It’s Not the Heat, It’s the Humidity: Wet Weather Increases Legionellosis Risk in the Greater Philadelphia Metropolitan Area

David N. Fisman, Suet Lim, Gregory A. Wellenius, Caroline Johnson, Phyllis Britz, Meredith Gaskins, John Maher, Murray A. Mittleman, C. Victor Spain, Charles N. Haas, and Claire Newbern

School of Public Health and Department of Civil, Architectural, and Environmental Engineering, Drexel University, and Philadelphia Department of Public Health, Philadelphia; Montgomery County Department of Health, Norristown; Pennsylvania Department of Health, Reading; Bucks County Department of Public Health, Doylestown, and Chester County Department of Public Health, West Chester, Pennsylvania; Harvard School of Public Health, Boston, Massachusetts; and Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, Georgia

Background. Legionella species are abundant in the environment and are increasingly recognized as a cause of severe pneumonia. Increases in cases of community-acquired legionellosis in the greater Philadelphia metropolitan area (GPMA) led to concern that changing environmental factors could influence occurrence of disease.

Methods. We evaluated the association between weather patterns and occurrence of legionellosis in the GPMA, using both traditional Poisson regression analysis and a case-crossover study approach. The latter approach controls for seasonal factors that could confound the relationship between weather and occurrence of disease and permits the identification of acute weather patterns associated with disease.

Results. A total of 240 cases of legionellosis were reported between 1995 and 2003. Cases occurred with striking summertime seasonality. Occurrence of cases was associated with monthly average temperature (incidence rate ratio [IRR] per degree Celsius, 1.07 [95% confidence interval {CI}, 1.05–1.09]) and relative humidity (IRR per 1% increase in relative humidity, 1.09 [95% CI, 1.06–1.12]) by Poisson regression analysis. However, case-crossover analysis identified an acute association with precipitation (odds ratio [OR], 2.48 [95% CI, 1.30–3.12]) and increased humidity (OR per 1% increase in relative humidity, 1.08 [95% CI, 1.05–1.11]) 6–10 days before occurrence of cases. A significant dose-response relationship for occurrence of cases was seen with both precipitation and increased humidity.

Conclusions. Although, in the GPMA, legionellosis occurred predominantly during summertime, the acute occurrence of disease is best predicted by wet, humid weather. This finding is consistent with the current understanding of the ecological profile of this pathogen and supports the contention that sporadic legionellosis occurs through contamination of water sources.

Legionella species are environmentally abundant bacteria that can be isolated from water and soil samples [1, 2]. Contaminated building water–distribution systems have been a recognized cause of outbreaks of severe respiratory illness (legionnaires disease) since an outbreak in a Philadelphia hotel in 1976 [3]. However, legionellosis is increasingly recognized as an important cause of severe, sporadic community-acquired pneumonia [4]. Pneumonia caused by Legionella pneumophila is associated with respiratory failure and alveolar hemorrhage [4, 5]. Notwithstanding improved treatment and advances in supportive care, case-fatality rates are 10%–20% in the United States [6]. Immunosuppressed hosts and smokers appear to be particularly susceptible to pneumonia associated with Legionella species [4, 7].

Increased numbers of cases of community-acquired legionellosis have been noted in the greater Philadelphia metropolitan area (GPMA) during the past decade. Although increased numbers might reflect improved diagnostic technologies [8, 9], public-health officials have considered the possibility that such an increase might...
reflect a change in the local epidemiological profile of this disease, perhaps as a result of increased rainfall and flooding [10, 11]. Better understanding of environmental factors associated with legionellosis could enhance recognition of cases by clinicians and might lead to improved preventive efforts as well.

Traditional approaches to establishing causal links between environmental factors and occurrence of communicable disease may be confounded when occurrence of disease is seasonal. For example, increased incidence of disease during hot or rainy seasons might reflect population risk behavior during those seasons, rather than increased risk from environmental factors per se. The case-crossover study design is a novel epidemiological design that is useful for generating unconfounded estimates of effect with rare outcomes and repeated, transient exposures, such as weather [12, 13]. This study design has previously been used to evaluate links between acute changes in air quality and cardiac and respiratory events [14, 15], but its use in the study of environmental influences on occurrence of acute communicable disease has been extremely limited [16].

Our objective was to assess temporal and seasonal trends in occurrence of legionellosis in the GPMA from 1995 to 2003. We sought to evaluate links between monthly incidence of legionellosis and weather patterns using traditional analytic methods and to evaluate the association between acute changes in weather and occurrence of legionellosis using a case-crossover approach.

