

# Comparison of temperature-mortality associations estimated with different exposure metrics

Kate R. Weinberger<sup>a,b,\*</sup>, Keith R. Spangler<sup>b,c</sup>, Antonella Zanobetti<sup>d</sup>, Joel D. Schwartz<sup>d,e</sup>, Gregory A. Wellenius<sup>b</sup>

**Background:** Studies of the short-term association between ambient temperature and mortality often use temperature observations from a single monitoring station, frequently located at the nearest airport, to represent the exposure of individuals living across large areas. Population-weighted temperature estimates constructed from gridded meteorological data may offer an opportunity to improve exposure assessment in locations where station observations do not fully capture the average exposure of the population of interest.

**Methods:** We compared the association between daily mean temperature and mortality in each of 113 United States counties using (1) temperature observations from a single weather station and (2) population-weighted temperature estimates constructed from a gridded meteorological dataset. We used distributed lag nonlinear models to estimate the 21-day cumulative association between temperature and mortality in each county, 1987–2006, adjusting for seasonal and long-term trends, day of week, and holidays.

**Results:** In the majority (73.4%) of counties, the relative risk of death on extremely hot days (99th percentile of weather station temperature) versus the minimum mortality temperature was larger when generated from the population-weighted estimates. In contrast, relative risks on extremely cold days (first percentile of weather station temperature) were often larger when generated from the weather station observations. In most counties, the difference in associations estimated from the two temperature metrics was small.

**Conclusions:** In a large, multi-site analysis, temperature-mortality associations were largely similar when estimated from weather station observations versus population-weighted temperature estimates. However, spatially refined exposure data may be more appropriate for analyses seeking to elucidate local health effects.

## Introduction

The relationship between ambient temperature and acute mortality risk is well-documented. Results from communities around the globe consistently describe a nonlinear relationship in which temperatures of either extreme are associated with a higher risk of death.<sup>1–3</sup> The studies giving rise to these findings often relate daily values of temperature observed at a single weather station (often located at an airport) to the daily number of deaths occurring within a city, county, or similar unit of geographic aggregation using a time series or case-crossover design. In such designs, temperature measurements from local weather stations are used as a proxy for the average exposure of the population of interest.<sup>4,5</sup>

Given that weather stations are often located at airports and that airports are often located towards the periphery of population centers, it is plausible that observed temperatures could systematically differ from the true population-average temperature in any given area. For example, residents of cities where weather stations are located near the coast while the majority of people live farther inland (e.g., Boston, Massachusetts; Los Angeles, California) may on average experience more extreme temperatures than those recorded at the station due to the moderating influence of the ocean. The use of spatially refined temperature estimates (such as those developed by Shi et al., Thornton et al., and the PRISM Climate Group<sup>6–8</sup>) may reduce measurement error resulting from systematic differences between weather station observations and population average temperatures. For example, gridded temperature estimates can be used to assign exposure on a finer spatial scale (e.g., at the postal code or county level) or can be used in combination with census data to estimate population-weighted daily average exposures. Of course, the degree to which such products reduce measurement

<sup>a</sup>School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada; <sup>b</sup>Department of Epidemiology, Brown University School of Public Health, Providence, Rhode Island; <sup>c</sup>Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, Rhode Island; <sup>d</sup>Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; and <sup>e</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts.

\*Corresponding author. Address: School of Population and Public Health, University of British Columbia, 2206 East Mall, Vancouver, British Columbia, V6T 1Z3, Canada. E-mail: kate.weinberger@ubc.ca (K.R. Weinberger).

The computing code and exposure data required to replicate the results reported in this manuscript can be obtained by contacting the corresponding author. The mortality data can be obtained by making a data request to the National Center for Health Statistics.

Copyright © 2019 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of Environmental Epidemiology. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

Environmental Epidemiology (2019) 3:e072

Received: 24 April 2019; Accepted 9 September 2019

Published online 30 September 2019

DOI: 10.1097/EE9.000000000000072

## What this study adds

Studies of the short-term association between temperature and mortality often use temperature observations taken at a weather station to represent the average exposure of individuals living across large geographic areas. In this study, we compared temperature-mortality associations estimated with daily temperature observations from a single weather station to associations estimated with daily population-weighted temperature estimates in each of 113 populous US counties. In most counties, we found little difference between associations estimated from these two exposure metrics, suggesting that the use of weather station observations has not led to systematic under or overestimation of the relationship between temperature and mortality. However, spatially refined exposure data may be more appropriate for analyses seeking to inform local adaptation efforts or in locations far away from weather stations.

error depends on the quality of the underlying prediction models, with results from gridded datasets yielding potentially more or less systematic error compared with station data, which reflect actual measurements.

In addition to systematic measurement error as described above, the use of both station observations and gridded estimates can lead to random measurement error. For example, the use of an unbiased but imperfect temperature measurement as a proxy for the true temperature at a weather station is subject to classical measurement error, which is expected to attenuate effect sizes.<sup>9</sup> Berkson error—which is expected to lead to decreased precision but not otherwise bias effect estimates—may arise from the use of a single value of temperature to represent the exposure of many individuals living across municipalities in which there is spatial variability in daily temperatures.<sup>10,11</sup> Results from studies using information from gridded weather datasets may have more or less random measurement error compared with station observations.

