Research Article

Modeling and Syndromic Surveillance for Estimating Weather-Induced Heat-Related Illness

Alexander G. Perry,1, 2 Michael J. Korenberg,2 Geoffrey G. Hall,1, 3 and Kieran M. Moore1, 4

1 Public Health Informatics Group, Kingston, Frontenac, Lennox & Addington Public Health, 221 Portsmouth Avenue, Kingston, ON, Canada K7M 1V5
2 Department of Electrical and Computer Engineering, Queen’s University, Kingston, ON, Canada K7L 3N6
3 Department of Civil Engineering, Queen’s University, Kingston, ON, Canada K7L 3N6
4 Department of Emergency Medicine, Queen’s University, Kingston, ON, Canada K7L 3N6

Correspondence should be addressed to Alexander G. Perry, alexander.perry@queensu.ca

Received 22 November 2010; Accepted 24 February 2011

Abstracted patient charts can provide accurate post-diagnosis identification of heat-related hospital emergency department visits. However, the delays in obtaining data from this source can be over a year due to the time required for abstraction of patient records, reporting, and dissemination. Sometimes only limited historical data are available for analysis. Relying on the results of retrospective analyses to inform heat response plans assumes a population’s response to heat does not change over time, neglecting the possibility of long-term acclimation [6]. Responses that occurred historically may not be representative of future responses to heat events if these future events have characteristics that differ from those in the past, such as heat of higher intensities or longer durations [7]. Furthermore, there is evidence that the effects of heat on population health vary geographically [8], raising concerns about the generalizability of study results from other locations.

To overcome the limitations of an approach using only retrospective analysis to inform public health heat responses, real-time population-level heat-related illness (HRI) surveillance could provide situational awareness, enabling better
decisions to be made. However, there is a cost associated with setting up and maintaining such systems, although the economic benefit of heat watch and warning systems has been argued [9]. Heat-related illness surveillance using emergency department (ED) and emergency services dispatch (911) has been suggested [10, 11]. Though not necessarily a comprehensive measure of the complete impact of heat on a population since it fails to capture all care sought for heat illness, heat-related emergency department visits potentially provide a timely indicator of its acute and serious effects. Heat-related mortality has been shown to be correlated with emergency department visits [12]. Since mortality can lag acute heat effects [13] and information on cause of death can take time to be established and disseminated, monitoring emergency department visits would be preferable to monitoring mortality for surveillance purposes. Data from hospitals can often be obtained electronically allowing them to be monitored in real-time using syndromic surveillance [11, 14]. There are unfortunately two potential problems with using emergency department syndromic surveillance. The first is misclassification of visits, a result of nonspecific and nonsensitive syndrome definitions and the use of prediagnosis information. This could lead to errors in visit counts. The second is that such systems do not provide the advance warning required for taking preventative actions.

Because of the complementary advantages and disadvantages of surveillance and population-level heat-response models constructed from retrospective data on heat-related emergency department visits, public health may consider using both of these approaches. The objective of this study was therefore to investigate and compare syndromic surveillance of heat-related emergency department visits and predictive models for heat-related emergency department visits using weather variables as predictors constructed using retrospective data in south eastern Ontario, Canada, from 2003 to 2008.

A question facing those developing heat response plans is what set of environmental variables or heat measures is best used for issuing warnings and for building models that might predict the expected number of individuals affected by heat-related illness. The direct effects of heat at the level of the individual have been extensively studied, both through physical modeling of heat exchange [15, 16] and physiological studies [17]. These studies indicate that the heat removed through the evaporation of perspiration and the effects of radiant sources of heat (e.g., the sun) play significant roles in the heat stress experienced by an individual. Therefore, in addition to the ambient air temperature, measures of heat exposure should consider the effects of sun, humidity, and wind. A variety of measures of heat exposure that combine these effects have been created and used to develop occupational health and safety exposure guidelines and public health heat warning thresholds including the heat index, humidex (a measure similar to the heat index used in Canada), and wet bulb globe temperature (WBGT) [15, 16, 18]. This study explores the question of which weather variables appear to have population-level significance rather than a priori selecting one of these exposure measures.

