Severe and sustained episodes of summer heat are associated with increased morbidity and mortality, particularly in temperate regions. The death toll in an unprepared region can be substantial. For example, during 1–14 July 1993, the eastern United States experienced a severe heat wave with high temperatures (33.9°C–38.3°C) and high humidity (36%–58%) (Centers for Disease Control and Prevention 1994). During 6–14 July, the Philadelphia, Pennsylvania, Medical Examiner’s Office determined that 118 deaths were heat related (either a core body temperature of 40.6°C or higher, or a body found in a hot, unventilated environment). This is certainly an underestimate because heat is associated with increased mortality from a number of causes other than heat stroke (Shen et al. 1998). Historically, cardiovascular diseases have accounted for 13%–90% of the increase in overall mortality during and following a heat wave, while cerebrovascular disease accounted for 6%–52%, and respiratory diseases for 0%–14% (Kilbourne 1997). Heat waves also increase the rate of nonfatal illnesses.

Partly in response to heat waves in 1993 and 1994, the Philadelphia Hot Weather–Health Watch/Warning System (PWWS) was developed in 1995 to alert the city’s population when weather conditions pose risks to health (Kalkstein et al. 1996; Mirchandani et al. 1996; Sheridan and Kalkstein 1998). At that time, the local National Weather Service (NWS) issued heat advisories that relied heavily on the heat index, which identifies oppressive conditions based on a combination of temperature and relative humidity. Limitations of the heat index include the assumption that people respond to a combination of only two meteorological variables. This ignores research demonstrating that people respond to the effect of all weather variables acting simultaneously, and it does not take into account the negative impact of several consecutive days of oppressive weather.

The PWWS was designed and continues to run as a stand-alone early warning system that forecasts...
airmass type for the current day and the coming 2 days during the summer season. The PWWS uses six readily available weather elements to identify major air masses in Philadelphia (Kalkstein et al. 1996). Groups (synoptic categories) are identified with relatively homogeneous characteristics, with each day categorized into one of these groups. Analysis then determined which air masses are associated with excess mortality during the summer season. The appendix includes a description of the PWWS, including the measures taken during a heat wave to reduce morbidity and mortality. This system is the basis for more than 20 other heat–health watch warning systems being instituted in cities worldwide (Kalkstein 2003; Sheridan and Kalkstein 1998).

Because the local NWS has a government-mandated responsibility to issue heat-related warnings, it decided to use the PWWS as one input into their decision-making process to advise the Philadelphia Department of Health of potentially dangerous weather. The Philadelphia Department of Health then implements emergency precautions and mitigation procedures to reduce the mortality risk (Kalkstein et al. 1996). Data are available from 1995 to 1998 on which days heat-related warnings were issued. This article is the first attempt to calculate the number of lives saved and the economic benefit of warnings in reducing heat-related mortality.

**METHODS.** Kalkstein et al. (1996) identified two types of hot-weather air masses—maritime tropical and dry tropical—that have been historically associated with elevated mortality in Philadelphia. We refer to these weather patterns as “heat waves.” Our analyses included daily data for all such heat waves during 1995–98, plus up to three additional days after the end of each heat wave. The 3 days following a heat wave were included because mortality effects can lag heat waves by up to 3 days (Curriero et al. 2002; Laschewski and Jendritzky 2002). Including days following heat waves brought the total number of days in the sample to 210.

For the days in our sample, mortality data were obtained from the National Center for Health Statistics for the 65-yr and older age group in the Philadelphia metropolitan statistical area (MSA) for 1995–98 (National Center for Health Statistics 2000). We chose to limit our analyses to this age group because they are vulnerable to excessive heat; therefore, the statistical evidence regarding the effectiveness of the warning system should be strongest for this age group (Semenza et al. 1996; Smoyer 1998; Centers for Disease Control and Prevention 1995). The analyses were based on excess mortality, where excess mortality was calculated as the difference between reported mortality and the underlying mortality trend estimated from years prior to 1995, as described in the appendix (Kalkstein et al. 1996). We believe that variations around such a mortality trend line are a better indicator of the effects of daily weather, and of efforts to counter those effects, than are medical examiners’ determinations that deaths were caused by extreme temperatures.

