variability in risk (Section 4), and public health responsed strategies (Section 5) are included. Section 6 summarizes the motivation for research and Section 7 presents descriptions of each of the four main chapters of the dissertation.

1. Heat and Human Health

The human body works to maintain a thermal equilibrium, eliminating heat at the same rate at which it is produced. Physiological processes attempt to balance heat gains and losses from respiration, evaporation, convection, conduction, work, and radiation. Evaporation is the only mechanism that always removes heat from the body, work (metabolism) the only process that always adds. The other three can result in a net loss or gain of heat depending on the surroundings. The body has a remarkable capacity to keep internal organs within a consistent temperature range, 37-39°C, despite exterior exposure to temperature variations of 30-40°C or greater. One of the principal mechanisms by which thermal regulation is achieved under high temperature conditions is the dilation of blood vessels closest to the skin. This circulation works to remove heat from the body's core and toward the skin where cooling mechanisms including conduction and evaporation transfer heat from the person to the environment. When further cooling is required, sweating begins from an average of 2.5 million endocrine glands scattered across the human skin (Winslow and Herrington 1949, Koppe et al. 2004). On an average

day, these glands produce a total of two to three quarts for a sedentary person (Kavaler 1981).

High temperatures and/or humidity levels can compromise physiological cooling mechanisms. High air temperatures reduce the temperature gradient between the skin and the environment making heat loss via conduction and convection more difficult. Sweating can be inhibited when the vapor pressure is high and the air is near saturation (Winslow and Herrington 1949, Kovats et al. 2009). In either of these cases, or a combination thereof, internal temperatures can increase and the thermoregulatory system can become strained (Basu and Samet 2002). Thermoregulatory strain can be enhanced under extreme conditions or with prolonged exposure (Braga et al. 2001).

When the thermoregulatory system fails to keep internal temperatures below the upper limits of the desirable range (37-39°C) illness and death may follow. Mild conditions include dehydration, when the body has insufficient fluids to replace those lost via sweating, and heat syncope, when the circulation cannot maintain sufficient blood pressure to supply oxygen to the brain. An individual suffering from heat syncope may feel light-headed or faint but symptoms are often relieved by sitting down for a short rest. A more severe diagnosis is heat exhaustion, characterized by an increase in body core temperature and high cardiovascular distress. As with heat syncope, there is insufficient circulation to maintain proper body function but in this case the oxygen deprivation is experienced throughout the body (Kavaler 1981). Heat stroke is an extremely severe diagnosis under which core temperatures exceed 40.5°C. This temperature level damages cellular structures and frequently leads to human mortality (Kovats et al. 2009).

2. Identifying impacts

a. Major heat events

In the United States, approximately 400 deaths annually can be directly attributed to heat (Basu and Samet 2002). Although those directly caused by exposure to high temperature account for a relatively small portion of total mortality, even mild exposures to hot environments can exacerbate other health problems because of the stress imposed on the entire human system (Kovats et al. 2009). This implies that although relatively few deaths are directly linked to hyperthermia, many other deaths and illnesses may be heat-related. Because there is no federal definition for heat deaths in the United States and different medical examiners interpret heat-related deaths different, it is probable that heat-related deaths are significantly underestimated (Changnon et al. 1996). A wide range of causes of death are believed to be weather-related (Kalkstein and Davis 1989), and using all-cause death totals has led to results comparable to those obtained using subdivisions such as weather-related deaths or non-accidental deaths (Sheridan and Kalkstein 2004). Thus the impact of high temperatures on human mortality may be far greater than suggested by the number of deaths labeled as directly caused by hyperthermia.

Several major heat events over the past century have been documented for their linkage with drastic elevations in mortality over relatively short time spans. These events have also served as reminders of the threat that high temperatures pose to human life and continue to motivate biometeorological research, including the present study. A major heat event can loosely be described as a prolonged period of elevated temperatures lasting from several days to several weeks. Amongst the more notable heat events in U.S.

history are the frequent heat waves of the 1930s associated with the dust bowl and Great Depression. Maximum temperatures are reported to have remained over 100°F in parts of Iowa for 53 consecutive days in 1936. Another severe heat wave in the 1980s reached into the Northern Plains states and persisted through much of the summer. Two hundred thirty excess deaths were reported over a 28-day period in St. Louis compared to the previous year's mortality including 24 heatstroke deaths on one single day. Other cities around the country reported similar spikes in the death count (Kavaler 1981).

As surveillance and communication have improved over the recent decades, so has the documentation of the impacts of severe heat events. The Chicago/Midwest heat wave of July 1995, for example, is the subject of several research articles (e.g., Semenza et al. 1996, Karl and Knight 1997, Whitman et al. 1997, Dematte et al. 1998, Palecki et al. 2001). The heat wave is believed to have directly led to 718 deaths across ten states over a four-day span (Changnon et al. 1996). The death toll of the August 2003 heat wave in Europe, however, dwarfs most of the other major events recorded to date. More than 14,000 deaths over a 20-day period are blamed on the excessive heat in France, and in Italy another 20,000 were recorded across July and August of the same summer resulting from high temperatures. Germany, the Netherlands, the United Kingdom, Portugal, and Belgium also reported significant elevations in mortality (Gosling et al. 2008; Kovats and Hajat 2008). Other heat events that have led to large increases excess deaths have been reported across much of the globe (Gosling et al. 2008).

b. Methodological approaches

The significant risk posed by exposure to high temperatures has motivated the application of a large suite of research techniques to better understand the relationship between heat stress and human mortality and morbidity. Across a wide range of methods, most researchers acknowledge that there is an optimal human thermal comfort range outside of which mortality increases with either cold or warmth (e.g., Kalkstein and Davis 1989, Martens 1998, Gosling et al. 2008). When mortality data are plotted against temperature at a given location, the resulting curve often appears in the shape of a "U," "V," or "J" (e.g., Curriero et al. 2002, Donaldson et al. 2003, Paldy et al. 2005, Laaidi et al. 2006).

Hindering the adoption of standardized methods for the evaluation of heat-mortality relationships are the many different ways in which one can definite mortality and heat. As discussed previously, there is no official definition of a heat-related death in the United States. This leads to variability in identification practices by medical examiners (Changnon et al. 1996). Further, during high temperature periods, deaths arising from a number of causes seem to increase rather than only those directly attributable to heat (Kalkstein and Davis 1989). This has led to some studies examining all-cause mortality, others examining only a selected set of deaths that are "likely" to be heat related, and others focusing solely on deaths that are specifically identified as heat-related by medical personnel. Heat risk research is also conducted with respect to morbidity via the examination of hospital admissions (e.g., Green et al. 2010) and ambulance dispatch calls (e.g., Bassil et al. 2009).

In addition to variability in results arising from differences in medical reporting procedures and selection of which subset of deaths to study, researchers have adopted different techniques to calculate the excess mortality that occurs during heat events. The expected or baseline mortality can be calculated, for example, as the average over a single month or entire year (Dessai 2002, Davis et al. 2003a, Dessai 2003, Gosling et al. 2007, cited in Gosling et al. 2008). Other studies compare each day's mortality to that observed on the same day in the previous year or average of several previous years (Conti et al. 2005, Michelozzi et al. 2006). Yet others define the baseline mortality based on that observed when temperatures fall in an optimal range (Donaldson et al. 2001, Donaldson et al. 2003). The choice of methods is largely dependent on the data available to the researcher and determines the applicability of results to other locations or time periods (Gosling et al. 2008).

The array of methods used to examine temperature-mortality relationships has been categorized into broad groups by Basu and Samet (2002) and Gosling et al. (2008) in recent reviews: descriptive and mapping epidemiological studies, approaches that use time series data and regression models, case-only or case-crossover studies, and synoptic climatological approaches.

The overall goal of the epidemiological approach is to explain patterns in mortality (or some other health outcome) using predictors including environmental variables (temperature, pollution) or demographic information such as socioeconomic status (Basu and Samet 2002, Gosling et al. 2009). In the case of environmental variables, a percentile

or threshold approach is often adopted that compares heat stress relative to typical conditions for a given location (e.g., Davis et al. 2003a). The environmental variables most often evaluated are the ambient temperature, relative humidity, and/or apparent temperature. The apparent temperature is a measure that accounts for both temperature and humidity and is commonly referred to as the Heat Index (Steadman 1979, Steadman 1984). Wind speed is occasionally included because of the role that moving air can play in modifying thermal stress. The treatment of these measures varies from one study to another, with certain authors using the daily mean, daily minimum, daily maximum, average over a multiple day period, lags of one day or more, etc. Furthermore, some authors examine the entire temperature time series throughout the year or an individual season, whereas others focus on specific time periods of interest such as heat waves. Increasingly, researchers are beginning to adopt both approaches simultaneously (Gosling et al. 2008).

One advantage of the regression approach is that confounding factors can sometimes be quantitatively accounted for by adding additional terms to the model. Commonly cited confounding (or correlated) factors include humidity, wind speed, air pollution, and demographic and social factors. Gosling et al. (2008) identified at least 40 variables that have been incorporated in recent temperature-mortality studies. Within the regression and modeling approach, some researchers have developed more complex biometeorological indices that take into account the location and time during the year at which extreme conditions are observed (Gosling et al. 2008). The Health Related Assessment of Thermal Environment (HeRATE) is one such index that includes recent conditions experienced by

a population over previous weeks (Koppe and Jendritzky 2005, Jendritzky and Tinz 2009). HeRATE builds upon the European tradition of developing biometeorological "scores" based on the human heat budget and physiological perception—previously developed variables include the Perceived Temperature (PT), Predicted Mean Vote (PMV), Physiological Equivalent Temperature (PET), and Standard Effective Temperature (SET) (Koppe et al. 2004, Jendritzky and Tinz 2009).

Another broad classification includes the case-only and case-crossover methods. These approaches typically focus on individual behavior over shorter time periods by comparing a period of interest such as a heat wave to some other control or reference period (Basu and Samet 2002, Gosling et al. 2008). Because the study centers on individuals, there is minimal need to control for confounding factors such as demographics and behavior (Basu and Samet 2002). The case-only approach was applied by Medina-Ramon et al. (2006) to investigate the characteristics of individuals dying on exceptionally warm days to those dying when temperatures were lower.

The final major methodological approach described by Gosling et al. (2008) is that of synoptic climatology, although synoptic air mass types can be integrated into any of the strategies previously discussed. With respect to identifying environmental conditions associated with elevated mortality risk synoptic climatology typically involves the identification of certain air mass types linked to oppressive conditions. The identification of homogeneous categories is often more desirable in the study of human health outcomes because people typically respond to multiple environmental variables

simultaneously (e.g., Knight et al. 2008). The recent history of synoptic climatology includes the development of the temporal synoptic index (TSI) (Kalkstein 1991) and Spatial Synoptic Classification (SSC) (Kalkstein and Greene 1997, Sheridan 2002). The current iteration of the SSC, the SSC2 (Sheridan 2002) includes a daily air mass categorization for hundreds of cities across North America and Europe. Synoptic types have been linked with daily mortality in variety of contexts (Kalkstein and Smoyer 1993, McMichael et al. 1996, Chestnut et al. 1998, Samet et al. 1998, Guest et al. 1999, Sheridan and Kalkstein 2004). Perhaps validating their usefulness, synoptic classification is a major component of heat watch/warning systems currently in place in cities across the globe (Gosling et al. 2008).

3. Spatial Variability

a. Environmental factors

The major environmental factors believed to influence heat-related risk at the local scale include temperature, air quality, and humidity. A number of researchers have compared the temperature-mortality relationship across larger groups of cities spanning entire countries or continents. As discussed previously, there is an optimal temperature range at which mortality reaches a minimum, outside of which mortality increases with elevated stress from cold or warm conditions. It is evident that the temperature of minimum mortality varies systematically from one location to another based on the acclimatization of people to their environment. In a study examining 48 cities across the United States, Kalkstein and Davis (1989) found that threshold temperatures for elevated mortality vary considerably from one city to another and that the effect is often much greater than

variation within cities arising from demographic factors. In warm locations like Phoenix, AZ, summertime mortality does not significantly increase until temperatures reach 43°C, whereas in cooler and maritime climates including San Francisco, Boston, and Pittsburgh, a significant increase occurred with temperatures near 30°C. Gosling et al. (2008) compiled the minimum mortality temperatures identified in 29 recent studies and similarly observed a logical progression whereby higher mortality thresholds (with respect to heat stress) are present in warmer climates. This theory is well corroborated (Braga et al. 2001, Braga et al. 2002, Donaldson et al. 2003, Pattenden et al. 2003, Medina-Ramon and Schwartz 2007, Jendritzky and Tinz 2009).

Far less straightforward is the manner in which urban areas modify thermal stress and mortality response. The complex structure of cities leads to large differences in temperatures between varying neighborhoods—differences which, when accumulated over the span of several days, may lead to certain areas exceeding physiological thresholds related to heat stress (Harlan et al. 2006). Elevated temperatures are commonly observed in cities when compared to surrounding areas because of the well-documented urban heat island (UHI) effect (e.g., Voogt and Oke 2003). In metropolitan areas, higher temperatures arise from a suite of processes including reduced radiant heat loss, lower wind velocities, increased exposure to radiation, anthropogenic heat sources, atmospheric contaminants, and reduced evaporation (Oke 1987, Jendritzky et al. 2000, Koppe et al. 2004). These factors combine to create environments that are potentially more stressful for occupants during high temperature events.

The magnitude of the UHI has been estimated at approximately a 10° difference between rural and urban areas for both Phoenix, AZ (Hawkins et al. 2004) and Tel Aviv (Saaroni 2000). By comparing the magnitude of the UHI for cities across the globe, Oke (1987) developed a relationship based solely on population and wind speed. An analogous relationship was later developed for Phoenix, where Brazel et al. (2007) estimate the temperature increase as 1.4 K (+/- 0.4K) per 1000 homes constructed. In similarly large cities, this effect far outweighs the temperature increases associated with predicted climate change for the next century (Jendritzky and Tinz 2009). Higher population densities have been linked with higher heat-related mortality (Medina-Ramon and Schwartz 2007), and urban heat risk may be enhanced in future decades as across the globe more people move to cities. One projection expects 60% of the world's residents to live in cities by the year 2025, up from approximately 50% presently (Koppe et al. 2004).

The spatial extent and magnitude of the UHI vary from city to city. In general, as one moves from the surrounding countryside toward the city, there is a sharp increase in temperature at the city limits, a temperature "plateau" that extends to nearly the city center, where another sharp increase can be observed. This theoretical pattern can be greatly influenced by open space, green space, and the varying designs of each individual city (Oke 1987). More specifically, higher temperatures are observed in areas with tall buildings, high building density, low fraction of green space, industrial land use, and sources of anthropogenic heat (Balling and Brazel 1987, Unger 1999, Matzarakis 2001, Stathopoulou et al. 2004). Cooler places are conversely associated with open spaces, plazas, lower building densities, and residential or commercial land use (Balling and

Brazel 1988, Saaroni 2000). The temperature pattern that results from a wide range of development and design characteristics in metropolitan areas can result in complex temperature patterns that evolve over both space and time (Nichol 1998, Kovats and Hajat 2008).

Meteorology and climate also play an important role in shaping the UHI. The strongest effects are typically observed under anticyclonic conditions when skies are clear and winds are weak (Oke 1987). Regional processes like mountain-valley and land-sea breezes can further modify the UHI (Clarke 1969, Oke 1987, Baumbach and Vogt 1999, Ichinose et al. 1999). These influences can be substantial: No UHI is observed for Kuwait City, likely due to its proximity to water and landscape patterns (Nasrallah et al. 2006). Further, the UHI commonly is strongest at night because of the substantial difference in cooling rates between urban and rural areas (Oke 1987).

The net effect of the UHI is that many city occupants are subject to higher temperatures than those living in rural areas during heat waves. The excessive heat can lead to more hot days and heat waves being identified in urban areas compared to rural ones (Tan et al. 2010). Further, differences within urban areas can lead to variations in thermal stress between proximate communities. Thus, based on temperature alone one would expect to see variability in heat-related mortality within metropolitan areas.

Metropolitan areas also influence other meteorological variables and processes. Wind speeds, for example, are generally lower within city centers, although in some cases air

flow can be enhanced when winds rush through city canyons. Under calm conditions, temperature differences between urban and rural areas can create a local-scale circulation. Urban areas have also been linked to increases in fog frequency, thunderstorms and precipitation (downwind of the city), and nighttime humidity levels (Oke 1987). Amongst these changes, elevated humidity should most directly impact thermal stress during high temperature events because of the role evaporation plays in the human heat budget.

Air quality is another environmental variable that may contribute to heat-related mortality during stressful periods, particularly in metropolitan areas that often have higher pollutant concentrations than rural areas. Determining the relative contribution of elevated temperatures and poor air quality to mortality has proved extremely difficult to date because of the high degree of collinearity between meteorological variables and pollutants (Kalkstein 1991, Sartor et al. 1995). Two of the most commonly examined pollutants that are known to impact human health are ozone and particulate matter. The formation of ozone relies on photochemical reactions that occur more readily with sunny conditions that often accompany heat waves. Particulate matter concentrations, along with ozone, can worsen with a stagnant air mass: this too is often the case when high temperatures persist for several days or more. Some studies have identified an interaction between pollutants and temperature (e.g., Shumway et al. 1988, Sartor et al. 1995, Sunyer et al. 1996, Touloumi et al. 1996, Roberts 2004, cited in Basu and Samet 2002) whereas others find little evidence for the modification of a mortality-temperature relationship by pollutants (Driscoll 1971, Kalkstein 1991, Samet et al. 1998, Green et al. 2010, cited in Basu and Samet 2002). Further complicating the study of pollutant-temperature-mortality

interactions are the infrequent measurement of small particulates in the U.S. (Braga et al. 2002), indoor air pollution to which individuals are more commonly exposed (Kilbourne 1997), and the variation of air quality within cities (Jerrett et al. 2005).

b. Demographic and social factors

Age is the most commonly cited demographic factor related to morbidity and mortality risk during heat events (e.g., Whitman et al. 1997, Danet et al. 1999, Smoyer et al. 2000, O'Neill et al. 2003, Conti et al. 2005, Medina-Ramon et al. 2006, Hajat et al. 2006, Johnson and Wilson 2009). Elderly populations are believed to be at higher risk to changes in temperature than the general population because of diminished thermoregulatory capacity and the potential interference of drugs with homeostasis mechanisms (Havenith 2001, Koppe et al. 2004).

More locally, communities with a high percentage of aged residents often show the highest mortality rates during heat events. In a 1993 heat wave in Philadelphia, most of the deaths were amongst isolated, elderly residents who did not have access to air conditioning (Mirchandani et al. 1996). For the same event, density of those aged 65 and above was identified as the strongest factor in a local-scale model of heat-related mortality (Johnson and Wilson 2009). Quantifying the elevated risk of elderly residents will continue to be important as the U.S. population continues to age (Changnon et al. 1996).

The impact of race on heat-related mortality is not as clear. Minority populations have often been linked to elevated mortality during heat events, including Native Americans

during the Philadelphia 1993 heat event (Johnson and Wilson 2009), African-Americans during the 1995 Chicago heat wave (Changnon et al. 1996), and African-American populations more generally across longer studies (O'Neill et al. 2003, Medina-Ramon 2006, Basu and Ostro 2008). Not all research has found a mortality effect based on race: no effect of race or ethnicity was found for heat-related hospital admissions in California (Green et al. 2010), and the difference between white and non-white groups was not found to be statistically significant for most U.S. cities (Kalkstein and Davis 1989).

Socioeconomic status is potentially linked to heat-related mortality because affluent residents may be more able to afford higher-quality housing and air conditioning. The high costs associated with energy required to cool homes may discourage air conditioning use amongst those with fewer financial resources (Smoyer 1998). In the U.S., economic status may be amongst the most important factors in determining heat-related risk, although this is not necessarily the case in other countries where air conditioning use is less common (Kovats and Hajat 2008). Areas with higher poverty rates were associated with statistically higher mortality rates in the 1993 Philadelphia heat event (Johnson and Wilson 2009).

A wide range of other factors have also been linked with elevated mortality rates during heat events. Women, for example, may be more at risk of dying during heat events (Stafoggia et al. 2006), although this pattern may arise because there are more elderly women than men in the population (Kavaler 1981). Other studies have found no gender effect (Basu and Ostro 2008, Green et al. 2010). Lower education levels have been associated with higher death risk in the U.S. (O'Neill et al. 2003) but this relationship

probably arises because of a linkage with wealth (Semenza et al. 1996). Those with higher levels of aerobic fitness experience less strain during high heat events (Havenith et al. 2001), whereas unhealthy individuals may have a more sedentary lifestyle with fewer exposure and acclimatization opportunities (Koppe et al. 2004). Obese individuals have more insulation near the skin and thus require increased blood flow for cooling, thereby straining the cardiovascular system (Koppe et al. 2004). Pre-existing medical conditions also elevate mortality risk during heat events, and relationships have been found with diabetes, ischemic heart disease, respiratory disease, cardiovascular disease, and chronic obstructive pulmonary disease (Medina-Ramon 2006, Gosling 2009).

Social characteristics of individuals and the community may also be related to mortality risk. Leaving windows and doors of homes open promotes ventilation; however, these measures may not be practical in places susceptible to crime (Changnon et al. 1996). Marital status has been linked to heat-related mortality as widows and widowers demonstrate a higher risk (Stafoggia et al. 1996), a characteristic that is likely closely linked to other findings related to isolated persons in general (e.g., Mirchandani et al. 1996).

Certain demographic and social factors can interact to create communities within metropolitan areas that demonstrate especially elevated risk during heat events. In Phoenix, for example, affluent whites tend to live in environments with more green space, and thus lower temperatures, whereas lower income Latinos live in places with higher building densities and more impermeable surfaces (Harlan et al. 2008). Thus the Latino community in Phoenix is subject to higher risk arising from both lower income

(resulting in less access to air conditioning) and higher temperatures in general. Places where these types of interactions may especially benefit from the adaptation of community-based warning systems and intervention measures. These relationships also shed light on potential inequities in health outcomes related to climate change (O'Neill et al. 2003).

c. Place-based studies

Spatial variability in the relationship between climate and heat risk has been well-studied at the global and region scale, especially for metropolitan areas. There is evidence for acclimatization across the United States, whereby populations in different cities experience different heat "thresholds" (Davis et al. 2003b). In the southeastern states, for example, mortality is elevated at much higher temperatures than in places where extreme heat events are relatively rare, like the Pacific Northwest. The same pattern can be observed in other cities across the globe (Gosling et al. 2008). Such knowledge of the relationship between the thermal environment and heat risk, however, does not yet extend to the sub-city scale where a range of thermal stresses exists (Oke 1987). Airport temperature measurements are commonly used for large-scale heat-mortality research, but these data may not accurately reflect the metropolitan exposure to thermal stress (de'Donato et al. 2007, Zauli Sajani et al 2008). Remote sensing techniques make possible a better representation of the varying thermal conditions people experience (Stathopoulou et al. 2005, Voogt and Oke 2003), yet despite this advance, no study has comprehensively examined the impact of local-scale climatic variations on heat-related mortality (de'Donato et al. 2008).

The spatial distribution of heat-related mortality within metropolitan areas has been examined in a limited number of studies over the past several decades, dating back to maps of deaths in St. Louis during a heat wave in 1966 by Henschel et al. (1969) (Smoyer 1998). This gap in the research was later highlighted by Smoyer (1998), who was the first to advocate for a geographically oriented analysis of extreme weather impacts, especially heat-related mortality. Her research focused on a spatial analysis of heat-related deaths for the St. Louis, Missouri metropolitan statistical area (MSA) during the summers of 1980, 1983, 1988, and 1995, each of which were characterized by high temperatures and at least one heat wave (defined as a period of at least two consecutive days with apparent temperatures above 40.6°C). The emphasis of her study was to identify characteristics of places, rather than people, which were associated with high risk. In more traditional population-based studies, she believed that the risk of committing "ecological fallacy"—drawing inferences about individuals based on aggregate—was high, a problem avoided by focusing on the spatial distribution of various risk factors. Explanatory variables were derived from census data to estimate thermal stress, economic resources, and isolation of residents. The results suggested that heat-related mortality is higher in portions of the city with less neighborhood stability and fewer economic resources.

Recent progress has been made in the spatial analysis of heat risk at the local scale via the examination of ambulance call data for Toronto, Canada (Dolney and Sheridan 2006, Bassil et al. 2009) and county-scale mortality data for the state of Ohio (Sheridan and Dolney 2003). The Canada studies revealed that areas within the downtown core showed

a high rate of heat-related dispatch calls but acknowledged that the reasons for the spatial pattern were unclear. Another area of elevated dispatch calls came from the Lake Ontario shoreline, where the researchers speculated that heat-stressed individuals seek relief from the weather (Dolney and Sheridan 2006, Bassil et al. 2009). In the Ohio study, the authors found that the urban population was at no greater risk than those who lived in suburban or urban areas. The impacts of their result are unclear, as it is well known that urban areas are warmer than their surroundings and thus should exhibit elevated heat stress during extremely warm periods (Sheridan and Dolney 2003). One of the important contributions of both projects was the visualization of local-scale heat-related risk, a product not widely available in the literature.

The most current and comprehensive effort to assess heat vulnerability was completed by Reid et al. (2009) who created a national map of estimated morbidity and mortality risk at the census-tract scale. Heat vulnerability was estimated from several risk factors including demographics, land cover, air conditioning use, and diabetes prevalence. No response variable was incorporated. The most vulnerable areas were shown to include the northeast and Pacific coast as well as inner city areas nationwide. Amongst the most significant findings was the implied variability of risk within metropolitan areas with generally higher risk in the downtown core. The authors note the importance of identifying elevated risk zones within cities for proper targeting of heat intervention strategies. The research proposed here will follow a more rigorous strategy to identify such zones by analyzing a larger set of risk factors, including critical measurements of atmospheric conditions, within a framework focused on health outcomes.

Two studies published within the past several years are the first to integrate satellite assessment of urban temperature variability with demographic and social information in a model built on mortality data. A Canadian team evaluated warm season daily mortality for Montreal over a 14-year period using Landsat imagery, ozone data, and property assessments, and found evidence that mortality risk is higher in places with higher surface temperatures (Smargiassi et al. 2009). Their dataset included some 51,000 summertime deaths over the period of record. Many of the authors' analyses point to a significant increase in risk for warmer postal codes; however, the variance in the results is not fully detailed. Only two thermal images were used to determine the local-scale variability in heat risk, and postal codes were classified into either a "warm" or "cold" group based on the satellite imagery. The treatment of demographic factors and air quality appear similarly broad. The study signifies research movement toward a comprehensive assessment of local-scale heat-related mortality, but the methodology applied leaves a high level of uncertainty, some of which is acknowledged by the authors.

The United States research team led by Daniel Johnson of IUPUI has produced multiple publications in the past several years focused on the Philadelphia heat wave of 1993. Their most recent contribution (Johnson et al. 2009), again examines this event but includes satellite imagery to add local-scale temperature variability to their model of mortality risk. The authors found that adding a land surface temperature estimate to their binary logistic regression model (which already included socioeconomic indicators) improved overall performance. More broadly, they concluded that thermal remote sensing imagery could be a key tool in examining heat-related risk. The universality of

their findings needs to be validated as the model is based on one single heat event with 118 attributed deaths. Furthermore, the adoption of a binary logistic regression may have been more appropriate in the Johnson et al. study because of the binary outcome variable (census tracts either had a mortality event or did not) than in other studies examining mortality risk over longer time periods where a range of mortality rates are present across metropolitan communities.

4. Temporal Variability

a. Heat-related mortality trends

Global mean temperature increased by 0.74°C between 1906 and 2005, with some of the hottest years on record observed within the most recent two decades (Trenberth et al. 2007). Although considerable debate continues regarding future climate change and underlying causal mechanisms, the increase in temperatures has led to more frequent exposure to thermally stressful conditions in some locations (e.g., Robinson 2001, McGregor et al. 2002, Yan et al. 2002). Combating the potential negative impacts of global climate change are rapid advancements in technology that create more comfortable environments for humans to live. In general, places with less prevalent air conditioning experience higher heat-related mortality (O'Neill et al. 2005, Medina-Ramon and Schwartz 2007). Across several United States cities, for example, every 1% increase in air conditioning has been shown to reduce mortality by over 1 death per year per million residents (Davis et al. 2003a). Numerous studies identify a decline in heat-related mortality across U.S. metropolitan areas across the past several decades (e.g., Marmor 1975, Davis et al. 2003a, Davis et al. 2003b, Barnett 2007, Sheridan et al. 2009).

Air conditioning is the likely cause of declining heat-related and cardiovascular mortality in the United States over the past several decades (Davis et al. 2003a, Barnett 2007), but it is unclear how mortality may change in a potentially warmer climate in the future as more cities approach air conditioning saturation. The benefits of air conditioning use also vary between individuals for socioeconomic reasons, and different responses are possible with the stress that air conditioners place on power grids and the heat they release to the environment (Klinenberg 2002, Gosling et al. 2008, Hess et al. 2009). Frequent exposure to air-conditioned living spaces may also inhibit one's ability to acclimatize to generally higher temperatures (Hajat et al. 2010).

b. Short-term factors: duration, lags, displacement

There are several factors operating on the scale of days to weeks that need to be considered when evaluating heat-related mortality. Heat event duration, lag in response, and mortality displacement are commonly cited variables that contribute to the overall response of both individuals and populations.

When extremely high temperatures are observed over several successive days, it is believed that there is a cumulative stress placed on the human thermoregulatory system (e.g., Semenza et al. 1996, Braga et al. 2001). This has inspired the inclusion of heat event duration as a potential predictor of mortality in many studies, a large number of which have found that prolonged heat events are often associated with the highest increases in mortality and that heat event duration is a variable that explains much of the

variance in mortality time series (e.g., Kalkstein 1991, Smoyer 1998, Huynen et al. 2001, Hajat et al. 2002, Qiu et al. 2002, Rocklov 2007, Tan et al. 2007).

Another important factor to consider is the potential time delay between high temperatures and an increase in mortality. Individuals in a population respond to thermal stress differently, and many may not suffer immediately when hot conditions develop. Accordingly, a lag is noticeable in many time series between the onset of extreme conditions and elevated mortality incidence. Gosling et al. (2008) review 24 different studies over roughly the past decade that have examined lag effects with respect to temperature-mortality relationships. Nearly all studies report a lag effect of three days or fewer, although the temporal component of mortality response may be different between hot and cold spells (Braga et al. 2001). A one-day lag has been shown to produce the most consistent response in the United States (Davis et al. 2003a).

While mortality increases are common in the few days immediately following extremely high temperatures, a decrease is often subsequently observed. This pattern has been attributed to a theory of "mortality displacement" whereby many of the individuals that die under extreme conditions were those already in poor health who would have died in the near future regardless of the heat event; the extreme conditions simply accelerated their death forward by a few days or weeks (Gosling et al. 2008). Following a heat event, mortality is lower than would typically be expected because those individuals that would have normally died during that time period were lost during the event. Mortality displacement has been observed as soon as three days after a heat events (e.g., Kunst et

al. 1993, Hajat et al. 2002, Pattenden et al. 2003, Hajat et al. 2005, 2006), and at longer time periods reaching 30 days or more (e.g., Sartor et al. 1995, Huynen et al. 2001, Le Tertre et al. 2006, Gosling et al. 2007, Gosling et al. 2008). Mortality deficits following heat events have been estimated between 15 and 70% over a suite of studies (Gosling et al. 2008), and recent work demonstrates that the magnitude of displacement effects varies geographically and between heat events of differing severity (Saha et al. 2013).

c. Seasonality

In addition to long-term trends over years and decades, in many locations, heat-related mortality varies based on the timing of individual heat events (e.g., Kalkstein and Davis 1989; Tan et al. 2009). It is often thought that early-season warm spells can have a greater impact on mortality than those that occur later in the season (e.g., Paldy 2005). For example, in London, warm-season mortality increases during extreme heat events averages 3.26%, but when heat spells strike in the cold season, the mortality increase is over 5% (Hajat 2002). These patterns may be related to mortality displacement as discussed above, short-term acclimatization, which can take place over the period of a few weeks (e.g., Koppe et al. 2004, Koppe and Jendrizky 2005, Hajat et al. 2010), or a combination thereof. Short-term acclimatization generally results in reduced core temperatures, earlier onset of sweating, increased blood flow to the skin, and lowered metabolism, which have been observed to take place in as little as 3-4 days after exposure to a new thermal environment (Koppe et al. 2004). Furthermore there is considerable seasonality with respect to mortality across the entire year (Davis et al. 2004) with higher

death rates in winter than in summer in many locations. No research describing mortality seasonality variability within communities was found.

d. Heat and climate change

One of the major environmental questions of our time is 'how will future global climate change affect human health and well-being?' Across continents, high uncertainty exists regarding the potential role that changing temperature and precipitation regimes may play related to the spread of disease, availability of food and other resources, and risk posed by natural hazards including floods, droughts, and heat waves. Some research suggests that the future climate will be one with heat waves that are more intense, more persistent, and more frequent (Meehl and Tebaldi 2004). The Intergovernmental Panel on Climate Change predicts that higher maximum temperatures, more hot days, higher minimum temperatures, and increases in heat index over land areas are all very likely in the 21st century (Trenberth et al. 2007). If this is to be the case, the effects on human mortality may be concentrated in localities where the temperature is higher, including metropolitan areas because of the urban heat island effect. One projection for New York City, NY, estimates a 70% increase in heat-related deaths over the next five decades from 1990s levels (Knowlton et al. 2007).

Estimates of future heat-related mortality widely vary. Many countries expect future increases because of rising temperatures, although because of varying practices in identifying heat-related deaths, it is not entirely certain how substantial the increase will be (Gosling et al. 2008). For example, in Finland is it believed that heat-related deaths comprise less than one-half of one percent of all annual deaths, and therefore despite a

projected 2°C warming, a recent assessment did not expect substantial mortality increase (Hassi and Rytkonen 2005). It is possible that an increase in heat-related deaths could be offset by a reduction in cold-related deaths: Davis et al. (2004) suggest a net annual decline of 2.65 deaths per million U.S. citizens with a seasonally-uniform 1°C increase in temperatures because of a greater reduction in cold-related mortality than increase in heat-related mortality. A winter-dominant warming, which is more in agreement with past observations and future projections, was anticipated to lead to 8.92 fewer deaths per million for the same reason (Davis et al. 2004). One global-scale assessment of thermal comfort showed that the net effect of a commonly-utilized climate change model will lead to an overall decrease in comfort in the highly-populated middle latitudes and subtropics; increases in comfort were limited to sparsely populated regions (Jendritzky and Tinz 2009).

Gosling et al. (2008) make a special point to acknowledge the considerable uncertainty in any projection of future heat-related mortality. More specifically, they identify five major areas of uncertainty that contribute to the challenge of forecasting such deaths in the future: emissions, processes and parameters within climate models, downscaling, temperature-mortality modeling, and population change and adaptation. Uncertainty related to future emissions has led researchers to adopt a variety of scenarios in climate change models which themselves are limited by the accuracy of parameters based on processes that are not perfectly understood. These models are run at a coarse resolution for the entire globe, thereby making it difficult to identify impacts at any one place, such as a highly-populated city. There is next a suite of temperature-mortality models

proposed in the literature, many of which focus on specific places or regions, all of which have some degree of error or inaccuracy. Finally, these mortality models are projected onto a population that may be of a different size in the future and be more or less adapted to climatic change. Gosling et al. (2008) advocate for the examination of where errors are most compounded throughout the process of projecting heat-related mortality to better understand where future research should be targeted.

5. Response Strategies

a. Intervention and mitigation

Humans are able to mitigate the impact of high-temperature events via natural mechanisms and technological advances. The human body has a physiological capacity to adapt to changes in environmental conditions, and sensitivity to heat exposure can be adjusted over both short-term periods such as 3-12 days (Koppe et al 2004, Hajat 2010), a few weeks (Koppe and Jendrisky 2005), or longer-term acclimatization periods on the order of several years (Frisancho 1991). This acclimatization is clearly evident in large-scale analyses of spatial variability in heat tolerance: places with higher overall temperatures have a higher mortality threshold temperature (e.g., Kalkstein and Davis 1989, Eurowinter Group 1997, Pattenden et al. 2003, Jendritzky and Tinz 2009). Despite this natural coping mechanism, in many cases, people choose to live and work in environments in which they are not physiologically prepared.

The impact of air conditioning on lessening heat-related mortality and morbidity has been substantial in many developed counties. A statistically significant inverse relationship

between air conditioning use/prevalence and heat mortality has been documented on multiple occasions (Braga et al. 2001, Davis et al. 2003, O'Neill et al. 2005; Medina-Ramón and Schwartz 2007). However, air conditioning has not mitigated risk equally for all stakeholders. Energy costs associated with air conditioning use may discourage or prevent use by those with fewer financial resources (Smoyer 1998). This problem can be exacerbated by the heat that air conditioning units release to the environment (Gosling 2009), and the overall stress that air conditioning systems place on the entire power grid (Hess et al. 2009).

When the threat of a dangerous heat event is identified in advance, individual citizens and larger community response groups can adopt a series of practices that reduce the risk of heat-related fatalities. Often the most needed efforts are communicating the risk to the elderly and isolated who may be unaware of the expected heat and can greatly benefit from relocation to a cooling center. Other action steps typically enacted once a dangerous heat event is forecast include the promotion of a "buddy system" that encourages neighbors to check on one another regularly, the activation of dedicated phone "heat-lines," home visits by city officials, high alert status at nursing care facilities, the cancellation of utility suspensions, increased staffing at hospitals, and outreach to the homeless (Koppe et al. 2004). A reduction in heat-related mortality in Milwaukee, Wisconsin in recent years has been attributed to effective public health preparedness rather than differences in heat levels (Weisskopf et al. 2002).

b. Warning systems

More recently, heat warning systems have been developed that trigger public alerts in advance of potentially stressful conditions. Much of the research in this field has been conducted by Laurence Kalkstein et al., who have developed warning systems for a suite of cities across the globe using synoptic air mass types. In general, these systems provide guidance for federal agencies rather than directly warning the public. They act as decision-support tools for weather forecasters and emergency managers to issue alerts and enact certain protocols. Specialized forecasts are issued by the warning system when the next day's weather is predicted to be an oppressive air mass type (typically associated with hot and dry or hot and moist conditions, depending on the location). When an oppressive air mass is predicted for the next day, a more specific forecast of the potential for elevated mortality is made (Kalkstein et al. 1996, Sheridan and Kalkstein 2004). Some locations that have not adopted air-mass-based warning systems use a temperature or combined temperature-humidity threshold for enacting emergency measures (Koppe et al. 2004).

Heat warning systems are believed to save lives. In the case of the Philadelphia Heat-Watch-Warning System (the original synoptic warning system developed by Kalkstein et al. in 1995), an estimated 117 lives have been saved since its creation. Using a risk measure called the value of statistical life, the financial benefit to Philadelphia for using the system can be estimated at over $400 million, while the costs for the system's operation were approximately $200,000 over its first three years (Ebi et al. 2004).

c. Potential for improvement

Advances in technology, surveillance, and reporting have made it possible to accurately assess the local-scale geographic variability in heat-related mortality (Ruddell et al. 2010). The biometeorological community has made significant progress in recent decades by improving our knowledge of the human-heat sensitivity, adaptation and acclimatization, and the development of heat warning systems customized for individual cities. There is a significant opportunity to continue working toward the welfare of those at risk by integrating knowledge of the microscale environmental conditions, characteristics of subpopulations, and heat-related mortality records. There are considerable variations in the urban microclimate that can have a major effect on biological outcomes, and these differences are not well-accounted for in current assessments of mortality risk (Kilbourne 1997, Ruddell et al. 2010). The efforts of Smargiassi et al. (2009) and Johnson et al. (2009) highlight the opportunity to combine environmental, demographic, and social characteristics of communities to identify places that can benefit from targeted response strategies. Such research can lead to the promotion of health community characteristics and more efficient allocation of resources (Smoyer 1998).

6. Motivation for study

At present, biometeorologists have developed a good understanding of the relationship between high temperatures and human physiology, the extent to which the built environment modifies thermal conditions, and have identified certain populations that demonstrate an increased risk to heat stress. Little or no work, however, integrates all of these factors into models that identify the most heat-sensitive places and explain why they are more susceptible. A cross-city comparison of the spatial response in mortality associated with elevated heat stress would lay the foundation for understanding the role that the built environment, microclimate, demographics, and land cover play in shaping the mortality patterns. Many heat-related deaths are preventable (Luber et al. 2006, Smoyer et al. 2000)—one of the practical results that could arise from this research would be identification of where to direct prevention efforts and understand why those places and populations are more sensitive to the impact of heat waves.

Environmental, demographic, and social factors vary across the metropolitan landscape and interactively shape the risk profile of individual communities. Quantitative assessment of this interaction is possible with the availability of high-resolution imaging of the urban heat island and long-term local-scale mortality records, but research to date has only suggested the possibility of such a project. Successful approaches have been adopted to measure the urban heat island and compare population characteristics to mortality, and evaluate trends and seasonality in mortality for individual cities. Largely missing from the literature is the application of these approaches at a local scale. By adopting such a framework, research into the geographic dimensions of heat-related

mortality within large U.S. cities can advance understanding of the factors that contribute to the location and timing of heat-related deaths.

7. Dissertation Structure

The subsequent four chapters of this dissertation are each based on scientific manuscripts intended for, or already appearing in, the peer-reviewed literature. Each chapter is intended to be a standalone product, and thus contains its own abstract, main text, figures, and references. Because these chapters were prepared for publication in different journals, there are also substantial formatting differences.

Chapter 2 represents a proof-of-concept analysis focusing on Philadelphia, Pennsylvania. In this manuscript, spatial variability in the heat-related mortality rate is explored across the 48 postal codes that comprise Philadelphia County. Subsequently, a number of potential explanatory variables are linked to the spatial pattern in mortality using principal components regression. The set of explanatory variables includes social and demographic indicators obtained from census data, zoning and land use information, and remote measurements of surface temperature within the city. This chapter was published in *Environmental Health* in 2012.

Hondula, D. M., Davis, R. E., Leisten, M. J., Saha, M. V., Veazey, L. M., & Wegner, C. R. (2012). Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis. *Environmental Health*, *11*(1), 1-11. doi: 10.1186/1476-069X-11-16

In Chapter 3, a new method is introduced for quantifying heat-mortality relationships for spatial units within cities where daily mortality counts are too low to permit the application of other approaches. The method takes advantage of advances in the environmental health field over the past ten or more years that include modeling relationships between mortality to temperature and time with non-linear, semi-parametric smoothing splines. This technique offers the advantage of accounting for interannual differences in the seasonal pattern in mortality. The method we present is applied to mortality data for Philadelphia. This chapter was published in the *Journal of Epidemiology and Community Health* in 2013.