Several recent studies suggest that the magnitude of the temperature–mortality association is either similar<sup>12,13</sup> or larger<sup>14</sup> when assessed using spatially refined exposure estimates versus weather station measurements. However, these studies were conducted in relatively limited geographic areas, leaving open the question of whether their findings are specific to these locations or more broadly generalizable. In this study of >100 counties in the contiguous United States, we evaluated whether the use of weather station temperature measurements leads to different inferences about the association between temperature and mortality when compared with the use of population-weighted temperature estimates, which are assumed to be a more accurate proxy for population average exposure. Specifically, we modeled the association between daily mean temperature and mortality in each county using (1) a daily time series of temperature observations from a single first-order weather station within the county boundaries and (2) a daily time series of

population-weighted temperature estimates constructed from the PRISM. PRISM is a publicly available, gridded dataset developed for the contiguous United States by spatially interpolating meteorological observations from a variety of observatory networks using a multivariate regression model accounting for elevation, topography, and other geophysical characteristics.<sup>7,15</sup>

## Methods

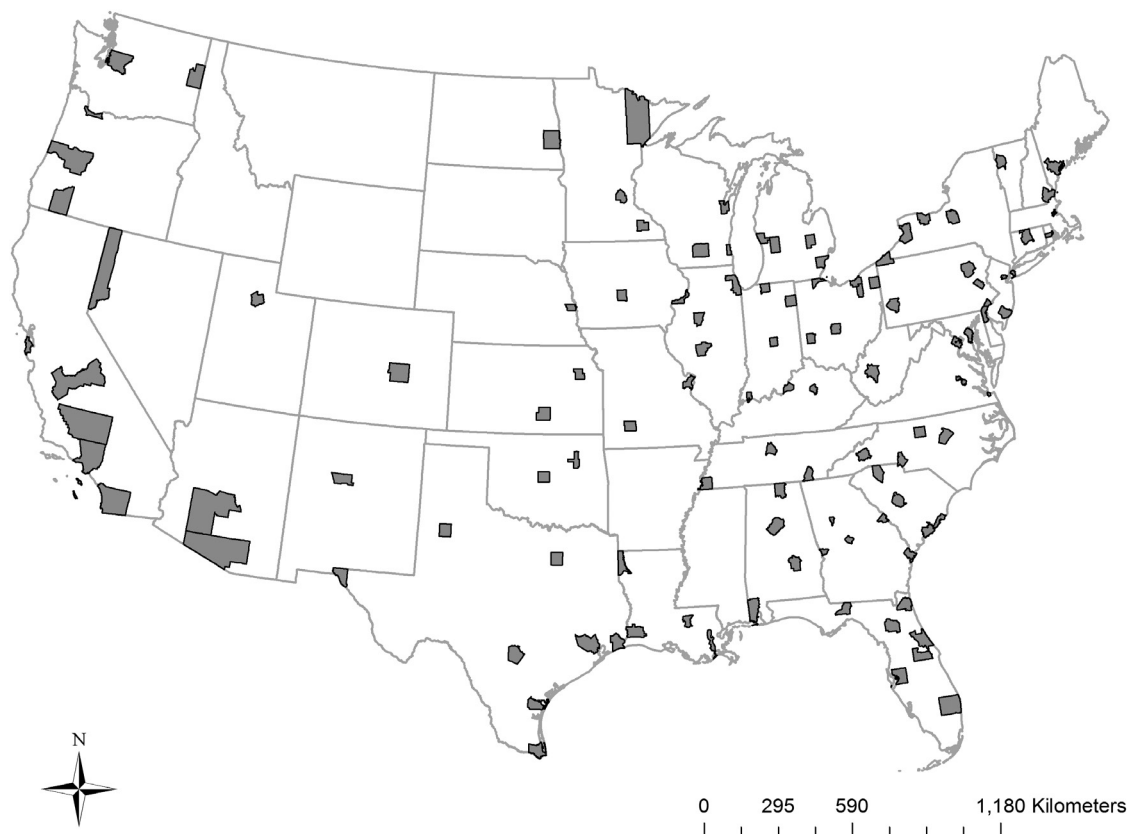
### Data sources

We obtained individual-level data on all deaths occurring in US counties with a population >100,000 from the National Center for Health Statistics. From these data, we constructed a time series of the daily number of deaths (excluding those due to external causes) in each county for which we had continuous mortality data during a 20-year study period of 1987–2006.