Studies of the population-level effects of heat suggest the existence of lagged relationships between heat exposure and illness or mortality, meaning that exposure over and up to several days in advance may impact the development of heat illness or be associated with heat-related mortality [8, 13]. Nonlinear and threshold effects have also been suggested [8, 19]. Population-level acclimation, through mechanisms such as behavioral (e.g., increased use of air conditioning) or physiological adaptation, and survivorship biases may seasonally modify population-level effects of heat [6, 8]. Because of the potential importance of these factors, this study uses nonlinear time series modeling techniques that allow the influence of lagged, nonlinear, and threshold effects of weather variables on heat-related illness to be captured.

To summarize, the objectives of the paper are (1) to construct predictive models for estimating heat-related visits, (2) to develop a simple syndrome definition for syndromic surveillance of heat-related visits, and (3) to compare the predictive models and syndrome definition to a gold standard measure of heat-related visits (NACRS).

2. Methods

2.1. Study Design. The study used an ecologic retrospective time-series analysis of heat-related hospital emergency department visits to ten hospitals across four health unit jurisdictions in south eastern Ontario, Canada, and weather data from five weather stations in the same area operated by Environment Canada from April 1, 2003 to December 31, 2008. The area covered included Peterborough in the west to Brockville in the east to Bancroft in the north to Lake Ontario in the south. Together these health units had a combined population of approximately 655 000, or roughly 5% of the 2006 Ontario population [20]. Only hospitals in the study area for which emergency department chief complaint data and National Ambulatory Care Reporting System (NACRS) data were available were included.

2.2. Outcome Definition and Measurement. Hospital emergency department visits were obtained from the NACRS database. All hospitals in Ontario are required to submit information on emergency department visits to the Canadian Institute for Health Information (CIHI) for inclusion in NACRS. The information in this database includes reasons for the visit, abstracted from patient charts and coded using the International Classification for Disease, 10th revision, Canadian Enhancement (ICD10-CA) coding system. CIHI routinely assesses the quality of data in this database using a number of methods [21]. Because it provides information on the post-diagnosis reason for visit, the NACRS data set was considered to reflect the true reason for visit and therefore served as a gold standard for the actual number of heat-related emergency department visits. Fields used from the NACRS database included ICD-10CA coded reason for visit, institution visited, age, sex, date of visit, partial postal code, and a nonidentifying encounter number that could be used to identify visits for the same individual on the same day, thereby preventing duplicate counting.

2 Journal of Environmental and Public Health
Chief complaint at time of patient registration at the emergency department was obtained for the hospitals included in the study. When an individual presents to the emergency department, the triage nurse enters a short (approximate five to ten word) free-text description of the patient’s chief complaint before the patient is diagnosed and treated. This information, visit date and time, age, gender, and partial postal code are sent from hospitals to a central database for use in a real-time syndromic surveillance system. This system has been described in detail elsewhere [22].

Previous epidemiological studies suggest that medical conditions that appear to be associated with hospital visits (admissions as well as to emergency departments) for heat-related illness include not only its direct effects but also malaise [11], some renal conditions [4, 5, 11], electrolyte and fluid imbalance [3, 5, 11], diabetes [5, 11], chronic respiratory conditions, cardiovascular disease, and cerebrovascular disease [4, 5, 8, 23, 24]. Heat has been associated with increased mortality for many of these conditions [4, 13, 23] and physiological evidence substantiates many of these associations [17]. However, because a large proportion of emergency department visits for many of these conditions is likely not heat-related, we defined heat-related illness as only those visits directly attributed to heat: ICD-10CA codes T67.0-T67.9, X30 [10], accepting the limitation that this may underestimate the true number of heat-related emergency department visits.