Hondula, D. M., Davis, R. E., Rocklöv, J., & Saha, M. V. (2013). A time series approach for evaluating intra-city heat-related mortality. *Journal of Epidemiology and Community Health*. doi:10.1136/jech-2012-202157

Chapter 4, emphasizes an inter-city comparison of the factors associated with spatial variability in heat-related mortality. Using guidance from earlier work (Chapter 2), a set of explanatory variables that includes social, environmental, and demographic factors is prepared for each city at the postal code scale. Multiple regression models are then constructed to determine those factors that are most closely associated with the mortality patterns. Determinants of high-risk spatial zones vary from one city to another, encouraging the development of city-specific, empirically-driven risk assessments for other locations where heat is associated with negative health impacts. Portions of this chapter will be submitted to *Environmental Health Perspectives*.

Hondula, D. M., Davis, R. E., (2013) (Draft). Geographic dimensions of heat-related mortality in seven U.S. Cities. *Environmental Health Perspectives.*

Temporal variability in intra-city mortality is the focus of the final manuscript, Chapter 5. We use the statistical method developed in Chapter 3 to identify locations within each of the study cities where mortality rates on hot summer days are significantly different than baseline mortality. This procedure is repeated in an iterative fashion, withholding one year of data from the model and then comparing the high-mortality zones from the training data versus those observed in the withheld year. Interannual consistency in intra-city mortality can increase confidence that places targeted for intervention strategies based on historical data will be those where the risk is highest in the future. This chapter will be submitted to the journal *Health and Place*.

Hondula, D. M., Davis, R. E., (2013) (Draft). The predictability of high-risk zones for heat-related mortality in seven U.S. Cities. *Health & Place.*

Chapter 6 concludes the dissertation with an overview of the major findings from each previous section, discussion of scientific contributions, and presents opportunities for future research.

References

Balling Jr, R. C. & Brazel, S. W., 1988. High-resolution surface temperature patterns in a complex urban terrain. *PHOTOGRAM. ENG. REMOTE SENS.*, 54(9), p.1289–1293.

Barnett, A.G., 2007. Temperature and cardiovascular deaths in the US elderly: changes over time. *Epidemiology (Cambridge, Mass.)*, 18(3), pp.369-372.

Bassil, K.L. et al., 2009. Temporal and spatial variation of heat-related illness using 911 medical dispatch data. *Environmental research*, 109(5), p.600–606.

Basu, R. & Samet, J.M., 2002. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic reviews*, 24(2), p.190.

Basu, Rupa & Ostro, B.D., 2008. A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *American Journal of Epidemiology*, 168(6), pp.632-637.

Baumbach, G. & Vogt, U., 1999. Experimental determination of the effect of mountain-valley breeze circulation on air pollution in the vicinity of Freiburg. *Atmospheric Environment*, 33(24-25), p.4019–4027.

Braga, A. L.F, Zanobetti, A. & Schwartz, J., 2001. The time course of weather-related deaths. *Epidemiology*, 12(6), p.662–667.

Braga, Alfésio L F, Zanobetti, Antonella & Schwartz, Joel, 2002. The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environmental Health Perspectives*, 110(9), pp.859-863.

Brazel, A. et al., 2007. Determinants of changes in the regional urban heat island in metropolitan Phoenix (Arizona, USA) between 1990 and 2004. *Climate Research*, 33(2), p.171.

Changnon, S.A., Kunkel, K.E. & Reinke, B.C., 1996. Impacts and Responses to the 1995 Heat Wave: A Call to Action. *Bulletin of the American Meteorological Society*, 77, p.1497–1506.

Chestnut, L.G., others, 1998. Analysis of differences in hot-weather-related mortality across 44 US metropolitan areas. *Environmental Science & Policy*, 1(1), p.59–70.

Clarke, J.F., 1969. Nocturnal urban boundary layer over Cincinnati, Ohio. *Monthly Weather Review*, 97(8), p.582–589.

Conti, S. et al., 2005. Epidemiologic study of mortality during the summer 2003 heat wave in Italy. *Environmental Research*, 98(3), p.390–399.

Curriero, F.C. et al., 2002. Temperature and mortality in 11 cities of the eastern United States. *American Journal of Epidemiology*, 155(1), p.80.

Danet, S. et al., 1999. Unhealthy effects of atmospheric temperature and pressure on the occurrence of myocardial infarction and coronary deaths: a 10-year survey: the Lille-World Health Organization MONICA project (Monitoring Trends and Determinants in Cardiovascular Disease). *Circulation*, 100(1), p.e1.

Davis, R.E. et al., 2004. Seasonality of climate—human mortality relationships in US cities and impacts of climate change. *Climate Research* 26, p61–76.

Davis, R.E. et al., 2003a. Changing heat-related mortality in the United States. *Environmental Health Perspectives*, 111(14), p.1712–1719.

Davis, R.E. et al., 2003b. Decadal changes in summer mortality in U.S. cities. *International Journal of Biometeorology* 47, p166-175.

de'Donato, F.K. et al., 2007. Airport and city-centre temperatures in the evaluation of the association between heat and mortality. *International Journal of Biometeorology*, 52(4), pp.301-310.

Dematte, J.E. et al., 1998. Near-fatal heat stroke during the 1995 heat wave in Chicago. *Annals of Internal Medicine*, 129(3), p.173.

Dessai, S., 2003. Heat stress and mortality in Lisbon Part II. An assessment of the potential impacts of climate change. *International Journal of Biometeorology*, 48(1), p.37–44.

Dessai, S., 2002. Heat stress and mortality in Lisbon Part I. model construction and validation. *International Journal of Biometeorology*, 47(1), pp.6-12.

Dolney, T.J. & Sheridan, S.C., 2006. The relationship between extreme heat and ambulance response calls for the city of Toronto, Ontario, Canada. *Environmental Research*, 101(1), p.94–103.

Donaldson, G.C., Keatinge, W.R. & Nayha, S., 2003. Changes in summer temperature and heat-related mortality since 1971 in North Carolina, South Finland, and Southeast England. *Environmental research*, 91(1), p.1–7.

Donaldson, G.C. et al., 2001. Heat-and cold-related mortality and morbidity and climate change. *Health effects of climate change in the UK*, p.70–80.

Driscoll, D.M., 1971. The relationship between weather and mortality in ten major metropolitan areas in the United States, 1962–1965. *International Journal of Biometeorology*, 15(1), p.23–39.

Ebi, K.L. et al., 2004. Heat watch/warning systems save lives–Estimated costs and benefits for Philadelphia 1995–98. *Bulletin of the American Meteorological Society*, 85(8), p.1067–1073.

Eurowinter Group, 1997. Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. *Lancet* 349, p. 1341–1346.

Frisancho, A.R., 1993. *Human adaptation and accommodation*, Univ of Michigan Pr.

Gosling, S. N, McGregor, G. R & Paldy, A., 2007. Climate change and heat-related mortality in six cities Part 1: model construction and validation. *International Journal of Biometeorology*, 51(6), p.525–540.

Gosling, Simon N. et al., 2008. Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change*, 92(3-4), pp.299-341.

Green, R. et al., 2010. The effect of temperature on hospital admissions in nine California counties. *International Journal of Public Health*, 55(2), pp.113-121.

Guest, C.S. et al., 1999. Climate and mortality in Australia: retrospective study, 1979-1990, and predicted impacts in five major cities in 2030. *Climate Research*, 13, p.1–15.

Hajat, S. et al., 2005. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology*, 16(5), p.613.

Hajat, S. et al., 2006. Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology*, 17(6), p.632.

Hajat, S. et al., 2002. Impact of hot temperatures on death in London: a time series approach. *Journal of Epidemiology and Community Health*, 56(5), p.367.

Hajat, Shakoor, O'Connor, M. & Kosatsky, Tom, 2010. Health effects of hot weather: from awareness of risk factors to effective health protection. *The Lancet*, 375(9717), pp.856-863.

Harlan, S. L et al., 2008. In the shade of affluence: The inequitable distribution of the urban heat island. *Research in social problems and public policy*, 15, p.173–202.

Harlan, S. L et al., 2006. Neighborhood microclimates and vulnerability to heat stress. *Social Science & Medicine*, 63(11), p.2847–2863.

Hassi, J. & Rytk\önen, M., 2005. *Climate warming and health adaptation in Finland*, Finnish Environment Institute.

Havenith, G., 2001. Individualized model of human thermoregulation for the simulation of heat stress response. *Journal of Applied Physiology*, 90(5), p.1943.

Hawkins, T.W. et al., 2004. The role of rural variability in urban heat island determination for Phoenix, Arizona. *Journal of Applied Meteorology*, 43(3).

Henschel, A. et al., 1969. An analysis of the heat deaths in St. Louis during July, 1966. *American Journal of Public Health*, 59(12), p.2232.

Hess, J.J. et al., 2009. Climate change and emergency medicine: impacts and opportunities. *Academic Emergency Medicine*, 16(8), p.782–794.

Huynen, M.M. et al., 2001. The impact of heat waves and cold spells on mortality rates in the Dutch population. *Environmental health perspectives*, 109(5), p.463.

Ichinose, T., Shimodozono, K. & Hanaki, K., 1999. Impact of anthropogenic heat on urban climate in Tokyo. *Atmospheric Environment*, 33(24-25), pp.3897-3909.

Jendritzky, G., Grätz, A. & Friedrich, M., 2000. The assessment of human thermal climates in cities. In *Biometeorology and Urban Climatology at the Turn of the Millenium (ed. by RJ de Dear, JD Kalma, TR Oke and A. Auliciems): Selected Papers from the Conference ICB-ICUC*. p. 65–69.

Jendritzky, G. & Tinz, B., 2009. The thermal environment of the human being on the global scale. *Global Health Action* 2: 10.3402/gha.v2i0.2005.

Jerrett, M. et al., 2005. Spatial Analysis of Air Pollution and Mortality in Los Angeles. *Epidemiology*, 16(6), pp.727-736.

Johnson, D.P. & Wilson, J.S., 2009. The socio-spatial dynamics of extreme urban heat events: The case of heat-related deaths in Philadelphia. *Applied Geography*, 29(3), p.419–434.

Johnson, D.P., Wilson, J.S. & Luber, G.C., 2009. Socioeconomic indicators of heat-related health risk supplemented with remotely sensed data. *International journal of health geographics*, 8(1), p.57.

Kalkstein, L. S, 1991. A new approach to evaluate the impact of climate on human mortality. *Environmental Health Perspectives*, 96, p.145.

Kalkstein, L. S & Davis, R.E., 1989. Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Annals of the Association of American Geographers*, 79(1), p.44–64.

Kalkstein, L. S & Greene, J. S, 1997. An evaluation of climate/mortality relationships in large US cities and the possible impacts of a climate change. *Environmental health perspectives*, 105(1), p.84.

Kalkstein, L. S et al., 1996. The Philadelphia hot weather-health watch/warning system: development and application, summer 1995. *Bulletin of the American Meteorological Society*, 77(7), p.1519–1528.

Kalkstein, L. S. & Smoyer, K. E., 1993. The impact of climate change on human health: some international implications. *Cellular and Molecular Life Sciences*, 49(11), p.969–979.

Karl, T.R. & Knight, R.W., 1997. The 1995 Chicago heat wave: How likely is a recurrence? *Bulletin of the American Meteorological Society*, 78(6), p.1107–1119.

Kavaler, L., 1981. *A matter of degree: heat, life, and death*, Harpercollins Childrens Books.

Kilbourne, E.M., 1997. Heat waves and hot environments. *The public health consequences of disasters*, p.245–269.

Klinenberg, E., 2002. *A social autopsy of disaster in Chicago: Heat wave*, Chicago: The University of Chicago Press.

Knight, D.B. et al., others, 2008. Increasing frequencies of warm and humid air masses over the conterminous United States from 1948 to 2005. *Geophysical Research Letters*, 35(10), p.L10702.

Knowlton, K., Lynn, B., Goldberg, R. A., Rosenzweig, C., Hogrefe, C., Rosenthal, J. K., & Kinney, P. L. (2007). Projecting heat-related mortality impacts under a changing climate in the New York City region. *Journal Information*, *97*(11).

Koppe, C. et al., 2004. *Heat waves: risks and responses*, Regional Office for Europe, World Health Organization.

Koppe, Christina & Jendritzky, Gerd, 2005. Inclusion of short-term adaptation to thermal stresses in a heat load warning procedure. *Meteorologische Zeitschrift*, 14(2), pp.271-278.

Kovats, R. S & Hajat, S., 2008. Heat stress and public health: a critical review. *Anunual Reviw of Public Health*, 29, pp.41-55.

Kovats, S. et al., others, 2009. Heat-waves: risks and responses. Published by World Health Organization.

Kunst, A.E., Looman, C.W.N. & Mackenbach, J.P., 1993. Outdoor air temperature and mortality in the Netherlands: a time-series analysis. *American journal of epidemiology*, 137(3), p.331.

Laaidi, M., Laaidi, K. & Besancenot, J.P., 2006. Temperature-related mortality in France, a comparison between regions with different climates from the perspective of global warming. *International Journal of Biometeorology*, 51(2), p.145–153.

Le Tertre, A. et al., others, 2006. Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*, 17(1), p.75.

Luber, G.E., Sanchez, C.A. & Conklin, L., 2006. Heat-related deaths—United States, 1999–2003. *MMWR Morb Mortal Wkly Rep*, 55(29), p.796–798.

Marmor, M., 1975. Heat wave mortality in New York City, 1949 to 1970. *Archives of environmental health*, 30(3), p.130.

Martens, W.J.M., 1998. Climate change, thermal stress and mortality changes. *Social Science & Medicine*, 46(3), p.331–344.

Matzarakis, A., 2001. Climate and bioclimate information for tourism in Greece. *on Climate, Tourism and Recreation.*, p.171.

McGregor, G. R. et al., 2002. An evaluation of the nature and timing of summer human thermal discomfort in Athens, Greece. *Climate Research*, 20(1), p.83–94.

McMichael, A.J., Haines, A. & Slooff, R., 1996. Climate change and human health: an assessment prepared by a Task Group on behalf of the World Health Organization, the World Meteorological Organization and the United Nations Environment Programme.

Medina-Ramón, M & Schwartz, J, 2007. Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occupational and Environmental Medicine*, 64(12), pp.827-833. Available at: [Accessed May 2, 2010].

Medina-Ramón, Mercedes et al., 2006. Extreme Temperatures and Mortality: Assessing Effect Modification by Personal Characteristics and Specific Cause of Death in a Multi-City Case-Only Analysis. *Environmental Health Perspectives*, 114(9), pp.1331-1336.

Meehl, G.A. & Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, 305(5686), p.994.

Michelozzi, P. et al., 2006. Temperature and summer mortality: geographical and temporal variations in four Italian cities. *Journal of epidemiology and community health*, 60(5), p.417.

Mirchandani, H.G. et al., 1996. Heat-related deaths in Philadelphia-1993. *The American Journal of Forensic Medicine and Pathology*, 17(2), p.106.

Nasrallah, H.A., Brazel, A.J. & Balling Jr, R. C, 2006. Analysis of the Kuwait City urban heat island. *International journal of climatology*, 10(4), p.401–405.

Nichol, J. E., 1998. Visualisation of urban surface temperatures derived from satellite images. *International Journal of Remote Sensing*, 19(9), pp.1639-1649.

O'Neill, M. S., Zanobetti, A., & Schwartz, J. (2003). Modifiers of the temperature and mortality association in seven US cities. *American Journal of Epidemiology*, *157*(12), 1074-1082.

O'Neill, M. S., Hajat, S., Zanobetti, A., Ramirez-Aguilar, M., & Schwartz, J. (2005). Impact of control for air pollution and respiratory epidemics on the estimated associations of temperature and daily mortality. *International Journal of Biometeorology*, *50*(2), 121-129.

Odland, J., 1988. *Spatial autocorrelation*, Sage Publications Newbury Park, CA, USA:

Oke, T.R., 1987. *Boundary layer climates*, Routledge.

Paldy, A. et al., 2005. The effect of temperature and heat waves on daily mortality in Budapest, Hungary, 1970–2000. *Extreme weather events and public health responses*, p.99–107.

Palecki MA, Changnon SA, Kumkel KE, 2001. The nature and impacts of the July 1999 heat

wave in the Midwestern United States: Learning from the lessons of 1995. *Bulletin of the American Meteorological Scoiety* 82 (7).

Pattenden, S., Nikiforov, B. & Armstrong, B. G., 2003. Mortality and temperature in Sofia and London. *Journal of Epidemiology and Community Health*, 57(8), p.628.

Qiu, D. et al., 2002. Relationship between a high mortality rate and extreme heat during the summer of 1999 in Hokkaido Prefecture, Japan. *Journal of Epidemiology / Japan Epidemiological Association*, 12(3), pp.254-257.

Reid, C. et al., 2009. Mapping Community Determinants of Heat Vulnerability. *Environmental Health Perspectives*. Available at: http://ehp.niehs.nih.gov/docs/2009/0900683/abstract.html.

Roberts, S., 2004. Interactions between particulate air pollution and temperature in air pollution mortality time series studies. *Environmental research*, 96(3), p.328–337.

Robinson, P.J., 2001. On the definition of a heat wave. *Journal of Applied Meteorology*, 40(4), p.762–775.

Rocklov, J., 2007. A Study of Heat and Heatwave Effects on Mortality in Sweden. *Epidemiology*, 18(5), p.S28.

Ruddell, D.M. et al., 2010. Risk and Exposure to Extreme Heat in Microclimates of Phoenix, AZ. In *Geospatial Techniques in Urban Hazard and Disaster Analysis*. pp. 179-202.

Saaroni, H. et al., 2000. Spatial distribution and microscale characteristics of the urban heat island in Tel-Aviv, Israel. *Landscape and Urban Planning*, 48(1-2), p.1–18.

Saha MV, Davis RE, Hondula DM, 2013. Mortality displacement as a function of event strength in seven U.S. cities. *American Journal of Epidemiology, submitted.*

Samet, J. et al., 1998. Does Weather Confound or Modify the Association of Particulate Air Pollution with Mortality?: An Analysis of the Philadelphia Data, 1973-1980. *Environmental Research*, 77(1), p.9–19.

Sartor, F. et al., 1995. Temperature, ambient ozone levels, and mortality during summer, 1994, in Belgium. *Environmental research*, 70(2), p.105–113.

Semenza, J.C. et al., 1996. Heat-related deaths during the July 1995 heat wave in Chicago. *New England Journal of Medicine*, 335(2), p.84.

Sheridan, S.C., 2002. The redevelopment of a weather-type classification scheme for North America. *International Journal of Climatology*, 22(1), p.51–68.

Sheridan, S.C. & Dolney, T.J., 2003. Heat, mortality, and level of urbanization: measuring vulnerability across Ohio, USA. *Climate Research*, 24(3), p.255–265.

Sheridan, S.C., Kalkstein, A. J & Kalkstein, L. S, 2009. Trends in heat-related mortality in the United States, 1975–2004. *Natural Hazards*, 50(1), p.145–160.

Sheridan, S.C. & Kalkstein, L. S, 2004. Progress in heat watch-warning system technology. *Bulletin of the American Meteorological Society*, 85(12), p.1931–1942.

Shumway, R.H., Azari, A.S. & Pawitan, Y., 1988. Modeling mortality fluctuations in Los Angeles as functions of pollution and weather effects. *Environmental research*, 45(2), p.224–241.

Smargiassi, A. et al., 2009. Variation of daily warm season mortality as a function of micro-urban heat islands. *Journal of epidemiology and community health*, 63(8), p.659.

Smoyer, K. E, 1998. Putting risk in its place: methodological considerations for investigating extreme event health risk. *Social Science & Medicine*, 47(11), p.1809–1824.

Smoyer, K. E., 1998. A comparative analysis of heat waves and associated mortality in St. Louis, Missouri–1980 and 1995. *International Journal of Biometeorology*, 42(1), p.44–50.

Smoyer, K. E., Rainham, D.G.C. & Hewko, J.N., 2000. Heat-stress-related mortality in five cities in Southern Ontario: 1980–1996. *International Journal of Biometeorology*, 44(4), p.190–197.

Smoyer, Karen E. et al., 2000. The impacts of weather and pollution on human mortality in Birmingham, Alabama and Philadelphia, Pennsylvania. *International Journal of Climatology*, 20(8), pp.881-897.

Stafoggia, Massimo et al., 2006. Vulnerability to Heat-Related Mortality. *Epidemiology*, 17(3), pp.315-323.

Stathopoulou, M., Cartalis, C. & Andritsos, A., 2005. Assessing the thermal environment of major cities in Greece. In *Proc. 1st International Conference on Passive and Low Energy Cooling for the Built Environment, Santorini, Greece*.

Steadman, R.G., 1979. The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *Journal of Applied Meteorology*, 18(7), p.861–873.

Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, *23*(12), 1674-1687.

Sunyer, J. & Castellsagué, J., others, 1996. Air pollution and mortality in Barcelona. *Journal of epidemiology and community health*, 50(Suppl 1), p.s76.

Tan, J. et al., 2007. Heat wave impacts on mortality in Shanghai, 1998 and 2003. *International Journal of Biometeorology*, 51(3), p.193–200.

Tan, Jianguo et al., 2009. The urban heat island and its impact on heat waves and human health in Shanghai. *International Journal of Biometeorology*, 54(1), pp.75-84.

Touloumi, G., Samoli, E. & Katsouyanni, K., 1996. Daily mortality and“ winter type” air pollution in Athens, Greece–a time series analysis within the APHEA project. *Journal of epidemiology and community health*, 50(Suppl 1), p.s47.

Trenberth, K.E. et al., others, 2007. *Observations: surface and atmospheric climate change. Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, D Qin, M Manning, Z Chen, M Marquis, KB Averyt, M Tignor and HL Miller (Cambridge: Cambridge University Press).

Unger, J., 1999. Urban-rural air humidity differences in Szeged, Hungary. *International Journal of Climatology*, 19(13), p.1509–1515.

Voogt, J.A. & Oke, T.R., 2003. Thermal remote sensing of urban climates. *Remote Sensing of Environment*, 86(3), pp.370-384.

Weisskopf, M.G. et al., 2002. Heat wave morbidity and mortality, Milwaukee, Wis, 1999 vs 1995: an improved response? *American Journal of Public Health*, 92(5), p.830.

Whitman, S. et al., 1997. Mortality in Chicago attributed to the July 1995 heat wave. *American Journal of Public Health*, 87(9), p.1515.

Winslow, C.E.A. & Herrington, L.P., 1949. *Temperature and human life*, Princeton University Press, Princeton, NJ.

Yan, Z. et al., others, 2002. Trends of extreme temperatures in Europe and China based on daily observations. *Climatic Change*, 53(1), p.355–392.

Zauli Sajani, S. et al., 2008. Bioclimatic characterisation of an urban area: a case study in Bologna (Italy). *International journal of biometeorology*, 52(8), p.779–785.

Chapter 2. Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis

Manuscript published in *Environmental Health* (2012), 11(1).

David M. Hondula[1*], Robert E. Davis[1], Matthew J. Leisten[1], Michael V. Saha[1], Lindsay M. Veazey[1] and Carleigh R. Wegner[1]

[1]Department of Environmental Sciences, University of Virginia, Charlottesville, USA

Abstract

Background

High temperature and humidity conditions are associated with short-term elevations in the mortality rate in many United States cities. Previous research has quantified this relationship in an aggregate manner over large metropolitan areas, but within these areas the response may differ based on local-scale variability in climate, population characteristics, and socio-economic factors.

Methods

We compared the mortality response for 48 Zip Code Tabulation Areas (ZCTAs) comprising Philadelphia County, PA to determine if certain areas are associated with elevated risk during high heat stress conditions. A randomization test was used to identify mortality exceedances for various apparent temperature thresholds at both the city and local scale. We then sought to identify the environmental, demographic, and social factors associated with high-risk areas via principal components regression.

Results

Citywide mortality increases by 9.3% on days following those with apparent temperatures over 34°C observed at 7:00 p.m. local time. During these conditions, elevated mortality rates were found for 10 of the 48 ZCTAs concentrated in the west-central portion of the County. Factors related to high heat mortality risk included proximity to locally high surface temperatures, low socioeconomic status, high density residential zoning, and age.

Conclusions

Within the larger Philadelphia metropolitan area, there exists statistically significant fine-scale spatial variability in the mortality response to high apparent temperatures. Future heat warning systems and mitigation and intervention measures could target these high risk areas to reduce the burden of extreme weather on summertime morbidity and mortality.

Key words

Biometeorology, heat waves, climatology, apparent temperature, spatial analysis, heat-health impacts, remote sensing, Landsat

Background

High heat and humidity pose a major health threat for residents of middle latitude climates during the warm season. In the United States, for example, heat ranks as the leading cause of weather-related mortality. Many of these deaths are believed to be preventable via the implementation of appropriate mitigation measures such as advanced notification of at-risk groups and availability of cooling shelters [1]. Research to date has examined the aggregate heat-health response for large metropolitan areas and identified robust relationships for many locales. When temperature-humidity measures rise above a geographically sensitive threshold, human mortality becomes greater than the typical seasonal pattern would suggest [2]. The consistency of this heat-mortality relationship has spawned heat warning systems across the globe for forecast zones comprising entire metropolitan areas or multiple counties [1]. Within these metropolitan areas, however, there exists considerable variability in environmental conditions and demographic and social characteristics of the population. Here we explore the relationship between high heat and humidity and human mortality at the local scale. Past work has primarily focused on a larger scale examining the response of an entire metropolitan area, but the allocation of resources intended to protect citizens from the dangerous effects of heat and humidity could be improved with a more specific knowledge of where the risk is highest within urban areas. The current state of reporting and data availability makes it possible to assess this risk with a multidecadal record of geographically-specific observations. In this manuscript we utilize such a record to evaluate intra-city mortality risk within a major United States metropolitan area.

Elevated temperatures are commonly observed in city centers when compared to surrounding areas because of the well-documented urban heat island (UHI) effect [3]. Higher temperatures are observed in areas with tall buildings, high building density, limited green space, industrial land use, and anthropogenic heat sources [4]. The complex nature of cities leads to large differences in temperatures between varying neighborhoods—differences that may lead to certain areas exceeding physiological thresholds related to heat stress while other locations maintain thermal comfort [5].

Demographic and social variability within metropolitan areas may also contribute to geographic variability in heat-related mortality risk. Age is the most commonly-cited demographic factor related to morbidity and mortality risk during heat events [6, 7]. Elderly populations are believed to be at higher risk to changes in temperature than the general population because of diminished or compromised thermoregulatory capacity [1].

The impact of race on heat-related mortality is not as clear. Minority populations have often been linked to elevated mortality during heat events [8-10], but other research found no significant mortality difference between races [11, 12]. Socioeconomic status is potentially linked to heat-related mortality because affluent residents may be more able to afford higher-quality housing and air conditioning [13]. Areas with higher poverty rates were associated with statistically higher mortality rates in the 1993 Philadelphia heat event [10]. A wide range of other factors have also been linked with, or suggested to be linked with, elevated mortality rates during heat events, including education level, gender, aerobic fitness, activity level, and pre-existing medical conditions [7].

Over roughly the past decade, researchers have started examining the spatial distribution of heat-related mortality within metropolitan areas. At the local scale, places with higher morbidity and mortality rates during heat events have been associated with lower neighborhood stability and income levels [10, 13] and proximity to the downtown core [14-16]. In addition, remote sensing imagery has been incorporated and places with higher thermal readings have also been linked with higher mortality rates [17, 18].

The present study aims to relate a multi-decade record of localized mortality data to a large suite of independent variables including demographic, social, and environmental components. Our goal is to determine if significant spatial variability in heat-mortality exists within Philadelphia County, PA, and if so, examine the underlying factors that may be responsible for that variability.

Methods

Data

Individual all-cause mortality records for Philadelphia County residents were obtained from the Pennsylvania Department of Health for the period 1983–2008. We limited our analysis to deaths that occurred within the County boundary. The dataset includes the zip code of residence and age of the decedent. Excluding cases where the zip code was not available, the record contains 409,554 deaths over the 26-year period.

Surface hourly meteorological data for the entire period of record were obtained from Philadelphia International Airport (see Figure 1), including measurements of air temperature, dew point temperature, and wind speed. We calculated hourly apparent temperature (AT) using a parameterization of the Steadman model

$$AT = -2.653 + (0.994*T) + [0.0153 * T_D^2)] + C$$

where T is the dry-bulb temperature, T_D is the dew point temperature, and C is a correction based on wind speed [9,19,20]. We linearly interpolated the temperature correction for each integer value of wind speed between 0 and 16 m/s because the table of corrections only provides values for coarse increments of wind speed [19]. The correction for 16 m/s was used in cases where the wind speed exceeded 16 m/s. The airport AT measurements serve as the basis for the identification of days associated with exceptionally high heat and humidity conditions across Philadelphia County.

The suite of variables we incorporated to compare to the spatial pattern in heat-related mortality includes demographic and social factors as well as characteristics of the buildings and land surface. Population counts by age were obtained for each census tract in the County for 1980, 1990, and 2000 from the United States Census Bureau and the National Historical Geographic Information System (NHGIS). Additional variables obtained from the census include year 2000 tract-level measures of race, education level, income, occupancy, and building age (see Table 1). For geographic analysis, we used boundary shapefiles for the census tracts and year-2000 Zip Code Tabulation Areas (ZCTAs, Figure 1) from the Census Bureau and NHGIS.

We used zoning maps to further assess the surface characteristics of the built environment. Zoning ordinances constrain the use, coverage, form, and spatial arrangement of urban development. These regulations can have significant effects on urban environments [21]. Thus, the zoning variables may serve as proxies for high-resolution thermal measurements given that air temperature sensors are not available at the same level of spatial detail. Building density and zoning information for the County were obtained from the Pennsylvania Spatial Data Access (PASDA) clearinghouse GIS database. Both the zoning and building files contain thousands of individual polygons identifying each of several dozen zoning categories and individual building elements.

To assess intra-city variability in thermal stress, we utilized imagery of the Philadelphia area from the Landsat 7 Enhanced Thematic Mapper Plus (ETM+). Landsat is a sun-synchronous satellite with a 16-day overpass interval. We downloaded 47 warm-season

relatively-cloud free images using the USGS Global Visualization Viewer "Glovis" spanning the period 2004–2010 and selected two images that corresponded with periods of extremely elevated air temperature measured at Philadelphia International Airport (Figure 2).

Modeling

To obtain fine-scale mortality counts, daily all-cause mortality is aggregated for each of the 48 ZCTAs comprising Philadelphia County for the periods 1983–2008 (9,490 days). These data are de-seasoned to remove any residual effects of the intra-annual mortality cycle, and age-standardized to account for temporal changes in population demographics. All-cause data are commonly employed in heat impact research because of the lack of a standardized definition for heat-related death and the potential for heat stress to contribute to other causes of mortality (especially respiratory and cardiovascular causes). We calculated the seasonality as the 30-day LOWESS-smoothed daily means of the County-wide mortality sum. We then scaled this seasonality model such that the mean of the seasonality curve for each ZCTA-year matched the background mortality rate observed for each ZCTA-year. We next age-standardized the mortality data based on the ZCTA-level population age structure obtained from the 1980, 1990, and 2000 census data using ten age classes (0 – 4, 5–14, 15–24,…,85 and above) and interpolated (by age class) within each ZCTA.

The de-seasoned, age-standardized daily mortality departures were sorted into AT groups based on the hourly airport data to examine the heat-mortality relationship. After testing

different combinations, we ultimately chose to use overlapping 3°C-wide AT bins with a 1.5°C interval between the midpoints of each bin. This particular bin size represented balance between being able to relatively accurately identify a threshold temperature while preserving enough samples in each bin for each to be a reliable indicator of the mortality response. We then calculated the mean mortality departure within each bin. A one-tailed randomization procedure was used to identify a significant response for a particular AT bin. The mean excess mortality for a given AT range is compared to the 95^{th} percentile mean of 10,000 randomly-drawn subsets of the same sample size as the test group. Samples were drawn exclusively from days falling within the warm season, defined here as between calendar days 150–275 (approximately June–September). If the observed mortality is above the 95^{th} percentile, we identify a statistically significant mortality elevation. We excluded any bins with a sample size of 5 or fewer cases from analysis. The randomization procedure is used in place of a traditional t-test because of the non-normal distribution of the daily mortality departures [22]. In all cases the mortality response is expressed as a percent difference relative to the mean warm-season mortality in Philadelphia County (or within each ZCTA) of 0.26 deaths per 10,000 residents per day.

We first used the randomization procedure to evaluate the mortality response by AT for the entire County (total daily sum). We tested the mortality response by AT and hour for 12:00 a.m. to 8:00 p.m. on the day of death and all 24 hours for the two days immediately prior (each of 68 hours was tested for each AT bin). (We did not examine AT impacts after 8:00 p.m. local time on the day of death.) Examination of the overall city response

was used to guide the local-scale analysis to an AT/time combination when the mortality signal was robust. The minimum value of the first bin above which mortality remains consistently significantly greater than zero was adopted as the threshold AT for the given hour [2]. We then calculated the overall mortality response when AT values in excess of the threshold were observed (instead of within each AT bin) for the entire County and for each of the 48 ZCTAs. The ultimate dependent variable in the regression model is the mean excess mortality for certain ATs at a given hour of the day and lag. The randomization test was again used to identify significantly elevated mortality at both the County-wide and ZCTA scale, except in this instance the test statistic was based on the cumulative response above the threshold AT instead of within a particular AT bin.

Because the mortality data were provided with zip code of residence, ZCTAs serve as the geographic unit of analysis. All explanatory variables were projected into year 2000 ZCTAs using the Hawth's Tools Polygon-In-Polygon feature within ArcGIS.

The zoning code for Philadelphia includes many different classification types, including over twenty categories of residential zoning alone. Because a large number of these zoning categories were constrained to only a handful of parcels throughout the city, we combined similar zoning categories into seven overall classifications (Table 1). This aggregation was aided by numerical information within the zoning code related to lot size, building heights, etc., as well as street-level photography of the structures present for each zoning type.

Landsat imagery was employed to assess local-scale variability in surface temperatures. We converted the pixel-by-pixel brightness numbers to spectral radiance and then to temperature using the ArcGIS raster calculator (Figure 2a) [23]. Once all images were converted to temperature values, we used the ArcGIS Zonal Statistics tool to calculate the mean surface temperature within each ZCTA (Figure 2b). We used a mid-morning image from May 15, 2004 (Air temperature = 295.4 K at time of image) and a mid-morning image from July 29, 2008 (300.9 K). These two days best met our criteria of having little or no cloud cover and high air temperature out of the 47 images we downloaded. We added the mean surface temperature by zip code for each image as separate variables into the overall pool.

Principal components analysis was then used to reduce the number of variables from the independent pool and eliminate colinearity. As there are only 48 "cases" (ZCTAs) in the study, reducing the number of independent variables is especially important to avoid over-fitting the regression model.

We used multiple linear regression (conducted in SPSS statistical software, version 19.0.0) to relate the principal components of our independent variables to the local-scale mortality response. Significant variables were deemed to be those with a partial p-value of less than 0.05. The residuals from the regression models were examined for spatial autocorrelation using Moran's *I* statistic in ESRI ArcMap 9.3 to determine if an additional term is needed in the regression model to properly account for the true degrees of freedom in a spatially autocorrelated field.

Results

Countywide mortality is significantly elevated on and following days with high ATs. A significant threshold AT is evident for each hour spanning the 3-day period leading up to and including the day of death (Figure 3a). The threshold temperature varies by hour such that higher afternoon ATs are associated with the same mortality response as lower morning ATs. The mean mortality exceedance when ATs occur above the threshold is 5.2%; however, the cumulative response varies based upon when during the day high ATs occur (Figure 3b). Over the 68-hour period, we observed three peaks in the mortality rate when ATs occurred above the threshold: most of the afternoon hours two days prior to death, the mid-morning hours (7:00 a.m. to 11:00 a.m.) on the day prior to death, and the late afternoon and evening hours (8:00 p.m. to midnight) on the day prior to death.

We focused our local-scale analysis on cases where the 1-day lag 7:00 p.m. local time AT occurred above 34°C. This specific AT/time combination was chosen because of the robustness of the significant elevation in citywide mortality (based on the randomization test for the 35.5°C bin) and the sample size, with 110 such occurrences over the period of record. Citywide mortality increases 9.3% following days with a 7:00 p.m. AT above 34°C (see Figure 2b). Significantly elevated mortality, however, is only observed in 10 of the 48 ZCTAs within Philadelphia County (Figure 1). The remainder of the ZCTAs do not show a significant increase. ZCTAs associated with higher mortality following hot days are located in the central and west portions of the County.

Six principal components were extracted from the pool of 25 variables by analyzing a scree plot (all have eigenvalues ≥ 1.0). Collectively they account for 84.3% of the variability originally present in the dataset (Table 2). PC1 (35.9% explained variance) is highly correlated with surface temperatures from the two satellite images and is also strongly related to socioeconomic status. Component 1 scores are also positive in ZCTAs with an abundance of high density residential housing and high percentages of residents below poverty thresholds and lacking a high school diploma.

Three principal components were significant in the regression model: PC1, PC5, and PC6 (Figure 4). PC5 is representative of high density housing and mixed-use zoning. PC6 is most closely related to age and mixed-use zoning (see Table 1). The regression model identified a significant relationship ($p < 0.001$) between the three components and ZCTA-level mortality

$$M^* = 0.019 + 0.017(PC1) + 0.012(PC5) – 0.008(PC6)$$

where M* is the predicted mortality rate within each zip code, and PC refers to the respective principal components. The partial *p*-values for the coefficients were <0.001, 0.002, and 0.027, respectively, and the overall model adjusted R^2 was 0.439. The model indicates that heat-related mortality is greatest in areas with 1) a high number of residents below poverty thresholds, 2) residents lacking a high school diploma, 3) residents living in high-density housing, 4) more elderly persons, 5) high surface temperatures, and 6) mixed-use zoning.

The regression model performs well in identifying high-mortality locations (Figures 5 and 6). Each ZCTA with observed significant mortality exceedances was predicted to have a high mortality rate by the regression model. The range of values predicted by the model is not as great as the range in observed values, and in a few cases, model error is of the same order of magnitude as the observed mortality departures themselves. Residuals from the regression model were found to be randomly spatially distributed as measured by Moran's I statistic (I index 0.08, Z-score 1.13), indicating that an additional term in the model to account for spatial autocorrelation is not needed.

Discussion

The proper definition of environmental conditions that cause heat-related morbidity and mortality is an unsettled question in human biometeorological research. Various studies have employed different threshold variables, such as maximum temperature, afternoon AT, and morning dew point temperature [7]. Here, we attempt to address this shortcoming by examining AT diurnality for Philadelphia County. Although there are significant mortality elevations when ATs exceed the threshold at any hour, we found three periods with exceedances greater than 10%. Mortality rates were highest when thresholds were exceeded in the morning or evening hours on the day immediately prior to death and in the afternoon two days prior. Examining effects by hour, rather than using more conventional metrics like daily maximum, minimum, or mean temperature suggests specific hazardous periods. These patterns may arise in part from the specific threshold chosen for each hour: the threshold temperature might be expected to follow an even smoother pattern that shown in Figure 3a. In particular, the threshold temperature seems to increase rapidly on lag 1 between roughly 6:00 and 10:00 a.m., and thus the relatively high values here might be leading to the spike in the response at the same time in Figure 3b. The other peaks in the response curve (Figure 3b) seem less likely to be influenced by variations in the threshold curve. Future work might examine the mortality response above various percentiles of hourly temperature rather than a mortality-based threshold. The lack of an especially high response on the day of death (Figure 3b) may arise from the absence of time for exposure and resultant physiological stress (i.e., the response is not “immediate.”)

Mortality following days with high ATs in Philadelphia County is not randomly spatially distributed but is concentrated in several distinct regions. Certain ZCTAs exhibit mortality that is more than 30% above the daily citywide average for particular AT-time combinations. Intra-County variability in heat-related mortality has been observed or suggested elsewhere [10, 13-18], but the majority of studies to date have focused on a larger spatial scale, single heat events, hot summers, or did not consider the actual mortality response. This study is among the first to quantify local-scale mortality responses over a multi-decadal period.

Several of the variables associated with higher local-scale mortality are consistent with observations and hypotheses in the literature, including high-density housing, low socioeconomic status, high surface temperatures, and elderly populations [13, 16]. The spatial distribution of heat-related mortality in Philadelphia County during the 1993 heat wave was previously examined and the same variables were associated with elevated risk [10, 17]. The lack of a strong relationship with recreational zoning is surprising because we expected places with more parks and green space to have lower surface temperatures, thereby reducing heat and heat-related mortality. Recreational zoning is highest in two ZCTAs along the Schuylkill River in the western portion of the County, one of which also has a high percentage of high density residential zoning. However, the two zoning types are not interspersed, and where green space is not intermingled amongst residential areas, the mitigating effect on temperature in dense residential areas may be diminished. Although a large body of research points to the advantages of adding green space to lower temperatures in the urban environment, we are not able to conclude that ZCTAs

with more parkland are associated with lower mortality rates. This does not indicate that green space is not beneficial, but rather that many other variables may confound the signal, especially at the scale of this analysis. We are continuing to investigate the relationship between zoning types, air and surface temperatures, and mortality outcomes.

This study also incorporates the relatively recent approach of including remotely-sensed measurements of surface temperature in the study of heat-related mortality. Individuals living in areas with higher surface temperatures are at greater risk following hot days. This finding is consistent with the expectation that individuals living in hotter places are under greater physiological stress [10]. We are encouraged that the results from a remote sensing approach are similar to those using other sensors or models of the UHI.

We did not directly identify race as a key factor in the spatial distribution of heat-related deaths. Principal component loadings for the racial variables were only high in one significant component (PC 1), but loadings for other variables (income, surface temperature, educational attainment, and density of development) were higher. As previously documented for Phoenix [5], minority populations in Philadelphia County live in areas that are associated with higher surface and air temperatures. We directly observed the relationship with surface temperature and can infer the relationship with air temperature because of the high density of residential development in these locales. Racial variable loadings are very small for the other two components included in the model (PCs 5 and 6). Thus, we cannot conclude that race alone is a key factor in the spatial distribution of heat-related deaths in Philadelphia.

There are a few limitations we faced in creating our model for Philadelphia that merit discussion. The sociodemographic and zoning variables were derived from data available at a fixed point in time (e.g., the year 2000 census). However, the underlying demographics and zoning ordinances both change over time, a process we were unable to capture using this approach. This introduces some uncertainty into the results, and future research should explore local-scale mortality patterns over both space and time.