We obtained observations of daily mean temperature at all first-order weather stations in the United States from the Integrated Surface Database Lite.<sup>16</sup> We obtained gridded estimates for daily mean temperature for the contiguous United States at a 4-km resolution from PRISM.<sup>7,15</sup> We used these gridded estimates to create a time series of population-weighted daily mean temperature estimates for each county in the contiguous United States, as previously described.<sup>17</sup>

### Site selection

From the 297 US counties in the contiguous United States for which we had mortality data in every year, we selected 113 counties in 42 states (Fig. 1) that met the following criteria throughout the study period of 1987–2006: (1) at least one first-order weather station reporting to the Integrated Surface Database Lite database located within the county boundaries, (2) temperature data reported from the weather station in every year, (3)



**Figure 1.** Location of the 113 study counties within the contiguous United States.

no change in the weather station identifier, and (4) <1% of days missing temperature values at the weather station. We classified days as missing if that day's value of daily mean temperature was calculated from fewer than 18 hourly temperature measurements. Two of the selected counties—Los Angeles County in California and Queens County in New York—contained more than one first-order weather station. We used temperature observations from Los Angeles International Airport and LaGuardia Airport for Los Angeles and Queens, respectively.

### Statistical analysis

In each county, we began by modeling the association between daily mean temperature observed at the county's first-order weather station (hereafter, "station temperature") and daily counts of mortality. Specifically, we used distributed lag nonlinear models<sup>18</sup> with a quasi-Poisson distribution to model the cumulative association between station temperature and mortality over 21 days of lag. We modeled the temperature–response curve with a quadratic B-spline and three internal knots at the 10th, 75th, and 90th percentiles of county-specific temperature, and the lag-response curve with a natural cubic B-spline with three knots at equally spaced values on the log scale (i.e., at values of approximately 1, 3, and 8). We adjusted each county-specific model for seasonal and long-term time trends (natural cubic B-spline with 8 degrees of freedom per year), day of week, and federal holidays. We then used the cumulative association between temperature and mortality in each county to fit a multivariate meta-analytic model, including the mean and range of the county-specific temperature distribution as observed at the weather station over the 20-year study period as predictors. From this model, we extracted the best linear unbiased prediction of the temperature–mortality association for each county, centering the spline of temperature at the county-specific temperature of minimum mortality. This modeling strategy is consistent with a previous large, multi-country study of temperature and mortality.<sup>2</sup>

Next, we repeated these steps, but using the daily population-weighted temperature estimates from PRISM (hereafter, "PRISM temperature") for each county in place of station temperature. While in some counties there were a very small number of days missing values of station temperature, there were no days with missing values of PRISM temperature. As the primary analysis, we used data from every single day in the PRISM temperature time series to estimate the temperature–mortality association. However, in sensitivity analysis, we additionally explored the impact of excluding days from the PRISM temperature time series that had a missing value in the station temperature time series.

For each county, the process described above yielded two exposure–response curves describing the nonlinear association between temperature and mortality, one constructed from station temperature and the other constructed from PRISM temperature. Hereafter, we refer to the exposure–response curves generated using the weather station temperature observations as "station curves" and the population-weighted temperature estimates as "PRISM curves." Both the shape and the temperature at which mortality risk is lowest (i.e., the minimum mortality temperature) may vary between the station and PRISM curves within a county.

To quantitatively assess differences between station and PRISM curves, we compared the relative risk (RR) for the temperature–mortality association generated from each county's station and PRISM curves at four values of temperature, using the county- and curve-specific minimum mortality temperature as the reference. Specifically, for each county, we made these comparisons at the 1st, 2.5th, 97.5th, and 99th percentiles of the station temperature distribution. For example, if the 99th percentile of the station temperature distribution in a given county was 35°C, we would compare the RR at 35°C on the

station curve to the RR at 35°C on the PRISM curve in that county. In a different county where the 99th percentile of station temperature was 30°C, we would compare RRs from the two curves at 30°C. Thus, within each county, we made comparisons between the two exposure–response curves at a constant value of temperature, but allowed the temperature at which this comparison was made to vary across counties to accommodate different climates.

In secondary analysis, we instead compared the RR for each county's station and PRISM curves at a constant percentile of the temperature distribution, rather than at a county-specific constant value of temperature. For example, in each county, we compared the RR at the 99th percentile of station temperature on the station curve to the RR at the 99th percentile of PRISM temperature on the PRISM curve. This approach differs from our main analysis, in which we made comparisons in each county at a constant value of temperature (i.e., at the 99th percentile of station temperature).

Lastly, we computed the fraction of deaths (and 95% empirical confidence intervals [eCI]) attributable to heat and cold for each curve in each county and across all counties.<sup>19</sup> For each exposure–response curve in each county, we defined heat as all temperatures above the county-specific minimum mortality temperature and cold as all temperatures below the county-specific minimum mortality temperature.