2.3. Exposure Measurement. Many measures of heat exposure have been developed: heat index [15, 25], humidex (Canada) [15, 25], and Wet Bulb Globe Temperature (WBGST) [15]. Because these measures combine different weather variables (some or all of temperature, humidity, wind speed, and solar radiation), and different weightings of each of these variables, this study considered each weather variable separately. This avoided any prior assumptions about the relative importance of individual weather variables in the exposure.

Hourly measurements of weather were obtained from Environment Canada for five weather stations located across Southeastern Ontario near the hospitals included in the study. Measurements included temperature, dew point, pressure, and wind speed. Unfortunately, a direct measure of solar radiation was not available, and therefore this variable was not considered in this study. Maximum, minimum, and average daily values of the temperature, dew point, wind speed, were calculated from the data from each of the five weather stations over the 24-hour period for each day. Missing observations were omitted from these calculations. The average value of the variables across weather stations was used in the analysis. In order to assess the error in doing this, variation in each variable across weather stations was calculated.

The humidex (HU) and heat index (HI), also referred to as apparent temperature, can be calculated as [15, 25]:

\[
HU = T_a + 0.5555(w - 10),
\]

\[
HI = -2.719 + 0.994T_a + 0.016(D)^2,
\]

where \(T_a\) is the air temperature in °C (shielded from ambient radiation), \(D\) is the dew point in °C, and \(w\) is the vapour pressure of water in hPa given below.

Maximum water vapor pressure, \(w\), was derived from maximum dew point by

\[
w = 6.11e^{5417.7530((1/273.16)-(1/(273.16+D)))}.
\]

This variable can be considered a transformation of dew point, and may be more directly related to the effects of humidity on an individual’s experience of heat [25, 26]. It was considered as a measure of humidity in the analysis.

2.4. Analysis. The association between visit counts, maximum temperature, average wind speed, and maximum water vapor pressure were assessed using a Poisson regression model accounting for possible overdispersion using SAS software. Same-day measurements of weather variables were examined in the Poisson regression model; lagged effects were not considered. Interactions were also not considered. Predictor variables in the model included an indicator variable differentiating spring and early summer (April/May/June) from other months of the year to investigate the possibility of acclimation over the period of one year. This choice was informed by preliminary exploratory analyses that had included a separate indicator variable for each month and by the hypothesis that acclimation may affect the response to heat. An indicator variable for weekends was also included in the model to control for the effects of the weekly pattern of emergency department visits [27].

Investigating the possibility of lagged and nonlinear relationships, threshold effects, and interactions between weather predictors in modeling their association with heat-related emergency department visit counts requires considering a large number of possible candidate terms in a regression model. To address this challenge, a method developed for nonlinear time-series modeling, Fast Orthogonal Search (FOS), was employed [28] (details of the method are given in this reference and not repeated here). This technique allows the construction of a very flexible nonlinear model describing the relationship between heat-related emergency department visits and weather variables by selecting terms from a potentially very large set of candidate terms. The objective is similar to forward selection in regression, except that the FOS algorithm implicitly expresses the selected candidate terms as a set of orthogonal functions. The reduction in mean square error associated with selecting any given candidate is found without having to actually find the orthogonal functions. FOS also allows the model coefficients for the selected terms to be found without finding these orthogonal functions. This minimizes the required computation and allows a large set of candidate terms (tens to hundreds of thousands) to be rapidly searched, which would otherwise require lengthy computation time.

The FOS model can be expressed as

\[
y[n] = \sum_{m=0}^{M} a_m p_m[n] + e[n],
\]
where $n$ is the time index, $a_m$ is a scalar coefficient, $e[n]$ represents the error, and $p_m[n]$ is a selected candidate function.

The method is potentially useful in examining the relationship between the time series of heat-related emergency department visits and environmental variables because in addition to handling nonlinearity and lagged effects, it allows threshold effects to be included in the candidate set. For example, if we only expect values of temperature over 25°C to be relevant, a function describing a step occurring at that temperature could be included as a candidate.