A heat wave warning was issued only if the local NWS office concurred with a PWWS recommendation that a warning should be issued. Before the development of the PWWS, the NWS issued most heat-health alerts based on guidelines that relied heavily on the computation of the heat index (NOAA 1994). An excessive heat warning was issued when the daytime heat index values were expected to exceed 40.5°C for more than 3 h a day on two consecutive days, or when the daytime heat index was expected to exceed 46°C for any length of time (Kalkstein et al. 1996).

During the time period of the study, the local NWS office agreed to use the PWWS for guidance in the issuance of heat advisories and excessive heat warnings. However, forecasters did not completely rely on the system and were often conservative in issuing advisories and warnings. The criteria used by a forecaster to decide whether (or not) to concur with a PWWS recommendation have not been archived. In general, forecasters usually called a warning if the heat index was greater than 40.5°C and rarely called a warning if the heat index was below 38°C (even if the PWWS called for such an issuance) (L. S. Kalkstein 2003, personal communication). The PWWS was primarily used for guidance when the heat index was between these values. Some forecasters felt that the system called too many advisories and warnings and were concerned that the public might become less responsive to subsequent warnings. An early evaluation of the system indicated that, because of this policy, warnings were not called frequently enough and that heat-related deaths occurred on days when the system called for a warning and the forecasters did not issue one (Kalkstein et al. 1996). More recently, forecasters are using the PWWS as the primary guidance, although they are advised to use discretion and not call advisories and warnings based exclusively on the system.

The NWS concurred with a PWWS recommendation that a warning be issued less than one-quarter of the time; during the period of 1995–98, there were 70 days when the PWWS suggested a heat warning,
RESULTS. Estimation of lives saved. Among the explanatory variables considered for inclusion in the regression model, only the time of season and warning indicator were convincingly associated with excess mortality. The omission of the daily weather variables might seem surprising, but is less so when one recognizes that all the days in our statistical sample were identified as potentially life threatening, based on weather variables. Thus, within this subset of days, weather differences might be relatively small in terms of their effects on mortality. Other variables that did not significantly contribute to the explanation of excess mortality were the duration of the heat wave and airmass type (maritime tropical or dry tropical). The latter means that there were no significant differences between the air masses.

Excluding variables not significantly associated with mortality left the following regression: Excess mortality = 3.27 – 0.049 × time of season – 2.58 × warning indicator. The t-statistic for the time-of-season variable was –2.55 and the p-value was 0.06. The t-statistic for the warning indicator was –1.43 and the p-value was 0.08. The R-squared value for the equation was 0.04.

The low R-squared value suggests that omitted disease processes and accidents are the primary determinants of mortality, as expected. The coefficient for the time-of-season variable means that excess mortality (assuming no warning) could range from +3.22 to –3.54 lives, depending on whether a heat wave occurred on the first or last day of the season.

The estimated coefficient of the warning indicator variable means that when a warning was issued, assuming no mortality displacement, 2.6 lives were saved, on average, for each day that a warning was issued, as well as for the 3 days immediately following the warning. Mortality displacement occurs when deaths are brought forward in time; that is, some individuals who died during a heat wave would have died anyway within a short period of time after the heat wave. Because there were 45 days during 1995–98 when warnings were issued (or the effect of warnings persisted over the 3-day lag after the warning ended), the estimated total number of lives saved was 117.

With a p value of 0.08 for the estimated warning indicator coefficient, we cannot (at the 5% level) reject the hypothesis that warnings do not save lives. The 95% confidence interval on the number of lives saved includes negative numbers, even though such a possibility is not credible. The 95% confidence interval on the estimated coefficient of the warning indicator ranges from about –1.0 to +6.1, which implies a confidence interval on the number of lives saved from about –45 to +275.

It is important here to understand that the key question is “did this warning system save any lives?” Because the p value of the warning indicator is about 0.08, there is a 92% chance that at least one life was saved by the system. Given the low costs of this warning system (considered below), one would surely want to continue it, given this statistical evidence. In fact, given the costs of the system, and the presumption that the system is at worst ineffective, one would want to continue using the system unless the evidence overwhelmingly indicated that the system does not save lives. Stated differently, this is a context where it does not make sense to withhold judgment until one is 95% confident that the system actually works; rather, it is a context where it makes sense to assume the system works, unless the evidence is very strong that it does not.
Because the serial correlation of errors is a potential statistical problem in time series analysis, we also estimated the association using heat waves as the sampling unit. For this purpose, excess mortality was defined as the average daily excess mortality for the days of the heat wave (and 3 days after), and the warning indicator was set to 1 whenever there was a warning issued on at least one of the heat wave days. The results obtained this way were very similar to those obtained with days as the sampling unit; in particular, the coefficient of the warning indicator was the same, that is, a saving of 2.6 lives per day.