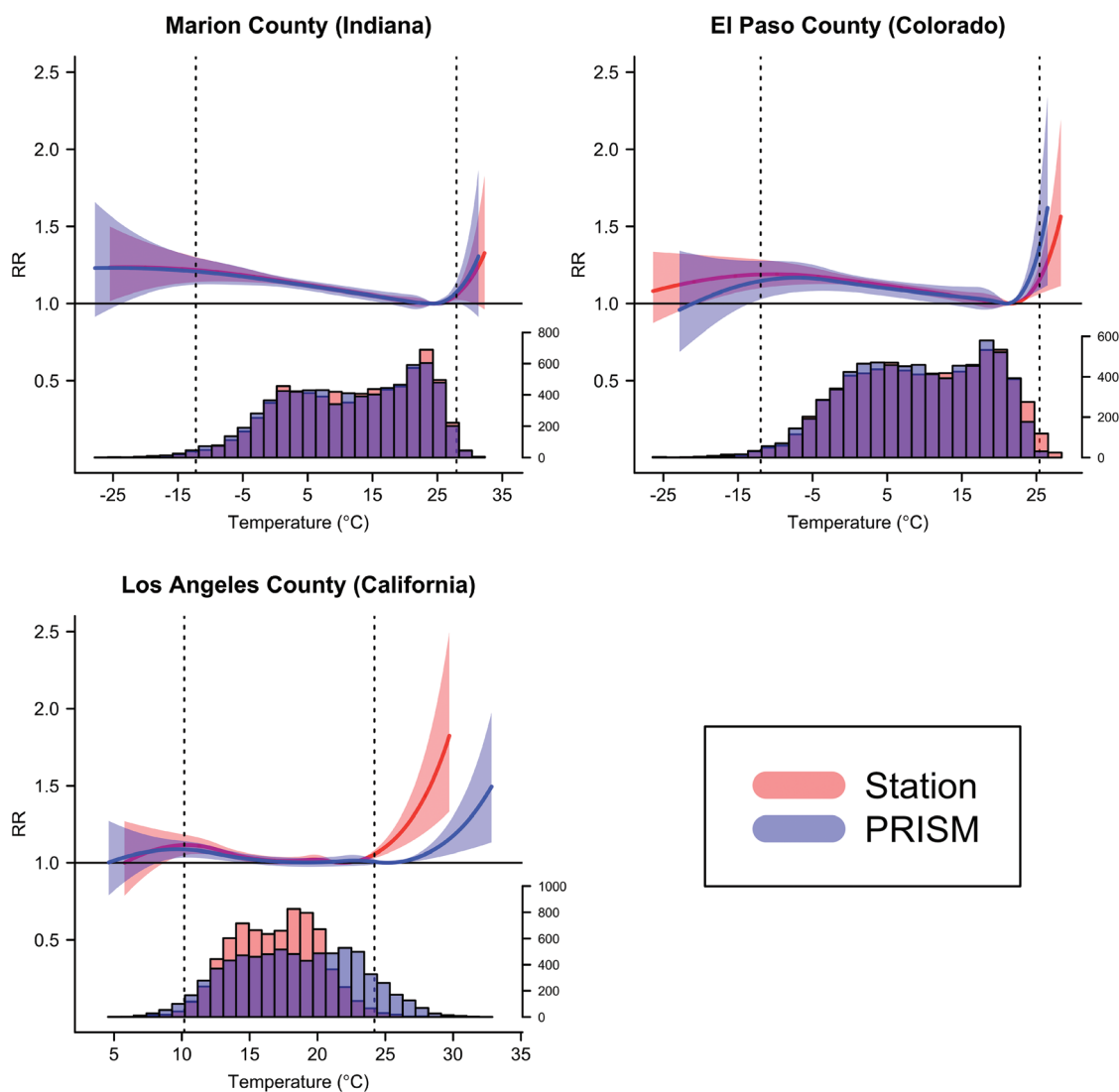
Analyses were carried out in R version 3.3.3<sup>20</sup> using packages "dlnm"<sup>21</sup> and "mvmeta."<sup>22</sup>

### Results

In the year 2000, ~82.2 million people—or about 29% of the US population—lived in the 113 study counties. Between 1987 and 2006, a total of 11,761,285 deaths occurred in these counties. The distribution of daily mean temperature varied substantially across counties, with the median value observed at first-order weather stations ranging from 5.6°C in Saint Louis County, Minnesota to 24.8°C in Palm Beach County, Florida (eTable 1; <http://links.lww.com/EE/A61>). The correlation between daily temperature values observed at weather stations and daily population-weighted temperature estimates constructed from PRISM in each county was generally high (eTable 1; <http://links.lww.com/EE/A61>).

Figure 2 superimposes exposure–response curves and temperature distributions derived from station observations versus PRISM estimates for three example counties: Marion County, Indiana (largest city: Indianapolis), El Paso County, Colorado (largest city: Colorado Springs), and Los Angeles County, California (largest city: Los Angeles). To facilitate comparisons of the magnitude of the temperature–mortality association on very cold and very hot days across temperature metrics, a dashed line is plotted at the 1st and 99th percentiles of the temperature distribution observed at each county's weather station. The RR for the 99th percentile of station temperature versus the minimum mortality temperature as estimated from the station curve was similar to the RR from the PRISM curve in Marion county, smaller than the RR from the PRISM curve in El Paso county, and larger than the RR from the PRISM curve in Los Angeles county (Table 1). Notably, in Los Angeles, the PRISM curve is characterized by a substantially higher minimum mortality temperature compared to the station curve (25.3°C vs. 21.9°C). Consequently, the value of temperature intended to represent extreme heat in Los Angeles (i.e., the 99th percentile of the station temperature distribution, which falls at 24.2°C) is below the minimum mortality temperature on the PRISM curve.