We were especially interested in exploring the relationship between the complex temperature patterns present in the metropolitan area and heat-related mortality. Satellite imagery has become much more accessible and makes this type of analysis possible using surface temperature measurements. The surface temperature field may be much different from the air temperature field over the same area, and we are not claiming that the two are identical, although some research indicates a high degree of correlation between the two fields during daylight hours [24]. There are many aspects of the urban heat island worthy of consideration in the context of urban health, including day/night variability and the contrast between the surface heat island and that of the canopy layer. We are investigating if residents of places with higher morning surface temperatures on hot days are at greater risk.

We are currently implementing a cloud-masking scheme that will increase the number of available images as well as extending our sample prior to 2004. We believe that the use of remote imagery in our study, and others, could be greatly enhanced if more surface temperature images were used. In just the two used in our study, there is variability in the

surface temperature pattern that may be linked to seasonal differences, synoptic-scale conditions, or other environmental controls. An additional concern with the satellite imagery is that many of the image pixels are measuring rooftop temperature, which may not be representative of the surface conditions experienced where individuals might be living, working, or spending time outdoors.

We observed a high correlation between the surface temperature field and several socioeconomic variables, as evident by the high loadings on the first principal component. The use of principal components analysis represented a tradeoff between examining potential effects of a larger suite of variables believed to be associated with risk and interpretability of results. Although we can definitively say that places with higher surface temperatures are associated with higher mortality risk, those places are also those with a high percentage of residents living in poverty and a high percentage of residents without a high school diploma. This pattern has been observed for other cities in the United States [5] and makes it difficult to pinpoint a causal relationship between the individual predictor variables and the health response. Even if it is difficult to separate the effects of individual variables, however, identifying characteristics of places associated with higher heat-related mortality can lead to improvements in the allocation of medical resources during dangerous conditions. Our future analysis in other cities in the United States where socioeconomic status and surface temperatures may not be as highly correlated may shed light on the relative impact of exposure, education, and income on heat-related risk.

The role of air quality in leading to increased mortality during heat waves is a topic of continued debate in the literature and is beyond the scope of this study. As heat waves are commonly associated with clear skies and stagnant air, conditions are ideal for the rapid buildup and accumulation of various pollutants. It is likely that during heat waves a portion of the excess deaths are attributable to the thermal stress whereas others might be linked to high concentrations of unhealthy atmospheric constituents. We did not incorporate air quality data into this study but encourage future study of the interactive effects of heat and air quality on summertime mortality as well as the potential for differential mortality over space as a result of local-scale air quality variability. Both are topics of active ongoing investigation by the authors and many others. We also note that we were unable to locate air conditioning use data at an appropriate resolution for this study. Air conditioning has become widely adopted in the United States and increases in availability have been linked to decreases in heat-related mortality [2]. However, we believe that air conditioning availability and usage is likely highly correlated with measures of socioeconomic status, and thus may be implicitly included in our analysis. Access and willingness to use medical care is a potential confounder at the individual level that we were not able to represent at the scale of this study, although it may be highly correlated with the socioeconomic variables included. Finally, the use of AT may not identify all of the critical physiological factors in evaluating the heat-mortality relationship and we intend to adopt this approach with other physiological indices in the future.

Intra-city variability in the response to high heat and humidity conditions indicates an opportunity for the improvement of heat-health watch-warning systems (HHWWS) that have been deployed in cities across the globe. When a dangerous event is forecast, for example, emergency managers might reprioritize allocation of medical resources to those geographic areas responsible for the largest portion of the heat-related deaths in the past. A more thorough effort to build and validate a predictive model of both the timing and placement of heat-related deaths is recommended prior to operational changes in any HHWWS. Longer-term strategies to reduce the heat stress and health burden in these localities might be considered as well, such as the implementation of building weatherization programs, adding green space to the city landscape, adoption of low-albedo and/or green building practices, and location of future healthcare facilities.

Conclusions

We have identified statistically significant fine-scale spatial variability in heat-related mortality within Philadelphia County, PA, over the period 1983–2008. Following days on which the 7:00 p.m. AT exceeds 34°C Countywide, mortality is significantly elevated, but the excess deaths are not randomly distributed throughout the metropolitan area. Instead, only 10 of the 48 individual ZCTAs that comprise the County exhibit significantly higher mortality. Compared with areas that do not have elevated mortality following hot days, these 10 ZCTAs have a higher percentage of elderly residents, a higher percentage of residents of low socioeconomic status, more high-density residential and mixed-use zoning, higher surface temperatures, and more recreationally-zoned area. A portion of the spatial distribution of heat-related mortality arises from underlying demographic, social, and environmental variability. The overall Countywide response varies based on the specific timing and intensity of high heat and humidity. Afternoon AT thresholds are higher than morning thresholds, but especially high mortality rates are observed when the threshold is crossed either during the mid-morning or late afternoon hours.

The significant local-scale variability in heat-related mortality identified for Philadelphia County suggests an opportunity for improved heat preparedness and management strategies. In the case of alerting the public, localities associated with excess mortality could receive additional notification or special forecasts when hot conditions are expected. These places are also prime candidates for facilities that can help residents

escape the impact of high ATs. The authors are adopting this approach for other United States cities in different climate zones to determine if certain factors are consistently associated with elevated risk during heat waves. Understanding the characteristics of places especially sensitive to the dangers of high heat and humidity may ultimately reduce the impact of extreme summertime conditions on human health.

Acknowledgements

Many thanks to Marina Matthew and the Pennsylvania Department of Health for providing the mortality data for this study, Bruce Hayden and John Porter at the University of Virginia, Laurence Kalkstein at the University of Miami, and Karen Smoyer-Tomic of HealthCore for their valuable input. David M. Hondula was partially supported by a United States National Science Foundation Graduate Research Fellowship and a United States Environmental Protection Agency Science to Achieve Results (STAR) Graduate Research Fellowship. Additional research support was provided by the Michael Garstang Award at the University of Virginia.

Figures

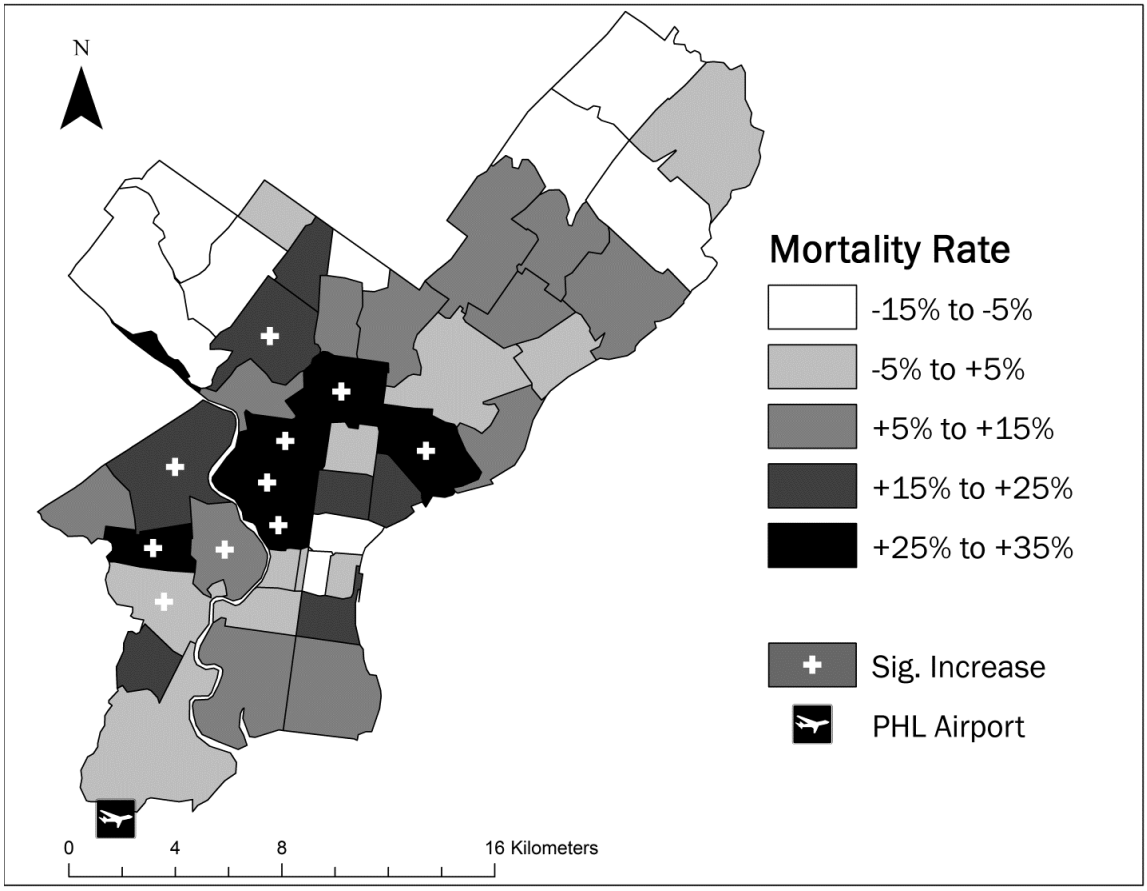


Figure 1.

Detailed legend: Mortality exceedance rates for Philadelphia county ZCTAs on days following those with 7:00p.m. local time apparent temperatures greater than 34°C. ZCTAs with mortality rates significantly greater than the background rate are identified by a white cross. The location of Philadelphia International Airport (PHL), the source of meteorological data employed in this study, is represented by the airplane symbol.

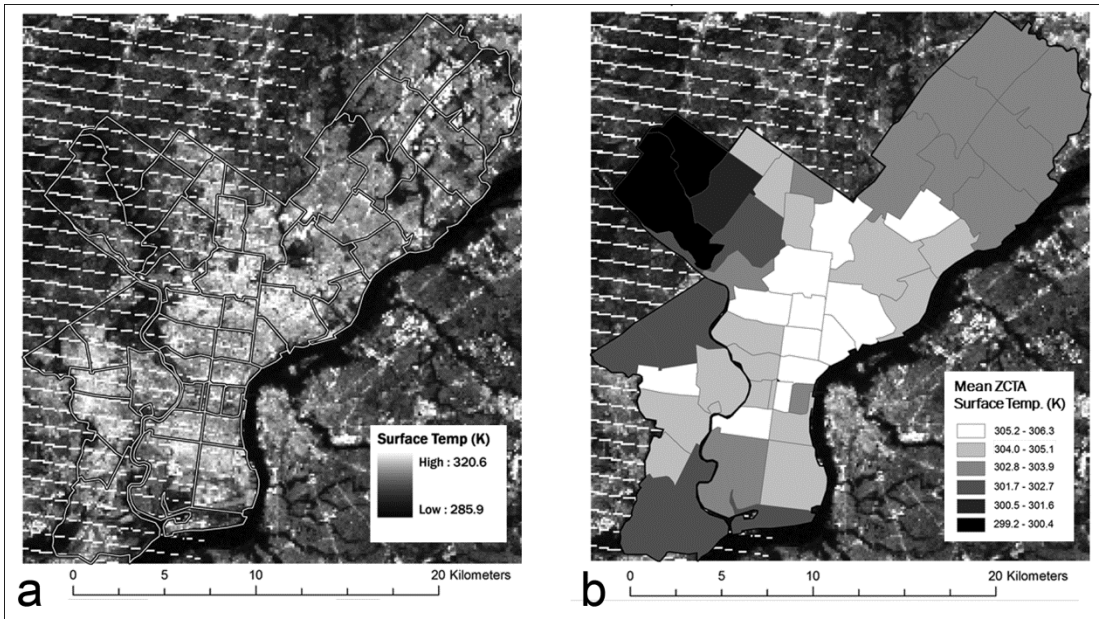


Figure 2.

Detailed legend: (a) Landsat thermal image of surface temperatures (K) in Philadelphia County, PA, from July 29, 2008. Zip code tabulation areas (ZCTAs) that comprise the county are shown with white outlines. (b) The mean surface temperature within each of the Philadelphia County ZCTAs is shown with grayshading.

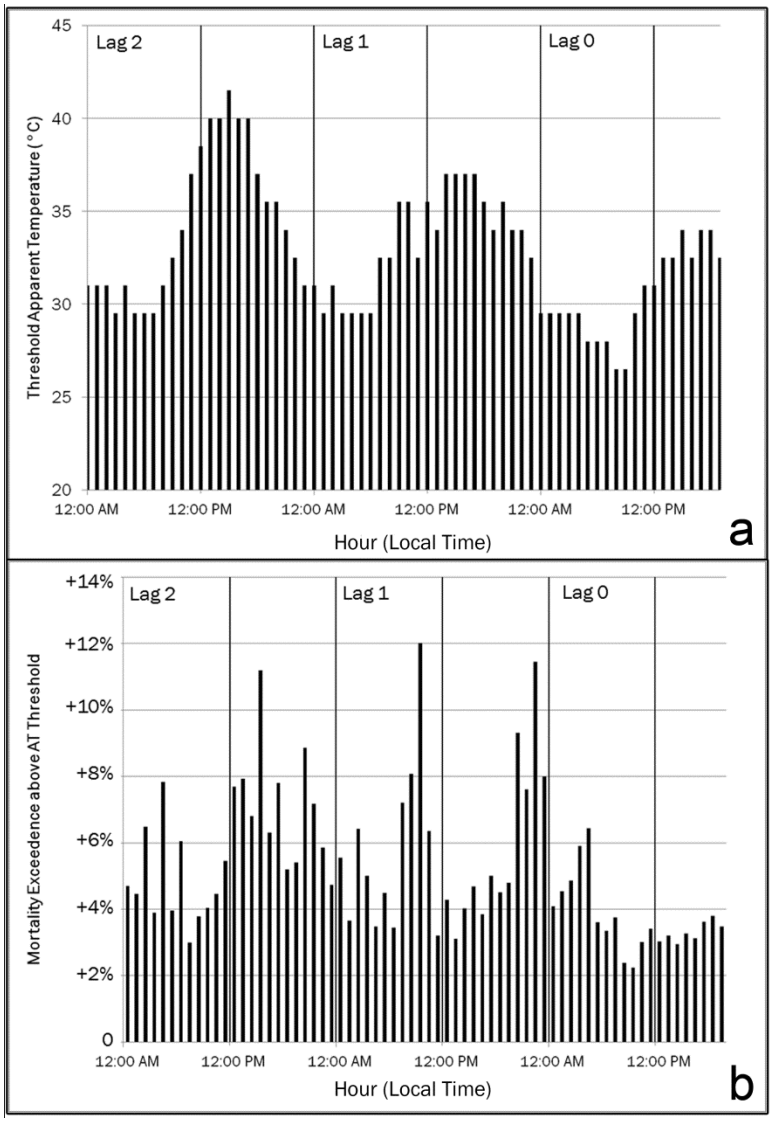


Figure 3.

Detailed legend: (a) Threshold apparent temperatures for heat-related mortality in Philadelphia county, PA. (b) Mortality exceedances for Philadelphia County, PA when apparent temperatures exceed the threshold between midnight and 8:00 p.m. local time on the day of death (lag 0), and for each of the 24 hours on the two days prior.

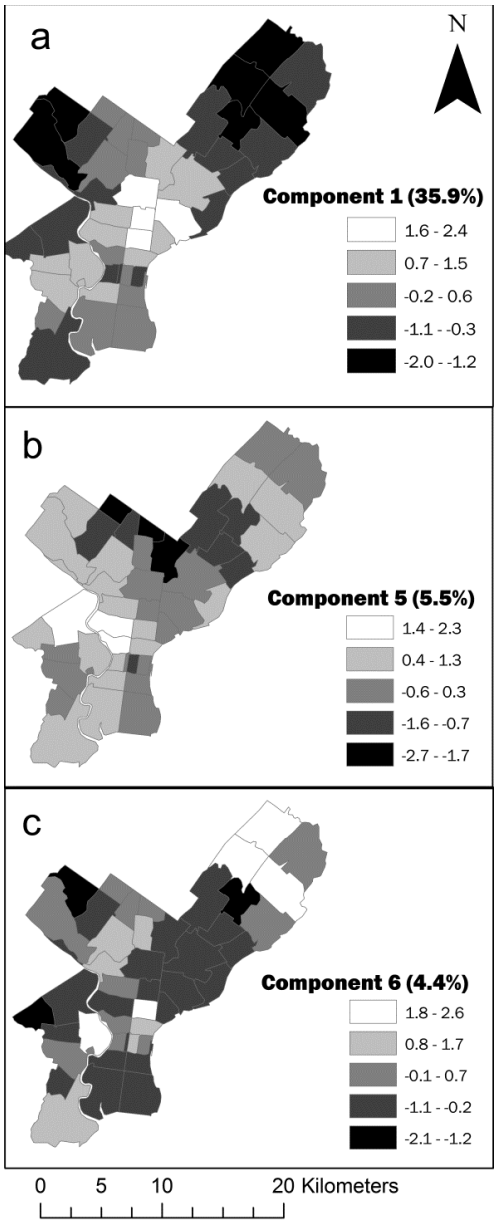


Figure 4.

Detailed legend: Principal component scores by ZCTA for three significant components (a: PC1, b: PC5, c: PC6) included in regression model relating explanatory variables to heat-related mortality. The percentage of variance explained by each component is shown in parentheses.

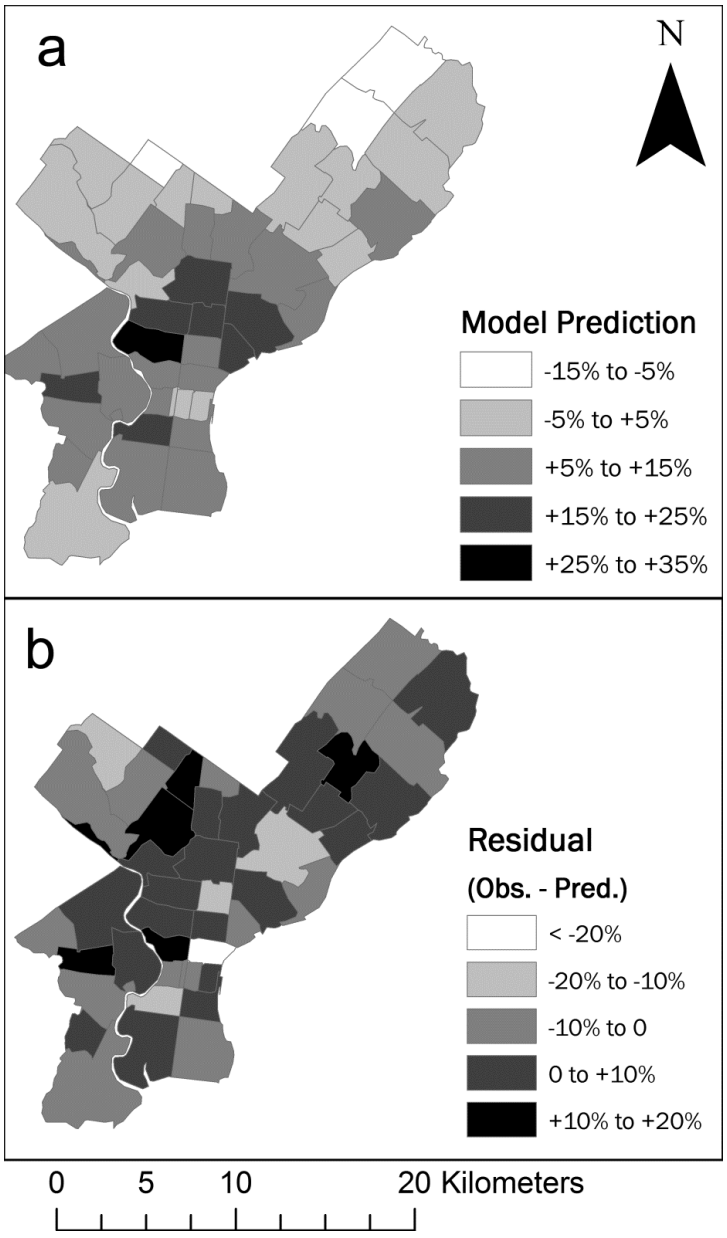


Figure 5.

Detailed legend: (a) Regression model predicted mortality exceedances for Philadelphia County ZCTAs on days following 7:00p.m. local time apparent temperatures above 34°C. (b) Differences between observed mortality rates and model predictions by ZCTA.

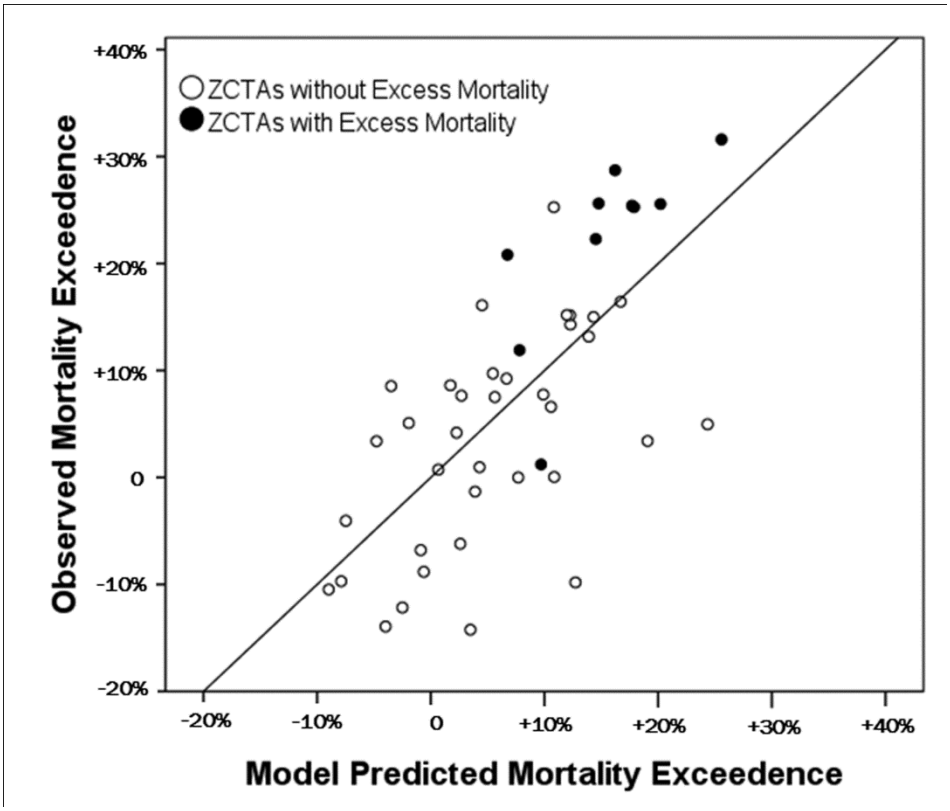


Figure 6.

Short title: Comparison of model predictions and observations.

Detailed legend: Scatterplot comparing observed mortality rates and model predictions on days following 7:00p.m. local time apparent temperatures above 34°C by ZCTA. The 1:1 line is included for reference. Filled circles identify the 10 ZCTAs with statistically significant mortality exceedances in Figure 1.

Tables.

Table 1. List of explanatory variables for assessing the spatial distribution of heat-related mortality exceedances and loadings for extracted principal components. Loadings greater than 0.6 or less than -0.6 are shown with bold text and significant components in the regression model are identified with an asterisk.

	Component (% Variance Explained)					
	1 (35.9)*	2 (16.4)	3 (10.2)	4 (7.7)	5 (5.5)*	6 (4.4)*
ZONING AND LAND USE						
% Low Density Residential	**-.636**	-.290	.122	.398	.017	.088
% Mid Density Residential	-.402	-.456	.012	-.015	-.578	.087
% High Density Residential	**.702**	.131	-.217	-.079	.449	-.129
% Recreational	-.404	-.438	-.058	.304	.396	-.323
% Industrial	.161	.020	**.646**	-.420	.229	.161
% Mixed Use	-.073	.142	-.256	.394	.434	.592
% Commercial	.045	**.797**	-.418	.015	.029	.036
% Building Coverage	.514	**.663**	-.442	-.053	-.049	-.078
DEMOGRAPHICS						
% White	**-.647**	**.610**	.361	-.129	.046	-.060
% Black	.471	**-.648**	-.513	.086	-.071	.051
% American Indian	**.655**	.039	.457	.287	.012	-.028
% Asian	-.017	**.695**	-.165	.144	-.244	.099
% Pacific Islander	.400	.271	.255	**.656**	.003	.181
% Other race	**.615**	.146	.541	.333	-.079	-.041
% Two or more races	.500	.324	.292	.572	-.334	-.110
% Nonwhite	**.647**	**-.610**	-.361	.129	-.046	.060
% Over age 65	-.345	.300	-.003	.142	.159	**-.620**
% Without hs diploma	**.799**	-.135	.402	-.191	.107	-.095
Median per capita income	**-.651**	.455	-.328	.181	.078	-.103
% Below Poverty Line	**.912**	-.005	.006	.066	.266	.050
% Below 2x Poverty Line	**.925**	-.115	.060	-.053	.202	-.033
% Living Alone over age 65	**-.648**	.044	.370	-.187	.095	.177
% Living Alone	**-.768**	.260	.086	-.085	.200	.247
SURFACE TEMPERATURE						
Surface Temp. Image (5/15/2004)	**.762**	.360	.001	-.332	-.201	.107
Surface Temp. Image (7/29/2008)	**.849**	.340	-.152	-.261	-.124	.042

Table 2. Land use zoning categories aggregated from Philadelphia Zoning Code.

Zoning Category	**Description**
Low Density Residential	Suburban, single-family, detached households with lawns
Mid Density Residential	Semi-detached or attached households with some green space
High Density Residential	Attached households with minimal or no green space
Recreational	Parks and protected natural areas
Industrial	All industrial complexes, stockyards, or ports
Commercial	Center-city office buildings, retails centers, corner shops
Mixed-Use	Strip malls, movie theaters, stadiums, hospitals, colleges, condominiums

References

1. Koppe C, Jendritzky G, Kovats RS, Menne B: Heat waves: risks and responses. In *Health and Global Environmental Changes, Series No. 2*. Copenhagen: WHO Regional Office for Europe; 2004.

2. Davis RE, Knappenberger PC, Michaels PJ, Novicoff WM: Changing heat-related mortality in the United States. *Environ Health Persp* 2003, 111: 1712-1719.

3. Voogt JA, Oke TR: Thermal remote sensing of urban climates. *Remote Sens Environ* 2003, 86:370-384.

4. Balling RC, Brazel SW: Time and space characteristics of the Phoenix urban heat island. *J AZ-NV Acad Sci* 1987, 21:75–81.

5. Harlan SL, Brazel AJ, Prashad L, Stefanov WL, Larsen L: Neighborhood microclimates and vulnerability to heat stress. *Soc Sci Med* 2006, 63:2847-2863.

6. Medina-Ramón M, Zanobetti A, Cavanagh DP, Schwartz J: Extreme Temperatures and Mortality: Assessing Effect Modification by Personal Characteristics and Specific Cause of Death in a Multi-City Case-Only Analysis. *Environ Health Persp* 2006, 114:1331-1336

7. Gosling SN, Lowe JA, McGregor GR, Pelling M, Malamud BD: Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change* 2008, 92:299-341.

8. Changnon SA, Kunkel KE, Reinke BC: Impacts and Responses to the 1995 Heat Wave: A Call to Action. *B Am Meteorol Soc* 1996, 77:1497-1506.

9. O'Neill MS, Zanobetti A, Schwartz J: Modifiers of the Temperature and Mortality Association in Seven US Cities. *Am J Epidemiol* 2003, 157:1074-1082.

10. Johnson DP, Wilson JS: The socio-spatial dynamics of extreme urban heat events: The case of heat-related deaths in Philadelphia. *Appl Geogr* 2009, 29:419-434.

11. Kalkstein LS, Davis RE: Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Ann Assoc Am Geogr* 1989, 79:44-64.

12. Green RS, Basu R, Malig B, Broadwin R, Kim JJ, Ostro B: The effect of temperature on hospital admissions in nine California counties. *Int J Public Health* 2010, 55:113–121.

13. Smoyer KE: Putting risk in its place: methodological considerations for investigating extreme event health risk. *Soc Sci Med* 1998, 47:1809-1824.

14. Dolney TJ, Sheridan SC: The relationship between extreme heat and ambulance response calls for the city of Toronto, Ontario, Canada. *Environ Res* 2006, 101:94-103.

15. Bassil KL, Cole DC, Moineddin R, Craig AM, Wendy Lou WY, Schwartz B, Rea E: Temporal and spatial variation of heat-related illness using 911 medical dispatch data. *Environ Res* 2009, 109:600-606.

16. Reid C, O'Neill MS, Gronlund CJ, Brines SJ, Brown DG, Diez-Rouz AV, Schwartz J: Mapping Community Determinants of Heat Vulnerability. *Environ Health Persp* 2009, 117:1730-1736.

17. Johnson DP, Wilson JS, Luber GC: Socioeconomic indicators of heat-related health risk supplemented with remotely sensed data. *Int J Health Geogr* 2009, 8:57.

18. Smargiassi A, Goldberg MS, Plante C, Fournier M, Baudouin Y, Kosatsky T: Variation of daily warm season mortality as a function of micro-urban heat islands. *J Epidemiol Commun H* 2009, 63:659-664.

19. Steadman RG: The Assessment of Sultriness. Part II: Effects of Wind, Extra Radiation and Barometric Pressure on Apparent Temperature. *J Appl Meteorol* 1979, 18:874–885.

20. Kalkstein LS, Valimont KM: An Evaluation of Summer Discomfort in the United States Using a Relative Climatological Index. *B Am Meteorol Soc* 1986, 67:842–848.

21. Wilson J: Evaluating Environmental Influences of Zoning in Urban Ecosystems with Remote Sensing. *Remote Sens Environ* 2003, 86: 303-321.

22. Sheridan SC, Dolney TJ: Heat, mortality, and level of urbanization: measuring vulnerability across Ohio, USA. *Clim Res* 2003, 24:255-265.

23. Xian G, Crane M: An analysis of urban thermal characteristics and associated land cover in Tampa Bay and Las Vegas using Landsat satellite data. *Remote Sens Environ* 2006, 104:147-156.

24. Nichol J: Remote Sensing of Urban Heat Islands by Day and Night. *Photogramm Eng* 2005, 71:613–621.

Chapter 3. A time series approach for evaluating intra-city heat-related mortality

Manuscript published in *Journal of Epidemiology and Community Health* (2013).

DM Hondula, RE Davis, J Rocklöv, MV Saha

KEYWORDS: Temperature, mortality, extreme heat, urban, time series

ABSTRACT

Extreme heat events are a leading cause of weather-related mortality. Most research to date has considered the aggregate response of the populations of large metropolitan areas, but the focus of heat-related mortality and morbidity investigations is shifting to a more fine-scale approach where impacts are measured at smaller units such as postal codes. But most existing statistical techniques to model the relationship between temperature and mortality cannot be directly applied to the intra-city scale because small sample sizes inhibit proper modeling of seasonality and long-term trends. Here we propose a time series technique based on local-scale mortality observations that can provide more reliable information about vulnerability within metropolitan areas. The method combines a generalized additive model with direct standardization to account for changing mortality rates in intra-city zones. We apply the method to a 26-year time series of postal code-referenced mortality data from Philadelphia County, USA, where we find that heat-related mortality is unevenly spatially distributed. Fifteen of 46 postal codes are associated with significantly elevated mortality on extreme heat days, a majority of which are located in the central and western portions of the County. In some cases the local-scale mortality rate is more than double the County average. Identification of high-risk areas can enable targeted public health intervention and mitigation strategies.

INTRODUCTION

Most studies examining the heat-health response of urban populations have used health data aggregated across cities or counties, but there is increasing interest in measuring the heat-health response at a finer spatial scale.[1–4] At the broadest level, this research is motivated by the significant health burden posed by extremely high heat and humidity levels in many middle-latitude cities each summer. On and immediately following days in which the temperature (or a temperature-humidity measure) exceeds a location-specific threshold, the mortality rate often increases several percent or more above what would normally be expected at that time of the year.[3,5] This overall relationship is robust to the particular statistical method chosen by the investigator and often the specific independent variable used for analysis.[5,6]

The extension of this research field to a finer spatial resolution is largely motivated by the heterogeneity in environmental and sociodemographic factors within metropolitan areas. There is general agreement that the major individual-level risk factors for heat-related mortality are age, income, and isolation;[7,8] others report additional factors including race, pre-existing medical conditions, activity level, and education.[9-10] Thus it is already possible to determine those zones within cities where the heat-related mortality risk might be *expected* to be highest.[11] Spatial information at this scale could be incorporated by public health officials to efficiently locate medical resources for emergency response measures taken before, during, and immediately following extreme heat events. In a growing number of cities, these protocols are associated with the activation of a heat-health warning system.[12] Geographic specification of high risk

zones could also enhance our understanding of the impacts of the urban heat island on human health, potentially providing empirical context for longer-term mitigation measures related to green space and building weatherization.[13-14]

Few studies, however, report intra-city zones where *observed* mortality is highest when extreme heat occurs. The limited exploration of this research area might be partially attributable to certain methodological challenges that we are aiming to overcome with the approach presented in this manuscript. Methods for evaluating the impact of extreme heat on human health tend to vary by discipline, such as studies designed around synoptic air-mass types and individual-level analysis using case-only and case-crossover methods.[6] In recent years, the use of various time series strategies including generalized additive models has become more common, most recently leading to the adaptation of distributed lag non-linear models designed to capture the impacts of a range of temperatures across a range of lags.[15,16]

This shift toward more flexible modeling strategies has not, to date, accommodated the intra-city scale where sample sizes become relatively small through subsetting of the original dataset across both space and time. For example, in the case of Philadelphia County, Pennsylvania, a metropolitan area in the United States home to nearly 2 million residents, the average daily mortality within each of the 48 postal codes that comprise the county is less than 1.0.[17] Conceptually one could model the temperature-mortality relationship for each individual unit using a Poisson regression time series framework, accounting for a number of confounding factors. However, at this scale there are not sufficient sample sizes large enough to model correctly the seasonality and long-term

temporal mortality trends that must be accounted for when evaluating the impacts of high temperature events.

Recent studies that do use area-based outcome data to map heat-related risk acknowledge limitations based on sample size or methodology, focus on single extreme events, make certain assumptions about seasonality or long-term trends (or do not account for them), combine days into certain larger subsets, or rely on aggregations of small spatial units to a coarser scale where existing methods can be used.[1,2,17-23] To the best of our knowledge, no published study uses a method for analyzing a fine-scale, long-term mortality time series using a technique that a) accounts for varying year-to-year mortality seasonality and long-term time trends and b) estimates the temperature dose-response curve for each individual geographic unit. At a minimum, the use of a long time series to estimate local-scale effects might increase the applicability of results, but it is also possible that further model specification would lead to a different estimation of heat effects. Here we present a two-stage method for evaluating local-scale heat risk that combines elements of various approaches previously presented for measuring risk at the aggregate scale. Our goal is to estimate the temperature-mortality relationship for small spatial units within cities to foster analysis of environmental and social factors that might be responsible for spatial variability in risk. Ultimately, estimation of temperature effects on health at the sub-city scale based on outcome data could enhance practical measures aimed at improving public health.

METHODS

We collected daily all-cause mortality records for Philadelphia County, Pennsylvania, USA, for the period 1983–2008. The data were obtained from the Pennsylvania State Department of Health and were de-identified, thus no IRB approval was necessary for this research. The records contain the postal code of residence of each decedent as well as the date of death. We aggregated the records to a matrix M with a sum of deaths within each postal code *z* on each day in the record *d* and a vector N with a sum of deaths across all postal codes for each day in the record. Two of the 48 postal codes (19108 and 19112) were excluded from the analysis because they represent commercial districts where only a few individuals reside. No age groups were excluded from the analysis. Additional details on these mortality data can be found elsewhere.[17]

Hourly weather station data were obtained from the National Climatic Data Center for Philadelphia International Airport located on the southern perimeter of Philadelphia County. Afternoon maximum temperature (T_{PM}) was the exposure metric used, calculated as the maximum temperature observed between 1300 and 1900 local standard time. This time series has less than 0.1% missing observations over the study period. The mean summertime T_{PM} during the time period was 28.0°C; the maximum was 39.4°C.

The analysis was completed in three stages: (i) estimation of seasonality and long-term trends for the entire county, (ii) estimation of seasonality and long-term trends for each individual postal code, and (iii) estimation of heat effects for each individual postal code.

Stage I: Estimation of county-level seasonality and trend

The first stage of analysis utilizes the daily mortality time series aggregated across all postal codes within the County (Figure 1). We employed a generalized additive model to estimate the effects of afternoon temperature (T_{PM}) and long-term trends and seasonality (Trend) on daily mortality counts (M) using the equation

$$Log[E(M)] = \alpha + s(T_{PM}, df = 6) + s(Trend, df = 130)$$

where α is the model intercept and *s* is a penalized smoothing spline.[24,25] This type of semiparametric model has been widely used to estimate the effect of environmental factors on mortality while accounting for confounding variables including seasonality. [20, 26-28]. The modeled seasonality (which varies from year to year, represented by the solid line in Figure 1) will be carried throughout the local-scale stage of the model, whereas the other components (temperature and long-term trends) will be flexible. The model utilized a quasi-Poisson link function and splines with five degrees of freedom per year for the trend/seasonality variable and six for temperature. Sensitivity analysis was performed to compare model deviance to the maximum number of degrees of freedom available for the smoothing terms. The use of five degrees of freedom per year for the Trend term provides the model sufficient flexibility to capture seasonal effects within each year as well as long-term time trends across the entire record (this term is represented by the solid line in Figure 1). We extracted the model predicted value *B* from the Trend term for each day in the period of record (often referred to as a "mortality baseline").

A new time series, the daily standardized mortality ratio (SMR), is calculated by dividing the observed mortality total *M* by baseline mortality *B* on each day *d*. *B* was generated

assuming a temperature of 10°C on every day in the period of record (the specific value chosen here does not impact the results because of latter calculations using a "relative to baseline" framework). An SMR of 1.0 indicates that the mortality was equal to expectations for time of year and date in the period of record; values greater than 1 indicate excess mortality and less than 1 a mortality deficit.

Stage II: Estimation of postal code-level seasonality and trend

The next step is to create an SMR time series for each postal code. These time series need to account for demographic changes that might lead to temporal trends in the daily mortality count. The other major component of the generalized additive model is the interannual variability in mortality seasonality. The specific shape of the seasonal mortality pattern may vary from one postal code to another, and each postal code's seasonality may differ from the County aggregate. Because of sample size limitations, however, it is not possible to generate a model within which the shape of the seasonal mortality curve within each postal code can vary from year to year. It is possible to generate an average seasonality curve for each postal code over the entire period of record, but this approach would miss any shifts in seasonality related to larger-scale occurrences such as the timing of the flu season. Accordingly, we wish to preserve the interannual variability in seasonality derived from the aggregate data (B), with the tradeoff that we do not specify the shape of seasonality within each postal code individually. Thus we are using inference to estimate local-scale effects based on the more reliable coarse-scale data. We calculate an SMR value for each postal code z and day d using the equation

$$SMR_{z,d} = M_{z,d}/[\frac{B_d}{\mu(B)} \times \theta(M_z)]$$

where *M* is the number of deaths within each postal code *z* on day *d*, μ() represents the mean daily death count over the entire period of record, and θ() represents the average mortality rate within each postal code for the calendar year containing day *d*. The denominator for the postal code SMR time series is calculated by scaling and shifting the seasonality curve from the citywide data to match each postal code's mortality rate on a year-to-year basis. This forces the citywide seasonality curve's shape on each postal code, but allows flexibility in the time series to accommodate the possibility of changing postal-code level population and mortality trends over the period of record. We then divide the observed daily mortality counts by the daily values calculated from this adjusted seasonality curve to produce the SMR time series for each postal code.

The resulting time series contains daily standardized mortality ratios for each postal code that exclude seasonality (based on the city-wide mortality records) and long-term population and mortality trends (specified for each postal code). In the next stage of analysis we calculate the dose-response curve for each postal code using a binning method previously used to evaluate aggregate city-wide response.[5]

Stage III: Estimation of heat effects within postal codes

First we test the binning method on the citywide data with overlapping 3°C intervals. A randomization test is used to determine if the mortality response for a subset of days we identify as "extreme heat" differs from the normal summertime mortality rate. For a given sample of n days associated with an extreme heat event, we draw 1,000 random subsets of size n from all summertime days in the period of record, where summer is defined as the 150^{th} through 275^{th} day of the calendar year. The mean SMR is calculated for each of the 1,000 subsets, and the 95^{th} percentile of this set of means is chosen as the critical value for a one-tailed test of the mean SMR of the subset of days associated with extreme heat. The test statistic T is the extreme heat sample mean divided by the critical value (T > 1 indicates a significant mortality increase). We also calculate the percent increase in the SMR on extreme heat days by dividing the sample mean over the mean SMR of all summer days in the period of record (the mean SMR of summer days in the period of record is not exactly 1.0 because the original models were calculated using the full year-data to model the complete seasonal cycle as accurately as possible). This procedure is then repeated for every postal code within each city using each postal code's respective de-seasoned time series. The complete local-scale workflow is demonstrated in Figure 2.

Generalized additive models were constructed using the gam function of package mgcv run in R version 2.15.1. All additional analyses were completed using MATLAB version 7.13. [29-31]

RESULTS

The generalized additive model for Philadelphia identified significant variability related to seasonality and long-term trends (captured in one variable) and temperature. Both terms in the model were significant ($p<0.001$), and the model fitted values accounted for 31.7% of the variance in the original data (R^2). Visual inspection suggested residuals were normally distributed and residual variance was not correlated with model predicted values. The seasonality-trend component of the time series (Figure 1) shows an overall trend toward lower mortality rates within the County over the time period, particularly within the last 20 years of the record. On average, mortality is highest in January (48.4 deaths per day) and lowest in August (39.5 deaths per day), although the amplitude and specific timing of the interannual mortality maximum and minimum are variable.

The dose-response curve for Philadelphia County (Figure 3) follows the U- or J-shaped mortality-temperature relationship that has been found for many other middle latitude cities.[6] Temperatures between 5–10°C are associated with minimum mortality, and maximum mortality is associated with the highest temperatures. When the curve is evaluated relative to the mortality rate associated with the mean summertime daily maximum temperature of 28°C, we find a statistically significant increase in mortality above 33°C (approximately the 90^{th} percentile of summertime afternoon maximum temperatures). This is the first point at which the lower bound of the mortality confidence interval occurs above the mortality rate associated with normal summertime conditions. The shape of the curve indicates mortality rates approaching up to 10% above baseline when afternoon maximum temperatures near 40°C. The seasonality and trend time series

for the individual postal codes showed variability relative to that constructed for the overall countywide data (e.g., Figure 2b).

The shape of the temperature-mortality curve for Philadelphia County based on the modeled relationship from the generalized additive model is quite similar to that obtained from the binning analysis (Figure 4). Although the curve from the binning analysis is not as smooth as that obtained from the generalized additive model, and there is some divergence at the low end of the temperature range, the two curves are nearly identical for the range of values one might be concerned with when investigating extreme heat effects. The relative risk values obtained from the binning analysis fall within the 95^{th} percent confidence interval from the generalized additive model. The width of the confidence intervals differs between the two methods but follows the same general pattern (see Supplementary Material Figure 1). The agreement between these two curves supports the use of the binning approach at the postal code level where a generalized additive model cannot be created because of the small sample sizes.