**Estimation of the benefits of the PWWS.** For some public policy decisions, lives are assigned monetary value. The Environmental Protection Agency (EPA), for example, places a value on lives [value of a statistical life (VSL)] to assess the benefits of policies that reduce pollution (Smith et al. 2001). VSLs are reported in units of dollars per life saved. The key piece of data underlying the VSL is the value of mortality risk reduction. It has been estimated by analyzing wage rates in occupations with differing mortality risks, to identify the wage differential due to mortality risk, and by contingent valuation, that is, by asking a sample of people to assess their own willingness to pay to reduce risks. The results of both methods are usually similar.

Two recent studies addressed the question of whether the value of a statistical life is lower for older people or people with health problems (Krupnick et al. 2000; Smith et al. 2001). This is of particular interest in connection with the PWWS, because the population most at risk from heat is older and frequently suffers from health problems. Smith et al. (2001) estimated VSLs for older Americans who were still working, and reported a range of $5.3–$6.6 million. While the people in the Smith et al. study were younger (generally 50–65 yr old) and probably healthier than the people most at risk from heat waves, they found no indication that the VSL fell with age over this range, nor that the VSL was lower for those with serious illness. Krupnick et al. (2000) estimated the VSL for Canadians 40–75 yr old, and reported a range of $0.82–$2.6 million (U.S. dollars). Krupnick et al. also found that poor health did not lower the VSL. However, they concluded that Canadians 70–75 yr old were willing to pay about one-third less than younger adults for mortality risk reductions.

We decided to use the EPA estimate of $6.12 million for the value of a statistical life (Smith et al. 2001) and accepted the Krupnick et al. (2000) finding that VSLs eventually fall off with age, being one-third lower for the 70–75 age group. Thus, we assumed $4 million for the VSL among people 65 yr of age or older in Philadelphia. This implies that the gross benefits of the Philadelphia heat wave warning system could be on the order of $468 million (117 lives saved times $4 million) over the 3-yr period.

**Estimation of the costs of the PWWS.** As discussed in detail in the appendix, when a heat wave warning is declared in Philadelphia, the city takes a number of actions to reduce the risk of heat-related mortality. Most of these actions do not have direct monetary costs, including actions taken by city employees as a normal part of their jobs, actions taken by volunteers, and delayed actions (i.e., halting service suspensions). A few of the actions taken do have direct costs, including the Heatline and additional Emergency Medical Service (EMS) crews. Costs for these actions are primarily additional wages. We estimated these wage costs to be $1,000 per day on weekdays and $3,000 per day on weekends for the Heatline, and $4,000 per day for EMS crews. We do not have information on other direct costs, but believe they would not raise the total beyond about $10,000 per day when a heat wave warning is issued. During 1995–98, there were 21 such days. This implies that the total cost of the Philadelphia system was on the order of $210,000 over the 3-yr period.

The Philadelphia system depends on weather forecasts generated by the NWS. One might want to allocate some portion of the costs of those forecasts, and count it as part of the costs of the system. Obviously, making such an allocation would be very difficult and we did not attempt to do so. Instead, the net benefit estimate reported below may be interpreted as one of many societal benefits derived from weather forecasting systems.

**Estimation of the net benefits of the PWWS.** Obviously, if the Philadelphia system saves any lives at all, it will have large estimated dollar benefits. This is because the VSL, for even one life, is bigger than the cost of running the system. In fact, system costs are so far into the “noise” of the estimate that they are essentially irrelevant. Thus, based on this analysis, we would conclude that the net benefits of the issued heat wave warnings were around $468 million over the 3-yr period of 1995–98.