The station and PRISM curves for all 113 counties are shown in eFigure 1; <http://links.lww.com/EE/A61>. In most counties, the shape of the exposure–response curve was similar when estimated with station versus PRISM temperatures, with both very hot and very cold temperatures giving rise to an elevated risk



**Figure 2.** Exposure–response curves showing relative risks (RR) for the 21-day cumulative association between daily mean temperature and all-ages mortality in three counties modeled using temperature observations from weather stations (red) and population-weighted temperature estimates from PRISM (blue), 1987–2006. Vertical dashed lines are placed at the 1st and 99th percentile of the county-specific temperature distribution as observed at the weather station. RR, relative risk.

of death. In the majority of counties (73.4%), we found that the PRISM curve yielded a larger estimate of the RR than the station curve at the 99th percentile of the station temperature distribution (Fig. 3). However, in many counties, the magnitude of the difference between RRs was small. Indeed, Figure 3 demonstrates that two of the counties shown in Figure 2—El Paso and Los Angeles—give rise to some of the most pronounced differences in RRs, while Marion County represents a more typical case where RRs are fairly similar across temperature metrics. Los Angeles was the only county in which the temperatures selected to represent very hot days (i.e., the 99th and 97.5th percentiles of the station temperature distribution) were cooler than the minimum mortality temperature as estimated on the PRISM curve.

Figure 4 shows the difference in the  $\log(\text{RR})$  estimated from each of the two curves at four values of temperature. Whether the association estimated from the PRISM curve is larger or smaller than the association estimated from the station curve depends on what point on the curve is considered. Specifically, when considering hot values of temperature (99th and 97.5th percentiles of the station temperature distribution) associations tended to be larger when estimated from the PRISM curve, but when considering cold values of temperature (1st and 2.5th percentiles of the station temperature distribution) associations tended to

be larger when estimated from the station curve. Results were similar when excluding from analysis any days with a missing value in the station temperature time series (eFigure 2; <http://links.lww.com/EE/A61>). In secondary analyses, we observed that the magnitude of the difference in associations was smaller when comparisons were made at a constant percentile of the temperature distribution within counties rather than a constant value of temperature (eFigures 3, 4; <http://links.lww.com/EE/A61>).

The fraction of deaths across the study counties attributable to heat, cold, and all temperatures was similar when estimated using station versus PRISM temperature (Table 2). For example, we estimate that 0.27% (95% eCI = 0.24, 0.29) and 0.31% (95% eCI = 0.26, 0.34) of deaths in the study counties were attributable to heat when estimated using station and PRISM temperatures, respectively.

## Discussion

In this study of 113 counties representing more than a quarter of the US population at the time, we found that RRs for the association between very hot temperatures and mortality tended to be somewhat larger when estimated using population-weighted temperature estimates constructed from a gridded dataset (PRISM) than when estimated using temperature

**Table 1**  
**RR of death (95% CI) at the 99th percentile of the weather station temperature distribution versus the minimum mortality temperature, shown for both the station curve and the PRISM curve in each of three example counties**

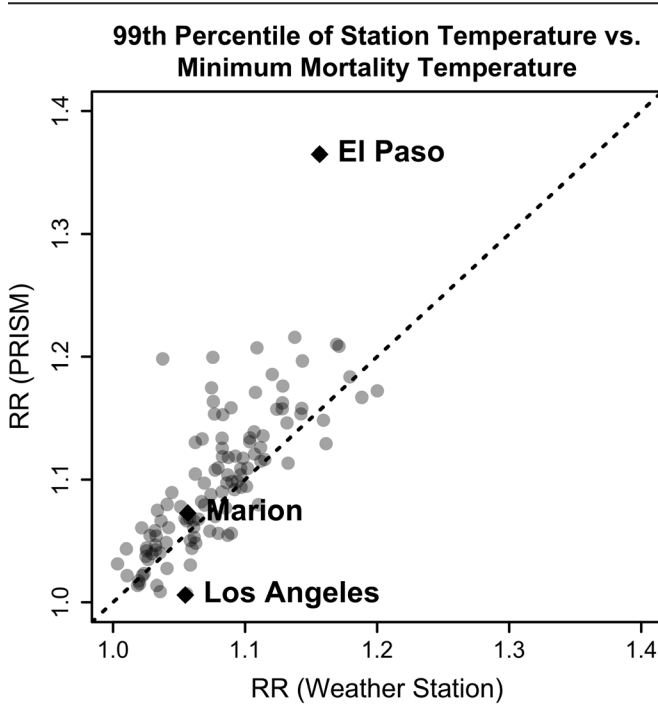
County	Station curve		PRISM curve	
	Contrast <sup>a</sup> (°C)	RR (95% CI)	Contrast <sup>b</sup> (°C)	RR (95% CI)
Marion (Indiana)	27.9 vs. 24.4	1.06 (1.02, 1.09)	27.9 vs. 24.3	1.07 (1.00, 1.15)
El Paso (Colorado)	25.4 vs. 21.7	1.16 (1.06, 1.26)	25.4 vs. 21.2	1.36 (1.09, 1.71)
Los Angeles (California)	24.2 vs. 21.9	1.05 (1.03, 1.08)	24.2 <sup>c</sup> vs. 25.3	1.01 (0.99, 1.02)

RR, relative risk.