In the analysis performed in this study, each candidate was of the general form

$$p_m[n] = x_1[n - q_{11}] \ldots x_1[n - q_{1m_1}]x_2[n - q_{21}] \ldots x_p[n - q_{p1}]
\ldots U_i(x_1[n - r_{i1}] - S_{li1}) \ldots U_p(x_p[n - r_{pl}] - S_{pl})
\ldots U_L(x_p[n - r_{lp}] - S_{lp}) \ldots U_R(x_p[n - r_{Rp}] - S_{Rp}),$$

where $x_p$ is the $p$th predictor. There could be up to $m_p$ factors of a given predictor, $p$, in each candidate function. The lag for each factor of each predictor could be different and is denoted $q_{ij}$ where $i$ refers to the predictor and $j$ refers to each individual factor in that predictor.

Additionally, step functions, triggered by the value of one of the predictors, were possible factors in a term. The steps allow threshold effects to be modeled. In a right step, denoted $U_R$, when the value of the argument of the step function is less than a threshold value, its value is zero; otherwise, it is one. Conversely, for a left step, $U_L$, when the value of the argument of the step function is greater than a threshold value, the value is zero; otherwise, it is one. The $S_{ij}$ parameter in the argument of the step function gives the threshold value of step transition; here the $i$ subscript refers to the step type and the $j$ subscript refers to the predictor. The value of the predictor used to trigger the step is compared to $S$. The lag of this predictor, $r_{ij}$, was taken to be the minimum lag allowed in that predictor (i.e., $r_{ij} = \min(k|q_{j1}, q_{j2}, \ldots, q_{jm_j}|)$). Only one step function was allowed per predictor in each candidate term. The FOS algorithm was programmed by the authors using MATLAB, and the implementation was verified by fitting known relationships.

Each candidate term in the model could contain multiple factors. Log-transformed visits were considered as the model output. Predictors included the same weather variables used for the Poisson regression model. Table 1 describes the component factors possible in a given candidate term by giving the limits on the parameter values for (4) for each predictor. Step functions could occur without any factors of the corresponding predictor occurring. In preliminary exploratory analyses, maximum temperature and water vapour pressure lagged at up to seven days were allowed as candidates. The values in Table 1 were selected based on previous literature and exploratory analysis. Visit counts were log-transformed before fitting the model. The set of candidate terms searched included all possible combinations of each of the factors, approximately 140 000 possible candidate terms, from which FOS obtained a concise model within minutes on a standard laptop computer.

2.5. Heat-Related Illness Syndrome Definition. To develop a syndrome definition for heat-related illness based on emergency department chief complaint, NACRS visits (the gold standard) were matched to emergency department chief complaint using date and time of visit, hospital visited, partial postal code, age, and gender since a unique identifier was not available for matching records between data sets. By evaluating the positive predictive value (PPV) of the most commonly appearing text strings in emergency department chief complaints relative to the NACRS gold standard, a set of strings identifying likely heat-related visits was created. The resulting time-series of heat-related visits counts was created by counting the daily number of emergency department chief complaints containing at least one of the text strings in the set. This time-series was compared to the NACRS time-series of heat-related visit counts.

2.6. Model Evaluation and Comparison. A second model for estimating emergency department visits using FOS which included as predictors both daily counts of various text strings in chief complaints found to be associated with heat-related illness in addition to the weather variables was created. If the surveillance data provides additional benefit beyond the weather variables, then such a model should give better estimates of the actual number of heat-related visits.

To compare the predictive ability of the three models described above (Poisson regression, FOS model containing weather variable predictors only, and FOS model containing weather variables, and chief complaint text strings as predictors), the data was divided into training and validation data sets. The training data set (April 1, 2003 to December 31, 2006) was used to fit the models, and the validation data was used to assess the model fit (January 1, 2007 to December 31, 2008). Model estimates were compared to the gold standard (NACRS heat-related emergency department visit time-series) using the mean square error (MSE) and Pearson correlation coefficient.