**Conclusions.** The results suggest that issuing an individual warning lowered daily mortality by about 2.6 lives on average, based on a number of assumptions. One is that there was no mortality displacement. To the extent that there is mortality dis-
placement, our estimate of the number of lives saved is inflated. It was not possible to estimate mortality displacement with the data available to us.

The literature contains a wide range of estimates of mortality displacement following heat waves, from very little to explaining much of the net excess mortality (Huynen et al. 2001; Kalkstein 1998; Kunst et al. 1993; Laschewski and Jendritzky 2002; Rooney et al. 1998). Based on this literature, one might reduce the estimated number of lives saved by 20%–50% to take mortality displacement into account. However, we did choose to use a $4 million value of life in our benefits estimate (rather than the $6 million value used by EPA), precisely because the population affected here is older and less healthy. One could argue, therefore, that we have already allowed mortality displacement in our value-of-life assumption. In any case, the net benefits of the system would remain strongly positive even if we also lowered the estimate of lives saved to account for possible mortality displacement.

While mortality displacement arguably means that we have estimated too high a value for this system, there are other considerations that suggest that our estimate could be too low. First, we only considered mortality reductions in the 65-yr and older population. If there were mortality reductions in younger populations too, this would add to the estimated benefits. Second, we did not estimate any increased morbidity associated with excess heat exposure. It is reasonable to assume that morbidity also was reduced by the PWWS. Hospitalizations were shown to increase in London, United Kingdom, and Chicago, Illinois, during the 1995 heat waves (Rooney et al. 1998; Semenza et al. 1999). Morbidity reduction is an additional benefit not included in our analyses.

Our results suggest that more lives might have been saved if more warnings had been declared. However, there is a question about whether the daily number of lives saved might have been lower on days when warnings might have been declared, but were not. To address this question, we looked at temperatures on different subsets of days. On the 70 days when the PWWS suggested a warning, the average temperature was 32.68°C. On 16 of these days, the NWS concurred and a warning was issued; on these days plus the additional 5 days on which the NWS issued a warning independent of the PWWS, the average temperature was 32.77°C. On the remaining 54 days when the PWWS suggested a warning and the NWS did not concur, the average temperature was 32.34°C.

While these data suggest that the NWS was marginally more conservative about declaring warnings, the temperature differences were very small. This supports the view that the number of lives saved per day might have been about the same if warnings had been declared on additional days within maritime tropical and dry tropical air masses. Moreover, the low estimated cost of warnings argues for being more liberal about issuing warnings. In support of this, recent analyses in London and Toronto, Ontario, Canada, suggest that excess mortality begins at relatively low temperatures (Hajat et al. 2002; Semenza et al. 1999; Smoyer-Tomic and Rainham 2001). However, there remains some risk that if too many warnings are declared they may become less effective, especially with regard to the voluntary actions taken when a warning is issued.

The $R^2$ squared for the linear regression is low, which implies that most of the variation in excess mortality is attributable to variables not included in our model. It is reasonable to expect that disease processes and accidents are the dominant causes of mortality in the 65-yr and older age group; we do not have variables representing these phenomena in our model. However, as long as these omitted variables are uncorrelated with our included variables (most importantly, uncorrelated with the warning indicator), the estimated coefficients of our included variables will be unbiased. Particularly in the case of the warning indicator, it is unlikely that this variable is statistically correlated with whatever myriad of other variables is driving the unexplained variation in daily excess mortality.

Evaluations are underway for similar heat watch warning systems recently established in other cities, including Toronto, Dayton, Ohio, and Phoenix, Arizona, although there currently are no direct estimates of the number of lives saved. However, an observation based on the first year of operation of the Philadelphia system during the hot summer of 1995 supports our results. It was noted that, as the summer progressed, the mortality algorithm within the system overestimated heat-related mortality with greater regularity and by a greater amount. This supports the suggestion that people respond to the heat advisories and warnings in ways that save lives (Kalkstein et al. 1996). Further, in a detailed evaluation of heat watch–warning system operation in a number of cities, it was clear that cities utilizing systems like the one in Philadelphia were developing and instituting more sophisticated heat intervention plans whenever warnings were called (Kalkstein 2003). These improved plans may be a direct result of system development.

The methods described here can be used to evaluate the costs and benefits of implementing heat wave early warning systems. Our results suggest that the
benefits in terms of lives saved will far outweigh the operational costs of such systems, at least for cities located in temperate regions.