During the study period there were 370 days (14 per year) when the afternoon maximum temperature exceeded the threshold of 33°C. On these days, the average mortality rate increased by an average of 6.4%. Only fifteen of the 46 postal codes examined were associated with statistically higher mortality rates in the same set of study days. Many of these postal codes comprise a nearly contiguous region near the center and western boundary of the County (Figure 5). The mortality rate in postal codes with significantly high mortality on extreme heat days ranged between 6.7 and 22.6% above baseline. Across the entire County, 39 postal codes had mortality rates above the baseline rate and

seven had rates below the baseline. Five postal codes were associated with increases of 6.7% (the lowest mortality elevation among significant postal codes) or more but these increases were not statistically significant based on the randomization test, which is sensitive to sample size. The places identified as “high mortality” are not necessarily those that are typically associated with higher mortality rates on non-extreme heat days because all calculations are relative to the baseline mortality specific to each postal code. Thus residents of certain postal codes in Philadelphia County have historically been at greater risk of mortality following extreme heat days compared to residents in neighboring areas.

DISCUSSION

Using a new method developed for the investigation of local-scale mortality effects associated with extreme heat, we were able to confirm a hypothesis that there is significant spatial variability in heat-related mortality within Philadelphia County, USA. This method provides a framework for future investigations into sub-city-scale modifiers of the impact of extreme temperatures on human health. The finding of differential mortality within a metropolitan area has implications for improvements to both short- and long-term measures aimed at reducing the public health burden associated with extreme heat. Future heat-warning systems, for example, could trigger a more localized series of intervention measures. Recent studies have documented the effectiveness of various longer-term strategies aimed at reducing thermal stress through the use of improved insulation, double-paned windows, reflective coverings for exterior surfaces, green space, and more.[32,33] Public programs that support such infrastructural enhancements might first focus on the regions where the health burden related to high temperatures is the greatest. We suggest that resources be proportionally allocated based on geographical patterns in susceptibility.

The approach we have presented is applicable for other cities with similar spatial subdivisions of the population and geographically-specified mortality records. In the United States, some mortality records may be available at the finer census tract scale. The results of this study are likely sensitive to the specific postal code boundaries, an issue commonly facing spatial geographers and ecological design.[34] There are other limitations of this work that could be differently addressed by future researchers. We did

not include modification of the temperature-health relationship by air pollutants. Previous work has documented that there may be some effect modification, namely by ozone and particulates, but the air pollution effect may be small compared to the temperature effect during extreme summer events.[35,36] This study has also focused on single day heat events without consideration of cumulative effects and/or lag structures, and there is an added health burden when high temperatures are observed several days in succession.[16,37,38] We did not examine whether the response differs by age groups, which could provide additional information regarding the appropriateness of various intervention strategies in different locations. Finally, these results differ somewhat from those previously presented for Philadelphia County where seasonality and long-term trends were modeled using direct standardization based on the changing population and age structure of the population. [17]

Quantifying the varying effect of temperature across small spatial units could also be addressed with a spatiotemporal Bayesian hierarchical model, as implemented in a wide range of other studies adopting an ecological design [39,40]. Given the contrasting theoretical frameworks and estimates of uncertainty, along with differences in necessary computer processing time, it would be useful to compare such an approach to the one we have presented here. Ongoing work by the authors is directed at applying this methodology for a suite of other cities in different climatic regions as well as relating the spatial patterns in heat-related mortality to various environmental and sociodemographic variables. Focusing on a higher spatial resolution can improve the understanding of the

factors linked to heat-related mortality, subsequently leading to more effective strategies to mitigate the adverse health effects of extreme heat.

ACKNOWLEDGEMENTS

The authors thank Bruce Hayden and John Porter at the University of Virginia, Larry Kalkstein at the University of Miami, Karen Smoyer-Tomic at HealthCore, Inc., and Adrian Barnett of Queensland University of Technology for many helpful contributions. The mortality data were supplied by the Bureau of Health Statistics & Research, Pennsylvania Department of Health, Harrisburg, Pennsylvania. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions.

FUNDING

This work was supported by a Science to Achieve Results Graduate Research Fellowship from the U.S. Environmental Protection Agency (Grant FP-91733701-1), the Virginia Space Grant Consortium's Graduate Fellowship program, and the U.S. National Science Foundation (NSF) and Swedish Research Council (Vetenskapsrådet) through the Nordic Research Opportunity for NSF Graduate Research Fellows.

FIGURES

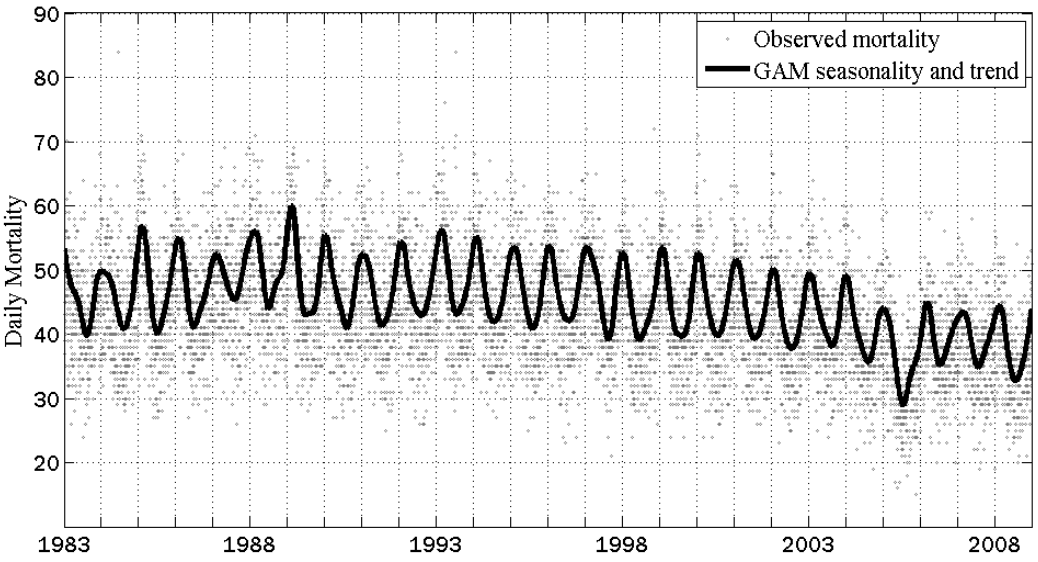


Figure 1. The time series of daily all-cause mortality for Philadelphia County, PA, USA, 1983–2008. The gray dots represent daily observations and the solid black line represents the seasonality and trend component (5 degrees of freedom per year) of a generalized additive model with the effect of temperature removed. On this plot the seasonality and trend component is plotted assuming a temperature of 10°C.

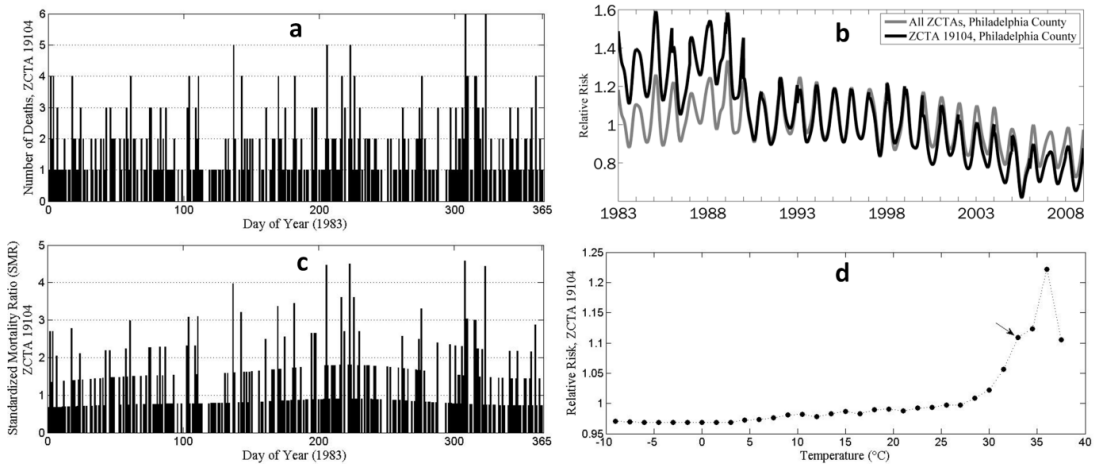


Figure 2. A graphical illustration of the described method for estimating postal code-level temperature effects. Panel (a) is a plot of the daily deaths in postal code 19104 in the first year of the period of record. In panel (b) the postal code-level seasonality and trend estimates are shown compared to those for the entire County. In panel (c) the raw death counts in (a) have been transformed to a standardized mortality ratio by dividing the raw count by the baseline mortality rate shown in panel (b). Panel (d) shows the estimated relative risk of mortality when the one-day lag afternoon maximum temperature exceeds a range of thresholds incrementing by 1.5°C. The dashed line was added to aid in visual inspection only. The arrow identifies the point on the postal code-level dose-response curve corresponding to the Countywide temperature threshold (and map of effects, Figure 5).

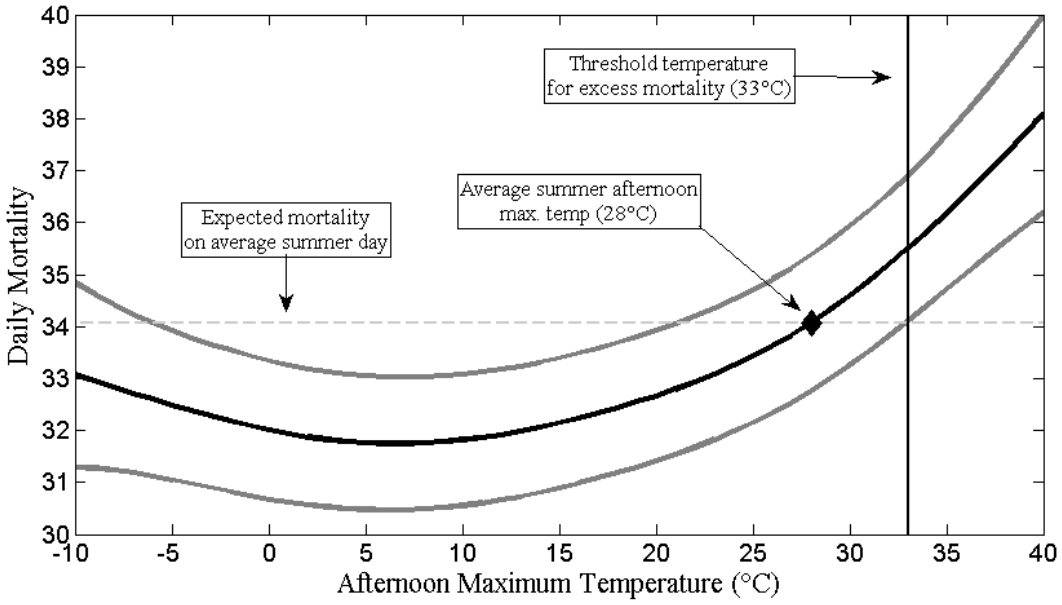


Figure 3. The one day lagged temperature-mortality relationship for Philadelphia County, PA, USA, derived from a generalized additive model using 26 years of daily mortality records with seasonality and long-term trends removed and daily afternoon maximum temperature. The solid black line represents the model predicted daily mortality and the gray lines represent the 95th percentile confidence interval for the predicted values. The vertical bar indicates the temperature threshold at which mortality becomes significantly higher than that associated with mean summertime conditions (represented by the horizontal dashed line). The predicted values shown assume a reference background mortality rate of 34 deaths per day, the average summer mortality over the last three years of the time series.

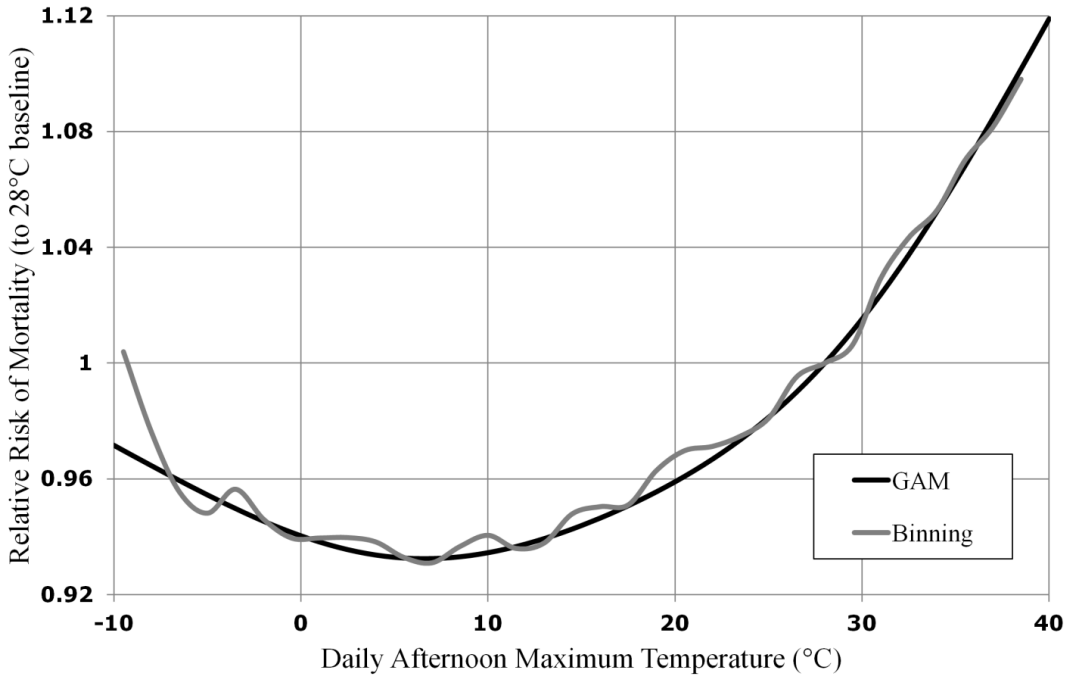


Figure 4. A comparison of the one day lagged temperature-mortality relationship for Philadelphia County estimated by a generalized additive model (black line) and the same relationship estimated by a binning method that examines the mortality response for overlapping temperature intervals. In both cases seasonality and long-term trends in the mortality record were removed with a generalized additive model.

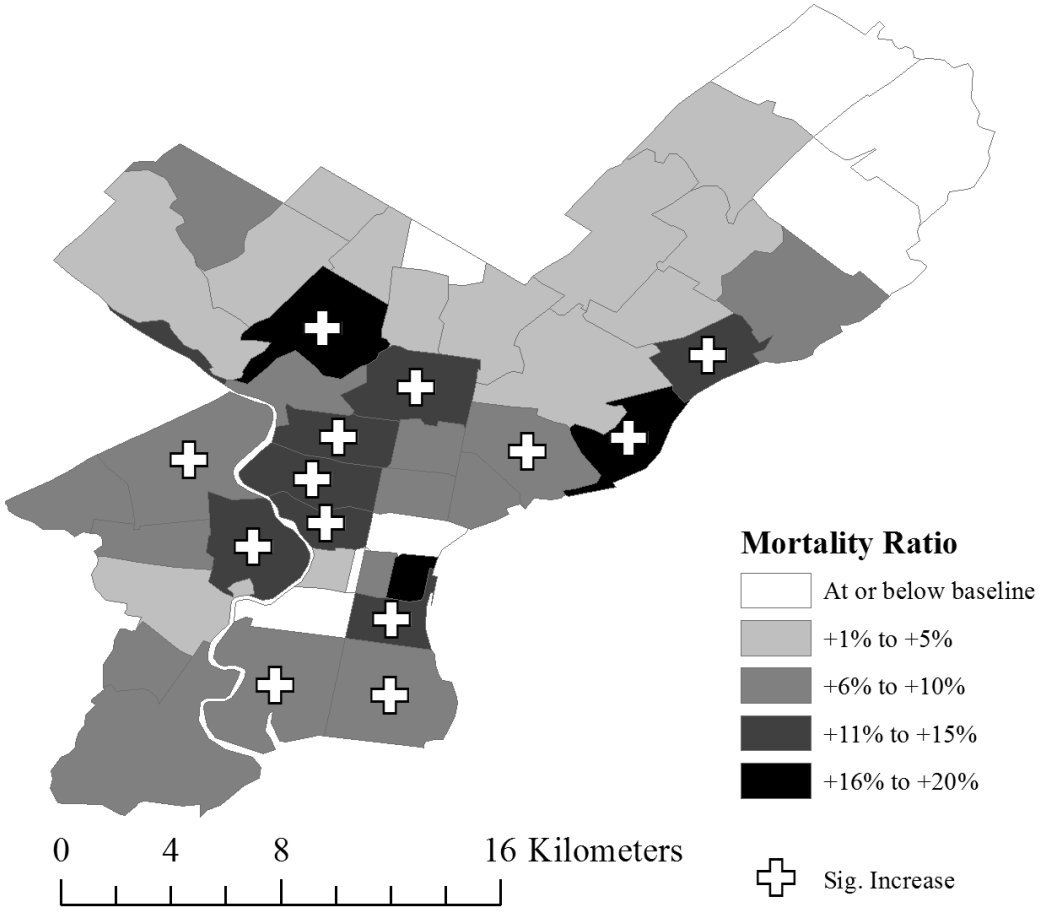


Figure 5. Mortality rates within forty-six postal codes comprising Philadelphia County for a sample of 192 days immediately following extreme heat days on which the afternoon maximum temperature exceeded 33°C. The mortality rate is shown as a percentage relative to normal summertime conditions (34 deaths/day for the entire county for the most recent three years, see Figure 2). ZCTAs with a statistically significant elevation in mortality on extreme heat days are identified with a plus sign (+).

References

1 Vaneckova P, Beggs PJ, Jacobson CR. Spatial analysis of heat-related mortality among the elderly between 1993 and 2004 in Sydney, Australia. Soc Sci Med 2010;**70**(2):293-304.

2 Goggins WB, Chan EYY, Ng E, et al. Effect Modification of the Association between Short-term Meteorological Factors and Mortality by Urban Heat Islands in Hong Kong. PLoS ONE 2012;**7**(6):e38551+.

3 Anderson GB, Bell ML. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. communities. Environ Health Persp 2011;**119**(2):210-218.

4 Barnett AG, Hajat S, Gasparrini A, et al. Cold and heat waves in the United States. Environ Res 2012;**112**:218-224.

5 Davis RE, Knappenberger PC, Michaels PJ, et al. Seasonality of climate-human mortality relationships in US cities and impacts of climate change. Clim Res 2004;**26**:61-76.

6 Gosling SN, Lowe JA, McGregor GR, et al. Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. Climatic Change 2009;**92**(3):299-341.

7 Semenza JC, Rubin CH, Falter KH, et al. Heat-Related Deaths during the July 1995 Heat Wave in Chicago. N Engl J Med 1996;**335**(2):84-90.

8 Åström D, Bertil F, Joacim R. Heat wave impact on morbidity and mortality in the elderly population: A review of recent studies. Maturitas 2011;**69**(2):99-105.

9 Stafoggia M, Forastiere F, Agostini D, et al. Vulnerability to Heat-Related Mortality. Epidemiology 2006;**17**(3):315-323.

10 Bouchama A, Dehbi M, Mohamed G, et al. Prognostic Factors in Heat Wave–Related Deaths: A Meta-analysis. Arch Intern Med 2007;**167**(20):2170+.

11 Reid C, O'Neill M, Gronlund C, et al. Mapping Community Determinants of Heat Vulnerability. Environ Health Persp 2009;117(11):1730–6.

12 Hondula DM, Vanos JK, Gosling SN. The SSC: A Decade of Climate-Health Research and Future Directions. Int J Biometeorol 2013, available online, print publication forthcoming.

13 Stone B, Rodgers MO. Urban Form and Thermal Efficiency: How the Design of Cities Influences the Urban Heat Island Effect. J Am Plann Assoc 2001;**67**(2):186-198.

14 Harlan SL, Brazel AJ, Prashad L, et al. Neighborhood microclimates and vulnerability to heat stress. Soc Sci Med 2006;**63**(11):2847-2863.

15 Armstrong B. Models for the relationship between ambient temperature and daily mortality. Epidemiology 2006;**17**(6):624-631.

16 Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. Statist Med 2010;**29**(21):2224-2234.

17 Hondula DM, Davis RE, Leisten M, et al. Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: A case-series analysis. Environ Health. 2012;**11**(1):16+.

18 Sheridan SC, Dolney TJ. Heat, mortality, and level of urbanization: measuring vulnerability across Ohio, USA. Clim Res 2003;**24**:255-265.

19 Johnson D, Wilson J, Luber G. Socioeconomic indicators of heat-related health risk supplemented with remotely sensed data. Int J Health Geogr 2009;**8**(1):57+.

20 Smargiassi A, Goldberg MS, Plante C, et al. Variation of daily warm season mortality as a function of micro-urban heat islands. J Epidemiol Commun H 2009;**63**(8):659-664.

21 Loughnan ME, Nicholls N, Tapper NJ. The effects of summer temperature, age and socioeconomic circumstance on Acute Myocardial Infarction admissions in Melbourne, Australia. Int J Health Geogr 2010;**9**(1):1-11.

22 Hattis D, Ogneva-Himmelberger Y, Ratick S. The spatial variability of heat-related mortality in Massachusetts. Appl Geogr 2012;**33**:45-52.

23 Reid CE, Mann JK, Alfasso R, et al. Evaluation of a Heat Vulnerability Index on Abnormally Hot Days: An Environmental Public Health Tracking Study. Environ Health Persp 2012;**120**(5):715-720.

24 Hastie TJ, Tibshirani RJ. Generalized Additive Models. Vol 43. Chapman & Hall, 1990.

25 Wood SN. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC Press, 2006.

26 Peng RD, Dominici F. Statistical Methods for Environmental Epidemiology with R: A Case Study in Air Pollution and Health. 1st ed. Springer; 2008.

27 de'Donato FK, Stafoggia M, Rognoni M, et al. Airport and city-centre temperatures in the evaluation of the association between heat and mortality. Int J Biometeorol 2008:**52**(4):301-310.

28 Tong S, Wang XY, Guo Y. Assessing the Short-Term Effects of Heatwaves on Mortality and Morbidity in Brisbane, Australia: Comparison of Case-Crossover and Time Series Analyses. PloS one 2012:**7**(5):e37500

29 The MathWorks, Inc. MATLAB release 2011b (version 7.13). Natick, Massachusetts, USA.

30 R Development Core Team. R: a language and environment for statistical computing. 2012. R Foundation for Statistical Computing, Vienna, www.R-project.org.

31 RStudio. RStudio: Integrated development environment for R (Version 2.15.1). 2012. Boston, MA, USA.

32 Solecki WD, Rosenzweig C, Parshall L, et al. Mitigation of the heat island effect in urban New Jersey. Environmental Hazards 2005;**6**(1):39-49.

33 Smith C, Levermore G. Designing urban spaces and buildings to improve sustainability and quality of life in a warmer world. Energ Policy 2008;**36**(12):4558-4562.

34 Openshaw S. The modifiable areal unit problem. Norwick, Norflok: Geo Books 1983.

35 Rainham DG, Smoyer-Tomic KE. The role of air pollution in the relationship between a heat stress index and human mortality in Toronto. Environ Res 2003:**93**(1):9-19.

36 Stafoggia M, Schwartz J, Forastiere F. Does Temperature Modify the Association between Air Pollution and Mortality? A Multicity Case-Crossover Analysis in Italy. Am J Epidemiol 2008;**167**(12):1476-1485.

37 Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. Epidemiology 2009;**20**(2):205-213.

38 Rocklov J, Barnett AG, Woodward A. On the estimation of heat-intensity and heat-duration effects in time series models of temperature-related mortality in Stockholm, Sweden. Environ Health 2012;**11**:23.

39 Lowe R, Bailey TC, Stephenson DB, et al. The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. Stat Med 2012:**32**(5):864-883.

40 Richardson S, Abellan JJ, Best N. Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). Stat Methods Med Res 2006:**15**(4):385-407.

SUPPLEMENTAL MATERIAL

	Population (Year 2000)	Mean Daily Deaths	Minimum	5th Percentile	25th Percentile	Median	50th Percentile	75th Percentile	95th Percentile	Number of Zeros (n=9490)
Philadelphia County	1,458,614	43.16	15	30	37	43	49	57	84	0
ZCTA 19102	4,094	0.06	0	0	0	0	0	1	2	8,966
ZCTA 19103	17,668	0.51	0	0	0	0	1	2	5	5,699
ZCTA 19104	47,924	1.04	0	0	0	1	2	3	10	3,482
ZCTA 19106	9,078	0.10	0	0	0	0	0	1	3	8,606
ZCTA 19107	11,025	0.17	0	0	0	0	0	1	3	8,024
ZCTA 19111	57,784	1.83	0	0	1	2	3	4	9	1,611
ZCTA 19114	30,041	0.82	0	0	0	1	1	3	6	4,123
ZCTA 19115	30,756	1.11	0	0	0	1	2	3	7	3,144
ZCTA 19116	31,949	0.77	0	0	0	1	1	2	6	4,411
ZCTA 19118	9,588	0.35	0	0	0	0	1	2	5	6,754
ZCTA 19119	28,817	0.85	0	0	0	1	1	3	6	4,054
ZCTA 19120	60,737	1.35	0	0	0	1	2	4	8	2,610
ZCTA 19121	33,429	1.42	0	0	0	1	2	4	8	2,413
ZCTA 19122	18,908	0.46	0	0	0	0	1	2	5	6,015
ZCTA 19123	9,427	0.32	0	0	0	0	1	1	4	6,920
ZCTA 19124	63,384	1.70	0	0	1	1	2	4	9	1,805
ZCTA 19125	23,345	0.71	0	0	0	1	1	2	7	4,649
ZCTA 19126	18,275	0.52	0	0	0	0	1	2	6	5,707
ZCTA 19127	5,946	0.16	0	0	0	0	0	1	3	8,121
ZCTA 19128	35,098	1.07	0	0	0	1	2	3	7	3,289
ZCTA 19129	12,293	0.26	0	0	0	0	0	1	3	7,365
ZCTA 19130	23,273	0.62	0	0	0	0	1	2	5	5,116
ZCTA 19131	46,366	1.53	0	0	1	1	2	4	9	2,048
ZCTA 19132	38,403	1.66	0	0	1	1	2	4	9	1,859
ZCTA 19133	28,362	0.69	0	0	0	0	1	2	5	4,779
ZCTA 19134	53,622	1.75	0	0	1	2	3	4	9	1,768
ZCTA 19135	28,542	0.90	0	0	0	1	1	3	6	3,877
ZCTA 19136	36,922	1.07	0	0	0	1	2	3	7	3,215
ZCTA 19137	6,563	0.29	0	0	0	0	1	1	4	7,091
ZCTA 19138	29,164	0.88	0	0	0	1	1	3	7	3,963
ZCTA 19139	42,611	1.56	0	0	1	1	2	4	9	2,066
ZCTA 19140	53,034	1.48	0	0	1	1	2	4	9	2,244
ZCTA 19141	32,058	0.97	0	0	0	1	2	3	6	3,665
ZCTA 19142	27,372	0.73	0	0	0	1	1	2	6	4,616
ZCTA 19143	68,447	1.97	0	0	1	2	3	5	10	1,421
ZCTA 19144	47,375	1.26	0	0	0	1	2	3	7	2,751
ZCTA 19145	43,053	1.45	0	0	1	1	2	4	8	2,226
ZCTA 19146	35,773	1.30	0	0	0	1	2	4	9	2,733
ZCTA 19147	32,263	0.94	0	0	0	1	1	3	7	3,751
ZCTA 19148	46,497	1.57	0	0	1	1	2	4	11	2,033
ZCTA 19149	48,603	1.27	0	0	0	1	2	3	7	2,705
ZCTA 19150	25,838	0.55	0	0	0	0	1	2	4	5,527
ZCTA 19151	29,703	0.84	0	0	0	1	1	3	7	4,133
ZCTA 19152	28,957	1.33	0	0	0	1	2	3	7	2,555
ZCTA 19153	12,183	0.25	0	0	0	0	0	1	3	7,400
ZCTA 19154	34,069	0.69	0	0	0	0	1	2	7	4,776

Table S1. Descriptive mortality statistics for the 46 postal codes in Philadelphia, PA examined in this study. The "Number of Zeros" refers to the number of days in the period 1983–2008 within each postal code with a daily death count equal to zero.

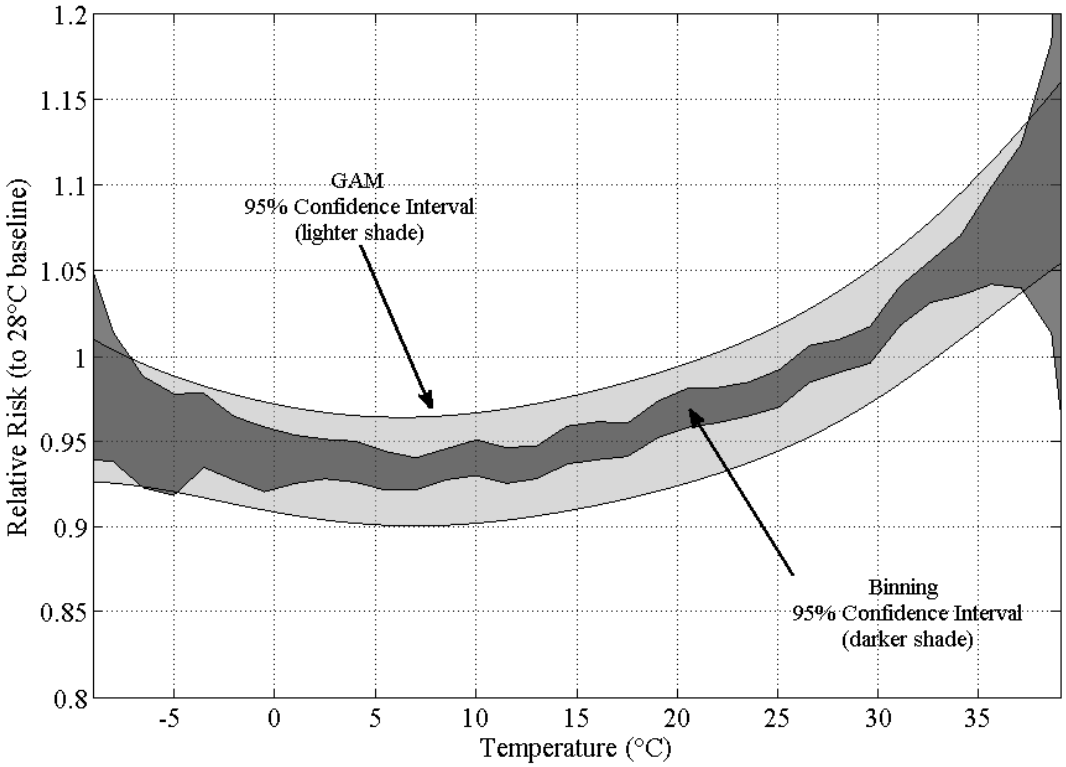


Figure S1. A comparison of the confidence intervals for the one day lagged temperature-mortality relationship for Philadelphia County. The darker shading indicates the confidence interval estimated using a binning method and the lighter shading indicates the interval estimated using a generalized additive model. Both models included a term to model the seasonality and long-term trends in mortality rates.

Chapter 4. Geographic Dimensions of Heat-Related Mortality in Seven U.S. Cities

DM Hondula, RE Davis, MV Saha, CR Wegner, LM Veazey

Manuscript to be submitted to Environmental Health Perspectives

Background: Heat-related mortality is a leading weather-related cause of death, and spatially targeted intervention measures may help protect the public when dangerous conditions occur. Although it is believed that heat-related risk varies from place to place within metropolitan areas, there is little quantitative evidence of such variability.

Objectives: We sought to identify places within large U.S. metropolitan areas where the mortality rate is highest when extreme heat occurs. We then aimed to identify key variables associated with high-risk areas and determine if the set of explanatory variables was consistent from one city to another.

Methods: Temperature-mortality relationships for seven large cities in the United States were determined with a Poisson regression model based on daily georeferenced mortality data from a twenty-year period. The models accounted for long-term time trends in mortality and mortality seasonality. We used these models to identify a statistical threshold temperature within each city when mortality becomes abnormally high, and then quantified mortality rates for the postal codes comprising each city on above-threshold days. All subsets and principal components regression were used to determine the demographic and environmental variables most closely associated with intra-city variability in heat-related risk.

Results: A threshold temperature, above which mortality statistically significantly increased above normal summer rates, was found for six of the seven study cities. Mortality increased by an average of 2.9–6.6% on above-threshold days. Intra-city mortality rates on above-threshold days were not spatially uniform. Statistically significant increases in mortality were found in 12–44% of the postal codes comprising each city. Postal codes associated with greater risk were associated with more developed land, lower income, more elderly, more young children, lower educational attainment, and prevalence of various races, but the specific variables included in regression models varied from one city to another. Multivariate regression models accounted for 14.1–33.5% of the spatial variability in heat-related mortality. Principal components regression models accounted for less variability and highlighted a different set key risk factors in some cities. A principal component reflective of general socioeconomic status was included as a key risk factor in models for five cities.

Conclusions: Historical data demonstrate considerable intra-city variability in mortality rates on warm summer days within six large U.S. metropolitan areas. There is quantitative evidence that residents of certain locations are at greater risk from extreme heat and could benefit from spatially targeted intervention strategies. As the characteristics of these high-risk locales vary from one city to another, public health plans focusing on extreme heat vulnerability should be tailored to individual jurisdictions.

Introduction

Forecasts of more severe and frequent heat waves in the future have captured the attention of public health officials and researchers. Extreme heat already ranks as the leading weather-related cause of death in the United States (Luber et al. 2006), and the possibility that the related public health burden will increase in the future has motivated a range of stakeholders to pursue new strategies to protect citizens (Lowe et al. 2011, Yardley et al. 2011). We use multidecadal, geographically referenced medical records to test two hypotheses: (1) that heat-related mortality is spatially variable within major metropolitan areas, and (2) that places with higher percentages of elderly residents, lower incomes, and a greater extent of built environment features will be those with the highest rates of heat-related mortality. If confirmed, such variability would encourage more fine-scale, geographically targeted intervention measures when dangerous conditions are forecast and occur.

Much of the existing knowledge on the relationship between high temperatures (and humidity) and human health is derived from studies using aggregated data representing the entire populations of large cities. The discovery that the relationship varies from one city to another (e.g., Kalkstein and Davis 1989, Curriero et al. 2002) was instrumental in advancing the city-specific heat-health warning systems currently operating across the globe (Sheridan and Kalkstein 2004, Hondula et al. 2013). An underlying premise of these systems is that the population's sensitivity to high temperatures varies spatially, and thus the threshold temperature for activating warning systems and deploying resources for interventions should also vary. Thus, information about geographic variability in the

response *between* cities is already motivating spatially targeted intervention and mitigation activities.

There is ample evidence supporting the notion that the response to high temperatures might also be spatially variable *within* cities. Important determinants of this intra-city variability in heat-related risk include vulnerability related to social and economic characteristics and fine-scale differences in exposure to high temperatures arising from urbanization effects.

A number of social and economic characteristics have previously been identified as contributing to elevated risk for heat-related mortality. Age is commonly cited as the most important risk factor, as the elderly have diminished thermoregulatory capacity related to a lesser ability to sweat and potential complications arising from co-morbidities and medications (Koppe et al. 2004, Gosling et al. 2009). Infants and young children have also been found to be at risk, in part due to underdeveloped thermoregulatory systems, but also from accidental overheating when left locked and unattended in vehicles (Green et al. 2001). Individuals with lower incomes have been found to be at higher risk, which may occur because they are less likely to own or be able to use air conditioning (Kilbourne et al. 1982, Naughton et al. 2002, Harlan et al. 2013). Those who live alone or are socially isolated from the community may not receive warning messages or healthcare assistance in a timely fashion when heat events occur (Semenza et al. 1996, Kalkstein and Sheridan 2007). Pre-existing medical conditions reduce the capacity of some individuals to physiologically manage stressful environmental conditions (Semenza et al. 1996, Naughton et al. 2002). When examined collectively, these variables and

others can lead to a spatially complex risk-scape for heat-related mortality strictly arising from characteristics of the population (Reid et al. 2009).

Environmental factors can also play a role in modifying health risks arising from extreme heat within urban areas. The differing thermal properties of built environment features, relative to the natural landscape, create microclimates within cities where temperatures can reach 5–10°C higher than surrounding areas under certain meteorological conditions. This contrast arises because of the manner in which buildings and paved surfaces absorb and release energy, anthropogenic heat output from buildings and vehicles, and differences in surface moisture and evaporation rates related to surface permeability (Oke 1982). When an extreme heat event occurs, the severity of the event thus varies for individuals living and working in different parts of the city. Those living in places most impacted by the urban heat island will likely experience higher temperatures and greater thermal stress. The potential negative health impacts of the urban heat island have motivated a wide range of infrastructure investments aimed at promoting more thermally tolerable urban environments (Corburn 2009). But there is little empirical evidence in the literature that intra-city temperature variations contribute to intra-city variability in morbidity or mortality rates during extreme heat events (Smargiassi et al. 2009, Laaidi et al. 2012).

Given the lack of information on intra-city variability in heat-related risk using observed mortality data, our goals are: (1) To identify locations within cities where the mortality rate is highest during extreme heat events; (2) To understand the socioeconomic and environmental factors associated with high-risk zones and their applicability from one

place to another; (3) To demonstrate quantitative linkages between built environment features and heat-related health outcomes. At a more general level, such information can facilitate more targeted and effective intervention measures by helping health and emergency management officials determine where and how they should allocate public resources to combat negative consequences of extreme heat events (Ebi and Schmier 2005).

Methods

Data sources

Daily mortality data including the postal code of residence of the decedent were obtained for seven major metropolitan areas in the United States (Atlanta, Georgia; Boston, Massachusetts; Minneapolis-St. Paul, Minnesota; Philadelphia, Pennsylvania; Phoenix, Arizona; Seattle, Washington; St. Louis, Missouri) that span multiple climate zones. Data were sourced from the respective state departments of health. On average, 22 years of data were available for each city; the period of records varied slightly based on data availability (Table 1). In total, 2,117,584 cases were examined. There were no periods of missing data. Spatial boundaries for each city were chosen to include the populated metropolitan core and immediate surrounds. The number of postal codes included per city ranged from 63 to 101.

Hourly meteorological data for each city were obtained for the appropriate time period from the archives of the United States National Climatic Data Center (http://www.ncdc.noaa.gov). The data were selected from the first-order weather station closest to each study city with complete periods of record, which typically came from a major airport proximate to the study boundaries. The air temperature, dew point temperature, and wind speed time series had few missing values (<1% of all observations for each station). There is a lack of consensus in the literature regarding the optimal exposure variable to use when examining warm-season temperature-mortality relationships (Barnett et al. 2010, Vaneckova et al. 2011). We calculated the daily maximum afternoon temperature, in which "afternoon" was defined as the five-hour

window centered on the hour of average maximum temperature. We also calculated the daily afternoon maximum apparent temperature, a variable that combines the effects of temperature, humidity and wind, following a parameterization of the Steadman model (Steadman 1984, Hondula et al. 2012).

Data representing the age, economic status, and other social and demographic indicators (Table 2) were obtained from the United States Census Bureau from the year 2000 census. The data were downloaded for year 2000 Zip Code Tabulation Areas (ZCTAs, henceforth "postal codes") from the National Historical Geographic Information System portal (https://www.nhgis.org) (Minnesota Population Center 2011). Geographic boundary files for the postal codes were obtained from the same location. Land cover information was obtained from the National Land Cover Database (NLCD) through the Multi-Resolution Land Characteristics Consortium (http://www.mrlc.gov/index.php). The NLCD includes 39 different classification types for 80 m^2 pixels spanning the United States, including three separate types representative of the built environment (low, medium, and high-intensity development). In these developed land cover types, 20– 49%, 50–79%, and more than 80% of the total cover are associated with impervious surfaces, respectively. The high intensity type includes apartment buildings, industrial and commercial use, and row houses, whereas the low and medium types typically reflect areas with single-family homes (Fry et al. 2011). We calculated the percentage of each land cover type within each postal code using the Zonal Statistics tool in ESRI ArcMap version 10.0. The non-developed NLCD types present in the study cities were subsequently aggregated into broader categories (Table 2).

Postal code temperature-mortality estimates

Postal-code level heat-related mortality estimates are derived from a multi-stage statistical procedure that accounts for seasonality and long-term time trends. In the first stage, the time and temperature components of the time series aggregated across all postal codes within each city are estimated with a generalized additive model (Wood 2006). Each term is estimated with a fixed natural smoothing spline; seven degrees of freedom per year are used for the time component and five for the temperature component. The model uses a quasi-poisson link function to account for potential overdispersion in the mortality time series and takes the form:

$$Log[E(M)] = \alpha + s(T_{PM}, df = 5) + s(Time, df = 7 * n(y)) \quad \text{[Equation 1]}$$

where daily mortality *E(M)* is a modeled as a nonlinear smoothed function of daily maximum afternoon temperature T_{PM} with five degrees of freedom and a temporal counter *Time* for seasonality and long-term trends with seven degrees of freedom per year *y*. The temporal component of the model is shown in Figure 1.

A threshold temperature is determined from each city's aggregate (non-spatial) model by determining the lowest temperature at which mortality is significantly greater than what is expected for mean summer conditions. The mean summer temperature is calculated for the months June–August. The relative risk of mortality is calculated for all temperatures above the summer mean, and the threshold is the lowest temperature at which the

confidence interval of the relative risk does not include 1.0. The relative risk and confidence intervals are calculated using the equations:

$$RR_{T_x} = e^{\left(\left(\hat{M}(T_x) - \hat{M}(T_m)\right)\right)} \quad \text{[Equation 2]}$$

$$95\%\ CI(RR_{T_x}) = e^{\left(\left(\hat{M}(T_x) - \hat{M}(T_m)\right) \pm 1.96 \times \sqrt{\varepsilon(T_x)^2 + \varepsilon(T_m)^2}\right)} \quad \text{[Equation 3]}$$

where the relative risk at a given temperature (RR_{Tx}) equals the exponential of the difference in model-predicted log-relative risk mortality at the given temperature T_x and the mean summer temperature T_m. The confidence interval accounts for the joint variance in these two estimates.