2.7. Ethics. Ethics approval for the study was obtained from the Queen’s University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board.

3. Results

Heat-related emergency department visits ranged from 67 to 161 per year for the complete years included in the study (2004 to 2008). The median age of individuals with heat illness was 29, and half of those presenting to the emergency department were between the ages of 18 and 49. Ten percent were over age 71, and ten percent were under age 11.

The variation in weather variables across the five weather stations in the study area is given in Table 2.

Table 3 presents the parameter estimates, associated standard error, and $P$ values for the Poisson regression.
model which included maximum temperature, maximum water vapour pressure, average wind speed, and weekend and spring/early summer indicator variables as predictors of visits. The results show that of the weather variables included in the model, only maximum water vapour pressure and maximum temperature were significant, while average wind speed was not.

Figure 1 shows the proportion of days with a given number of emergency visits for various Heat Index ranges out of the total days over the study period with that Heat Index range. Heat Index was chosen because it is a function of only temperature and humidity, the only weather variables found to be significantly related to heat in the above model.

Equation (5) gives the model estimating heat-related emergency department visits, $y(k)$, using weather variables developed using FOS:

$$
\log[y(k)] = 1.237 \times 10^{-6} U_R \left[ t_{\text{max}}(k) - 15 \right] \left[ t_{\text{max}}(k) - 15 \right]^2 \\
+ 4.356 \times 10^{-4} U_L \left[ t_{\text{max}}(k) - 15 \right] \left[ t_{\text{max}}(k) - 15 \right] \\
+ (k_{\text{max}}(k) - 15) v(k - 1) v(k - 1) U_L [s(k) - 20] \\
- 4.476 \times 10^{-4} U_R \left[ t_{\text{max}}(k) - 30 \right] \left[ t_{\text{max}}(k) - 30 \right] \\
- 4.476 \times 10^{-4} U_R \left[ t_{\text{max}}(k) - 30 \right] \left[ t_{\text{max}}(k) - 30 \right] \\
- 4.476 \times 10^{-4} U_L [s(k) - 10] [s(k) - 10] [v(k - 1)]^2,
$$

where the symbols for the variables in the model above are given in the headings in the first column of Table 1.

Of the 1 444 519 emergency department surveillance records containing any chief complaint available over the study period, 95 945 (6.6%) visits could be matched to visits in the NACRS database based on the variables available. In this matched dataset, 225 records represented visits for heat-related illness as defined by the ICD10-CA codes given earlier. The most frequently occurring strings were identified from these confirmed heat-related visits. Table 4 presents the number of visits from the matched dataset containing each string, for both visits that were for heat-related illness and not for heat-related illness based on the associated ICD-coded reason for visit. This table is sorted according to the positive predictive value (PPV) of each string. The strings “sun”, “exhaust”, and “heat” had the highest positive predictive value of visits for heat-related illness of the frequently occurring strings and therefore defined the heat-related syndrome. However, using these three strings combined flagged only 81 of the 225 (36%) true visits for heat-related illness, while flagging another 40 visits that were not heat-related.

When the time-series of counts for each of the 14 most frequently occurring strings were included in a candidate set used by FOS in addition to the weather variables used above, the results showed that the best performing model included maximum temperature, maximum water vapour pressure, average wind speed, and weekend and spring/early summer indicator variables as predictors of visits. The results show that of the weather variables included in the model, only maximum water vapour pressure and maximum temperature were significant, while average wind speed was not.

Figure 1 shows the proportion of days with a given number of emergency visits for various Heat Index ranges out of the total days over the study period with that Heat Index range. Heat Index was chosen because it is a function of only temperature and humidity, the only weather variables found to be significantly related to heat in the above model.