<sup>a</sup> 99th percentile of the county-specific weather station temperature distribution versus the county-specific minimum mortality temperature from the station curve.

<sup>b</sup> 99th percentile of the county-specific weather station temperature distribution versus the county-specific minimum mortality temperature from the PRISM curve.

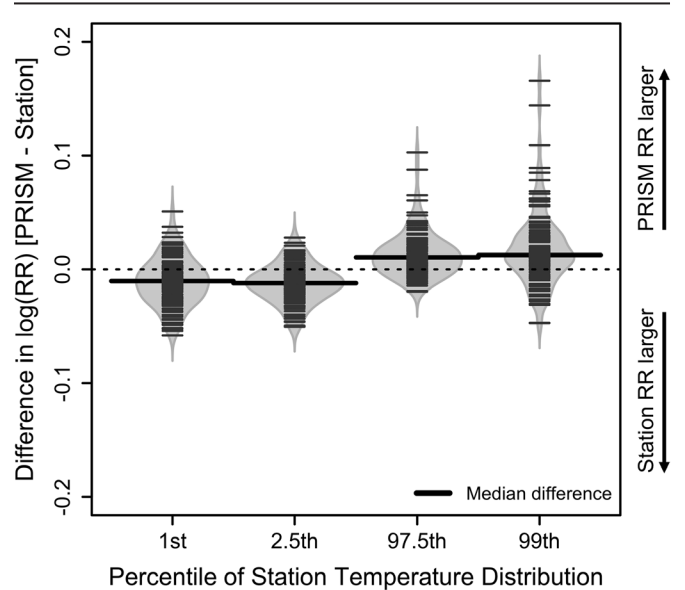
<sup>c</sup> Note that in Los Angeles county, the 99th percentile of the weather station temperature distribution is lower than the minimum mortality temperature from the PRISM curve.



**Figure 3.** Scatterplot of RRs for the 21-day cumulative association between daily mean temperature and mortality modeled using temperature observations from weather stations (x-axis) versus population-weighted temperature estimates from PRISM (y-axis) in each of 113 study counties. For each county, the value of temperature for which associations are plotted is held constant across datasets (i.e., at the 99th percentile of each county's weather station temperature distribution). RR, relative risk.

observations from weather stations. In contrast, RRs for the association between very cold temperatures and mortality tended to be somewhat smaller when estimated from PRISM versus station temperatures in the majority of counties. However, in most counties, the magnitude of the difference between the associations generated from each of the two exposure metrics was small for both very hot and very cold temperatures. Additionally, the shape of the exposure-response curve across the full range of temperatures was generally similar when estimated with station versus PRISM temperatures.

Our results are broadly consistent with previous studies comparing temperature-mortality associations generated with weather station versus spatially refined exposure data in more limited geographic areas. In a study of three states in the eastern United States, Lee et al<sup>14</sup> found a 2.05% increase (95% CI = 0.52, 2.91) in mortality for every 1°C increase in temperature above 28°C when assigning temperature exposures at the zip code level using a spatiotemporal model. This estimate was reduced to a 1.14% increase (95% CI = 0.08, 1.57) when temperature exposures were assigned based on the weather station



**Figure 4.** Distribution of the difference between the log(RR) for the 21-day cumulative association between temperature and mortality as estimated using PRISM versus station temperature in each of 113 counties. In each county, differences are calculated for the log(RR) at four values of temperature (i.e., the 1st, 2.5th, 97.5th, and 99th percentile of that county's weather station temperature distribution) versus the minimum mortality temperature. Differences are calculated such that counties where the PRISM curve yields a larger estimate of the temperature-mortality association receive a positive value.

**Table 2**  
**Fraction of deaths (95% eCI) attributable to temperature across all 113 counties generated using (1) the exposure-response curves generated from weather station temperature observations and (2) the exposure-response curves generated from population-weighted PRISM temperature estimates**

	Attributable Fraction (95% eCI) (%)	
	Station <sup>d</sup>	PRISM <sup>e</sup>
Heat <sup>a</sup>	0.27 (0.24, 0.29)	0.31 (0.26, 0.34)
Cold <sup>b</sup>	5.76 (5.44, 6.06)	5.51 (5.09, 5.88)
Total temperature <sup>c</sup>	6.03 (5.70, 6.34)	5.82 (5.38, 6.20)

<sup>a</sup> Temperatures above the county-specific minimum mortality temperature.

<sup>b</sup> Temperatures below the county-specific minimum mortality temperature.

<sup>c</sup> All temperatures.

<sup>d</sup> Calculated using the exposure-response curve generated from weather station temperature observations in each county.

<sup>e</sup> Calculated using the exposure-response curve generated from population-weighted PRISM temperature estimates in each county.

closest to the centroid of each zip code. Similarly, we found that population-weighted temperature estimates tended to yield slightly larger RRs for the association between extreme heat and mortality compared with station temperatures. However, the difference between estimates was small in the majority of locations, suggesting that on average across locations in the United States, weather station data likely provides a reasonable estimate of population average exposure.