In the next stage of the model, seasonality and long-term time trends are removed from the postal-code level data. As the low daily mortality counts within each sample do not permit reliable estimation of these effects by directly modeling them analogously to the city aggregate models, a combination of city-scale and local-scale pieces of information are used to construct an approximation. The shape of the seasonality curve within each year is obtained from the city aggregate model and scaled and shifted to match each postal code-year's average summer mortality count.

The original postal code daily mortality counts are divided by the resultant estimated baseline mortality time series to yield a daily mortality ratio, which serves as the primary variable for analysis henceforth. The final stage of modeling involves determining the temperature-mortality impacts for each postal code. The average mortality ratio on all days June–September exceeding the city-specific threshold temperature (M*) is

calculated for each postal code. A randomization test is used to identify those postal codes where average mortality on above-threshold days statistically differs from what would be expected under normal summer conditions. For a given number of above-threshold days, 10,000 equally-sized subsets of postal code mortality ratios are drawn with replacement (from summer days only) and the mean of each subset is calculated. A statistically significant high-temperature effect is deemed for postal codes where M* exceeds the 95th percentile mean ratio from the randomly-generated subsets (R_{crit}) (Sheridan and Dolney 2003). If there was no statistical difference between baseline mortality and that on above-threshold days, and above-threshold day mortality at the postal code scale was spatially independent and randomly distributed, one would expect 5% of postal codes to have "statistically significantly high" mortality on above-threshold days because of making multiple comparisons. More complete details on the development of this procedure for use in spatial units with low daily mortality counts can be found elsewhere (Hondula et al. 2013).

Spatial Regression Analysis

Potential associations between spatial variability in heat-related mortality and socioeconomic and environmental factors were explored with multivariate regression models. The dependent variable was the ratio of M* to $R_{crit,}$ (new variable T). This ratio was used instead of M* because the variance in the mortality ratio time series markedly differs from one postal code to another, making it difficult to draw meaningful comparisons from M* alone. T provides a standardized estimate for each postal code regarding how exceptional mortality is on above-threshold days for that specific location.

If mortality data for each postal code were normally distributed, the dependent variable for the analysis would be the t-statistic comparing each postal code's mean mortality to a certain baseline value. Here, T is analogous but for the randomization test accounting for non-normality.

Postal codes with low T values (classified as outliers based on standard deviation criteria and visual inspection) were excluded from the regression to avoid the likelihood of isolated leverage points having unduly large impact on the overall analysis. A majority of postal codes removed had very low populations and mortality counts. No more than five postal codes were excluded from the analysis in any given city.

All subsets linear regression was used to select variables for each city that are associated with spatial variability in heat-related mortality. An exhaustive set of all possible multivariate linear regression models was generated for each city with 1–10 independent variables included. The optimal model was selected from all possible subsets using Schwarz's information criterion (BIC) (Schwarz 1978). The model with the lowest BIC was then examined for spatial autocorrelation in the residuals (Moran's I), collinearity (Jarque-Bera test), and heteroskedasticity (Koenker-Bassett test). Where necessary, this procedure was iteratively repeated for each city as variables were excluded based on non-normality heteroskedasticity, and/or collinearity with other terms included in the model. We also performed linear regression using unrotated principal components of the original pool of explanatory variables. Components were identified separately for each city, and those with eigenvalues greater than one were saved as new independent variables. We

then generated a multiple regression model using the principal components by including all components with partial significance values less than 0.05.

Finally, we examined data from all cities simultaneously by merging all postal-code level data into one single data set. The same procedures as listed above for each city were followed to generate multiple regression models using the original pool of explanatory variables and principal components. Here, the principal components analysis was completed using all cities' data combined.

The regression models were generated using the *leaps* package in RStudio version 0.96.304 (RStudio 2012) and final modeling, diagnostics, and spatial corrections were performed with GeoDa version 1.4.0 (Anselin et al. 2006). Principal components analysis was completed with IBM SPSS Statistics version 20.0.

Results

Intra-city heat-related mortality

A statistically significant positive association between high temperatures and mortality was evident in six of the seven study cities (Table 1 and Figure 2). No threshold temperature was evident in Atlanta, so Atlanta is excluded from the remainder of the analysis. Threshold temperatures for statistically significant increases in heat-related mortality varied from 1.64 ºC (Philadelphia) to 3.79 ºC (St. Louis) above the summer mean temperature. Threshold temperatures were exceeded on 13.0–27.9% of summer days during each city's study period. Model-predicted relative risks at the threshold temperature were consistent across cities, varying from 1.015 (Philadelphia) to 1.020 (Minneapolis). Larger inter-city differences were evident in the average mortality rate on days when the threshold temperature was exceeded, ranging from a 2.9% elevation in mortality rates in St. Louis to a 6.6% elevation in Philadelphia. Results were similar for apparent temperature, and as air temperature is simpler to measure and calculate, it is the exposure variable chosen for the remainder of the analysis (Barnett 2010).

Significant intra-city spatial variation in mortality rates on above-threshold days was evident in each of the six cities examined (Figure 3). Intra-city differences in mortality rates far exceeded inter-city differences. Significant increases over baseline summer mortality were present for 15 of 60 examined postal codes in Boston (25%), 13 of 100 in Minneapolis (13%), 20 of 45 in Philadelphia (44.4%), 12 of 93 in Phoenix (12.9%), 11 of 60 in Seattle (18.3%), and 7 of 60 in St. Louis (11.7%).

Spatial Regression with Original Variables

From the original set of 25 demographic and environmental variables included as potential predictors of spatial variability in heat-related mortality, eight were included in multiple regression models across the six cities (Table 3). The number of variables used for each city varied from one to four. All models and individual terms were statistically significant and all models but one (Phoenix) passed diagnostic tests for lack of spatial autocorrelation in residuals. Two other diagnostic tests did not meet statistical criteria (residual heteroskedasticity in Minneapolis and residual normality in St. Louis), but in both cases, this was related to individual outlier residuals that did not substantively impact the regression. The models accounted for 14.1–33.5% of the spatial variability in heat-related mortality.

The specific set of independent variables included in the regression models varied from one city to another. Demographic variables associated with higher risk included lower per capita income (in Boston), higher percentages of elderly residents (in Minneapolis, Philadelphia, and Seattle), higher percentages of residents of Asian heritage (in Philadelphia), higher percentages of children under age five (in Philadelphia), lower percentages of white residents (in Philadelphia), lower median housing values (in Phoenix), higher percentages of residents of Pacific Islander heritage (in Seattle), and higher percentages of residents without a high school education (in St. Louis). The environmental variable associated with higher risk was higher percentages of medium-intensity development (in Boston, Minneapolis, and Phoenix). A spatial lag term for the dependent variable was included in the model for Phoenix only, as the original model

without a spatial lag term showed significant spatial autocorrelation in the residuals. This effect was reduced but not completely removed by adding a spatial lag term to the model.

Spatial Regression with Principal Components

Significant associations between principal components of the demographic and environmental factors and intra-city heat-related mortality were also evident in all cities (Table 4). On average, six principal components were extracted for each city that represented 81% of the variance of the original pool of 25 variables (Tables A1–A6). The principal components regression models explained 6.7– 27.7% of the intra-city mortality patterns. In all six cities the first principal component was reflective of socioeconomic status, and this component was included in regression models in five of the six locations.

Three principal components (PCs) were included in the model for Boston. A positive relationship was found for Boston PC2, which had strong positive loadings for percentage of children under age 5 and low intensity development, and strong negative loadings for one-person households and high intensity development. Boston PC1 was also included in the model with a positive association. This component loaded positively on percent below poverty, percent without a high school education, and percent of several non-white races, and loaded negatively on open space, per capita income, and percent over age 65. The third and final term included in the model was Boston PC4, which had strongest positive loadings on medium intensity development and old housing. This component was also positively related to areas with higher heat-related mortality. The model for Boston explained 24.3% of the heat-related mortality pattern.

Four PCs were included in the model for Minneapolis, which accounted for 27.7% of the spatial variability in heat-related mortality. Minneapolis PC1 had a positive regression coefficient and loaded positively on percent below high school education, with public assistance, below poverty, and medium intensity development, and loaded negatively on percent white. Minneapolis PC2 loaded positively on percent elderly and percent living alone and negatively on percent under age 5; this component also had a positive regression coefficient. Minneapolis PC3 was inversely related to areas with high mortality. PC3 is low in places with high percentages of open space and low intensity development. Minneapolis PC5 loaded strongly on percent from other races and water, wetland, and forest land cover types and was positively associated with high mortality zones.

No model was generated for Philadelphia using only PCs selected from the original search criteria as none of the partial significances were less than 0.05. However, when additional PCs were tested, Philadelphia PC9 (2.9% of original variance, $\lambda = 0.676$) was significantly positively related to the mortality pattern and explained 17.6% of the spatial variability. Loadings for this PC were strongest and positive for percent Asian, percent Pacific Islander, percent over age 65, and percent of homes built before 1970 or earlier.

Three PCs were included in the model for Phoenix and accounted for 21.9% of the spatial variability in heat-related mortality. Phoenix PC2 had high positive loadings for percent living alone, percent living alone over age 65, and medium intensity development. PC2 had a positive regression coefficient. PC1, which had high positively loadings with percent below high school education, percent with public assistance, percent below

poverty, and percent nonwhite, also had a positive coefficient. A negative relationship was found with Phoenix PC4, which had positive loadings for median housing value, per capita income, and forest land cover type.

The model for Seattle included only one variable, Seattle PC1. This component was strongly positive loaded on percent with public assistance, percent below high school education, and percent of nonwhite races, and negative loaded on per capita income and median housing value. PC1 had a positive regression coefficient and explained 6.7% of the spatial variability in heat-related mortality.

The model for St. Louis also included only one variable, St. Louis PC1. This model explained 11.3% of the spatial mortality pattern. PC1 was strongly positively loaded on percent below poverty, percent with public assistance, and percent below high school education. It was negatively loaded on median housing value, per capita income, and percent white.

All-City Models

All postal codes from all cities were combined for the final set of regression models (Table 5). Using all subsets regression with the suite of original demographic and environmental variables, the optimal multiple regression model included five variables and explained 25.4% of the variance. Terms with a positive coefficient included percent of homes built prior to 1970, percent of elderly living alone, percent Pacific Islander, and percent medium intensity development. Percent white was included in the model with a negative coefficient. The strongest cross-correlation between these independent variables

was between percent white and percent of homes built before 1970 (-0.426); most of the correlations had absolute values less than 0.25.

Analogously to the individual city models, we extracted principal components from the all-city data set. Seven of the components (AllCity PCs) had eigenvalues greater than one and these PCs accounted for 75.4% of the variance in the original data set. Five AllCity PCs were included in a multiple regression model that explained 26.3% of the variance in heat-related mortality across the entire data set. Four terms in the model, AllCity PC1, PC2, PC4, and PC5 had positive regression coefficients. AllCity PC1 had strong positive loadings on percent below high school education, percent below poverty, and percent with public assistance. AllCity PC2 was loaded most strongly on percent under age 5 (negative) and percent living alone (positive). AllCity PC4 had a high positive loading for low intensity development. PC5 had strong negative loadings on percent elderly and percent grass and crops land cover type. The fifth term included in the model, AllCity PC7, had a negative regression coefficient. This component had highest positive loadings for median housing value, per capita income, and percent American Indian, and a strongest negative loading for percent forest land cover type.

Discussion

In cities where high summer temperatures lead to elevated mortality rates, there is significant spatial variability in sensitivity to heat. Mortality records from recent time

periods spanning 14–26 years show that residents of certain portions of cities have been at greater risk of dying when extreme heat occurs. To the best of our knowledge, this study is one of the first to document such intra-city variability in risk.

The modeled relationship between temperature and mortality (Figure 2) for the cities evaluated is similar to that reported elsewhere—a U- or J-shaped curve where mortality rates increase at the lowest and highest temperatures (Curriero et al. 2002, Davis et al. 2003). Unsurprisingly, we found that the threshold temperature, defined as the lowest temperature at which mortality is significant different than observed for normal summer conditions, varies geographically. The highest threshold temperature was found for the warmest study location, Phoenix, while the lowest thresholds were in the coolest locations, Seattle and Boston. No threshold temperature was found for Atlanta, Georgia, a city with high temperatures and humidity that persists throughout much of the summer. Although the modeled relationship for Atlanta (not shown) was similar in shape to the other cities, the confidence interval for the estimated effect widened considerably at the highest temperatures. No relationship (or a weak one) between summer temperature and mortality has previously been reported for Atlanta and other locations in the southeastern United States (e.g., Curriero et al. 2002). Geographical variability in the threshold temperature is consistent with previous research, and it is believed that this variability arises because people in different locations physically and technologically adapt to their climate (e.g., Davis et al. 2003).

The threshold temperatures that we identified are lower than temperatures commonly used in heat-warning systems at which various public health intervention strategies are

activated. In some of the cities we examined, the threshold temperature was only two degrees (C) above the summer mean temperature, which resulted in a large sample of days included as "hot." This is an important contrast to draw between this research and others that use more stringent criteria to identify extreme heat days such as the 95th or 99th percentile summer temperature (e.g., Gosling et al. 2007, Anderson and Bell 2009). From a statistical standpoint, temperatures only a few degrees above normal summer conditions are associated with elevated mortality rates and should be considered when evaluating the total health burden related to high temperatures and when projecting future health impacts under climate change. However, we do not advocate that these lower thresholds be used in public health alert systems, as alarming the public too frequently may result in diminished alert effectiveness.

There is strong evidence in support of the hypothesis that heat-related mortality is spatially variable within urban areas. In each of the six cities examined, significant increases in mortality when temperature exceeded the city-specific threshold were confined to only a portion of the postal codes comprising each study area. This study joins a small but growing body of research documenting such inter-city variability based on historical medical data (e.g., Schuman 1972, Smargiassi et al. 2009, Vaneckova et al. 2010, Laadi et al. 2012, Reid et al. 2012, Johnson et al. 2012, Harlan et al. 2013). A portion of this previous work has focused on single locations and/or single heat events, whereas here we have included long records of mortality data for multiple locations. The maps and results in this study are developed using the entire time period, and temporal trends in heat-related mortality were not considered in the study design. However,

ongoing work by the authors (Hondula et al. 2013) is addressing this issue to determine if certain portions of the city are consistently associated with high mortality during heat events. More research in this area would help understand those places where interventions are most needed in the coming years.

Spatial Regression with Original Variables

We found a significant relationship between the spatial pattern in mortality and various potential explanatory variables in each city, but the specific variables included in the optimal multiple regression model varied from one city to another. Only two variables were included in the model for more than one city (percent land cover with medium-intensity development and percent of elderly residents). Thus, although certain variables may be important determinants of heat-related risk regardless of location, at this spatial scale, the strongest associations between socioeconomic and environmental factors and mortality outcomes are inconsistent from place to place. Because patterns in socioeconomic and environmental factors vary between cities, there may be interactions and/or competing effects in one city that are not present elsewhere.

The key risk factors for Boston were per capita income and percent medium-intensity development. Six contiguous postal codes were associated with high heat-related mortality rates in the southern part of the study area that include areas in and near Roslindale and Mattapan. These postal codes had both high percentages of developed land and low per capita incomes. There are sharp contrasts between these areas in terms of percentage of high-intensity land cover type, indicating that high heat-related mortality rates are possible even in places that are not characteristic of a central business district.

Moving northward into the city center, two adjacent postal codes between Roxbury and the city center were associated with significant increases. Both areas had high percentages of medium- and high-intensity developed land, but they were markedly different in terms of per capita income (more than $60,000 in the more eastern of the two versus below $22,000 in the more western). In the higher-income area, approximately 15% of residents live below the poverty line, and nearly 60% of households have only one person, so it is possible that these variables are contributing to higher risk in this area despite higher overall wealth. Per capita income in the East Boston area in the postal code containing Logan airport is among the lowest in the entire city. In the remaining postal codes with high mortality rates to the north and west of the city center, incomes are typically slightly below the citywide mean and percentage developed land is higher than the citywide mean, although there are other areas with similar characteristics without high risk. Model-predicted mortality in these locations was below observations. Thus there may be a separate risk factor here that is not captured in the multivariate regression.

Statistically significant elevations in mortality on above-threshold days are evident in areas scattered throughout the Minneapolis region, but the regression model indicates two predominant covariates: percent medium intensity development and percent elderly. Postal codes that feature the characteristics indicated by the model and high heat-related mortality include those immediately south of the central business district of Minneapolis, the postal code to the south and east of Edina, and the area between Minneapolis and Columbia Heights to the north. The other postal codes with significant elevations in mortality on above-threshold days do not fit the model well, as they have both percent

elderly and percent medium intensity land cover at or below the regional mean. Two areas that particularly poorly fit include the postal code southeast of Andover in Anoka and Ramsey counties, and the postal code including the lake district in the far west of the study region. It is difficult to build a hypothesis for sources of high risk in these regions based on the literature and data examined in this study. With respect to the westernmost postal code, it is possible that recreation on hot days is a driver of elevated risk, as one study from Toronto showed higher ambulance call-outs near the lakeshore where people sought relief from the heat (Bassil et al. 2009).

Nearly half of the postal codes in Philadelphia County are associated with statistically high mortality on above-threshold days, and the density of high-risk locations is greatest in the southern and central portions of the County. Four factors were found to be associated with mortality risk, including percent elderly, percent Asian, percent under age five, and percent white (protective). Postal codes with high percentages of elderly residents and higher heat-related mortality rates are located on the perimeter of the study region, including Northwest Philadelphia and residential portions of South Philadelphia. The two postal codes with the highest percentage of elderly residents, however, located in the most northern part of the County, were not associated with elevated risk. The five postal codes with the highest percentages of Asian residents were all associated with statistically high heat-related mortality, three of which are located in the southeasternmost portion of the County. The postal codes in the center of the county associated with high mortality had highest rates of children under age 5 and lower percentages of white residents. These postal codes also have the lowest per capita

incomes in the County, but income alone was not found to be a significant predictor of mortality. The westernmost postal code, north of Upper Darby Township, and the postal code containing Southwest Philadelphia did not fit the model well.

The variable that explained the greatest amount of variability in heat-related mortality rates across Phoenix was percent medium intensity land cover; the addition of other variables did not significantly improve the model and led to higher BIC. Of the 19 postal codes with 50% or more medium intensity land cover, four are associated with high heat-related mortality rates located in and near Glendale. This contrasts with only one of the 33 postal codes with 20% or less medium intensity land cover having high risk. The highly developed, high-risk areas are also associated with low income, and previously a high correlation has been found for Phoenix between development/urban heat island effects and poverty status (Harlan et al. 2007). Three of the remaining high-risk postal codes located near Paradise Valley, east Gilbert, and Sun City West, are associated with high percentages of elderly residents. The high-risk area between Mesa and Phoenix has among the lowest mean per capita incomes in the region. Three other high-risk zones located in and near Scottsdale do not fit the model well, nor do they seem to be associated with high risk related to any other variables we considered. It is important to acknowledge that the model for Phoenix was the only one that also included a term to account for spatial autocorrelation in the dependent variable, and these spatial effects were not completely removed with the addition of a spatial lag term, leaving the possibility that the model is not properly specified.

Postal codes in Seattle with higher percentages of elderly residents and Pacific Islander residents were associated with higher risk. High-risk postal codes in the central and western portions of the study region all are associated with some of the highest rates of Pacific Islander residents in the Seattle area. However, it should be noted that Pacific Islanders represent no more than 2.5% of the population in any single postal code. The areas near Bellevue, east of Seattle, and Fort Lawton to the northwest have higher percentages of elderly residents and higher heat-related mortality rates. High-risk areas near the central business district and to the south in eastern Auburn have lower average per capita incomes and more medium- and high-intensity development, which could be contributing to higher risk, but there is little or no association between risk and either of these variables across the entire Seattle area. The high-risk area east of Bellevue has a high percentage of Asian residents, which was also found for certain high-risk areas in Philadelphia.

Two clusters of high-risk zones are evident in the St. Louis area, one including postal codes in and immediately proximate to the city proper on the eastern edge of the study region, and another in outlying areas west of Chesterfield. The only explanatory variable included in the model was percent of residents with less than a high school education, which explained over 14% of the spatial variability. Three of the high-risk postal codes located near the city proper are among the ten lowest in terms of educational attainment. Other high-risk postal codes in this area include high rates of poverty (25% or more), and high prevalence of homes constructed before 1940 (60% or more). Throughout the city proper, there are many postal codes with low educational attainment that are not

associated with higher risk, and future work might help understand those factors that account for this differentiation. On the western edge of the study region, the postal code just west of Chesterfield has the highest level of educational attainment in the region and is thus an outlier for the model, causing the normality in error diagnostics to return a significant value (Table 3). This is a highly affluent area with few elderly and newer homes, and thus the source of risk is unclear. The postal code immediately to the south has lower educational attainment and income, although both values are close to the regional mean.

Across all cities, the associations between the model-selected variables and spatial mortality patterns are generally consistent with expectations one would derive from the literature. Environmental factors were included in models in three of the six study cities. Mortality is higher in places with greater percentages of developed land in Boston, Minneapolis, and Phoenix, which aligns with other studies investigating Montreal, Paris, Phoenix, and Chicago (Smargiassi et al. 2009, Laaidi et al. 2012, Johnson et al. 2012, Harlan et al. 2013). The associations we found for built environment characteristics were for medium-intensity development; high-intensity development was not included in the model for any city. For some cities this variable was intentionally excluded because of non-normality (many postal codes have zero high-intensity development) and/or high collinearity with other variables that had stronger associations with the mortality pattern. High-intensity development may not be a strong covariate with heat-related mortality because many of the most highly developed areas within cities have commercial districts with few residents or sometimes feature very expensive residences. More research is

required to understand how inter-city differences in the built environment are associated with differences in thermal exposure for residents, but the results of this study support the hypothesis that built environment factors contribute to spatial variations in heat-related mortality in certain locales. Some evidence to the contrary in the literature was conducted at coarser spatial scales (Sheridan and Dolney 2003, Hattis et al. 2011).

Demographic factors were included in models for all six study cities and were also generally consistent with previous research. Income and age, variables that commonly appear as risk factors in the heat-related literature, also appeared as key predictors in this study. Per capita income was inversely associated with heat-related mortality rates in Boston. Income is believed to be an important determinant of risk related to heat because those living with low incomes and/or in poverty may not be able to afford air conditioning, which is among the best means of protecting oneself during periods of extreme heat (Semenza et al. 1996, Naughton et al. 2002). Air conditioning data at the postal code scale was not available for this study, and thus income may be serving as the most appropriate proxy measure. It is surprising that income or income-related measures were not included in the models for more cities. Postal codes with higher percentages of elderly residents or higher percentages of elderly residents living alone were associated with higher heat-related mortality in Minneapolis, Philadelphia, and Seattle. Postal codes with more young children also had higher mortality rates in Philadelphia. These findings are also consistent with previous research (e.g., McGeehin and Mirabelli 2001, Basu and Ostro 2008).

A number of other demographic factors were also included in the multiple regression models that have received less attention in the literature. Postal codes with a lower percentage of white residents were associated with higher heat-related mortality rates in Philadelphia, as were postal codes with higher percentages of Asian residents. In Seattle, postal codes with higher mortality rates had higher percentages of residents with Pacific Islander heritage. The effect of race on heat-related risk has been examined in some research that has drawn mixed conclusions (e.g., Kalkstein and Davis 1989, O'Neill et al. 2003, Hattis et al. 2011). In Philadelphia, we found that the percentage of white residents was highly correlated with a number of other variables including income and educational attainment, and thus for this city the variable selected may be serving as a proxy for larger-scale demographic patterns (Hondula et al. 2012). Nevertheless, the finding that the postal codes with higher heat-related mortality rates have a higher percentage of non-white residents may be useful for government officials designing intervention strategies. Lower percentages of educational attainment were found to be associated with higher risk in St. Louis, which may be reflective of either the benefits of an educated public in understanding risk related to heat and the appropriate precautionary measures to take. We note that education and income were well correlated in St. Louis.

Comparison to Principal Components Regression

Principal components regression yielded statistical models that explained, on average, less variance in each city's spatial pattern in heat-related mortality than those generated using the original explanatory variables. The variance explained from the principal

components models was greater in Minneapolis (by 1.7%) and Phoenix (by 5.2%). In the other four cities the principal components models explained 2.8%–9.2% less variance. Thus while principal components analysis offers a theoretical statistical advantage in that all of the explanatory variables are orthogonal, there was no strong evidence that these terms were any more closely associated with spatial variability in mortality than combinations of the original variables that were tested for collinearity.

The principal components included in the regression models do provide a different perspective on the postal-code level characteristics associated with variability in heat sensitivity, as they more comprehensively capture demographic and environmental patterns that are not well-represented by any single variable. For example, Boston PC2, which had high positive loadings for low intensity development and percent children under five and strong negative loadings for percent living alone and higher density development types, seems to be clearly capturing a contrast between single urban dwellers and suburban families. In the regression model the latter of these two patterns was associated with increasing risk of heat-related mortality. This association was not evident from the single variable analysis, which yielded medium intensity development and low incomes as key predictors.

Minneapolis was one of the two cities where the principal components regression had a higher adjusted R^2 than the original multiple regression, but the key predictors were found to be largely the same. Areas with more elderly residents and areas with higher development are associated with higher risk. The PC regression did yield another pattern not evident from the original regression, in that places with old housing, uneducated

residents, and residents living in poverty, were also at higher risk. Based on the Minnesota model, PC regression appears to be advantageous because it allows for different combinations of variables to be associated with variability in risk. A similar result was found for Phoenix, where the original model included only one variable but the PC regression yielded three distinct spatial patterns associated with higher risk.

A contrasting example, however, emerged in Philadelphia, where none of the principal components with eigenvalues greater than one were significantly related to the intra-city heat-related mortality pattern. Compared with the original multivariable model for Philadelphia, this PC outcome demonstrates that it is possible that none of the predominant spatial patterns in demographic or environmental variables are associated with spatial variability in risk. Only in the more subtle spatial pattern represented by PC9 was an associated with heat-related risk found. Interestingly, the variables represented by Philadelphia PC9 (percent Asian, percent elderly, and old housing) share many commonalities with the variables included in the original regression (percent elderly, percent Asian, percent under age 5, and percent nonwhite).

As was true for the original multivariable models, the portion of variance explained by principal components in Seattle and St. Louis was the lowest of the six cities examined. The key factors associated with higher risk were found to be different in both cases. In Seattle, the PC most closely associated with the spatial variability in risk was representative of general deprivation and prevalence of minority races, whereas the original model included percent elderly and percent of one specific race. In St. Louis, education was the key factor from the original models, but was not highly loaded on the

component included in the PC regression. Instead, the component captured a pattern related to collocated poverty, old houses, and prevalence of minority races. As both the original and PC models accounted for relatively low percentages of the variance, it is unsurprising that the set of key predictive factors differs, as there appeared to be no predominant spatial pattern to capture regardless of collinearity among the independent variables.

The principal components also provide additional perspective on the results of the regression using the original set of variables. The procedure included tests for collinearity among independent variables included in the models, but there was no consideration for collinearity of excluded variables. Thus, the possibility exists that variables included in the model are highly collinear with excluded variables and may be representative of a different pattern than might be suspected simply from examining only the included variables.

Originally included in the model for Boston were per capita income and medium-intensity development. Both of these variables have high loadings on Boston PC1, which is representative of general socioeconomic status and racial variability. Additional variables with high loadings on Boston PC1 include educational attainment, public assistance, poverty, and race. All of these factors could contribute to areas with lower incomes and more intense development being associated with higher risk. Boston PC1 had a strong negative loading for percent elderly, providing an example of a relationship where places with fewer elderly have higher risk because of other factors. Per capita income also had a high loading on Boston PC2, representative of the suburban/urban

contrast previously discussed. As the set of variables with high loadings on PC2 substantively differs from those with high loadings on PC1, there are at least two separate income-related patterns associated with higher risk in Boston.

Variables included in the original model for Minneapolis were percent over age 65 and medium-intensity development. Percent elderly had a high loading on Minneapolis PC2, which was the predominant variable contributing to that component. Minneapolis PC2 was also included in the principal components regression, and accordingly spatial variability in the prevalence of elderly is a pattern related to risk largely independent from other variables. Conversely, percent medium-intensity development had a high loading on Minneapolis PC1, as was the case for median housing value, per captia income, educational attainment, public assistance, poverty, and race. Medium-intensity development may be included in the Minneapolis model as an indicator of overall socioeconomic status.

Of the original variables included for Philadelphia, three (percent elderly, percent children, and percent white) have high loadings on Philadelphia PC1, which like Minneapolis is an indicator of overall socioeconomic status. The loading for elderly on this component is negative, however—there are few elderly in the most economically disadvantaged areas of Philadelphia. But risk is high in these places, captured by in the inclusion of percent under five years old in the original regression, as this variable is highly correlated with many other socioeconomic indicators. Philadelphia PC1 may not have been included in the principal components regression because, with respect to heat-related mortality, it includes variables with a contradictory relationship to risk. The same

could be said for Philadelphia PC4, which has high loadings on percent elderly and percent white, two variables with opposite signs in the original multivariate regression. Philadelphia was the only city where percent Asian was included as a predictor variable; percent Asian highly loads on Philadelphia PC2. Places with high scores for PC2 are highly urbanized (high-intensity development has a positive loading while medium- and low-intensity have a negative loading), higher median housing values, more people living alone, and few children. If being Asian is not a true driver of heat-related mortality risk, the inclusion of percent Asian in the regression model may be highlighting built environment effects or living alone as important factors.

The only variable included in the model for Phoenix was percent medium-intensity development, which is highly loaded on Phoenix PC1 and PC2. As is the case in other cities, Phoenix PC1 is representative of general socioeconomic status. The finding that risk is high in places with both more intense development and lower socioeconomic status is consistent with previous research (Harlan et al. 2007), but also hinders attributing risk to either factor independently in this location. A different spatial pattern, evident in Phoenix PC2, associates medium-intensity development, percent living alone, and percent elderly. There may be two separate spatial patterns at work that both associate medium-intensity development with heat-related risk even if built environment characteristics are not the underlying cause. The reverse is also possible, that those who are more socioeconomically disadvantaged and those living alone may not be at higher risk for those reasons, but are at greater risk because they live in more developed

locations where their exposure to heat is likely to be more severe because of urban heat island effects.

Percent elderly and percent Pacific Islander were the two variables originally associated with risk variability in Seattle. As was true in Minneapolis, the highest loading for percent elderly is on a component (Seattle PC4) that is mostly reflective of spatial variability in percent elderly. Thus percent elderly is a unique spatial pattern in Seattle. On the other hand, percent Pacific Islander has a high loading on PC1, which, like other cities, is largely a socioeconomic indicator.

The original model for St. Louis included percent below high school education, which loads highly on St. Louis PC1. As is the case elsewhere, places with high PC1 scores in St. Louis have lower socioeconomic status and a greater prevalence of older dwellings. Percent medium and high-intensity development are also highly loaded on St. Louis PC1, indicating similarity to Phoenix that the most socioeconomically disadvantaged are living in places with more intense urbanization.

Common to five of the six cities examined was the inclusion of the first component in the principal component regression model. High scores for this component were associated with lower socioeconomic status and higher mortality rates on threshold-exceeding days. Thus, in general, lower socioeconomic status was associated with higher risk in a majority of the locations we examined. However, the specific variables contributing to this component varied from one city to another, which makes it difficult to construct an overall socioeconomic indicator appropriate for each city. For example, places with high scores on PC1 tended to have more homes built before 1940 or 1970 in most locations,

but the relationship was weaker in Boston and hardly evident in Seattle. Isolation rates were also highly variable across PC1: there was a strong positive loading (0.65) for percent living alone on Minneapolis PC1, but a negative loading (-0.40) for percent living alone on Philadelphia PC1. It may be for this reason that Philadelphia PC1 was not included in the Philadelphia principal components regression model, whereas individual variables that covary with socioeconomics (percent under five, percent nonwhite) were included in the original models. Spatial variability in heat-related mortality is related to predominant spatial patterns in the independent variables across most of the cities examined, but an even more detailed understanding of interactive effects of the independent variables on heat-related risk could help generate different statistical models that account for greater portions of the spatial variability in mortality on hot days.

All-City Models

The all city models initially yielded a surprising result, in that the percentage of variance explained by both the original multivariate model and the principal components model was superior to that of many of the individual city models. This seemed to indicate universality in the risk factors of postal codes with high heat-related mortality, sharply contrasting with the finding from the city-specific models that the key risk factors varied from one place to another.

Subsequent analysis, however, revealed that the strength of the all city models likely emerges from inter-city differences in the independent variables rather than true universal relationships between certain risk factors and heat-related mortality for individual postal codes. Analysis of variance revealed that model predicted values and residuals

significantly varied by city from both the original multivariate regression and the principal components regression (Table 6). Thus, it seems that the all-city models are not properly specified, and the outcome of this analysis is analogous to an ecological fallacy, drawing inferences about variability between locations or individuals based on the results of one ecological study at one location (Piantadosi et al. 1988). The notion of an all-city model remains an interesting point for future exploration, but likely requires a different modeling approach, more contiguous spatial regions, and/or standardization of variables from one city to another. We cannot conclude from this analysis that there are informative universal predictors of places of high heat-related mortality across the cities we examined.

Synthesis

A key point arising from the results is that the specific predictor variables that best explain the spatial variability in heat-related risk vary from one city to another. This was evident by comparing models across cities using the original explanatory variables as well as those generated from city-specific principal components. This confirms the findings of a number of other researchers who have investigated intra-city (or inter-region) variability in heat-related risk, generally focusing on only one location. In Massachusetts, percent African-American and percent elderly population were found to be associated with variability in heat-related mortality rates across 29 districts comprising the state (Hattis et al. 2011). Johnson et al. (2009) found that poverty and urban heat island effects were key predictor variables for one heat event in Philadelphia. In a separate study of heat-related mortality in Philadelphia, home values and percentage of

African-American residents were identified as the best predictors (Uejio et al. 2011). No relationship was found between socioeconomic status, vegetation cover, and variability in heat sensitivity among the elderly in Sydney, Australia (Vaneckova et al. 2010). Key risk variables reported in one study of Phoenix heat-related ambulance calls included intra-city temperature variability, percentage of African-American and Hispanic residents, isolation, and household vacancy (Uejio et al. 2011). This sample of studies represents the diverse methodologies, dependent variables, time periods, scales of analysis, and conclusions reached regarding the variables most closely associated with spatial variability in heat-related risk. Here we show that different variables best account for the spatial pattern in mortality across different cities using consistent methods.

A separate, related line of inquiry has been developed that incorporates heat vulnerability indices. In an evaluation of one such index across several different locations, Reid et al. (2012) found that the association between hospitalizations and the index varied from place to place and across disease types. Johnson et al. (2012) demonstrated that a heat vulnerability index developed specifically on spatial patterns in social and environmental data for Chicago explained a large portion of spatial variability in mortality arising from the Chicago 1995 heat wave. A similar association was found using a city-specific index developed for Phoenix focusing on deaths directly related to heat exposure over a nine-year time period (Harlan et al. 2013). The indices used in these two studies, and thus the role of key predictor variables, varied, as they were developed based on each location's demographic and environmental data. We have made no assumptions regarding the potential interactive or competing effects of variables that may have a similar spatial

pattern, as is done when calculating a vulnerability index *a priori*, and we encourage more research aimed at understanding the best manner in which to design such indices and their value and reliability. The results of the present study support the general notion of Johnson et al. (2012), that heat risk mapping may be most effective when developed at a local level.

There are several components of this study that could be addressed differently in future work. To start, this study examined all single days that occurred above a given temperature threshold. We did not treat periods of consecutive above-threshold days ("heat waves") any differently, but research has demonstrated some added effect on mortality when high temperatures occur several days in succession (e.g., Anderson and Bell 2009). There was also no consideration of the date when high temperatures occurred, but it may be the case that intra-city variability in risk is different for early-season versus late-season heat events if there is geographical variability in acclimatization rates.

We did not examine modification of effects by air pollutants, which is a topic of continued active discussion in the literature (Theoharatos et al. 2010, Ren et al. 2011). If the spatial pattern in mortality risk on hot days varies between periods of good and poor air quality, health officials may be able to more efficiently target their intervention strategies. Future work may also explore how different temperature thresholds, spatial units, exposure variables, and lagged and displacement effects are manifest with respect to spatial variability in heat-related mortality. The methodological framework we have presented could serve as a starting point for such research and the extension of this type of investigation to other locations.

Other modeling techniques like geographically weighted regression or regression trees could reveal in more detail variability in the importance of certain variables across space within cities. The fact that, in some cities, there were contrasts in the key factors that emerged from the original multivariate regression and those from principal components regression, suggests that there are likely complex interactions and spatial variability in the importance of certain predictors of risk within each city. Different modeling approached like those suggested above may be able to better capture some of these patterns and help identify different combinations of factors associated with high risk. There may also be absolute and relative thresholds in the relationships between certain factors and risk of heat-related mortality, warranting a modeling approach that can account for nonlinear relationships. Public health officials in the jurisdictions we examined might consider the variables selected using the procedure we have adopted in this manuscript as initial indicators of the types of strategies that might be helpful in reducing heat-related mortality in the high-risk portions of their cities.

Conclusions

Over the past two decades, days with temperatures 2ºC or more above mean summer conditions have been associated with statistically significant increases in mortality rates in six major U.S. metropolitan areas. On these hot days, mortality rates increase by several percent in each city, with higher increases associated with higher temperatures. However, the mortality rate on hot days within each city is variable, and statistically significant increases in mortality are confined to less than 50% of the postal codes

comprising each municipality we investigated. Spatially targeted short- and long-term intervention measures may be effective in reducing the public health burden related to extreme heat. Furthermore, demographic and environmental variables are associated with spatial variability in risk, enabling public health officials and city planners to accurately target certain populations and design strategies with heat-related mitigation measures. The specific variables most closely associated with spatial variability in heat-related risk vary from one city to another. Postal codes with more elderly residents, lower per capita income, more Pacific Islanders, more Asians, fewer whites, more children, and more developed area were found to be associated with higher risk, although no more than four of these factors were found to be the most closely associated variables in any one city. Principal components regression yielded an overlapping but non-identical set of predictor variables, but these models generally accounted for lower portions of the intra-city variance.

The use of daily georeferenced health data offers the opportunity to directly identify places within large metropolitan areas where the risk has historically been greatest on hot days. The areas within cities that appear to merit targeted intervention measures may differ from those one would identify using traditional notions of the drivers of spatial variability in heat-related risk that have not been verified by the examination of fine-scale mortality records within urban areas.

Figures

Figure 1. The daily mortality time series for each city included in the study. Each observation is shown as a gray circle. Superimposed on the figure is a smoothing spline showing the temporal component of a generalized additive model constructed for each city modeling temperature and time effects.