Equation (5) gives the model estimating heat-related emergency department visits, $y(k)$, using weather variables developed using FOS:

$$
\log[y(k)] = 1.237 \times 10^{-6} U_R \left[ t_{\text{max}}(k) - 15 \right] \left[ t_{\text{max}}(k) - 15 \right]^2 \\
+ 4.356 \times 10^{-4} U_L \left[ t_{\text{max}}(k) - 15 \right] \left[ t_{\text{max}}(k) - 15 \right] \\
+ (k_{\text{max}}(k) - 15) v(k - 1) v(k - 1) U_L [s(k) - 20] \\
- 4.476 \times 10^{-4} U_R \left[ t_{\text{max}}(k) - 30 \right] \left[ t_{\text{max}}(k) - 30 \right] \\
- 4.476 \times 10^{-4} U_R \left[ t_{\text{max}}(k) - 30 \right] \left[ t_{\text{max}}(k) - 30 \right] \\
- 4.476 \times 10^{-4} U_L [s(k) - 10] [s(k) - 10] [v(k - 1)]^2,
$$

where the symbols for the variables in the model above are given in the headings in the first column of Table 1.

Of the 1 444 519 emergency department surveillance records containing any chief complaint available over the study period, 95 945 (6.6%) visits could be matched to visits in the NACRS database based on the variables available. In this matched dataset, 225 records represented visits for heat-related illness as defined by the ICD10-CA codes given earlier. The most frequently occurring strings were identified from these confirmed heat-related visits. Table 4 presents the number of visits from the matched dataset containing each string, for both visits that were for heat-related illness and not for heat-related illness based on the associated ICD-coded reason for visit. This table is sorted according to the positive predictive value (PPV) of each string. The strings “sun”, “exhaust”, and “heat” had the highest positive predictive value of visits for heat-related illness of the frequently occurring strings and therefore defined the heat-related syndrome. However, using these three strings combined flagged only 81 of the 225 (36%) true visits for heat-related illness, while flagging another 40 visits that were not heat-related.

When the time-series of counts for each of the 14 most frequently occurring strings were included in a candidate set used by FOS in addition to the weather variables used above,
The difference equation found by FOS was:

\[
\log[y(k)] = 9.3114 \times 10^{-7} U_R [t_{\text{max}}(k) - 15][t_{\text{max}}(k) - 15]^2 \times (t_{\text{max}}(k - 1) - 15)\nu(k - 0)\nu(k - 1)U_L[s(k) - 20] \\
+ 4.2143 \times 10^{-4} U_R [t_{\text{max}}(k) - 15][t_{\text{max}}(k) - 15] \times (t_{\text{max}}(k - 1) - 15)U_L[s(k) - 20]g(n)\nu(k - 1) \\
+ 1.359 \times 10^{-3} U_R[t_{\text{max}}(k) - 15] \\
\times U_L[s(k) - 20](\nu(k))^2(w_{\text{heat}}(k)),
\]

where \(w_{\text{heat}}(k)\) is the time-series of counts of emergency department chief complaints containing the string "heat" and the other symbols are again defined in the headings of Table 1. Note that this was the only string of the 14 included in the candidate set found by FOS to reduce the model MSE when weather variables were also included in the model.

Table 5 compares the various methods for estimating heat-related visits counts presented above over the validation time period. The FOS-derived models produced slightly better estimates than the regression model over both training and validation data sets, with the best model for estimating the true visits being the model which included both emergency department chief complaint string count time-series and weather variables as predictors.

Figures 2, 3, 4, and 5 present provide additional visual comparisons of the time series of the syndrome counts and FOS models to each other and the gold standard NACRS ICD-coded heat-related emergency department visits over the time period used to validate the models.

4. Discussion

The results of the regression model indicate that temperature and humidity are most useful in estimating emergency department heat-related visits, although wind speed also appears predictive in the better-performing FOS model. These findings make sense when considering the fact that individuals are more likely to be exposed to humidity and
temperature which can be present even indoors compared to wind which can easily be shielded by physical barriers and are more likely to be highly variable across a given geographic region. Inaccurate measurement of wind by insufficiently capturing its variation across the study region may have contributed to this result. Analogously, it may be that heat exposure measures that include variables that have high variability across individuals and or geography, and therefore cannot be adequately measured, may not provide any better measure of population risk of heat-related illness than simpler measures that omit these variables. Unfortunately, this study did not have precise enough measures of solar radiation to allow this variable to be assessed.