Other studies have found little difference in effect estimates when employing exposure data sources with varying degrees of spatial information. For example, Guo et al<sup>12</sup> estimated the association between temperature and mortality in the large city of Brisbane, Australia using three exposure definitions: a single time series consisting of observations from a central weather station, a single time series consisting of daily values averaged across three weather stations, and spatially interpolated (via kriging) temperature estimates for each of 163 administrative areas within Brisbane. Models fit using each of these three exposures yielded similar effect sizes for both hot and cold temperatures. In the Paris area, Schaeffer et al<sup>13</sup> found little difference in temperature–mortality associations across a variety of exposure definitions, including values from a single monitoring station as well as population-weighted daily estimates derived from multiple monitoring stations. Again, this is consistent with our results.

While in most counties we observed similar RRs using the station versus PRISM curves, in a small number of locations the degree to which the PRISM curve estimate was larger on very hot days was more substantial. In Colorado's El Paso County, the RR comparing a 25.4°C day to the minimum mortality temperature was 1.16 (95% CI = 1.06, 1.26) for the station curve but 1.36 (95% CI = 1.09, 1.71) when based instead on the PRISM curve. Other counties with relatively large differences in the RR at the 99th percentile of the station temperature distribution include Maricopa County in Arizona (largest city: Phoenix), Union County in New Jersey (largest city: Elizabeth), and Madison County in Alabama (largest city: Huntsville). In these or other locations with more pronounced differences in RRs, our results may have implications for heat early warning and prevention activities that are triggered based on forecast temperatures for that location. For example, in some US communities, the temperatures at which heat advisories and excessive heat warnings are issued are based on epidemiologic studies that use exposure data from local weather stations.<sup>23–26</sup> In communities where risks are smaller when estimated using station temperatures, alert thresholds set based on studies using weather station measurements may be too low.

Studies of the association between temperature and morbidity and/or mortality often report absolute measures of health impacts, such as the attributable fraction or attributable number.<sup>2,27–29</sup> For example, in a study of 135 US communities in which community-specific associations were each estimated using a single monitoring station, Gasparrini et al estimated that 0.25% (95% eCI = 0.30, 0.39) and 5.51% (95% eCI = 5.17, 5.82) of deaths were attributable to hot and cold temperatures, respectively.<sup>2</sup> Although our study did not include an identical set of locations, we estimated similar attributable fractions using station temperature across our 113 study counties (heat: 0.27% [95% eCI = 0.25, 0.29], cold: 5.76 [95% eCI = 5.44, 6.06]). Results were also quantitatively similar when using PRISM temperature.

Strengths of this study include the use of established statistical methods commonly used in the temperature–mortality literature in order to provide evidence that is relevant to current practice, the comparison of RRs across temperature metrics for both hot and cold temperatures, and the inclusion of a large number of counties with diverse climates that encompass a large proportion of the US population. Specifically, we included in our study every populous county in the contiguous United States for which we had continuous mortality data and which

included a first-order weather station within its borders. It is important to note that this selection strategy excludes less populous counties, as well as those located farther from a first-order weather station. Thus, our results may not be generalizable to other counties.

Our study also has several other limitations. First, while we hypothesized that the population-weighted temperature estimates constructed from PRISM are a better proxy for population average exposure than weather station observations, we are not able to verify this assumption. However, in a previous validation study, we found that PRISM grid cell values exhibited a high degree of agreement with colocated weather station observations,<sup>17</sup> suggesting that the PRISM model itself performs well. Second, we are unable to empirically determine whether station observations and PRISM estimates have similar or different amounts of random measurement error, classical or Berkson. Third, information on<sup>24</sup> daily deaths was not available at a spatial scale smaller than a county. As a result, we were unable to assess potential differences in risk when allowing daily temperature exposures to vary spatially within a county as some previous authors have done.<sup>12,14</sup> Fourth, our results may not be applicable to other metrics of heat (e.g., heat index), other health outcomes (i.e., measures of morbidity), or more recent years. Finally, we did not address the related question of whether different measures of heat or cold stress (e.g., minimum temperature, maximum temperature, heat index) differ in their ability to predict mortality.<sup>30</sup> Instead, we assessed whether the source of the exposure estimates for a single measure (i.e., mean temperature) influenced the magnitude of our results.

In summary, we found that the overall association between temperature and mortality across a large number of US counties was similar when estimated based on two different exposure metrics. However, in a limited number of locations, population-weighted temperature estimates yielded substantially larger associations for extreme heat versus observations from weather stations. These results suggest that the use of observed temperature from first-order weather stations in prior studies of temperature health effects across many US counties has not led to systematic under or overestimation of the adverse health impacts of temperatures. However, studies seeking to establish the health effects of specific temperature thresholds in specific locations may benefit from leveraging available spatially refined meteorological data. Moreover, such data allow the estimation of the health effects of temperature even in areas located far from weather stations.

## Conflicts of interest

The authors declare that they have no conflicts of interest. This work was supported by grants F32-ES027742, R01-ES029950, and P30-ES000002 from the National Institute of Environmental Health Sciences (NIEHS), NIH, and by the Institute at Brown for Environment and Society (IBES). The contents of this report are solely the responsibility of the authors and do not necessarily represent the official views of the sponsoring organizations.