(figure follows on subsequent page)

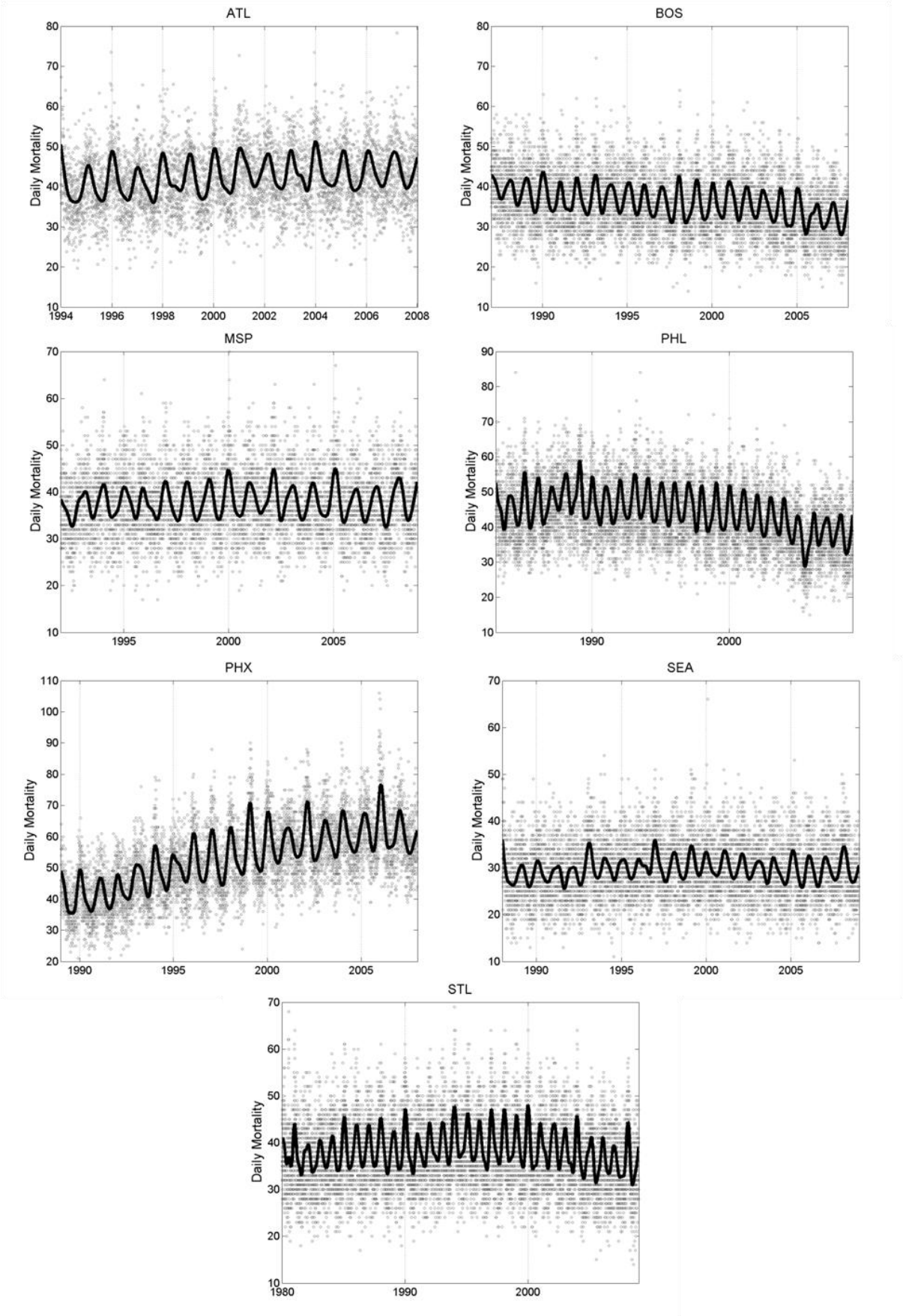
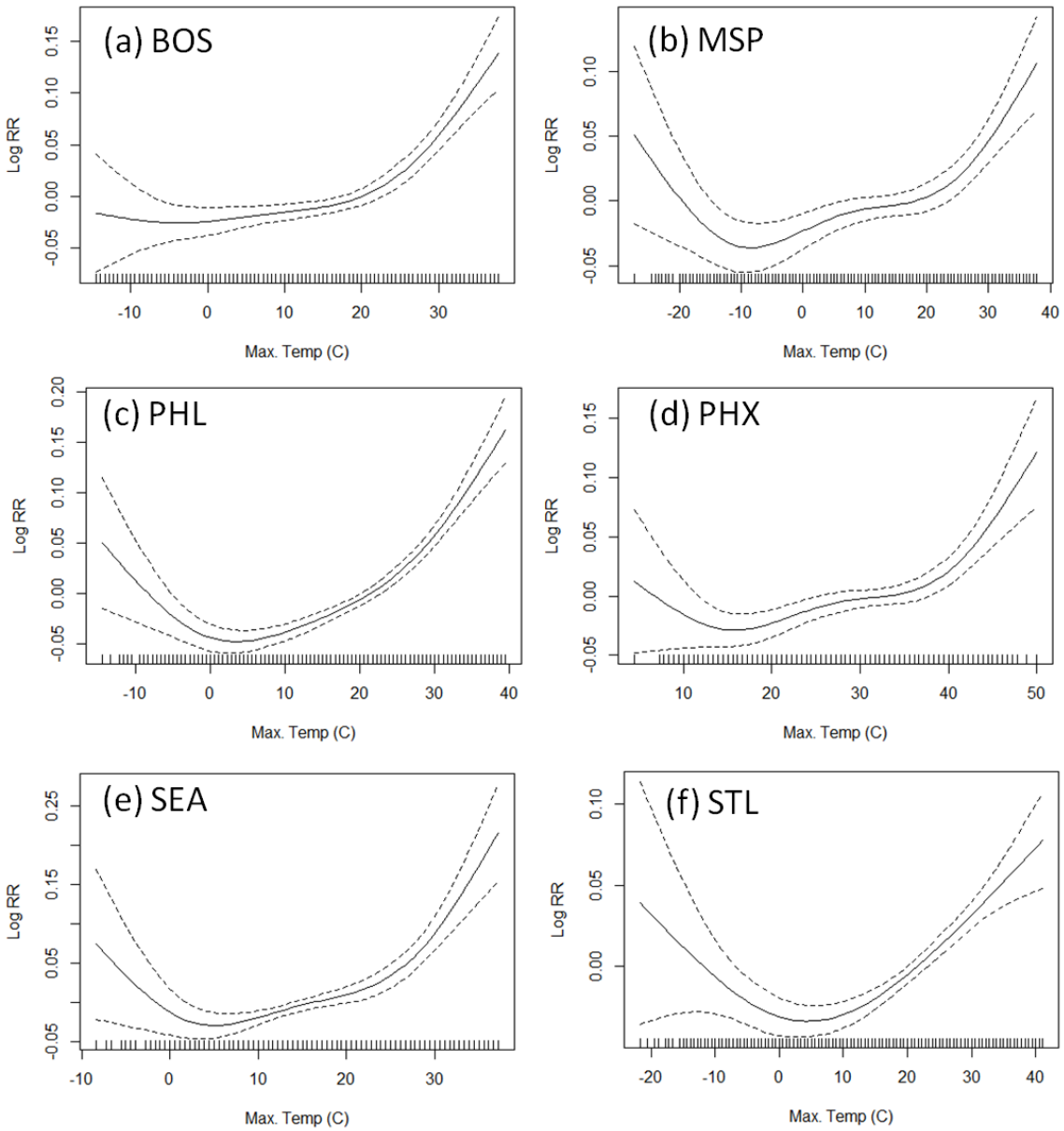
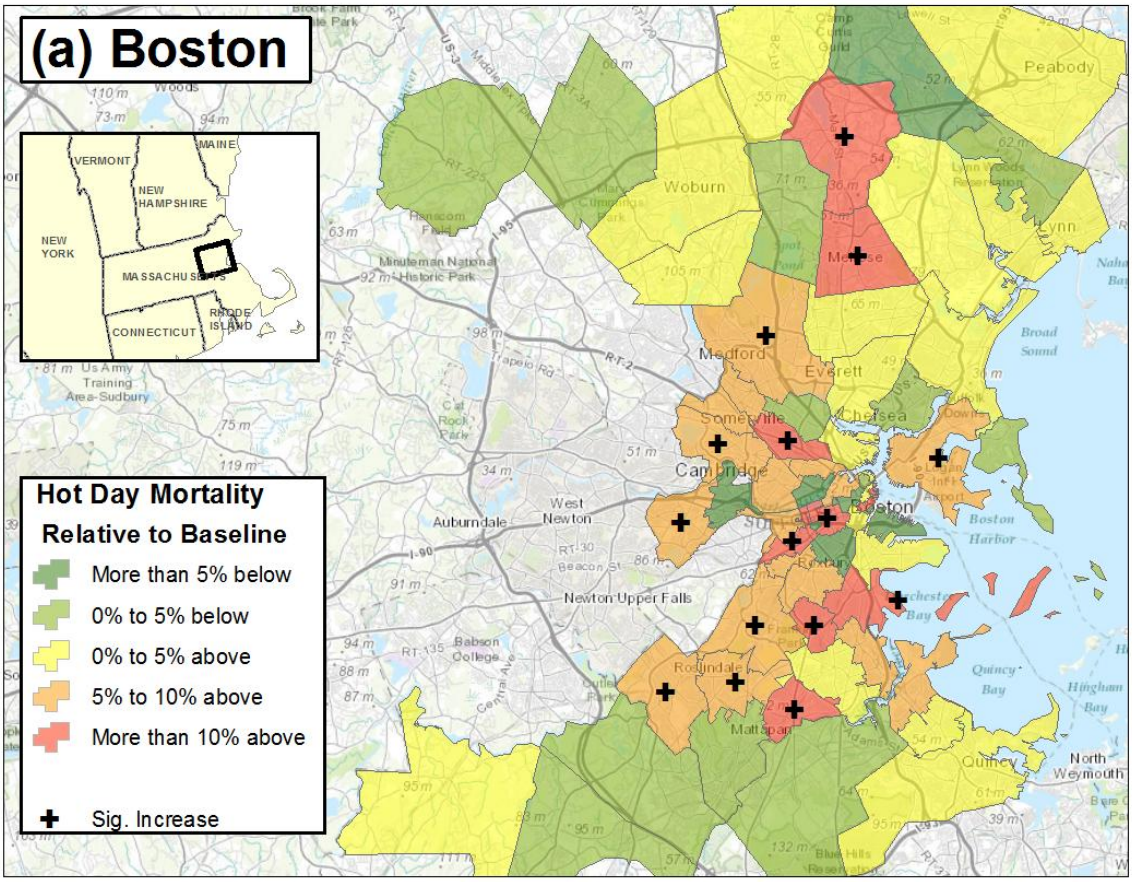
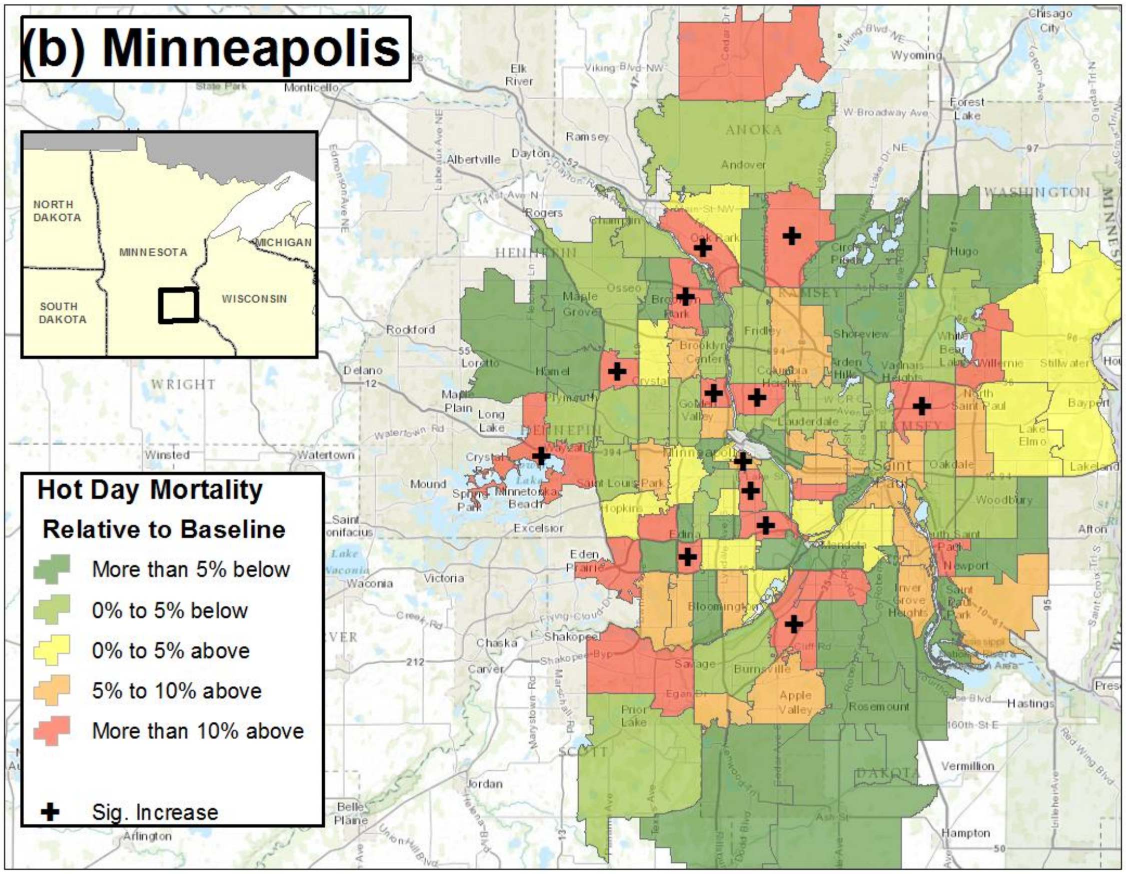

ATL
BOS
MSP
PHL
PHX
SEA
STL
Daily Mortality

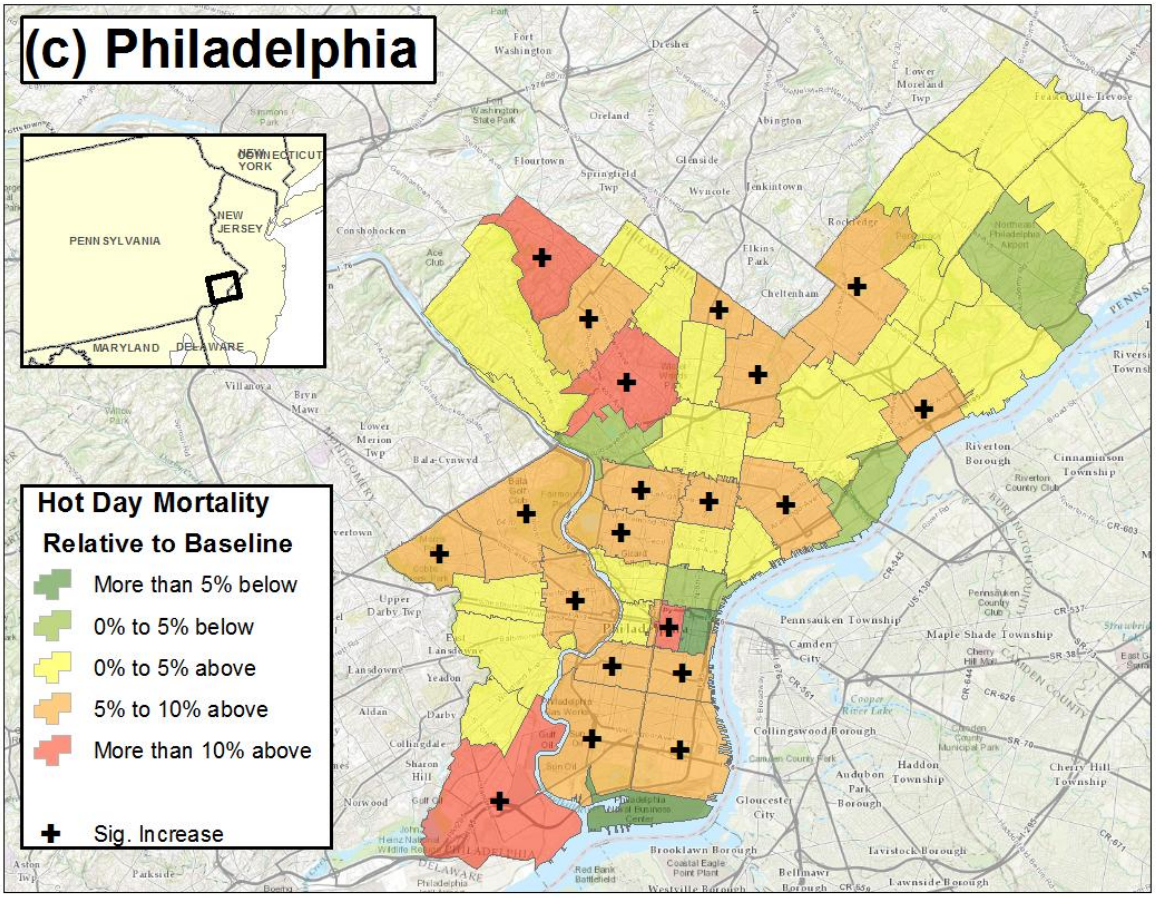
Figure 2. The modeled temperature-mortality relationship for each of the study cities based on historical data. The solid line is the model estimate and the dashed lines represent the confidence intervals. The tick marks on the horizontal axis show the distribution of temperature observations for each city.

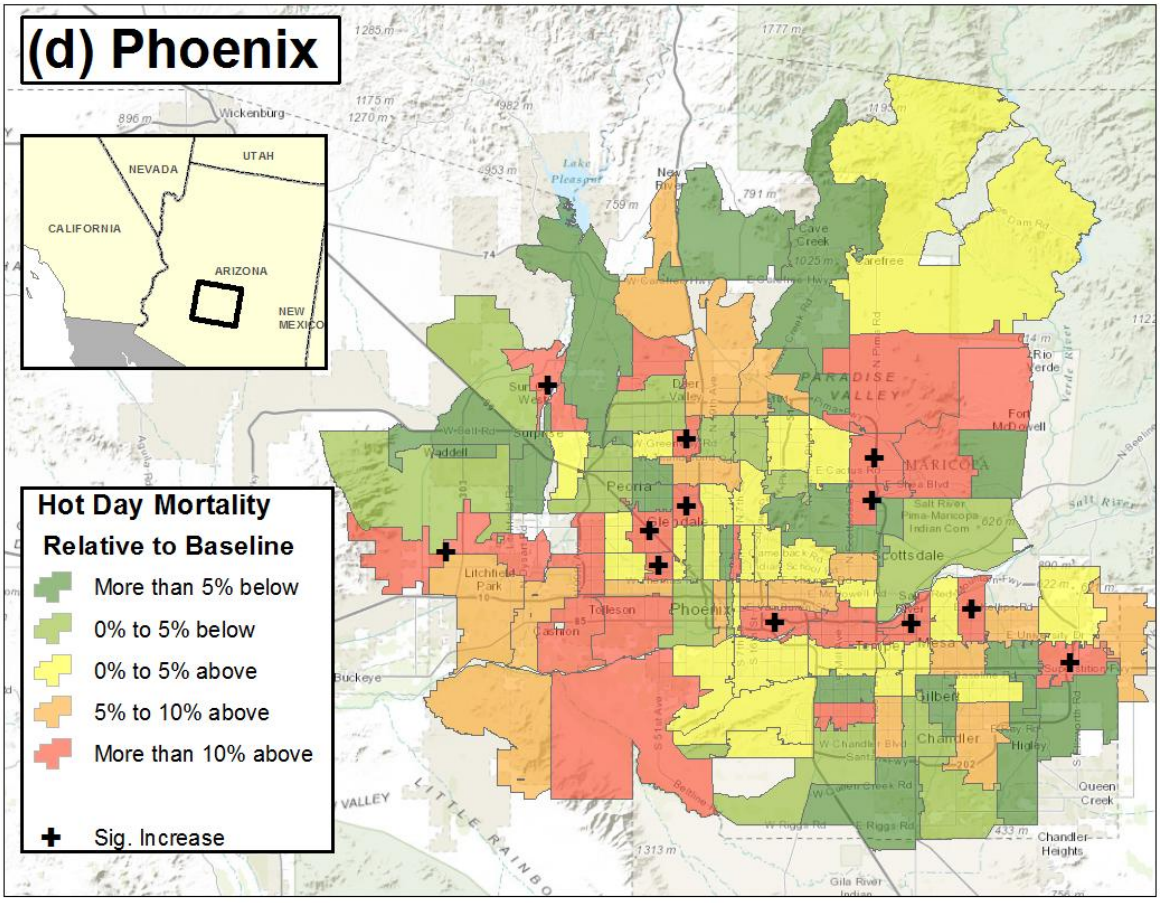


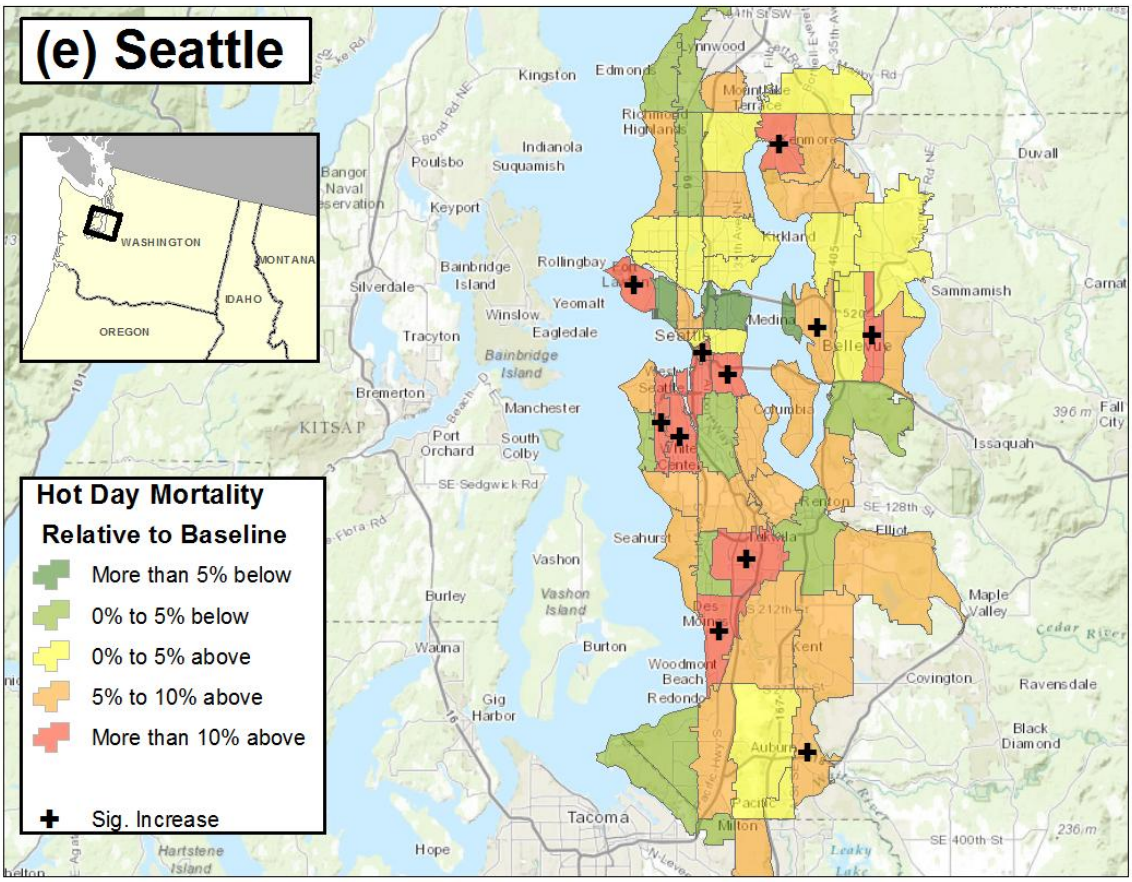
Figures 3 a–f. Maps showing intra-city variability in heat-related risk. The spatial units shown on the map are year 2000 Zip Code Tabulation Areas (postal codes). Each postal code is shaded according to the percent change in mortality on days that exceed a city-specific temperature threshold relative to a baseline summer mortality rate. Those postal codes where the mortality rate on threshold-exceeding days is statistically significantly greater than the baseline are identified with a plus (+) sign.

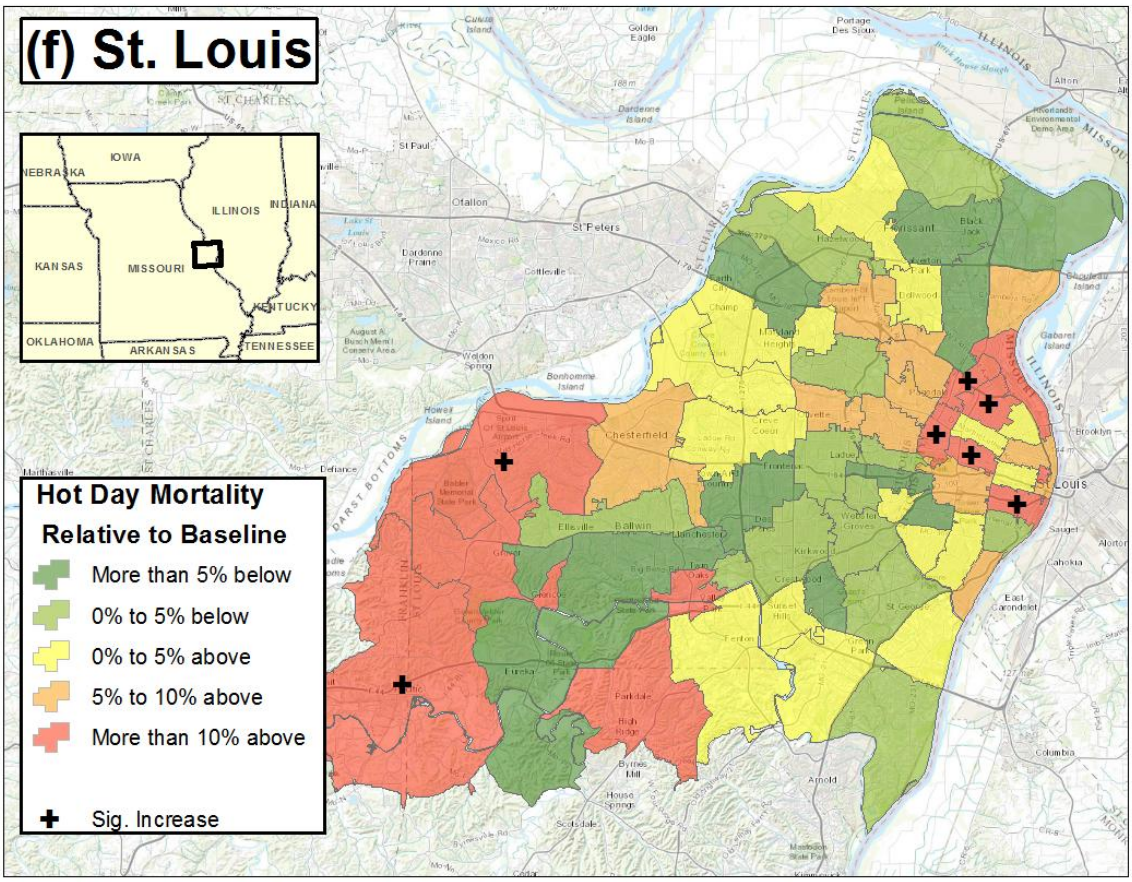



(b) Minneapolis
NORTH DAKOTA
MINNESOTA
MICHIGAN
SOUTH DAKOTA
WISCONSIN
Hot Day Mortality
Relative to Baseline
More than 5% below
0% to 5% below
0% to 5% above
5% to 10% above
More than 10% above
Sig. Increase


(c) Philadelphia
CONNECTICUT
NEW YORK
NEW JERSEY
PENNSYLVANIA
MARYLAND
DELAWARE
Hot Day Mortality
Relative to Baseline
More than 5% below
0% to 5% below
0% to 5% above
5% to 10% above
More than 10% above
Sig. Increase


(d) Phoenix
NEVADA
UTAH
CALIFORNIA
ARIZONA
NEW MEXICO
Hot Day Mortality
Relative to Baseline
More than 5% below
0% to 5% below
0% to 5% above
5% to 10% above
More than 10% above
Sig. Increase
Wickenburg
Cave Creek
Carefree
PARADISE VALLEY
Fort McDowell
MARICOPA
Peoria
Glendale
Phoenix
Scottsdale
Tolleson
Litchfield Park
Buckeye
Gilbert
Chandler
Queen Creek
Chandler Heights
Gila River Indian
Waddell
Surprise
Deer Valley
Salt River Pima-Maricopa Indian Com
LITTLE RAINBO
VALLEY


(e) Seattle
WASHINGTON
MONTANA
IDAHO
OREGON
Hot Day Mortality
Relative to Baseline
More than 5% below
0% to 5% below
0% to 5% above
5% to 10% above
More than 10% above
Sig. Increase
Kingston
Edmonds
Indianola
Suquamish
Poulsbo
Keyport
Bainbridge Island
Rollingbay
Silverdale
Winslow
Yeomalt
Eagledale
Tracyton
Bremerton
Manchester
KITSAP
Port Orchard
South Colby
SE Sedgwick Rd
Vashon
Vashon Island
Burley
Wauna
Burton
Gig Harbor
Tacoma
Hope
Hartstene Island
Kirkland
Bellevue
Sammamish
Issaquah
Kent
Auburn
Seahurst
Woodmont Beach Redondo
Covington
Maple Valley
Ravensdale
Black Diamond
Duvall
Fall City
Enumclaw
Milton
SE 400th St
SE 128th St
Cedar River


(f) St. Louis
IOWA
NEBRASKA
ILLINOIS
INDIANA
KANSAS
MISSOURI
KENTUCKY
OKLAHOMA
ARKANSAS
TENNESSEE
Hot Day Mortality
Relative to Baseline
More than 5% below
0% to 5% below
0% to 5% above
5% to 10% above
More than 10% above
Sig. Increase

Tables

Table 1. Descriptive statistics of the data used for this study from each of the six cities where a statistical relationship between summer temperature and mortality was found. The table also includes information about the city-specific temperature-mortality models including the relative risk (RR) predicted by the mortality at the temperature threshold (T*), and the average RR when the threshold is exceeded.

	Boston	Minneapolis	Philadelphia	Phoenix	Seattle	St. Louis
Descriptive Statistics						
Period of record	1987-2007	1992-2008	1983-2008	1989-2007	1988-2008	1980-2008
Average daily mortality	35.750	37.580	43.160	52.380	29.680	37.480
Average summer max. temp.	25.500	26.400	29.060	39.920	22.660	30.310
City-wide Temperature Model						
Threshold temperature (T*)	27.900	30.000	30.700	42.500	25.900	34.100
Model predicted RR at T*	1.017	1.022	1.015	1.019	1.021	1.016
% of summer days T* exceeded	26.5%	17.9%	27.9%	14.8%	22.0%	13.0%
Average RR when T* exceeded	1.059	1.042	1.066	1.037	1.062	1.029

Table 2. Independent variables included in the study as potential factors related to spatial variability in heat-related mortality at the postal code scale. The data sources are the U.S. Census Bureau via the National Historical Geographical Information Systems (NHGIS) (Minnesota Population Center 2011) and the National Land Cover Database (NLCD) (Fry et al. 2011).

Variable (by postal code)	**Source**
Percent of residents over age 65	NHGIS/U.S. Census
Percent of residents under age 5	NHGIS/U.S. Census
Percent of residents over age 25 without a high school diploma	NHGIS/U.S. Census
Percent of households with one resident	NHGIS/U.S. Census
Percent of households with one resident over age 65	NHGIS/U.S. Census
Percent of residents with public assistance for disability	NHGIS/U.S. Census
Percent of residents living in poverty	NHGIS/U.S. Census
Percent of residents who are white	NHGIS/U.S. Census
Percent of residents who are black	NHGIS/U.S. Census
Percent of residents who are American Indian	NHGIS/U.S. Census
Percent of residents who are Asian	NHGIS/U.S. Census
Percent of residents who are Pacific Islander	NHGIS/U.S. Census
Percent of residents of another race	NHGIS/U.S. Census
Percent of residents of two or more races	NHGIS/U.S. Census
Mean per capita income	NHGIS/U.S. Census
Median housing value	NHGIS/U.S. Census
Percent of dwellings built 1940 or earlier	NHGIS/U.S. Census
Percent of dwellings built 1970 or earlier	NHGIS/U.S. Census
Percent land area with open space	NLCD
Percent land area with low-intensity development	NLCD
Percent land area with medium-intensity development	NLCD
Percent land area with high-intensity development	NLCD
Percent land area with water or wetlands	NLCD
Percent land area with forest	NLCD
Percent land area with grass or crops	NLCD

Table 3. Details from multivariate linear regression models generated for each city relating intra-urban variability in heat-related risk to demographic and environmental factors. The Model Summary portion of the table includes several diagnostic tests for residual normality, heteroskedasticity, and autocorrelation. *The diagnostic information shown for Phoenix is for the regression model excluding the spatial lag term; the coefficients are shown for the final model that does include the spatial lag term. # The significant tests for residual normality in St. Louis and residual heteroskedasticity in Minneapolis are caused by one single outlier value that did not have a substantive impact on the regression.

	Boston	**Minneapolis**	**Philadelphia**	**Phoenix**	**Seattle**	**St. Louis**
Model summary						
N	60	100	45	92	60	60
F	15.867	18.378	4.528	19.182*	6.300	10.669
p	<0.01	<0.01	<0.01	<0.01*	<0.01	<0.01
Adj. R2	0.335	0.260	0.243	0.167*	0.152	0.141
Collinearity	8.480	5.280	16.825	3.44*	7.881	3.479
p (Jarque-Bera)	0.340	0.193	0.612	0.103*	0.261	<0.001#
p (Koenker-Bassett)	0.111	0.048#	0.753	0.007*	0.759	0.130
p (Moran's I)	0.492	0.755	0.268	0.019*	0.940	0.564
Coefficients						
% Medium-Intensity Development (per 10%)	0.019	0.049		0.018		
Per Capita Income (per $10000)	-0.015					
% Over Age 65 (per 10%)		0.605	0.035		0.044	
% Asian (per 10%)			0.039			
% Under Age 5 (per 10%)			0.097			
% White (per 10%)			-0.005			
% Pacific Islander (per 10%)					0.436	
% Below High School Education (per 10%)						0.282
Other Terms in Model						
Spatial Lag				0.297		
Constant	0.932	0.715	0.886	0.576	0.876	0.870

Table 4. Summary information for regression models generated for each city using city-specific principal components of the suite of explanatory variables shown in Table 2. The complete loadings matrix for the principal components, each of which briefly characterized in this table, can be found in the appendix.

	F	*p*	Adj. R2	Terms Included	Coefficient
Boston	7.311	<0.01	0.243	PC1 (Low education, high poverty, high-intensity development, nonwhite)	0.022
				PC2 (Low intensity development, children)	0.022
				PC4 (Older housing, medium and low-intensity development)	0.021
Minneapolis	10.494	<0.01	0.277	PC1 (Old housing, low education, poverty)	0.041
				PC2 (Elderly)	0.035
				PC3 (Low-intensity development)	-0.034
				PC5 (Natural land cover, other races)	0.024
Philadelphia	10.388	<0.01	0.176	PC9 (Asian, Pacific Islander, elderly, older homes)	0.019
Phoenix	9.503	<0.01	0.219	PC1 (Low education, high poverty, high public assistance, high nonwhite)	0.030
				PC2 (High living alone, medium-intensity development)	0.032
				PC4 (High home values, high income, forest land cover)	-0.020
Seattle	5.227	0.03	0.067	PC1 (High public assistance, low education, high nonwhite, low income)	0.018
St. Louis	8.52	<0.01	0.113	PC1 (High poverty, high public assistance, low education, low income)	0.029

Table 5. Summary information for regression models generated for data combined from all cities, analogous to the city-specific models shown in Tables 3 and 4.

	All Cities (orig.)	**All Cities (PCs)**
Model summary		
N	417	417
F	29.371	30.738
p	<0.01	<0.01
Adj. R2	0.254	0.263
Coefficients		
% Medium-intensity development (per 10%)	0.014	
% Living alone over age 65 (per 10%)	0.033	
% Homes built before 1970 (per 10%)	0.517	
% White (per 10%)	-0.690	
% Pacific Islander (per 10%)	0.531	
PC1 (low education, high poverty)		0.039
PC2 (low children, high living alone)		0.016
PC4 (low-intensity development)		0.023
PC5 (high elderly, high grass and crops)		0.011
PC7 (high home values, high incomes, high Amer. Indian, high forest)		-0.013
Constant	0.869	0.919

Table 6. Analysis of variance comparing regression model predicted values and residuals by city. The top half of the table reflects the all-city regression model using the original demographic and environmental variables. The bottom half of the table reflects the all-city principal components regression model.

Multiple Regression Models (no PCs)		Sum of Squares	df	Mean Square	F	Sig.
Unstandardized Predicted Value * City	Between Groups	.237	5	.047	24.070	.000
	Within Groups	.811	411	.002		
	Total	1.048	416			
Unstandardized Residual * City	Between Groups	.098	5	.020	2.843	.015
	Within Groups	2.835	411	.007		
	Total	2.933	416			
Multiple Regression Models (with PCs)		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Unstandardized Predicted Value * City	Between Groups	.229	5	.046	22.033	.000
	Within Groups	.854	411	.002		
	Total	1.083	416			
Unstandardized Residual * City	Between Groups	.151	5	.030	4.508	.001
	Within Groups	2.747	411	.007		
	Total	2.897	416			

References

Anderson, B. G., & Bell, M. L. (2009). Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology (Cambridge, Mass.)*, *20*(2), 205.

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical analysis*, *38*(1), 5-22.

Barnett, A. G., Tong, S., & Clements, A. C. A. (2010). What measure of temperature is the best predictor of mortality?. *Environmental research*, *110*(6), 604-611.

Bassil, K. L., Cole, D. C., Moineddin, R., Craig, A. M., Wendy Lou, W. Y., Schwartz, B., & Rea, E. (2009). Temporal and spatial variation of heat-related illness using 911 medical dispatch data. *Environmental Research*, *109*(5), 600-606.

Basu, R., & Ostro, B. D. (2008). A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *American journal of epidemiology*, *168*(6), 632-637.

Corburn, J. (2009). Cities, climate change and urban heat island mitigation: localising global environmental science. *Urban Studies*, *46*(2), 413-427.

Curriero, F. C., Heiner, K. S., Samet, J. M., Zeger, S. L., Strug, L., & Patz, J. A. (2002). Temperature and mortality in 11 cities of the eastern United States. *American Journal of Epidemiology*, *155*(1), 80-87.

Davis, R. E., Knappenberger, P. C., Michaels, P. J., & Novicoff, W. M. (2003). Changing heat-related mortality in the United States. *Environmental Health Perspectives*, *111*(14), 1712.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States, *PE&RS*, Vol. 77(9):858-864.

Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M., & Malamud, B. D. (2009). Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change*, *92*(3-4), 299-341.

Gosling, S. N., McGregor, G. R., & Páldy, A. (2007). Climate change and heat-related mortality in six cities Part 1: model construction and validation. *International journal of biometeorology*, *51*(6), 525-540.

Green, H., Gilbert, J., James, R., & Byard, R. W. (2001). An analysis of factors contributing to a series of deaths caused by exposure to high environmental temperatures. *The American journal of forensic medicine and pathology*, *22*(2), 196-199.

Harlan, S. L., Brazel, A. J., Jenerette, G. D., Jones, N. S., Larsen, L., Prashad, L., & Stefanov, W. L. (2007). In the shade of affluence: The inequitable distribution of the urban heat island. *Research in social problems and public policy*, *15*, 173-202.

Harlan, S. L., Declet-Barreto, J. H., Stefanov, W. L., & Petitti, D. B. (2013). Neighborhood effects on heat deaths: social and environmental predictors of vulnerability in Maricopa County, Arizona. *Environmental health perspectives*,*121*(2), 197.

Hondula, D. M., Davis, R. E., Leisten, M. J., Saha, M. V., Veazey, L. M., & Wegner, C. R. (2012). Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis.*Environmental health*, *11*(1), 1-11.

Hondula, D. M., Davis, R. E., Rocklöv, J., & Saha, M. V. (2013). A time series approach for evaluating intra-city heat-related mortality. *Journal of epidemiology and community health*.

Hondula, D. M., Vanos, J. K., & Gosling, S. N. (2013). The SSC: a decade of climate–health research and future directions. *International journal of biometeorology*, 1-12.

Kalkstein, A. J., & Sheridan, S. C. (2007). The social impacts of the heat–health watch/warning system in Phoenix, Arizona: assessing the perceived risk and response of the public. *International journal of biometeorology*, *52*(1), 43-55.

Kalkstein, L. S., & Davis, R. E. (1989). Weather and human mortality: an evaluation of demographic and interregional responses in the United States.*Annals of the Association of American Geographers*, *79*(1), 44-64.

Kilbourne, E. M., Choi, K., Jones, T. S., & Thacker, S. B. (1982). Risk factors for heatstroke. *Jama*, *247*(24), 3332-4.

Koppe, C., Kovats, S., Jendritzky, G., Menne, B., & Breuer, D. J. (2004). *Heat waves: risks and responses*. Regional Office for Europe, World Health Organization.

Laaidi, K., Zeghnoun, A., Dousset, B., Bretin, P., Vandentorren, S., Giraudet, E., & Beaudeau, P. (2012). The impact of heat islands on mortality in Paris during the August 2003 heat wave. *Environmental health perspectives*, *120*(2), 254.

Lowe, D., Ebi, K. L., & Forsberg, B. (2011). Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *International journal of environmental research and public health*, *8*(12), 4623-4648.

Luber, G. E., Sanchez, C., & Conklin, L. (2006). Heat-related deaths: United States, 1999–2003. *Morbidity and Mortality Weekly Report*, *55*(29), 796-798.

McGeehin, M. A., & Mirabelli, M. (2001). The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives, 109*(Suppl 2), 185.

Minnesota Population Center. *National Historical Geographic Information System: Version 2.0*. Minneapolis, MN: University of Minnesota 2011.

Naughton, M. P., Henderson, A., Mirabelli, M. C., Kaiser, R., Wilhelm, J. L., Kieszak, S. M., ... & McGeehin, M. A. (2002). Heat-related mortality during a 1999 heat wave in Chicago. *American journal of preventive medicine*, *22*(4), 221-227.

O'Neill, M. S., Zanobetti, A., & Schwartz, J. (2003). Modifiers of the temperature and mortality association in seven US cities. *American Journal of Epidemiology*, *157*(12), 1074-1082.

Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society, 108*(455), 1-24.

Piantadosi, S., Byar, D. P., & Green, S. B. (1988). The ecological fallacy.*American Journal of Epidemiology, 127*(5), 893-904.

Reid, C. E., O'Neill, M. S., Gronlund, C. J., Brines, S. J., Brown, D. G., Diez-Roux, A. V., & Schwartz, J. (2009). Mapping community determinants of heat vulnerability. *Environmental Health Perspectives*, *117*(11), 1730.

Ren, C., O'Neill, M. S., Park, S. K., Sparrow, D., Vokonas, P., & Schwartz, J. (2011). Ambient temperature, air pollution, and heart rate variability in an aging population. *American journal of epidemiology*, *173*(9), 1013-1021.

RStudio. *RStudio: Integrated development environment for R (Version 2.15.1)*. Boston, MA, USA, 2012.

Schuman, S. H. (1972). Patterns of urban heat-wave deaths and implications for prevention: data from New York and St. Louis during July, 1966. *Environmental Research*, *5*(1), 59-75.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Semenza, J. C., Rubin, C. H., Falter, K. H., Selanikio, J. D., Flanders, W. D., Howe, H. L., & Wilhelm, J. L. (1996). Heat-related deaths during the July 1995 heat wave in Chicago. *New England Journal of Medicine*, *335*(2), 84-90.

Sheridan, S. C., & Kalkstein, L. S. (2004). Progress in heat watch-warning system technology. *Bulletin of the American Meteorological Society*, *85*(12), 1931-1942.

Smargiassi, A., Goldberg, M.S., Plante, C., Fournier, M., Baudouin, Y., Kosatsky, T. (2009). Variation of daily warm season mortality as a function of micro-urban heat islands. *Journal of Epidemiology and Community Health, 63(8),* 659-664.

Theoharatos, G., Pantavou, K., Mavrakis, A., Spanou, A., Katavoutas, G., Efstathiou, P., ... & Asimakopoulos, D. (2010). Heat waves observed in 2007 in Athens, Greece: synoptic conditions, bioclimatological assessment, air quality levels and health effects. *Environmental Research, 110*(2), 152-161.

Vaneckova, P., Beggs, P. J., & Jacobson, C. R. (2010). Spatial analysis of heat-related mortality among the elderly between 1993 and 2004 in Sydney, Australia. *Social Science & Medicine, 70*(2), 293-304.

Vaneckova, P., Neville, G., Tippett, V., Aitken, P., FitzGerald, G., & Tong, S. (2011). Do biometeorological indices improve modeling outcomes of heat-related mortality?. *Journal of Applied Meteorology and Climatology*, *50*(6), 1165-1176.

Wood, S. N. (2006). *Generalized additive models: an introduction with R* (Vol. 66). Chapman & Hall.

Yardley, J., Sigal, R. J., & Kenny, G. P. (2011). Heat health planning: The importance of social and community factors. *Global Environmental Change*,*21*(2), 670-679.

Appendix – Tables

Tables A1–A6. Loadings and percent variance explained for the first ten principal components extracted for Boston (BOS), Minneapolis (MSP), Philadelphia (PHL), Phoenix (PHX), Seattle (SEA), and St. Louis (STL), from an original set of 25 demographic and environmental variables.