The results of this study suggest the possibility of an “early heat effect”, whereby heat experienced in the spring or early summer (defined as April, May, June) results in more “early heat effect” (whereby heat experienced in the spring or early summer (defined as April, May, June) results in more...
Figure 2: Comparison of syndrome counts for heat-related emergency department visits and NACRS ICD-coded heat-related emergency department visits.

Figure 3: Comparison of estimated visits for heat-related illness from FOS-generated model using weather variables as predictors and NACRS ICD-coded heat-related emergency department visits over validation data.

Figure 4: Comparison of syndrome counts for heat-related emergency department visits and estimated heat-related emergency department visits using FOS-generated model with weather-variable predictors over validation data.
Another weakness of the assessment of the relationship between heat and emergency visits performed in this study was failure to include covariates known to modify the effects of heat (such as age) or underlying medical conditions (such as diabetes), which may vary spatially. This may also bias the observed effect of heat. While these limitations impact the generalizability of the results, they do not impact the study objectives which rely on comparisons within the same population.

A perfect measure for the true number of heat-related emergency department visits was not available and the amount of error present in the NACRS data is not measurable. It has been suggested that some administrative data sources may underestimate heat-related cases: cases are underreported either because of failure of the clinicians assessing patients to recognize heat as a cause, failure to record this on patient records, or a failure to code heat as an underlying cause when medical records are being abstracted for inclusion in health databases such as NACRS [25, 30]. Since we defined heat-related emergency department visits as only those heat-specific ICD-10CA codes, it is likely that we underestimated the true number of visits. This may explain the low number of heat-related emergency department visits seen in the results of this study. To overcome these problems, other researchers have examined the excess numbers of deaths with heat [25] calculated by removing visits for known sources variability from visit totals; however, we chose to be conservative rather than run the risk of over-estimating numbers. In addition to missing visits, it was also possible that we included some visits not related to environmental heat exposure: the T67 group of ICD codes may include exposure to nonnaturally related heat sources (e.g., occupational exposure), providing a possible explanation for visits seen during winter months. Because of the restricted outcome definition of heat-related illness used in our study, the results may represent only the “tip of the iceberg” of heat effects on population.

Only a small number of chief complaints could be matched with the NACRS data (less than 7%), and some of these visits may have been mismatches since unique individual identifiers were not available. Difficulty matching and the fact that heat-related visits were relatively rare, reduced the amount of data available for creating a syndrome definition. Because of the lack of data, the syndrome definition was not validated on a set of data different from the one used to create it. This impacts the internal validity of the study: specifically the performance of the heat-related illness syndrome definition may be even less accurate than reported in Table 5. However, there would be no impact on validity of the model given by (6) since all strings in Table 4 were considered equally when constructing the model using FOS which was subsequently validated over a separate data set.

5. Conclusion

Temperature and humidity and were significantly associated with increased heat-related emergency department visits while there was less evidence supporting an association with wind speed. Heat in the spring and early summer appear to be associated with more visits, possibly due to subsequent acclimation. Very short lags (0 and 1 day) appeared to be important in explaining visits. Even with potential measurement errors in weather variables, time series models using weather variable predictors, fit using regional historic data, can be constructed that are highly correlated with future emergency department visits not used to fit the models. These models could be used with regional weather forecasts to predict visits and therefore could serve as an evidence-based population risk indicator for issuing heat warnings. Syndromic surveillance of heat-related illness might complement predictive models for estimating and monitoring population-level heat-related illness.

Acknowledgments

This work was funded in part by a Natural Sciences and Engineering Research Council of Canada (NSERC) CGS Scholarship held by A. G. Perry and by a grant from Health Canada.
References