## References

1. Guo Y, Gasparrini A, Armstrong B, et al. Global variation in the effects of ambient temperature on mortality: a systematic evaluation. *Epidemiology*. 2014;25:781–789.
2. Gasparrini A, Guo Y, Hashizume M, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet*. 2015;386:369–375.
3. Medina-Ramón M, Schwartz J. Temperature, temperature extremes, and mortality: a study of acclimatization and effect modification in 50 US cities. *Occup Environ Med*. 2007;64:827–833.
4. Zeger SL, Thomas D, Dominici F, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect*. 2000;108:419–426.

5. Schwartz J, Dockery DW, Neas LM. Is daily mortality associated specifically with fine particles? *J Air Waste Manag Assoc.* 1996;46:927–939.
6. Shi L, Liu P, Kloog I, Lee M, Kosheleva A, Schwartz J. Estimating daily air temperature across the Southeastern United States using high-resolution satellite data: a statistical modeling study. *Environ Res.* 2016;146:51–58.
7. PRISM Climate Group OSU. *Descriptions of PRISM Spatial Climate Datasets for the Conterminous United States.* Available at: [http://prism.oregonstate.edu/documents/PRISM\\_datasets.pdf](http://prism.oregonstate.edu/documents/PRISM_datasets.pdf). Accessed 16 January 2018.
8. Thornton PE, Thornton MM, Mayer BW, et al. *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3.* Available at: [https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds\\_id=1328](https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1328). Accessed 10 July 2017.
9. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med.* 1998;55:651–656.
10. Zhang K, Oswald EM, Brown DG, et al. Geostatistical exploration of spatial variation of summertime temperatures in the Detroit metropolitan region. *Environ Res.* 2011;111:1046–1053.
11. Harlan SL, Brazel AJ, Prashad L, Stefanov WL, Larsen L. Neighborhood microclimates and vulnerability to heat stress. *Soc Sci Med.* 2006;63:2847–2863.
12. Guo Y, Barnett AG, Tong S. Spatiotemporal model or time series model for assessing city-wide temperature effects on mortality? *Environ Res.* 2013;120:55–62.
13. Schaeffer L, de Crouy-Chanel P, Wagner V, Desplat J, Pascal M. How to estimate exposure when studying the temperature-mortality relationship? A case study of the Paris area. *Int J Biometeorol.* 2016;60:73–83.
14. Lee M, Shi L, Zanobetti A, Schwartz JD. Study on the association between ambient temperature and mortality using spatially resolved exposure data. *Environ Res.* 2016;151:610–617.
15. Daly C, Halbleib M, Smith JI, et al. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *Int J Climatol.* 2008;28:2031–2064.
16. Smith A, Lott N, Vose R. The integrated surface database: recent developments and partnerships. *B Am Meteorol Soc.* 2011;92:704–708.
17. Spangler KR, Weinberger KR, Wellenius GA. Suitability of gridded climate datasets for use in environmental epidemiology. *J Expo Sci Environ Epidemiol.* 2018.
18. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med.* 2010;29:2224–2234.
19. Gasparrini A, Leone M. Attributable risk from distributed lag models. *BMC Med Res Methodol.* 2014;14:55.
20. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2017.
21. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlmm. *J Stat Softw.* 2011;43:1–20.
22. Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat Med.* 2012;31:3821–3839.
23. Wellenius GA, Eliot MN, Bush KF, et al. Heat-related morbidity and mortality in New England: evidence for local policy. *Environ Res.* 2017;156:845–853.
24. US National Weather Service. *National Weather Service is Lowering Heat Advisory Thresholds for Northern New England and Much of New York.* Available at: <https://www.weather.gov/media/car/DSS/NEWHeatThresholds.pdf>. Accessed 31 March 2019.
25. Metzger KB, Ito K, Matte TD. Summer heat and mortality in New York city: how hot is too hot? *Environ Health Perspect.* 2010;118:80–86.
26. Kingsley SL, Eliot MN, Gold J, Vanderslice RR, Wellenius GA. Current and projected heat-related morbidity and mortality in Rhode Island. *Environ Health Perspect.* 2016;124:460–467.
27. Weinberger KR, Kirwa K, Eliot MN, Gold J, Suh HH, Wellenius GA. Projected changes in temperature-related morbidity and mortality in Southern New England. *Epidemiology.* 2018;29:473–481.
28. Fu SH, Gasparrini A, Rodriguez PS, Jha P. Mortality attributable to hot and cold ambient temperatures in India: a nationally representative case-crossover study. *PLoS Med.* 2018;15:e1002619.
29. Tian L, Qiu H, Sun S, Lin H. Emergency cardiovascular hospitalization risk attributable to cold temperatures in Hong Kong. *Circ Cardiovasc Qual Outcomes.* 2016;9:135–142.
30. Barnett AG, Tong S, Clements AC. What measure of temperature is the best predictor of mortality? *Environ Res.* 2010;110:604–611.