Table A1. Component Matrix - BOS

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	36.011	17.683	10.413	6.086	5.530	4.438	3.908	3.250	2.861	1.908
PctOverAge65	-.681	.164	-.513	.351	.115	.071	-.202	.048	.064	.125
PctUnderAge5	.050	.897	-.086	-.027	.136	.173	-.221	.058	-.002	-.079
PctBuiltBefore1940	.414	-.467	.467	.378	.388	.181	.014	.015	.020	.058
PctBuiltBefore1970	.287	-.005	.619	.471	.418	.170	.209	.035	-.058	.000
MedianHouseValue	-.230	-.440	.408	-.138	-.171	.622	.082	.103	-.011	-.098
PerCapitaIncome	-.580	-.453	.231	-.256	-.122	.236	-.402	-.021	-.003	.079
PctBelowHSEduc	.780	.149	-.462	.148	.109	.222	.014	.101	.038	-.104
PctLiveAlone	.157	-.892	.118	.033	-.132	.056	-.044	.097	.128	.209
PctLiveAloneandOverAge65	-.281	-.032	-.723	.446	.055	.256	.005	.188	.081	.235
PctwPublicAsst	.861	.298	-.213	-.036	-.009	.145	-.084	.036	.145	.067
PctBelowPoverty	.847	-.205	-.114	-.214	-.148	.037	.247	.058	-.032	.065
PctTwoRace	.881	.200	.032	-.061	.142	-.023	-.017	.098	-.052	-.129
PctWhite	-.848	-.256	.014	.033	.338	-.209	.004	.103	-.101	-.062
PctBlack	.661	.417	.234	-.066	-.337	.137	-.185	-.116	.150	.237
PctAmerInd	.875	.204	.158	-.181	.027	.129	-.079	.120	.017	.051
PctAsian	.203	-.398	-.504	.161	-.375	.220	.446	-.214	.013	-.213
PctPacIsl	.684	.025	-.011	-.194	.160	-.292	.374	.055	-.048	.333
PctOtherRace	.799	.271	-.064	-.155	.209	.110	-.105	.248	-.069	-.145
NLCDOpenSpace	-.634	.558	.149	.049	-.213	.008	.224	.244	-.156	-.004
NLCDLowIntensity	-.405	.572	.368	.281	-.307	.076	.189	.250	.065	.044
NLCDMedIntensity	.516	.209	.334	.463	-.186	-.220	-.064	-.293	.280	-.033
NLCDHighIntensity	.532	-.696	-.295	-.065	.172	-.012	-.097	.072	-.093	.011
NLCDWaterWetlands	-.512	-.111	.036	-.306	.128	-.103	.182	.373	.625	-.083
NLCDForest	-.597	.451	.007	-.295	.139	.271	.235	-.118	-.188	.225
NLCDGrassCrops	-.379	.293	-.086	-.263	.479	.258	.162	-.459	.275	.003

Table A2. Component Matrix - MSP

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	43.981	16.263	6.805	5.257	4.485	4.125	3.428	2.939	2.615	2.014
PctOverAge65	.138	.862	-.017	-.135	.036	-.323	.294	.083	.039	.003
PctUnderAge5	-.048	-.818	-.300	-.092	-.092	.066	.214	.140	.135	.143
PctBuiltBefore1940	.712	.080	.198	-.187	.121	.467	.044	-.294	.049	.077
PctBuiltBefore1970	.697	.420	-.020	-.331	.106	.152	.188	-.287	.049	-.023
MedianHouseValue	-.542	.334	.354	.315	-.060	.427	.230	.260	-.007	.051
PerCapitaIncome	-.632	.462	.129	.284	.096	.377	.232	.073	.076	.041
PctBelowHSEduc	.889	-.234	-.043	-.066	.162	-.202	.139	.079	-.043	-.045
PctLiveAlone	.650	.591	.339	.102	-.013	.005	-.133	-.018	-.069	.035
PctLiveAloneandOverAge65	.334	.779	.064	-.124	.036	-.349	.301	.086	.053	.046
PctwPublicAsst	.887	-.213	.040	.173	.070	-.068	.206	.080	-.105	-.026
PctBelowPoverty	.897	-.126	.217	.185	-.019	.015	-.005	-.023	-.095	-.096
PctTwoRace	.945	-.179	-.010	.038	.145	.050	-.035	.104	.007	-.008
PctWhite	-.911	.233	.114	-.188	.007	-.064	-.184	-.028	.094	.013
PctBlack	.809	-.215	-.129	.139	-.045	.167	.241	-.022	-.198	-.252
PctAmerInd	.719	-.168	.151	-.125	.367	.070	-.147	.285	.001	-.262
PctAsian	.646	-.201	-.200	.425	-.301	-.138	.222	-.086	-.069	.262
PctPacIsl	.577	-.160	.098	.373	-.040	-.100	-.037	-.307	.561	-.076
PctOtherRace	.680	-.123	.067	-.121	.445	.037	-.041	.234	.209	.366
NLCDOpenSpace	-.358	.351	-.575	.430	.208	.017	-.122	.196	.050	-.068
NLCDLowIntensity	.330	.481	-.685	-.150	-.050	.249	.017	-.110	.017	-.072
NLCDMedIntensity	.800	.206	-.198	.033	.000	.106	-.247	.094	.055	.170
NLCDHighIntensity	.656	.267	.416	.244	-.268	-.108	-.271	.037	-.045	.004
NLCDWaterWetlands	-.523	-.137	.051	.228	.529	-.112	-.012	-.345	-.368	.199
NLCDForest	-.660	-.124	.070	.283	.430	-.146	.156	-.112	.186	-.195

Table A3. Component Matrix - PHL

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	36.154	18.227	9.172	8.598	5.716	4.712	3.760	3.053	2.938	2.007
PctOverAge65	-.684	-.194	-.157	.498	.165	.302	.065	.029	.198	.097
PctUnderAge5	.705	-.524	.062	.230	-.144	-.035	.187	.013	.101	-.199
PctBuiltBefore1940	.484	.453	-.372	-.226	-.465	.224	.123	-.043	.071	.218
PctBuiltBefore1970	.501	-.048	-.454	-.067	-.529	.060	-.153	.257	.326	-.021
MedianHouseValue	-.724	.484	.229	-.124	.059	.200	-.013	.098	-.066	-.123
PerCapitaIncome	-.761	.359	.212	-.189	-.136	.314	.056	.200	-.029	-.121
PctBelowHSEduc	.887	-.131	-.131	.340	.067	-.031	.121	-.053	-.005	.083
PctLiveAlone	-.397	.734	-.128	-.306	.265	.217	-.027	-.003	-.081	.096
PctLiveAloneandOverAge65	-.418	.020	-.319	.657	.273	.339	.148	-.002	.201	.057
PctwPublicAsst	.909	-.112	-.009	.100	.228	.157	.110	-.078	-.071	-.001
PctBelowPoverty	.859	.227	-.070	-.067	.331	.100	.052	-.156	-.075	.115
PctTwoRace	.641	.265	.571	.040	-.038	-.002	-.169	.007	.272	-.059
PctWhite	-.618	.343	.108	.548	-.316	-.203	.034	.017	-.129	.083
PctBlack	.449	-.453	-.248	-.602	.287	.235	.002	-.003	.106	-.091
PctAmerInd	.771	.103	.233	.065	-.209	.207	-.153	-.051	-.266	.218
PctAsian	-.144	.641	.142	-.063	.195	-.464	-.097	-.262	.389	.122
PctPacIsl	.462	.309	.464	.142	.079	.340	-.269	.017	.249	-.120
PctOtherRace	.707	.111	.492	.338	-.017	.115	-.082	.099	-.178	-.023
NLCDOpenSpace	-.441	-.656	.293	-.098	-.162	.047	.212	-.329	.027	-.139
NLCDLowIntensity	-.498	-.563	.277	-.160	-.165	.229	-.194	-.246	.089	.299
NLCDMedIntensity	-.117	-.673	-.100	.050	.308	-.195	-.443	.376	-.048	.138
NLCDHighIntensity	.534	.738	-.259	.069	.090	-.128	.109	.068	-.027	-.105
NLCDForest	.117	-.144	.572	-.254	.103	-.099	.547	.393	.152	.260

Table A4. Component Matrix - PHX

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	34.963	15.140	12.573	6.773	5.466	4.700	3.792	3.129	2.370	2.216
PctOverAge65	-.413	.517	.532	-.342	-.156	.008	.049	.189	.166	-.055
PctUnderAge5	.685	-.534	-.282	.008	-.195	-.043	-.051	-.032	-.105	.019
PctBuiltBefore1940	.491	.351	.241	.447	-.113	.332	-.049	-.305	.032	-.034
PctBuiltBefore1970	.686	.477	.162	.293	.018	.024	.018	-.062	-.057	.098
MedianHouseValue	-.635	-.080	.070	.606	.153	.161	-.137	.069	.105	.224
PerCapitaIncome	-.795	.042	.053	.480	.182	.093	-.061	.101	.045	.099
PctBelowHSEduc	.890	-.081	.238	.023	-.275	.039	.048	.103	-.024	.072
PctLiveAlone	.096	.845	.193	.005	.273	.052	-.007	-.129	.079	-.012
PctLiveAloneandOverAge65	-.169	.713	.510	-.343	-.090	.049	-.008	.128	.122	-.012
PctwPublicAsst	.880	-.134	.285	.017	.174	.146	-.132	.044	.005	.010
PctBelowPoverty	.901	.058	.248	.170	.039	.167	.018	-.028	.054	.014
PctTwoRace	.755	-.362	.023	-.186	.392	.005	-.099	.119	.005	-.050
PctWhite	-.907	.301	-.101	.024	-.091	-.060	.001	-.130	-.094	.027
PctBlack	.671	-.002	-.152	.229	-.299	.015	.091	-.044	.322	-.258
PctAmerInd	.353	-.352	.408	-.219	.601	.202	-.279	.102	.078	-.011
PctAsian	-.202	-.059	-.696	.062	.400	.075	.201	-.197	.111	-.324
PctPacIsl	.350	.090	-.539	-.083	.171	-.283	.231	.092	.479	.351
PctOtherRace	.875	-.174	.025	.057	-.266	-.050	.138	.151	-.016	.084
NLCDOpenSpace	-.406	-.108	-.394	-.143	-.275	.535	-.225	.248	.206	-.193
NLCDLowIntensity	.150	.334	-.706	.044	-.133	.220	-.314	.203	-.089	.196
NLCDMedIntensity	.459	.651	-.339	-.052	.211	-.252	.011	.116	-.227	-.069
NLCDHighIntensity	.644	.530	-.118	.175	.087	.058	.140	.007	.005	-.109
NLCDWaterWetlands	-.125	-.089	-.065	-.221	.119	.627	.661	.060	-.167	.165
NLCDForest	-.252	-.126	.279	.539	.089	-.198	.257	.552	-.063	-.221
NLCDGrassCrops	-.392	-.619	.593	.038	-.064	-.077	.135	-.221	.135	.033

Table A5. Component Matrix - SEA

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	31.687	21.924	9.363	6.785	5.853	4.346	3.371	3.065	2.708	2.120
PctOverAge65	-.208	.407	.483	.683	.059	.101	-.096	-.102	.010	-.008
PctUnderAge5	.260	-.882	.042	-.029	.203	.040	-.040	.000	-.037	-.053
PctBuiltBefore1940	.103	.802	.078	-.353	.149	-.014	.325	.150	-.049	.062
PctBuiltBefore1970	.162	.630	.525	-.120	.273	.022	.196	.091	.137	.134
MedianHouseValue	-.631	.477	.252	-.170	-.242	.166	.047	.252	.109	-.082
PerCapitaIncome	-.745	.413	.107	-.049	-.269	.153	-.071	.193	.154	-.103
PctBelowHSEduc	.936	-.045	.051	.182	.007	.003	.120	-.026	-.048	.065
PctLiveAlone	.117	.872	-.295	.041	-.095	-.050	-.052	.134	.110	-.054
PctLiveAloneandOverAge65	.072	.666	.241	.652	.062	.019	.013	-.108	-.023	.027
PctwPublicAsst	.911	-.051	.011	.156	-.044	.025	.141	.071	-.103	.023
PctBelowPoverty	.706	.500	-.252	.114	-.154	-.050	.083	.028	-.087	.099
PctTwoRace	.897	-.165	.093	-.163	.046	-.056	-.073	.047	.037	-.044
PctWhite	-.857	.008	-.323	.127	.323	-.133	-.007	.034	.064	.105
PctBlack	.758	.197	.249	-.132	-.264	.072	.154	-.010	-.165	-.081
PctAmerInd	.737	.266	-.378	.245	.280	-.037	-.048	.036	-.064	.142
PctAsian	.634	-.085	.417	-.138	-.469	.175	-.056	-.118	-.111	-.138
PctPacIsl	.735	-.289	.155	-.011	.136	.092	-.145	.197	.432	.086
PctOtherRace	.790	-.280	.038	.036	.081	.135	-.137	.153	.378	-.041
NLCDOpenSpace	-.034	-.484	.059	.224	-.095	-.523	.532	.265	.099	-.106
NLCDLowIntensity	-.357	-.367	.597	.210	.001	.108	.192	-.199	.157	.089
NLCDMedIntensity	.113	.391	.274	-.425	.572	.220	.009	-.028	-.150	.083
NLCDHighIntensity	.377	.666	-.483	.122	-.245	.049	-.095	.013	.117	-.033
NLCDWaterWetlands	-.149	-.238	-.482	-.067	-.273	.462	.402	-.284	.210	.283
NLCDForest	-.253	-.497	.130	.187	-.257	.202	-.127	.515	-.277	.396
NLCDGrassCrops	-.064	-.255	-.320	.304	.340	.584	.234	.225	-.122	-.371

Table A6. Component Matrix - STL

	Component									
	1	2	3	4	5	6	7	8	9	10
Variance Explained (%)	35.955	15.791	12.035	7.236	4.981	4.652	3.865	3.069	2.748	1.974
PctOverAge65	.016	.721	-.453	.369	-.113	-.196	.116	.151	.135	.054
PctUnderAge5	.247	-.724	.082	-.289	.136	.028	.093	.330	.124	.126
PctBuiltBefore1940	.724	.206	.098	-.131	-.122	.277	-.321	.246	-.204	.116
PctBuiltBefore1970	.751	.401	-.065	-.252	.019	-.188	.008	.007	-.298	.033
MedianHouseValue	-.683	.378	-.056	-.067	.183	.531	-.021	.166	-.045	-.006
PerCapitaIncome	-.747	.445	-.023	-.065	.145	.399	.025	.118	-.026	-.003
PctBelowHSEduc	.882	-.281	-.185	.084	-.019	.015	.109	.150	.104	.061
PctLiveAlone	.600	.528	.063	.461	-.065	.010	-.184	-.035	-.047	-.036
PctLiveAloneandOverAge65	.367	.576	-.381	.530	-.124	-.040	.116	.213	.084	.002
PctwPublicAsst	.849	-.288	-.243	-.003	.164	.193	.018	.115	.135	.007
PctBelowPoverty	.888	-.156	-.200	.108	.146	.246	.002	.047	.106	-.044
PctTwoRace	.548	.113	.690	-.088	.024	-.218	-.086	.224	-.068	-.197
PctWhite	-.774	.276	.384	-.030	-.300	-.075	.066	.077	.011	.205
PctBlack	.756	-.298	-.428	.031	.280	.068	-.047	-.077	-.034	-.190
PctAmerInd	.572	-.019	.551	-.010	-.157	.053	.412	-.133	.000	.052
PctAsian	-.258	.429	.514	.010	.258	.194	-.366	-.109	.380	-.084
PctPacIsl	.119	.360	.433	.183	.512	.087	.325	.101	-.374	.011
PctOtherRace	.365	.180	.728	-.052	-.039	-.142	.136	.253	.278	-.214
NLCDOpenSpace	-.516	.283	-.323	-.286	.341	-.097	.454	.068	.204	.052
NLCDLowIntensity	.132	.504	-.195	-.627	.244	-.305	-.176	-.082	-.036	.027
NLCDMedIntensity	.765	.117	.268	-.100	-.035	.123	-.073	-.013	.108	.472
NLCDHighIntensity	.639	.137	.304	.147	.067	.242	.218	-.507	.049	.022
NLCDWaterWetlands	-.441	-.385	.174	.454	.419	-.258	-.169	-.022	.088	.205
NLCDForest	-.601	-.548	.095	.094	-.332	.272	.108	.077	-.101	-.091
NLCDGrassCrops	-.432	-.472	.246	.469	.290	-.169	-.093	.088	-.169	.061

Chapter 5. The predictability of high-risk zones for heat-related mortality in seven U.S. cities

DM Hondula, RE Davis

Manuscript to be submitted to Health & Place

Abstract

Heat-related mortality remains a public health challenge in the United States. One potentially beneficial strategy to reduce mortality is to focus intervention efforts on locations within metropolitan areas where the risk is greatest—where mortality rates have historically been the highest on the most severely hot summer days. The objective of this study was to determine the temporal consistency high-risk zones using historical georeferenced mortality data from seven U.S. cities. A generalized additive model was used to identify city-specific threshold temperatures associated with increased mortality, and then the mortality rate on threshold-exceeding days was calculated for each postal code comprising each study city. This process was iterated by withholding subsets of data from the model, enabling comparison of places with high mortality in a training data set to the outcomes on testing data. In all cities, the average mortality rate in postal codes targeted for intervention by the statistical model was higher than that in non-targeted areas. Targeted areas for interventions in the study data accounted for 50% of excess heat-related deaths despite only accounting for 25% of total mortality. Statistically significant interannual consistency in the spatial mortality pattern was found in Boston, Minneapolis, Philadelphia, and Seattle. In these cities, the relative risk of a postal code having high heat-related mortality in the testing data set was 32% higher for those identified as targets. Focusing intervention measures at certain geographical zones within these large urban areas could be an effective means of combating heat-related mortality because there is temporal consistency in places where the death rate is most sensitive to heat.

Introduction

The human response to high heat and humidity varies from individual to individual, but over large populations there is clear evidence that extreme summertime conditions lead to elevated mortality and morbidity rates (Davis et al. 2003, Sheridan et al. 2009, Guo et al. 2012). This information has contributed to the adoption of a number of measures aimed at protecting public health when dangerous conditions are forecast, including heat-health watch-warning systems (Hondula et al. 2013b). There is now a growing capacity to improve specificity in assessments of heat-related risk across both time and space. New knowledge about the human response to heat could lead to more efficient allocation of resources associated with the activation of heat emergency management protocols as well as those associated with longer-term mitigation and adaptation strategies. Recent research has contributed to increased awareness of geographic variability in the response to heat within large metropolitan areas (e.g., Vaneckova et al. 2010, Hondula et al. 2012) and of intra-annual variability in the metropolitan aggregate-scale response to heat (e.g., Guo et al. 2012). But little effort has been directed at the intersection of these sources of variability—exploring how the intra-city human response to heat, in particular morbidity and mortality, changes from one year to another.

Within metropolitan areas the heat-health risk is expected to vary spatially as the urban heat island and related microclimates create different levels of exposure for residents of different communities. Similarly, the underlying risk of the population varies as the demographic profile of city residents is spatially heterogeneous. Heat vulnerability indices have been proposed and mapped at fine spatial resolutions to highlight geographic

variability in expected risk arising from factors that are believed to impact heat sensitivity (e.g., Reid et al. 2009). Other research has mapped health outcomes to identify regions within cities that have historically been associated with higher rates of heat-related mortality and then deduced the association with certain variability factors from the observations (e.g., Smargiassi et al. 2009, Hondula et al. 2012, Laiidi et al. 2012, Reid et al. 2012).

In addition to exploring spatial variability in intra-city risk it is also important to consider the temporal dynamics in the response. It is known that the heat-mortality response varies across time (e.g., Davis et al. 2003, Rocklov et al. 2009, Guo et al. 2012). Although considerable debate continues regarding future climate change and population adaptation, the increase in temperatures in recent years combined with urban heat island impacts has resulted in individuals being more frequently exposed to thermally stressful conditions in some locations (Arnfield et al. 2002, Wilby et al. 2003). At the same time, infrastructure improvements including the use of air conditioning and building design might be decreasing the risk for some (Davis et al. 2003, Sheridan et al. 2009). Aside from these long-term trends, there is evidence that warm-season mortality might be influenced by antecedent wintertime conditions, and further evidence suggests that the characteristics of individual heat events including their timing, duration, and intensity, impact the mortality rate (Rocklov et al. 2009, Anderson and Bell 2011) . In light of these sources of variability at different time scales, it is possible that specific populations within cities most affected by heat might vary based on heat event characteristics, time in season, or

longer-term trends. From a geographical perspective, then, the *location* of high risk zones (e.g., Reid et al. 2009, Hondula et al. 2012) may not be consistent over time.

An examination of temporal variability in the intra-city response to heat is warranted because of the insights such a study could provide regarding the effectiveness of various intervention strategies, particularly those where resources are directed toward particular locations. Here we investigate the predictability of high-risk zones for heat-related mortality using multi-decadal geo-coded mortality records from seven climatically diverse cities in the United States. The goal of the study is to determine if spatial patterns in heat-related mortality are non-random temporally.

Data

We obtained mortality records from seven metropolitan areas across the United States: Atlanta, Georgia (period of record 1994–2007), Boston, Massachusetts (1987–2007), Minneapolis, Minnesota (1992–2008), Philadelphia, Pennsylvania (1983–2008), Phoenix, Arizona (1989–2007), Seattle, Washington (1988–2008), and St. Louis, Missouri (1980–2008). The mortality records contained the postal code of residence of each decedent, enabling us to determine the daily number of deaths within each postal code over the entire period of record. The average number of daily deaths within each city ranged from 30 to 53, and each city contained between 48 and 101 postal codes. Postal codes were selected as the spatial unit for this study because they were the smallest subdivision available across all cities. As these units are designed for mail delivery, the population is roughly evenly distributed between them (interquartile range: 16,728–35,407) and there is no bias from one city to another. Descriptive statistics of the mortality data sets used are available in (Supplementary) Table 1. These data were obtained from the respective state departments of health. Because the data were de-identified, no IRB approval was necessary for this research under Title 45 Part 46 exemption category 4.

Hourly meteorological data were obtained from the first-order weather stations closest to the study cities through the U.S. National Climatic Data Center. These data are publicly available online (www.ncdc.noaa.gov). We calculated the daily maximum afternoon temperature, defined as the maximum temperature observed in the five-hour window centered at the hour of the average daily maximum temperature. There were 43 days

(<0.1% of available data) where no data were available from this time window; these days were given a missing value for temperature.

Geographic boundary files for the study cities were sourced from the United States Census Bureau accessed through the National Historical Geographical Information System portal online (www.nhgis.org) (Minnesota Population Center 2011). We downloaded year-2000 Zip Code Tabulation Areas (ZCTAs) for the states comprising the study cities and extracted the boundaries corresponding to the postal codes available from the mortality records.

Methods

Model Training and Threshold Temperatures

We adopted a cross-validation framework to examine the predictability of mortality within postal codes in each of the study cities. The general procedure was to identify target zones within each city using all years of data except year y_w, and then evaluate how mortality in the target zones compares to that in the withheld year. This procedure was iterated for all years *y* in each city's period of record and across all zip codes.

In the first iteration, we calculated the relationship between temperature and mortality while accounting for seasonality and long-term time trends using a generalized additive model of the form:

$$Log[E(M)] = \alpha + s(T_{PM}, df = 5) + s(Time, df = 7 * n(y)) \quad \text{[Equation 1]}$$

where *E(M)* is the predicted daily mortality count, α is the model intercept, T_{PM} represents daily afternoon maximum temperature, *Time* is a variable to account for long-term time trends and seasonality, *n(y)* is the number of years in the time series (excluding the withheld year y_w), and *s()* represents thin plate smoothing splines for the temperature and time terms (Wood 2003, Wood 2006). Five degrees of freedom was set for the temperature term and seven per year for the time term for all cities after testing a number of other options for each variable and comparing generalized cross validation scores and visually examining plots of model predicted values for each term. All values for the withheld year were set to a missing value indicator. We then extracted model predicted values and partial error for the temperature component of the model.

As we are interested in examining mortality when summertime weather deviates from normal conditions, our entire analysis is conducted relative to the mean summer afternoon maximum temperature, where we defined summer as June–August inclusive. We then examined the model predicted values and standard errors and determined the first temperature higher than the summer mean at which the confidence interval of the relative risk between a given temperature T_x and the mean temperature T_m did not include one. The relative risk (*RR*) between a given temperature and the mean was calculated as:

$$RR_{T_x} = e^{\left(\left(\widehat{M}(T_x) - \widehat{M}(T_m)\right)\right)} \qquad \text{[Equation 2]}$$

where $\widehat{M}$ is the model predicted mortality at a given temperature T_x or the summer mean temperature T_m. The 95% credible interval of the relative risk was calculated as:

$$95\%\ CI(RR_{T_x}) = e^{\left(\left(\widehat{M}(T_x) - \widehat{M}(T_m)\right) \pm 1.96 \times \sqrt{\varepsilon(T_x)^2 + \varepsilon(T_m)^2}\right)} \qquad \text{[Equation 3]}$$

where ε is the partial model error in the temperature term at a given temperature. Thus the temperature threshold is defined here as the temperature at which model-predicted mortality significantly differs from that expected for mean summertime conditions. We then calculated the average mortality rate on all above-threshold days.

To this point we have calculated the effect of temperature for the city as an aggregate, but our focus here was on the smaller postal code units that comprise each study city. As smoothing spline estimates of seasonality derived directly from small samples of data can be unreliable, we examined the relationship between temperature and mortality within each postal code using a multi-stage model (Hondula et al. 2013a). In the first stage,

seasonality was removed from the city-wide data using a generalized additive model that includes terms for seasonality, long-term trends, and temperature. The seasonality and long-term time trends in the mortality data represent a mortality "baseline" – long term variability that is (presumably) unrelated to short-term variability in temperature. This baseline curve was then adjusted to match the changing mortality rates within each postal code based on population growth and demographic changes. A daily mortality ratio (*MR*) was calculated for each postal code by calculating the ratio between each day's sum of deaths within the postal code and the expected number of deaths based on the mortality baseline. We then group the postal code-level ratios into different "bins" based on temperature thresholds (as defined above).

This method was adapted to the current investigation by omitting the withheld year. To ensure that intra-annual, rather than inter-annual differences in mortality rates are emphasized in the analysis, an additional standardization was made to set the mean summertime *MR* equal to exactly one for each year in the period of record. No adjustment was made to the variance. We obtained estimates of within-postal code mortality for two separate groups of days: those with temperatures above and below the city-wide estimated threshold. We then used a randomization test (Sheridan and Dolney 2003, Hondula et al. 2013a) to identify postal codes where mortality was statistically significantly high on above-threshold days. These postal codes were labeled as "targets" for intervention activities.

Model Testing

The target zones were determined using all years except the one withheld year, excluded from the training model to assess *predictability*. We quantified predictability by comparing the mortality in the target zones in the withheld year to that in the non-target zones. In general, if mortality is higher in the target zones than the non-targets in the withheld year, we concluded that there is predictability in the spatial response. As each postal-code year is mean-standardized, there is no bias introduced from the methods that would cause a postal code identified as a target to be more likely to have high or low mortality in the testing year.

With only one year of withheld data and small sample sizes within each postal code, there were not enough data to generate a robust statistical model of within-postal code mortality for the withheld year analogously to what we have done for the rest of the time period. Instead, we calculated seasonality for the entire city in the withheld year using the model:

$$Log[E(M)] = \alpha + s(Time, df = 7) \qquad \text{[Equation 4]}$$

Temperature was not included in the model for the withheld year because effect estimates based on one year of data were found to be unstable (i.e., the magnitude, significance, and sign of the effect varied from one year to another). Although there is some concern that the time component of this model is incorrectly capturing temperature-related variability, in cases we examined there was no visual evidence of short-term deviations away from seasonality that might indicate over fitting.

Seasonality, as determined from equation 4, was removed from the postal code-level data by scaling the seasonal curve so that the mean mortality rate matches that within each postal code. We then calculated a mortality ratio *MR* for each day within each postal code for the withheld year y_w by dividing the observed mortality count by the scaled seasonality. As before, we then re-standardized so the mean summertime *MR* in the withheld period is equal to one. A postal code was labeled as "high mortality" during the withheld year if the mean *MR* on above-threshold days exceeded the critical value for significance as determined from the randomization test described above.

Model Evaluation

This entire procedure was repeated for each year y_j in the data set. Thus for each year we identified a set of target zip codes based on (*n(y)*-1) years of data and those postal codes that were associated with high mortality rates on above-threshold days in year *j*. We used 2x2 contingency tables to quantify the performance of the models in identifying postal codes with high mortality. For each testing iteration we tabulated the number of postal codes with high and low mortality and the number of target and non-target zip codes. Cross-tabulation of these quantities yields four distinct possibilities: (A) Target zones with high mortality in the withheld year, (B) non-target zones with high mortality in the withheld year, (C) Target zones without high mortality in the withheld year, and (D) non-target zones without high mortality in the withheld year (e.g., Table 2). Cells "A" and "D" in the contingency table are associated with correct forecasts, "B" and "C" with missed forecasts. We then calculated the relative risk (RR) of a postal code being associated with high mortality in the withheld year if it was classified as a target zone

compared to the risk of a postal code being associated with high mortality if it was classified as a non-target using the equation:

$$RR = \frac{A/(A+C)}{B/(B+D)} \qquad \text{[Equation 5]}$$

where the letters A–D refer to specific cells in the 2x2 contingency table (Table 2). A relative risk greater than 1.0 indicates that target zones are more likely to be associated with high mortality in the withheld years. We determined that high-mortality zones within each city were statistically predictable if the lower bound of the confidence interval for relative risk (Gardner and Altman 1994) was greater than 1.0.

Finally, we estimate the overall mortality burden attributable to heat for each city over the entire time period as well as for target postal codes exclusively during years they were identified as targets. "Excess mortality" is defined as the difference between the observed and expected number of deaths on a given day or set of days. We calculated the excess mortality ε for each postal code *i* within each year *j* on above-threshold days using the equation:

$$\varepsilon_{i,j} = M_{i,j} \times n(T_{PM} \geq T^*)_j \times (\mu_{i,j} - 1) \qquad \text{[Equation 6]}$$

where M is the mean daily summertime mortality specified for each postal code-year, T^* is the city-specific threshold temperature for each respective year, *n* is a count of the number of days the threshold temperature is reached, and μ is the average mortality ratio *MR* for each postal code and year on above-threshold days. Because of the procedures discussed above to derive *MR*, μ represents average mortality ratios on above-threshold

days standardized for seasonality and long-term time trends (see Equation 4 and subsequent text). Summation of ε across all postal codes and years yields the total excess mortality on above-threshold days for each city in aggregate.

We next sought to determine the portion of excess mortality that occurred in target zones. We repeated the calculation in Equation 6 for observed mortality in target zones (ε^t) only using observations from each year when a postal code was identified as a target. The percentage of excess mortality occurring in the target zones is then determined by multiplying the ratio of ε^t (summed across all postal codes and years) to ε (summed across all postal codes and years) by 100. For comparative purposes, we also calculate total summertime mortality (all days regardless of temperature) for all postal codes and for target postal code-years only using a variation of Equation 6.

Example of Training and Testing Methods

For illustrative purposes, we present the approach for Minneapolis, MN, for which data are available 1992–2008—thus 17 separate numerical "experiments" were conducted for that city. In the first iteration, the model was run using only data from 1993–2008, omitting 1992 (all values for 1992 were set to a missing value indicator). The mean summer temperature in the years included in the training period was 26.6°C, and the threshold temperature determined to be 30.2°C. This threshold corresponded to a model predicted increase in mortality of 2.3% (Figure 1). On the 296 days in the training period when the threshold temperature was exceeded, mortality increased 4.5% relative to baseline.

Based on the 1993–2008 data, 15 postal codes in Minneapolis were identified as target zones (i.e., mortality in 15 postal codes was significantly greater than baseline on above-threshold days in the training period). The mortality rate on above-threshold days in target zones averaged 23.4% above baseline, compared to 3.6% below baseline in non-target zones.

In the withheld year there were five days on which the temperature exceeded the threshold of 30.2°C. On these days, after accounting for seasonality (Figure 2), mortality was 15.6% above baseline across the entire city. The average mortality rate across the 15 target zones on these six days was 23.2% above baseline, whereas in the other postal codes mortality rates averaged 5.4% above. Of the 15 target zones, nine were associated with high mortality in the withheld year (cell "A" in the contingency table) and six were not (cell "C"). Of the remaining 86 postal codes, 31 were associated with high mortality in the withheld year (cell "B") and 55 were not (cell "D"). The relative risk for this iteration was 1.66 (95% confidence interval 1.01, 2.74), indicating that high-mortality locations in 1992 could have been statistically reliably "predicted" using 1993–2008 data.

To demonstrate the calculation of excess mortality, we focus on postal code 55112, where the average daily mortality rate during the withheld year was 0.77 deaths per day. Mortality on the six above-threshold days was, in this postal code, 10.3% above baseline, 0.85 deaths per day. The difference in these two daily mortality rates (0.08 deaths per day) is multiplied by the number of above-threshold days (five) for excess mortality in this postal code during the withheld year of 0.4 deaths. Across all target postal codes in 1992 there were 12.2 excess deaths.

Results

Model Training

Statistically significant relationships between temperature and mortality were established for six of the seven study cities in all training iterations. The relationship between temperature and mortality was only significant in a portion of the training models for Atlanta and thus a threshold could not be established for every year; Atlanta is excluded from the remainder of the analysis. The threshold temperatures (Table 1) for the six other cities were between 1.6°C (Philadelphia) and 4.1°C (St. Louis) above the respective city mean summer afternoon maximum temperatures. The model-predicted increase in mortality at the threshold temperatures range from a 1.5% increase (Philadelphia) to 2.3% increase (Minneapolis-St. Paul). On all threshold-exceeding days mortality rates were 3.2% (St. Louis) to 6.6% (Philadelphia) above mortality rates on summer days with below-threshold temperatures. Depending on the year and city, between 4 and 31 postal codes were identified as targets (i.e., associated with high mortality during the training period). Across all cities, the average number of target postal codes ranged between 12% and 39% of the total number of postal codes within the city.

For most cities there was a clear separation between postal codes that tended to be identified as targets during training iterations and those that were not. In Philadelphia, for example, ten postal codes were associated with high mortality in every training iteration, and another five were labeled as targets on over 90% of the training runs. At the other end of the distribution 26 postal codes were never associated with statistically significantly high mortality in any training model and one area was identified in only one

iteration. There were six remaining postal codes identified as targets in 20–80% of the training runs. A similar pattern was seen in the other cities where the distribution was strongly bimodal with peaks near 0% and 100%. Mortality rates in the target zones on above-threshold days was markedly higher than that in non-threshold zones. On average mortality rates in the target zones were between 4.6% (Philadelphia) and 20.5% (Minneapolis) greater that in non-target zones during the training iterations (Figure 3).

Model Testing

After identifying target zip codes for all but one withheld year in each city's period of record, we next evaluated how mortality in the target zones compared to that in the non-target zones during the single withheld year. As was the case during the training periods, mortality was higher in the target zones in the testing years relative to the non-target zones on average (Figure 3). The ratio of target to non-target zone mortality was similar between the training and testing periods in Boston (8% higher), Minneapolis (20% higher), and Philadelphia (4% higher). In Phoenix, Seattle, and St. Louis there was a larger difference. In St. Louis, for example, mortality was 11.9% higher in target zones using the training data but only 2.7% higher than the non-target zones in the testing period. Of the six cities Minneapolis showed the greatest difference in mortality between target and non-target zones in the withheld year and the difference was smallest in St. Louis. In all six cases the difference in mortality rates between target and non-target zones on above-threshold days was smaller in the testing period when compared to the training period (see Figure 3).

Unlike the training periods, however, there was considerable variability evident in the difference between target and non-target zones in each sample of testing (Figure 4). In total we examined 130 city-years of withheld data; in 87 of these cases (67%) mortality in the target zones was higher than that in the non-target zones. Mortality rates were higher in target zones in 16 of 17 (94%) testing cases in Minneapolis, which was also the city associated with the greatest difference in mortality between target and non-target zones. St. Louis target zones were found to be associated with higher mortality in only 14 of 26 (54%) testing cases, and this was the city associated with the smallest difference in mortality rates between targets and non-targets. Rates in the other four cities were between 58% and 76%.

Contingency tables were employed for comparison of the likelihood of a postal code being associated with high mortality in the testing period if it was classified as a target relative to the likelihood of it being associated with high mortality if not classified as a target (Table 2). In all six cities we examined the relative risk of high mortality in a target zone was greater than 1.0, but the 95% percent confidence interval for Phoenix and St. Louis included one. Therefore there was statically significant model skill in identifying high-risk locations in Boston, Minneapolis, Philadelphia, and Seattle. Among these locations, the city-specific relative risks varied from a low of 1.18 in Seattle to a high of 1.49 in Minneapolis; the overall average relative risk among these cities was 1.32. This means that, in these locations, the probability that a postal code will have high mortality in a given year is nearly 32% higher for zones classified as targets compared to non-targets. The average for Phoenix and St. Louis was 1.05.

Finally, we compared two of the key metrics of this study: (1) the likelihood that each postal code would be labeled as a target for each iteration using training data, and (2) the likelihood that each postal code would be associated with high mortality in the withheld year. As evident from city-specific scatter-plots (Figure 5) of these quantities compared to each other and the contingency tables, the overall tendency was for target zones to be associated with high mortality in the testing year. Correlations between the percentage of years each postal code was identified as a target and the percentage if withheld years that the postal code had high mortality were statistically significant for all cities based on the non-parametric Kendall tau test. Correlation coefficients ranged between 0.425 and 0.598.

Excess Mortality

We calculated the number of excess deaths on above-threshold days to estimate the portion of the total heat-related mortality burden that could potentially be alleviated by targeting specific postal codes. Estimates of average annual excess mortality per city ranged from 7.9 for St. Louis to 46.1 in Philadelphia. Over the study periods, we estimated 3,490 excess deaths on above-threshold days. Of these, nearly half (1,741: 49.9%) occurred in postal codes labeled as targets. It should be reiterated that excess deaths were calculated using data from withheld years only; training data excluding the withheld year were used to identify target locations. The percent of excess deaths occurring in target postal codes ranged from a low of 11.2% in Phoenix to a high of 103.4% in Minneapolis. (Note: a percentage greater than 100% indicates that mortality in non-target zip codes was below the summertime mean on above-threshold days).

To provide a point of comparison for the percentage of excess mortality found in target zones, the total number of summertime deaths and total number of summertime deaths in target zone-years is also shown in Table 3. If the percentage of excess mortality found in target zones mirrored the percentage of total mortality found in target zones, the excess mortality result might simply be reflecting different population sizes or overall mortality rates at the postal code scale. Although target postal codes represent 49.9% of excess deaths on above-threshold days, they account for only 25.0% of total summertime mortality. In five of the six cities target zones accounted for a disproportionate fraction of the excess mortality on above-threshold days. In Minneapolis, which showed the greatest contrast, target zones were associated with 103.4% of excess mortality but only 18.6% of total mortality. In Boston and St. Louis the difference was approximately a factor of two. Philadelphia and Seattle were associated with slight increases over total mortality, and in Phoenix target zones accounted for 11.2% of excess mortality and 15.4% of total mortality.

Discussion

A common theme in the heat-health literature is that heat-related mortality should be preventable. Our goal was to investigate whether prevention of such deaths could be facilitated by a geographically targeted approach, and our results indicate that there are locations within the cities we examined where the mortality rate is consistently high when extreme heat occurs. Our main findings were that postal codes identified as targets using a sample of training data were associated with higher mortality rates and were more likely to have high mortality in testing periods. There was statistically significant predictive ability in identifying high-mortality postal codes in four of the study cities. Overall, predictability was strongest in Minneapolis and weakest in St. Louis and Phoenix. We believe this study is the first of its kind to explore the predictability of zones of heat-related mortality.

We found that mortality in target postal codes was between 2.6% and 19.9% higher than in non-targets in withheld years, and that the likelihood of a postal code being associated with high mortality in a withheld year was anywhere between 1.6% and 49.3% higher if that postal code was identified as a target. These results support the notion that geographically-targeted heat intervention activities within cities may lead to greater benefits than a uniform city-wide approach. If heat-related mortality could be prevented in the target zones only—a relatively small number of postal codes in most of the study cities—the benefits for public health could be quite substantial. We found that almost half of the 3,490 excess heat-related deaths occurred in target postal codes where high mortality rates would have been expected.

There were considerable differences in the study results between cities. Regardless of the metric used to quantify predictability, Minneapolis consistently exhibited the greatest difference in heat impacts between target and non-target zones. Depending on the specific model iteration, only 10–15 of the 101 postal codes in the Minneapolis study area were identified as targets. On above-threshold days, these locations were associated with 20% higher mortality rates than non-target zones. More strikingly, however, these locations were associated with 103% of the *excess* mortality on above-threshold days. As the target postal codes only account for 18.6% of the total summertime mortality during the study period, the results provide strong evidence that there are specific localities within the Minneapolis area that are especially sensitive to heat.

In Phoenix and St. Louis, target postal codes were the least different from non-target zones. Target zones in both cities were associated with higher mortality rates and a higher relative risk of being associated with high mortality in any given year than non-targets, but the relative risk was not significantly different than random. The fact that the response is more homogeneous across Phoenix may not be surprising, as the hot desert climate has forced Phoenix residents to adopt a range of adaptation measures to cope with high temperatures throughout much of the year. We did not include enough cities in this study to determine if there is systematic variability in the predictability of high-risk zones related to geography and climate, but our results are consistent with such a hypothesis (generally higher predictability in Boston, Seattle, and Minneapolis, intermediate in Philadelphia, and lower in St. Louis and Phoenix). With respect to excess mortality, we found that the percentage of excess deaths within target postal codes in Phoenix (11.2%)

was lower than the percentage of total deaths (15.4%) that occurred within those regions. This result seems to contrast the fact that we found overall higher mortality rates on above-threshold days within target postal codes in Phoenix on hot days; the difference arises because the excess mortality calculations are weighted by the population of each postal code and the number of above-threshold days each year. The finding is consistent with the fact that there was not statistically significant model skill in identifying the places where heat mortality was more likely to occur in withheld years.

The use of a prediction-based framework for analysis using separate training and testing data sets is especially relevant and important for problems where there may be significant costs associated with a missed forecast (e.g, wasted resources or underserved vulnerable populations). Heat-health researchers have only recently begun to use this type of study design (Hajat et al. 2010). Although there is a very strong statistical association between heat and human mortality in many locations, the correlation between temperature and mortality is imperfect. Not all hot days are associated with excess mortality, and some high mortality are unaffected by heat (Hajat et al. 2010, Zhang et al. 2012, Pascal et al. 2013). This inherent noisiness should be accounted for when designing intervention and mitigation activities.

An additional strength of this study is the extension of an intra-city, outcome-based approach to identifying high-risk locations for heat-related mortality to six cities that have not been previously examined in such a manner. In an era of finite public resources, the identification of zones where elevated mortality rates can reliably be anticipated in advance offers a useful tool for planning and response. Practical measures that could

benefit from this work include public warning systems, the opening of cooling centers, activation of phone lines and buddy networks, and allocation of additional medical resources and personnel (Sheridan and Kalkstein 2004). There are also longer-term strategies employed to reduce the impacts of extreme heat that could benefit from identification of areas with temporally predictable risk; municipalities are adopting educational campaigns and building weatherization and modification programs aimed at reducing urban heat island effects (e.g., Solecki et al. 2005, Smith et al. 2008).