ACKNOWLEDGMENTS. This work was funded by the Electric Power Research Institute under Contract EP-P5431/C2688 and by NOAA under Contract 40AANA1A1110. The authors would like to thank Dr. Andrew Solow for his useful suggestions on the statistical analysis, and the helpful comments of the reviewers.

APPENDIX: PHILADELPHIA HOT WEATHER–HEALTH WATCH/WARNING SYSTEM. L. S. Kalkstein designed the Philadelphia Hot Weather–Health Watch/Warning System (PWWS) to predict periods when there is high risk of heat-related mortality (Kalkstein et al. 1996). Past weather was classified into more or less homogenous categories (synoptic categories), referred to as airmass types, based on air temperature, dewpoint temperature, cloud cover, sea level pressure, wind speed, and wind direction. Associations were determined between airmass types and mortality. Cause of death, place and date of death, age, and race were extracted from National Center for Health Statistics files for the Philadelphia metropolitan statistical area for the years 1964–66, 1973–76, 1978, and 1980–88; these were the years for which the date of death was available (Kalkstein et al. 1996). The analysis used the total number of deaths per day. Direct standardization was used to adjust all mortality data for changes in the total population over time. A mortality trend line was constructed for each time period, and daily mortality was expressed as deviations around this baseline. The mean daily mortality for each synoptic airmass category was determined, along with the standard deviation, to ascertain whether particular categories were associated with high or low mortality. Potential lag times of 1–3 days were accounted for in the analysis; the highest mortality was found with zero lag time.

Two airmass types were associated with elevated mortality in Philadelphia—maritime tropical and dry tropical. Because not all days within these air masses resulted in elevated mortality, a stepwise multiple regression analysis was used to identify which days within the air mass were associated with increased mortality. The regression identified several variables that were predictive of elevated mortality; these were as follows: the number of consecutive days that the air mass was present, the maximum temperature, and the time of season (e.g., whether the oppressive air mass occurs early or late within the season) (Kalkstein et al. 1996).

The accuracy of the forecast data and the performance of the categorizations were verified by backcasting, with archived model forecasts used as predictors. For the summer of 1988, the 24-h forecast identified 89% of the oppressive days and the 48-h forecast identified 71%. The system was designed to operate from 15 May through 30 September (139 days). Each day the PWWS forecasts the airmass category for the current day and the subsequent two (Kalkstein et al. 1996). The beginning and end of a heat wave are determined by airmass type and by the mortality predictive equations associated with each air mass. The system was designed to generate a health watch, a health alert, or a health warning, based on airmass type. The criteria for a health warning are that a maritime tropical or dry tropical airmass type is forecast for either that afternoon or the following morning, and the predictive model associated with that particular air mass forecasts four or more heat-related deaths. The local NWS office then determines whether (or not) to issue a warning based on the PWWS forecasts, the heat index, and other information. For example, in 1995 the NWS issued a heat warning on 9 of the 15 days recommended by the PWWS.

The city of Philadelphia and other agencies and organizations institute a series of intervention activities when the NWS issues a warning. Television and radio stations and newspapers are asked to publicize the oppressive weather conditions, along with information on how to avoid heat-related illnesses. In addition, these media announcements encourage friends, relatives, neighbors, and other volunteers (“buddies”) to make daily visits to elderly persons during the hot weather. These buddies are asked to ensure that the most susceptible individuals have sufficient fluids, proper ventilation, and other amenities to cope with the weather. A “Heatline” is operated in conjunction with the Philadelphia Corporation for the Aging to provide information and counseling to the general public on avoidance of heat stress. The Heatline telephone number is publicized by the media and by a large display seen over much of the center of Philadelphia. When a warning is issued, the Department of Public Health contacts nursing homes and other facilities boarding persons requiring extra care to inform them of the high-risk heat situation and to offer advice on the protection of residents. The local utility company and water department halt service suspensions during warning periods. The Fire Department Emergency Medical Service increases staffing during warnings in anticipation of increased service demand. The agency for homeless services activates increased daytime out-
reach activities to assist those on the streets. Senior centers extend their hours of operation of air-conditioned facilities during warning periods.

REFERENCES