This research concerns the *likelihood* that a particular postal code will experience heat-related mortality in a given year; this complements information about the magnitude of the response within each postal code aggregated over longer time periods (e.g., Vaneckova et al. 2010, Hondula et al. 2012). Geographic zones where the heat-related risk is consistently high are ideal candidates for focused intervention measures. Although the temperature-mortality relationship is indeed "noisy" as discussed above, in these particular locations the response is more consistent from one year to another. The methodology we presented in this manuscript could readily be applied to many other locations where heat may pose a threat to public health. In our experience the limiting factor in the feasibility of conducting such as study is access to geographically-referenced health outcome data; we advocate for the accessibility of such data to the research community. Furthermore the results of studies like this one merit re-evaluation on a regular basis, as new temporal and spatial patterns in risk may emerge in the coming years.

There are also a number of limitations, assumptions, and methodological choices that could be further explored in subsequent research. The leave-one-out study design was chosen based on the available periods of record for each of the study cities. This study design creates a disconnect from the proper chronological sequencing of training and testing data; in practice data, from 2007 could never be used to generate a model to predict outcomes in 1995. Splitting the data available to use into chronological periods simply did not allow for the generation of robust models and/or evaluation data. As surveillance data continue to be accumulated, we recommend that a different subsetting model be applied to determine if a model generated in one time period can be used to predict outcomes in a subsequent period. This is particularly important for the problem at hand because of the long-term time trends in the response that could be impacted by acclimatization and climate change.

Other elements of the study that might be approached in a different manner in the future are related to the identification of extreme heat and the modeling of effects. We identified an absolute temperature threshold (i.e., one that does not vary across months or years), but research efforts continue be directed to improve the identification and definition of high-risk extreme heat days (e.g., Barnett et al. 2012, Zhang et al. 2012, Antics et al. 2013). There may be both theoretical and model performance advantages to using more complex metrics for quantifying the relationship between human health and heat, and the optimal metric could vary from one city to another. We did allow the temperature threshold to vary across cities, but chose to hold all other model parameters constant. The thresholds we found were based on the model-predicted mortality (and error) at given

temperatures relative to the mean summer temperature, and for some cities this resulted in the threshold temperature falling several degrees or more below those linked to activation of various warnings and advisories by weather forecast offices (National Weather Service 2006) and temperatures commonly perceived as being dangerous (Sheridan 2007). This is not necessarily a weakness of our study, as we likely included all days where heat could have been a contributing factor to elevated summertime mortality rates, but a more stringent criterion could be applied by others. We also included *all* days where the temperature exceeded a city-specific threshold, but it would be worthwhile to examine if the response is more predictable when limiting the study to heat waves where dangerous temperatures are experienced over several consecutive days (Anderson and Bell 2011). We did not consider the effects of mortality displacement (Hajat et al. 2005, Saha et al. 2013) or potential confounding by air pollutants (e.g., Roberts 2004).

Conclusion

Using multi-decadal high-resolution mortality records from cities spanning the climate zones of the United States. we have found evidence that geographic zones associated with higher heat-related mortality are predictable. The relative risk of a given postal code being associated with high mortality in one withheld year if identified as a target zone from a training model was statistically better than random in Boston, Minneapolis, Philadelphia, and Seattle. In these cities there was a 32% greater chance that target zones for intervention activities had high heat-related mortality in a data set unseen by the predictive model. The relative risk was also greater than one in Phoenix and St. Louis, but predictability was not statistically significant. These cities are associated with greater temporal heterogeneity in the spatial mortality pattern on hot summer days. Across all six cities, 50% of excess mortality on above-threshold temperature days was confined to postal codes identified as targets. As these locations represented only 25% of total mortality during the study period, they were disproportionately impacted by extreme heat. Study results indicate significant reductions in heat-related mortality could be achieved with proper intervention programs aimed at specific localities within major metropolitan areas, particularly in Boston, Minneapolis, Philadelphia, and Seattle.

Figures

Figure 1. The temperature-mortality relationship for Minneapolis, MN, based on a generalized additive model using data from 1993–2008. The threshold temperature is the lowest temperature at which model-predicted mortality significantly differs from that predicted at the summer mean temperature. The model-predicted value is the solid center line and the confidence intervals are represented by the dashed lines.

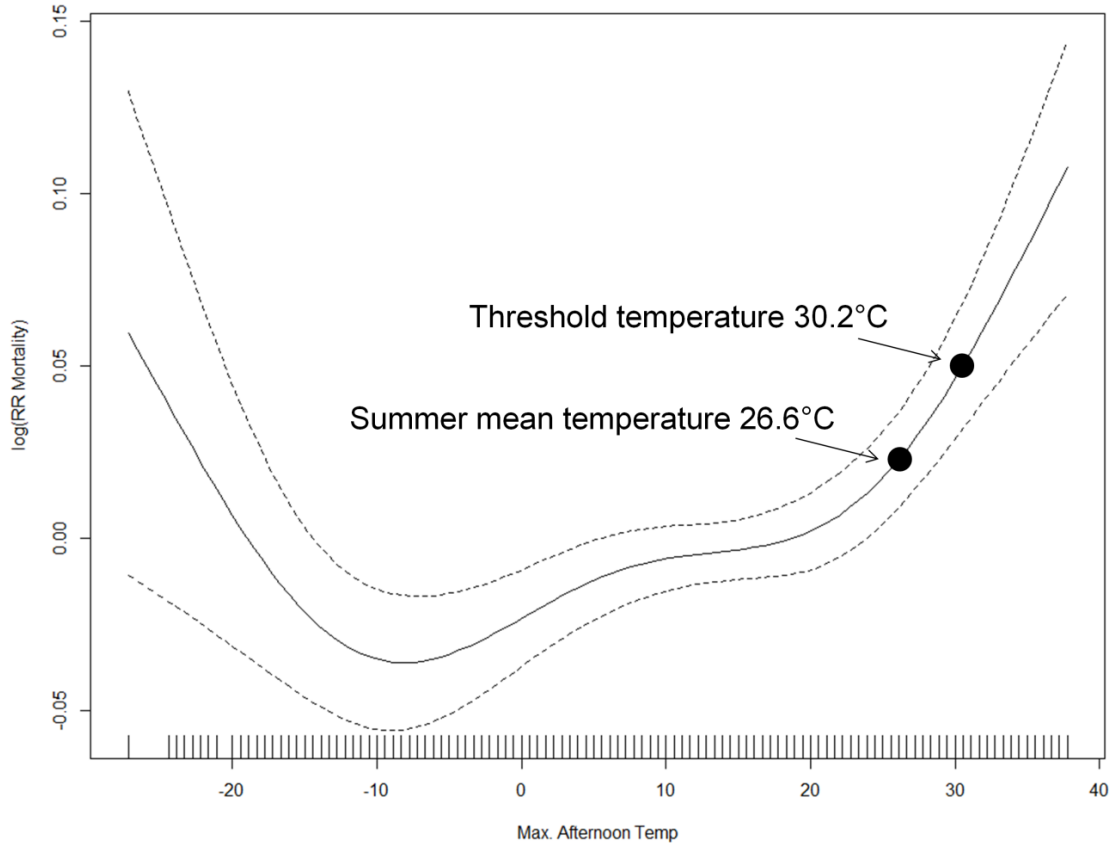


Figure 2. The seasonality in mortality in Minneapolis, MN, during 1992. The pattern emerges from a generalized additive model using daily mortality data. The model-predicted value is the solid center line and the confidence intervals are represented by the dashed lines.

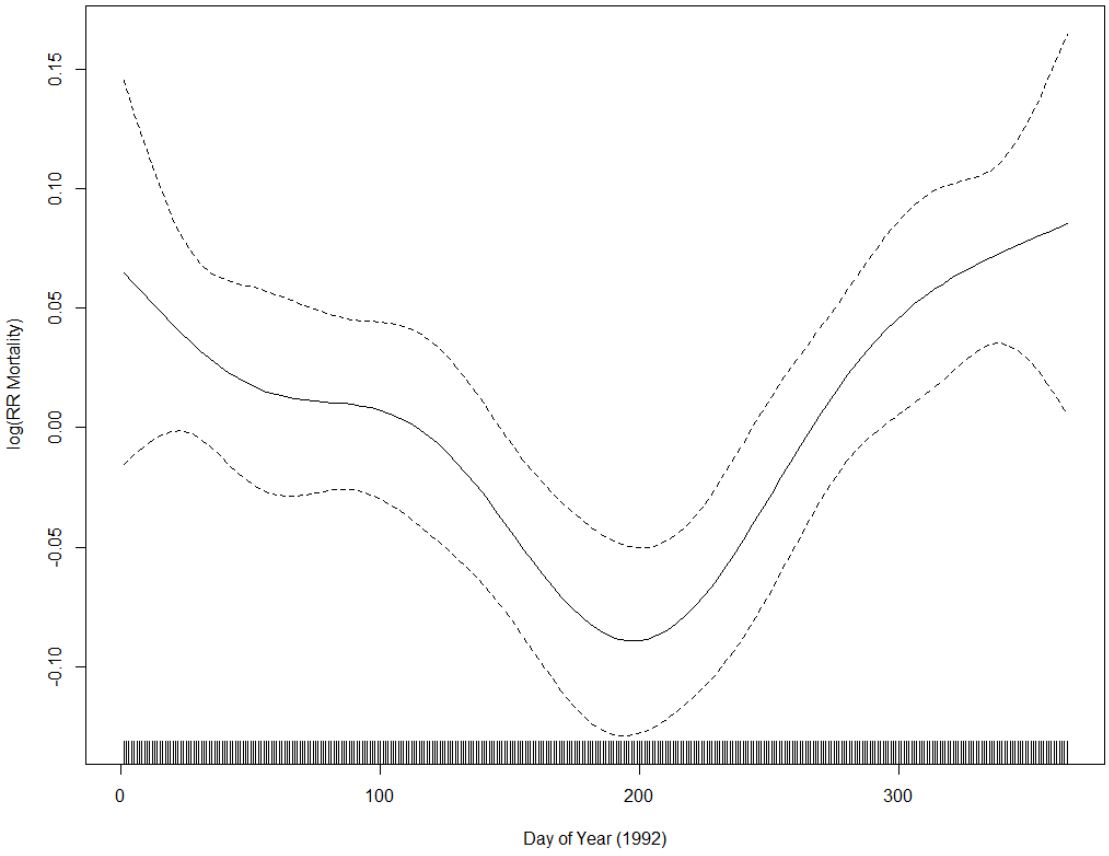


Figure 3. The ratio of mortality rates on high temperature days in target postal codes to non-target postal codes in each of the six cities examined. The ratios shown are the means across all training data sets (gray bars) and testing data sets (black bars). A ratio of 1.0 indicates that the mortality rate on hot days is equal between the two groups.

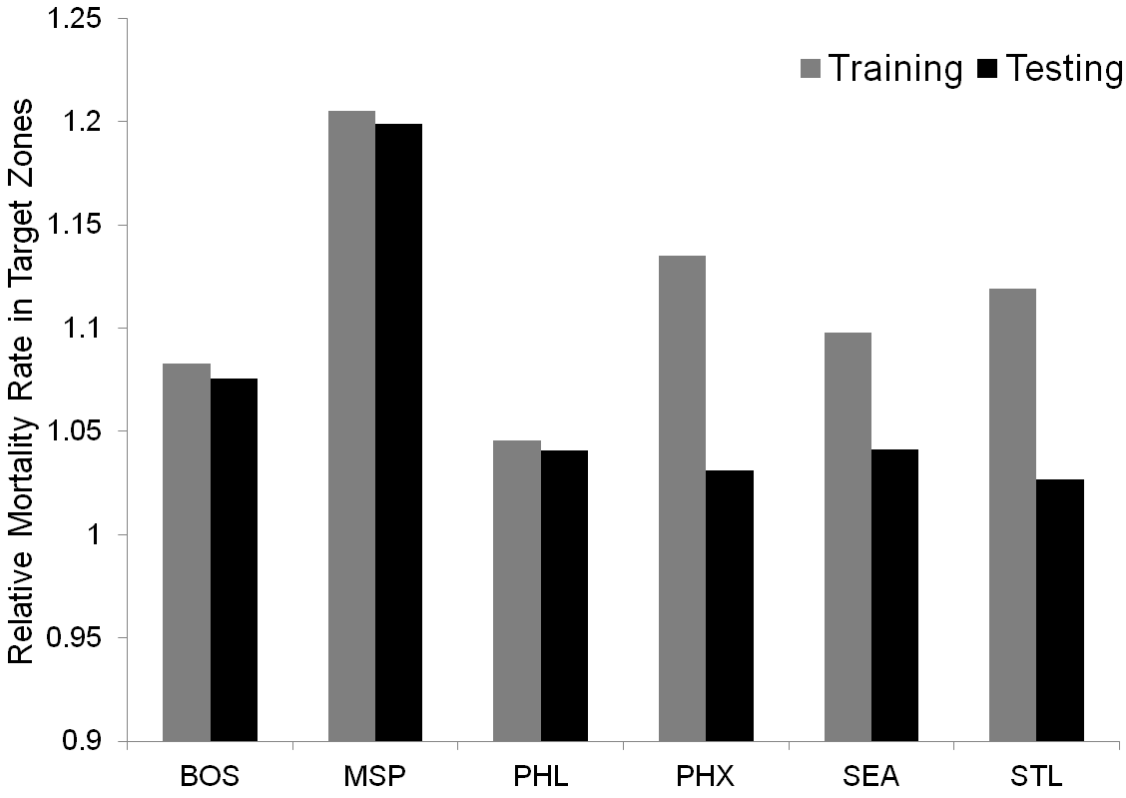


Figure 4. A year-by-year comparison of mortality ratios on high temperature threshold-exceeding days in target zones (black lines) and non-target zones (gray lines) in six study cities. A mortality ratio of 1.0 indicates that mortality on hot days was equal to the rate that would be expected for normal summer conditions.

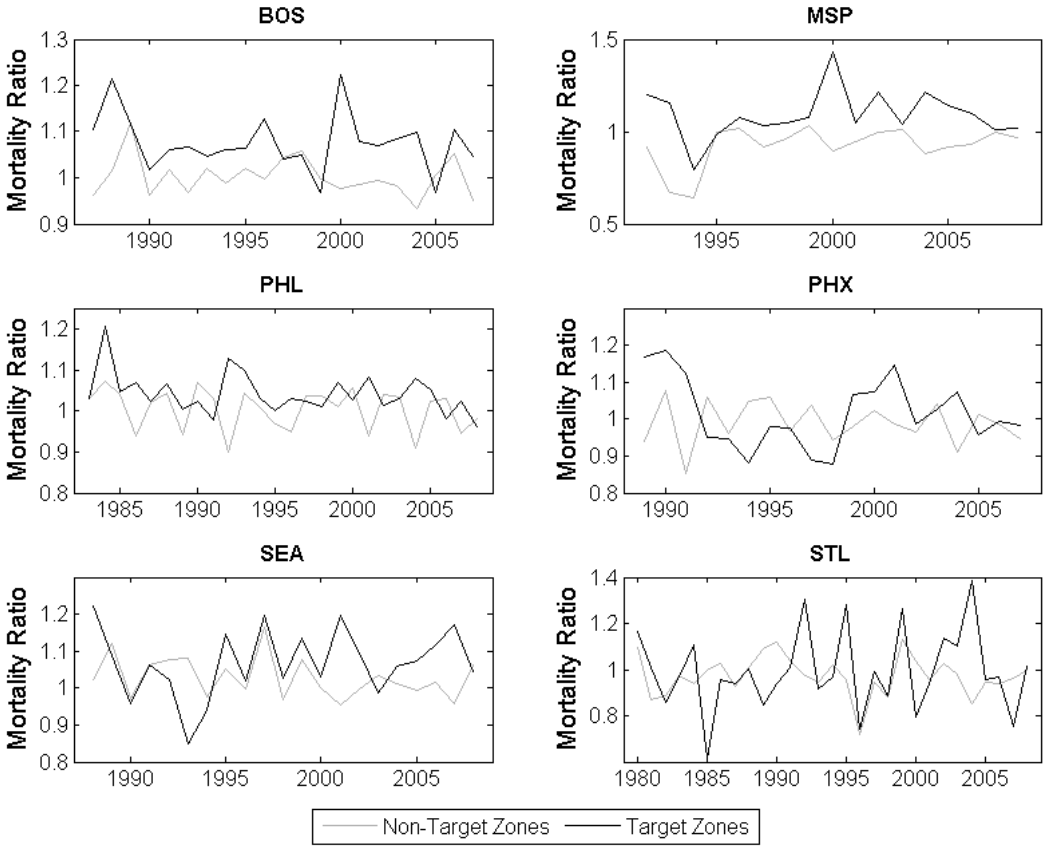
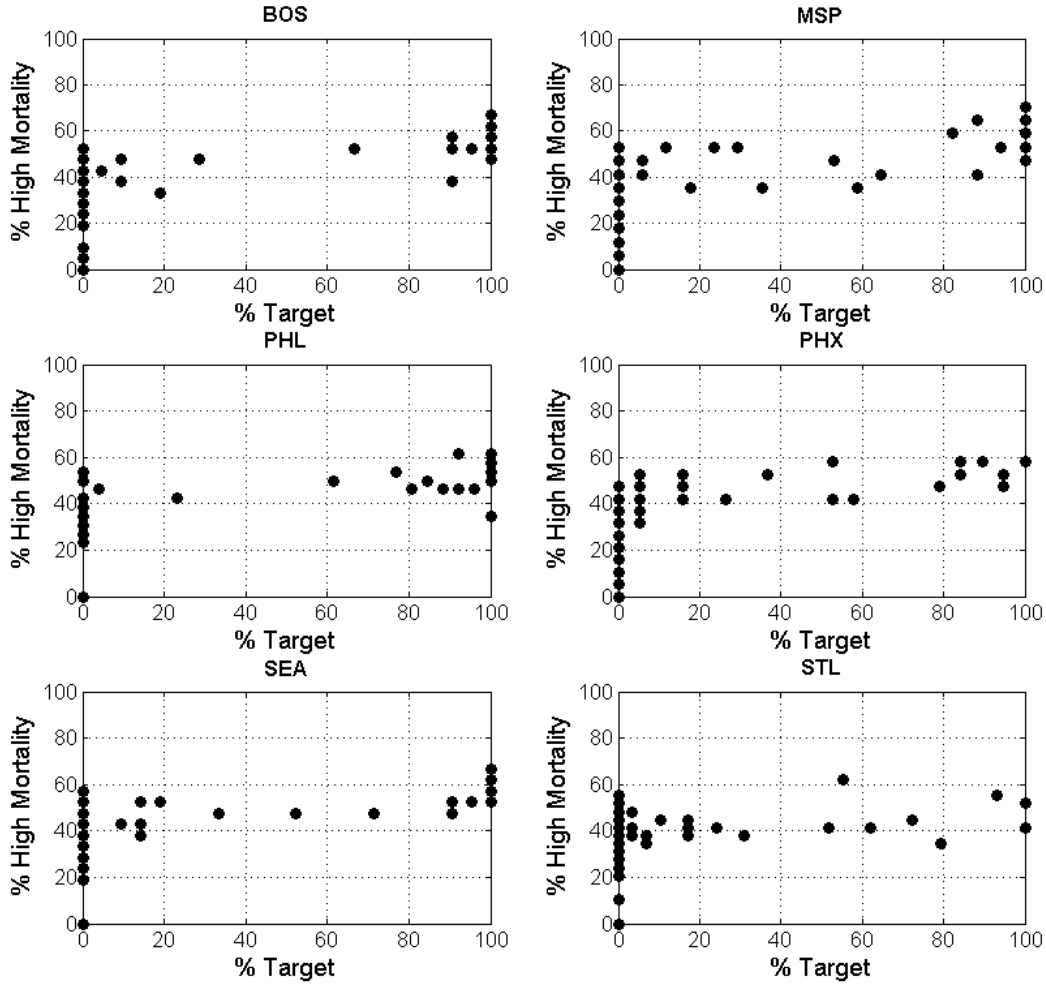


Figure 5. Scatterplots comparing the frequency that each postal code (each represented by one black circle) was identified as a target zone during a model training iteration and identified as a high-mortality zone during a model testing evaluation. A circle in the bottom left hand corner of each panel indicates places that were never identified as targets based on training data and were never associated with high mortality in testing data; the top right hand corner places that were always identified as targets and always had high mortality.



Tables

Table 1. Descriptive statistics of the meteorological and mortality data sets used in the study for six different United States cities. The table also includes information regarding the modeled temperature-mortality relationship for each city, the modeled threshold temperatures, and the number of spatial units identified as targets during model training iterations.

City	Boston, MA	Minneapolis-St. Paul, MN	Philadelphia, PA	Phoenix, AZ	Saint Louis, MO	Seattle, WA
Period of Record	1987–2007	1992–2008	1983–2008	1989–2007	1980–2008	1988–2008
Summer mean temperature in training periods (Min-Max)	25.5 (25.4-25.6)	26.4 (26.3-26.6)	29.1 (29.0-29.1)	39.9 (39.9-40.0)	30.3 (30.2-30.4)	22.7 (22.6-22.8)
Threshold temperature (Min-Max)	28.0 (27.8-28.2)	30.2 (30.0-30.6)	30.7 (30.6-30.8)	42.6 (42.4-43.0)	34.4 (33.8-35.8)	25.9 (25.7-26.1)
Model predicted mortality increase at threshold (95% CI)	1.018 (1.000-1.036)	1.023 (1.000-1.046)	1.015 (1.000-1.031)	1.019 (1.020-1.040)	1.016 (1.000-1.033)	1.021 (1.000-1.042)
Mean mortality on all above-threshold days relative to below-threshold days (Min-Max)	1.059 (1.056-1.063)	1.046 (1.041-1.050)	1.066 (1.060-1.070)	1.040 (1.035-1.055)	1.032 (1.026-1.046)	1.063 (1.059-1.067)
Number of postal codes	64	101	48	101	63	65
Mean number of target postal codes (Min-Max)	15.0 (13-16)	13.5 (10-15)	18.9 (16-21)	11.9 (10-15)	7.7 (4-11)	11.3 (10-14)

Table 2. Contingency tables comparing the relative frequency at which postal codes were identified as targets and non-targets during model training and associated with high mortality during model testing. The data shown are summed across all years in the study period for each respective city. Relative risk values greater than 1.0 indicate that postal codes identified as targets are more likely to be associated with high mortality in model testing compared to non-targets.

BOS		Training Sample	
		Target	Non-Targets
Withheld Year	High	158	373
	Not High	158	655
Relative Risk 1.378 (1.202, 1.580)			

MSP		Training Sample	
		Target	Non-Targets
Withheld Year	High	106	459
	Not High	124	1028
Relative Risk 1.493 (1.273, 1.751)			

PHL		Training Sample	
		Target	Non-Targets
Withheld Year	High	228	288
	Not High	264	468
Relative Risk 1.216 (1.067, 1.387)			

PHX		Training Sample	
		Target	Non-Targets
Withheld Year	High	82	565
	Not High	145	1127
Relative Risk 1.082 (0.899, 1.302)			

STL		Training Sample	
		Target	Non-Targets
Withheld Year	High	79	571
	Not High	140	1037
Relative risk 1.016 (0.842, 1.226)			

SEA		Training Sample	
		Target	Non-Targets
Withheld Year	High	108	434
	Not High	129	694
Relative Risk 1.184 (1.012, 1.386)			

Table 3. Descriptive statistics of total summer mortality and estimated excess heat-related mortality during the study period for each city examined. The column to the right represents summation across all six cities.

	BOS	MSP	PHL	PHX	SEA	STL	All Cities
Total Summer Deaths	63996	54204	96423	84536	53912	93317	446388
Total Summer Deaths (Targets Only)	19330	10090	47268	13020	10287	11800	111795
Total Excess Mortality	770.55	301.88	1198.80	456.05	533.29	229.08	3489.65
Total Excess Mortality (Targets Only)	439.79	312.06	766.54	50.85	113.12	58.94	1741.31
Percent of Total Deaths in Targets	30.21	18.62	49.02	15.40	19.08	12.65	25.04
Percent of Excess Deaths in Targets	57.08	103.37	63.94	11.15	21.21	25.73	49.90
Number of years	21	17	26	19	21	29	

References

Anderson, G. B., & Bell, M. L. (2011). Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. *Environmental Health Perspectives*,*119*(2), 210.

Antics, A., Pascal, M., Laaidi, K., Wagner, V., Corso, M., Declercq, C., & Beaudeau, P. (2013). A simple indicator to rapidly assess the short-term impact of heat waves on mortality within the French heat warning system. *International Journal of Biometeorology*, *57*(1), 75-81.

Arnfield, A. J. (2003). Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, *23*(1), 1-26.

Davis, R. E., Knappenberger, P. C., Michaels, P. J., & Novicoff, W. M. (2003). Changing heat-related mortality in the United States. *Environmental Health Perspectives*, *111*(14), 1712.

Gardner, M. J., & Altman, D. G. (1989). Statistics with confidence: confidence intervals and statistical guidelines.

Guo, Y., Barnett, A. G., & Tong, S. (2012). High temperatures-related elderly mortality varied greatly from year to year: important information for heat-warning systems. *Scientific reports*, *2*.

Hajat, S., Armstrong, B. G., Gouveia, N., & Wilkinson, P. (2005). Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology*, *16*(5), 613-620.

Hajat, S., Sheridan, S. C., Allen, M. J., Pascal, M., Laaidi, K., Yagouti, A., ... & Kosatsky, T. (2010). Heat-health warning systems: a comparison of the predictive capacity of different approaches to identifying dangerously hot days. *American Journal of Public Health*, *100*(6), 1137-1144.

Hondula, D. M., Davis, R. E., Leisten, M. J., Saha, M. V., Veazey, L. M., & Wegner, C. R. (2012). Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983-2008: a case-series analysis. *Environmental Health*, *11*(1), 1-11.

Hondula, D. M., Davis, R. E., Rocklöv, J., & Saha, M. V. (2013a). A time series approach for evaluating intra-city heat-related mortality. *Journal of Epidemiology and Community Health*.

Hondula, D. M., Vanos, J. K., & Gosling, S. N. (2013b). The SSC: a decade of climate–health research and future directions. *International Journal of Biometeorology*, 1-12.

Laaidi, K., Zeghnoun, A., Dousset, B., Bretin, P., Vandentorren, S., Giraudet, E., & Beaudeau, P. (2012). The impact of heat islands on mortality in Paris during the August 2003 heat wave. *Environmental Health Perspectives*, *120*(2), 254.

Minnesota Population Center. National Historical Geographic Information System: Version 2.0. Minneapolis, MN: University of Minnesota 2011.

National Weather Service. (2006). Heat wave: a major summer killer. http://www.nws.noaa.gov/om/brochures/heat_wave.shtml

Pascal, M., Wagner, V., Le Tertre, A., Laaidi, K., Honoré, C., Bénichou, F., & Beaudeau, P. (2013). Definition of temperature thresholds: the example of the French heat wave warning system. *International Journal of Biometeorology*,*57*(1), 21-29.

Reid, C. E., O'Neill, M. S., Gronlund, C. J., Brines, S. J., Brown, D. G., Diez-Roux, A. V., & Schwartz, J. (2009). Mapping community determinants of heat vulnerability. *Environmental Health Perspectives*, *117*(11), 1730.

Reid, C. E., Mann, J. K., Alfasso, R., English, P. B., King, G. C., Lincoln, R. A., ... & Balmes, J. R. (2012). Evaluation of a heat vulnerability index on abnormally hot days: an environmental public health tracking study. *Environmental Health Perspectives, 120*(5), 715.

Roberts, S. (2004). Interactions between particulate air pollution and temperature in air pollution mortality time series studies. *Environmental Research*, *96*(3), 328-337.

Rocklöv, J., Forsberg, B., & Meister, K. (2009). Winter mortality modifies the heat-mortality association the following summer. *European Respiratory Journal*,*33*(2), 245-251.

Saha, M.V., Davis, R.E., Hondula, D.M. (2013). Mortality displacement as a function of event strength in seven U.S. Cities. *American Journal of Epidemiology, submitted.*

Sheridan, S. C. (2007). A survey of public perception and response to heat warnings across four North American cities: an evaluation of municipal effectiveness. *International Journal of Biometeorology*, *52*(1), 3-15.

Sheridan, S. C., & Dolney, T. J. (2003). Heat, mortality, and level of urbanization: measuring vulnerability across Ohio, USA. *Climate Research*,*24*(3), 255-265.

Sheridan, S. C., Kalkstein, A. J., & Kalkstein, L. S. (2009). Trends in heat-related mortality in the United States, 1975–2004. *Natural Hazards*, *50*(1), 145-160.

Sheridan, S. C., & Kalkstein, L. S. (2004). Progress in heat watch-warning system technology. *Bulletin of the American Meteorological Society*, *85*(12), 1931-1941.

Smargiassi, A., Goldberg, M. S., Plante, C., Fournier, M., Baudouin, Y., & Kosatsky, T. (2009). Variation of daily warm season mortality as a function of micro-urban heat islands. *Journal of Epidemiology and Community Health*, *63*(8), 659-664.

Smith, C., & Levermore, G. (2008). Designing urban spaces and buildings to improve sustainability and quality of life in a warmer world. *Energy Policy*,*36*(12), 4558-4562.

Solecki, W. D., Rosenzweig, C., Parshall, L., Pope, G., Clark, M., Cox, J., & Wiencke, M. (2005). Mitigation of the heat island effect in urban New Jersey.*Global Environmental Change Part B: Environmental Hazards*, *6*(1), 39-49.

Vaneckova, P., Beggs, P. J., & Jacobson, C. R. (2010). Spatial analysis of heat-related mortality among the elderly between 1993 and 2004 in Sydney, Australia. *Social Science & Medicine, 70*(2), 293-304.

Wilby, R. L. (2003). Past and projected trends in London's urban heat island.*Weather*, *58*(7), 251-260.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(1), 95-114.

Wood, S. (2006). *Generalized additive models: an introduction with R*. CRC Press.

Zhang, K., Rood, R. B., Michailidis, G., Oswald, E. M., Schwartz, J. D., Zanobetti, A., ... & O'Neill, M. S. (2012). Comparing exposure metrics for classifying ‘dangerous heat’in heat wave and health warning systems. *Environment International*, *46*, 23-29.

Chapter 6. Review

1. Summary

This body of research has made multiple contributions to the fields of applied climatology and environmental health. A statistical method for calculating the temperature-mortality relationship for spatial units with sample sizes too small for conventional methods was developed. This method was applied to long-term daily mortality records, which revealed statistical differences in risks related to extreme heat events within a suite of large metropolitan areas in the United States over the past two decades. An ecological framework was the adopted to determine factors most closely associated with variability in intra-city risk. To the best of our knowledge, this dissertation represents the first comprehensive effort to assess intra-city variability in heat-related mortality using long-term data collected from multiple cities. The major findings from each chapter of the main text (2–5) are reported below, followed by general conclusions integrating across the chapters and identification of promising areas for continued and new exploration.

Chapter 2. Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983–2008: A case-series analysis.

- Georeferenced mortality records including the postal code of residence of the decedent show large differences in the mortality rate on hot summer days in one large U.S. city

- Intra-city variability in heat-related mortality is associated with spatial variability in a number of socioeconomic and environmental factors, and risk is highest in locations with high surface temperatures, more elderly, lower per capita income, and more intensely-developed land.
- Estimates of the impact of high temperatures on mortality rates vary depending on the specific hourly temperature measurement chosen for analysis

Chapter 3. A time series approach for evaluating intra-city heat-related mortality

- A statistical technique is introduced that enables estimates of heat-related mortality for spatial units with small sample sizes, combining ideas from different disciplines
- Low daily mortality rates inhibit direct modeling of seasonality and long-term time trends using a single smoothing spline and failure to accurately model these patterns can lead to misspecification of temperature-mortality associations
- Application of this new method to Philadelphia data identifies a set of high-risk locations not identical to those presented in Chapter 2

Chapter 4. The predictability of high-risk zones for heat-related mortality in seven U.S. cities

- There is sufficient temporal consistency in intra-city heat-related mortality patterns such that high-risk zones can be identified with data not used by a predictive model

- For the six study cities where a significant association between temperature and mortality was found, over half of excess deaths on above-threshold temperature days occurred in locations predicted to be high risk
- The predictability of high-risk zones varied considerably between study cities

Chapter 5. Geographic dimensions of heat-related mortality in seven U.S. cities

- Heat-related mortality is spatially variable within large U.S. metropolitan areas where there is a statistical link between high temperatures and elevated mortality
- Intra-city variability in heat-related mortality is substantially greater than inter-city differences, and statistically significant high mortality on hot days is confined to 11-44% of the postal codes comprising the cities examined
- Postal codes with higher heat-related mortality rates are associated with many characteristics that would be expected based on previous research, including more elderly residents, lower incomes, and greater extent of developed land, but the specific combination of factors that explain the spatial pattern varies from city to city
- The areas within cities where mortality rates are the highest on hot summer days may contrast with those expected using traditional vulnerability assessments, warranting a data-driven, city-specific approach to understand spatial variability in risk

2. Conclusions

There has been little formal evidence of intra-city variability in heat-related mortality rates based on health outcome data published to date. In the major U.S. cities we examined where high temperatures are associated with elevated mortality rates, significant variability in sensitivity to heat exists at the postal code scale. This dissertation presents the first comprehensive assessment of intra-city variability in heat-related mortality rates that utilizes daily, multidecadal, georeferenced health records from climatically diverse locations. The results add a finer-scale layer to our understanding of risks related to extreme heat, as most previous research has focused on the aggregate response of the populations of large metropolitan areas. For decades it has been known that there are important inter-city differences in sensitivity to heat, thereby encouraging the development of city-specific warning systems. This research shows that intra-city differences also exist and can be quite large. This evidence should provide additional motivation for public officials to improve the level of spatial specificity in their planning and response efforts focused on extreme heat.

The statistical method presented in Chapter 3 offers other researchers in the field an accessible means of determining heat-health associations for small spatial units in other jurisdictions. The method could also be readily applied for other environmental health hazards where the analysis is complicated by important inter- and intra-annual variability in the response variable. The methods currently used to estimate temperature-mortality associations for time series with higher daily mortality counts simply are not reliable for the sample sizes investigated in this dissertation—often three or fewer deaths per postal

code per day. Borrowing more reliable information from larger-scale units is a method that produces model estimates of temporal and temperature-related effects that are more plausible than those generated simply by applying established methods to these smaller-sample data sets.

The characteristics of places associated with higher heat-related mortality rates varied across the cities examined in this dissertation (Chapter 4). This is vital information for the research community, public health officials, and government planners involved in identifying and helping vulnerable populations in metropolitan areas. Although the characteristics of high-risk places were largely consistent with previous research, the specific combination of factors included in statistical models for each city varied. This means that if one were to apply a consistent scheme for identifying vulnerable locations within a suite of cities, there would likely be high-risk locations missed and low-risk locations receiving potentially unnecessary aid.

This dissertation provides support for the notion that built environment features can contribute to elevated risk of heat-related mortality that may arise because of urban heat island effects. The intensity of developed land cover was associated with intra-city variability in heat-related mortality in several of the cities examined and was more closely related to the variability than various demographic indicators. This encourages the consideration of—at least in some cities—the adoption of design features aimed at reducing thermal stress for residents.

The demographic characteristics of high-risk locations identified from this research can be used by emergency management personnel and planners in assessing the types of

mitigation and intervention strategies that may be most effective for sensitive populations within their jurisdictions. This research demonstrates that these personnel should not assume that established risk factors necessarily translate to their areas of governance, and where, possible, efforts should be made to empirically understand risk variability within their service region.

The fact that there is some temporal consistency in the year-to-year spatial pattern in heat-related mortality (Chapter 5) is encouraging for the future application of the outcomes of this dissertation. For the majority of the cities we examined, the places associated with high heat-related mortality in any given year are often those that would be predicted by an empirical model based on training data. Across the study sites, more than half of excess deaths on above-threshold days were found in places that would have been identified as target locations for intervention activities based on quantitative evidence. As heat-related deaths are believed to be largely, if not fully, preventable, effective intervention measures solely focused on these particular locations could go a long way toward reducing the unnecessary health burden associated with extreme heat.

3. Future Research Opportunities

The possibility of increasingly frequent and severe heat waves in the future, combined with an increasingly urban population, indicates that heat-related morbidity and mortality may be a persistent challenge in the coming decades. Research in the areas represented in this dissertation offers the means of providing new information to public health officials, city planners, policymakers, and other researchers to help develop the most effective and efficient strategies for saving lives and reducing healthcare costs. As the chapters of this dissertation represent the first comprehensive, empirically-driven, local-scale assessment of heat-related mortality across multiple U.S. cities, it is the author's hope that this body of work serves as the starting point for research in the coming years spanning multiple disciplines that are ultimately connected to understanding how natural hazards like extreme heat impact people and how any negative impacts can be minimized. A number of scientific and application-driven questions have emerged during the course of this research, and many doors have been opened for future opportunities from the compilation of data and initial set of analyses. A sample of potential future projects is discussed below, some of which have already been started by collaborators at the University of Virginia and elsewhere.

Understanding drivers of variability in the surface urban heat island

Follow-up work to Chapter 2 included the creation of a collection of remote imagery of the Philadelphia surface urban heat island. Over the study period, 21 images were found for clear summer days with relatively high temperatures, and the collection includes images from a variety of months and years. Preliminary analysis of this image collection revealed significant variability in the surface urban heat island during warm days, which

discourages the approach adopted in Chapter 2 (and elsewhere in the peer-reviewed literature) of using only one or two satellite images to represent intra-urban microclimatic variability. It was determined that there is low confidence that any one available image will be representative of the mean surface urban heat island on hot days.

Subsequent analysis revealed an interesting pattern in which there is an association between time of year and the range of surface temperatures across the County boundaries. More specifically, the contrast between the city center and surrounding suburbs was larger in the late spring and early summer and smaller in late summer and early fall. This decreasing range mostly was driven by increasing relative minimum temperatures—the suburban heat island signature more closely resembled that of the city center late in the season. If this collection of images represents a true seasonal signal in spatial variability in urban heat island intensity, it would be useful to understand the physical drivers of this seasonal pattern. One hypothesis is that there is greater moisture content in soil and vegetation in suburban areas in late spring compared to late summer, and evaporation of this moisture provides cooling relief.

The finding also has potential implications for heat stress differences between suburban and city-center residents, in that those living in the urban core experience higher temperatures than those living in surrounding areas early in the season. Previous research has shown that early-season heat waves have a disproportionate effect on mortality, and also that those who live in more highly developed areas are more vulnerable to heat. It could be the case that early-season heat poses a greater threat to city-center residents, adding a temporal dimension to targeted intervention strategies. Understanding the health

impacts of this seasonal variability in the heat island can also provide additional justification and motivation for adoption of design features aimed at reducing urban heat island effects.

Quantifying personal thermal exposure in urban landscapes

One of the major assumptions of this dissertation adopted in Chapters 2 and 5 is that the environment in which people reside serves as a good representation of the conditions they experience. In Chapter 2, for example, the surface temperature of the postal code of residence of decedents was compared to intra-city variability in mortality rates to determine if places with higher temperatures are also associated with higher mortality rates. It is not necessarily the case, however, that the surface (or air) temperature of a given postal code of residence accurately portrays the thermal experience of residents of that postal code, and thus it is difficult to draw definitive conclusions, particularly at the individual level, about the contribution of the built environment to heat-related health risks. The behavior of individuals moving through complex urban landscapes is largely unaccounted for in present research. Many individuals spend the majority of their time in climate-controlled residences, vehicles, and workplaces. Others do not have access to air conditioning or work in settings where heat exposure is common. Urban residents, the elderly, outdoor workers, those in poverty, and children are most susceptible to extreme heat, but it is unclear how variability in heat exposure contributes to this susceptibility. Patterns of exposure might also vary by season, weather, and time of day. This variability necessitates measurement of personal exposure to harmful levels of heat that is largely absent in present research and practice.

We have proposed a pilot project to instrument individuals moving throughout urban landscapes with wrist-worn temperature sensors with the goal of understanding temperature from a person-based, rather than place-based, perspective. Related future work could also involve capturing data through mobile phones, vehicles, and other fixed and mobile sources. In the coming years, individuals could theoretically receive personal alerts when their individual level of exposure reaches a certain threshold.

Mapping heat-health impacts at multiple scales

This dissertation shifts the focus of analysis of heat-related mortality from the city-wide scale to the intra-city (postal code) scale, and this shift makes it possible to draw different conclusions regarding where the most vulnerable populations are located that one can reach using coarsely-aggregated data from large populations. However, meaningful and helpful research in this area is possible along the entire spectrum of scales ranging from inter-city down to the level of individual residences. As they are derived from an ecological design, the results of this dissertation must be treated appropriately, in that the patterns and associations seen in certain locations and at this scale of analysis do not necessarily extend to other locations and other scales. For example, while this dissertation has shown that in some cities postal codes with higher percentages of developed land are associated with higher mortality rates on hot days, this does not necessarily mean that cities with more developed land should expect higher heat-related mortality than those with more open space. Thus, more research is needed to understand the spatial scales at which certain factors are important determinants of risk. Exploring spatial variability at

different scales can also provide a more comprehensive perspective on vulnerable populations within the service regions of public officials at various levels of government. The data compiled in this dissertation make it immediately possible to conduct analysis comparing heat-related mortality rates across counties of certain states, as well as exploring census-tract-level variability within certain locations. In some cases, individual-level residential information is available. Other stakeholders may have interests at different spatial scales that are appropriate for their planning, intervention, and mitigation efforts.

Assessing emergency managers' perceptions of spatial variability in heat-related risk and willingness to use empirically-driven results

The author and contributors to this dissertation have intended from the start of the project to see that this work is communicated to public officials in the study cities in a manner that makes improvements to their heat-health response measures possible. Directly engaging these stakeholders to share project results and learn end-user needs may be an effective means of doing so. Based on the scientific literature, as well as documentation produced by health and emergency departments in the study cities, there is very little evidence that geographical dimensions of heat-related risk are an integral part of the response efforts in the study cities. We have developed and are in the process of conducting an interview-based project to learn emergency managers' understanding and perceptions of spatial variability in risk within their service regions. Interviews conducted to date suggest that while these public officials perceive differential risk within their

cities, formal mechanisms for targeting certain locations with tailored interventions do not exist. Further, in some cases the locations perceived to be high-risk for heat-related mortality by public officials have not been historically associated with higher risk on hot days (perhaps because of successful interventions), and in other cases locations with higher historical risks were unexpected. While initially engaging officials in the conversation has proven challenging thus far, we believe that the conversations that have occurred have provided new perspective and can help shape future response efforts. A continued, persistent effort to ensure that this research reaches a range of stakeholders in each of the study cities is imperative to see that it achieves its maximum societal impact.