**METHODS**

The GPMA is a 2165-square-mile area located in southeastern Pennsylvania that had a population of 3,872,783 in 2000 [17] (figure 1). The region includes 5 counties (Bucks, Chester, Delaware, Montgomery, and Philadelphia), 4 of which have county health departments. Public-health services in Delaware County are provided by the Pennsylvania Department of Health. Demographic data for each county and for the US population as a whole were obtained from the 2000 US Census [17].

**Case identification.** Legionellosis is a notifiable condition in the Commonwealth of Pennsylvania. The Commonwealth uses the uniform case definition endorsed by the National Notifiable Diseases Surveillance System [18]. Cases are considered to be confirmed if a compatible clinical illness is associated with isolation of *Legionella* species or detection of *L. pneumophila* serogroup 1 antigens in appropriate clinical specimens or if a compatible illness is associated with increased titers of antibody to *L. pneumophila* serogroup 1. After the report of a case, appropriate epidemiological investigations are conducted by the
relevant public-health authority, to identify additional cases and/or possible environmental sources of the infection. Information on cases of legionellosis—including the date of onset, age and sex of the patient, presence of underlying medical conditions, smoking status, fatal outcome (if known), and whether the case was part of an identified outbreak—was obtained from county and state health authorities serving the GPMA. Individuals were considered to be immunocompromised if they were documented to have diabetes, a solid cancer, or hematological malignancy or to be taking immunosuppressive medications.

**Meteorological data.** Complete meteorological data for the study period were available from airport weather stations in Montgomery County (Willow Grove Naval Air Station) and Philadelphia County (Philadelphia International Airport) [19]. Because of within- and between-county differences in elevation and population density, we used average measures of maximum and minimum temperature, relative humidity, wind speed, atmospheric pressure, and precipitation obtained at the 2 weather stations and used average or total values as exposure variables (figure 1).

**Statistical methods.** Crude and age- and sex-standardized rates of occurrence of legionellosis were calculated using the estimated population from the 2000 US Census as the reference population [17]. Seasonal and temporal trends in occurrence of cases were evaluated using an approach similar to that described by Jensen et al. [20], with construction of Poisson regression models incorporating monthly and yearly terms:

\[
E(Y) = \exp \{ \alpha + \beta_1 \text{year} + \beta_2 \sin(2\pi \times \text{month}/12) + \beta_3 \cos(2\pi \times \text{month}/12) \}.
\]

E(Y) is the expected case count during a given month and year, α is a constant, and each \( \beta \) represents a regression coefficient for year or month. The association between monthly case counts and meteorological exposures was explored through construction of univariable and multivariable Poisson regression models. Multivariable models were created using a backwards-elimination algorithm, with covariates retained for P ≤ .20 [21], and year of occurrence was controlled for. Sine and cosine of month were not included in multivariable models, because of strong colinearity with temperature (\( r^2 = 0.95 \)). SEs were adjusted for clustering by county. When the date of onset of symptoms was not available, the month of occurrence of the case was based on the date of the first positive diagnostic test result for legionellosis.

Acute associations between meteorological exposures and occurrence of cases were evaluated using a case-crossover approach [12, 13]. The case-crossover design is analogous to the case-control design but is characterized by self-matching, with individuals serving as their own controls. In the context of case-crossover studies of environmental exposures, the hazard period is composed of person-time during which an event of interest occurred, whereas control periods are composed of person-time during which an event of interest did not occur. To avoid the introduction of bias due to temporal and seasonal trends, we employed random directionality of control selection [22, 23].

We used a time-stratified 2:1 matched case-crossover design, with hazard periods identified as the date of onset of symptoms of legionellosis in public-health records. Person-time at risk was divided into 3-week strata beginning on 1 January 1995. Control periods were the 2 days within each stratum that could be matched to the hazard period by day of the week and could precede, follow, or both precede and follow the hazard period.

Estimates of plausible effect periods [12, 13] (i.e., lags between acute weather occurrence and onset of disease) were based on a usual incubation period of 3–5 days for legionellosis, with a range of 2–10 days [24, 25]. We used both daily exposures and aggregate or average exposures. Aggregate and average exposures were defined as “likely during incubation” if they occurred 1–5 days before onset of disease, as “likely preceding incubation” if they occurred 6–10 days before onset of disease, or as “preceding incubation” if they occurred 11–15 days before onset of disease. Incidence rate ratios (IRRs) for occurrence of cases, based on weather effects, were approximated through construction of conditional logistic regression models [26].

We assessed dose-response relationships between exposures and outcomes by assigning within-stratum (i.e., within a 3-week period) quintile ranks to environmental exposure variables and incorporating the resulting exposure estimates into regression models, both as dummy variables and as 5-level ordinal variables. Evidence for a linear dose-response relationship was assessed using the Wald test for trend [27]. Modification of effects by subject- and disease-related factors was assessed through the creation of interaction terms, which were incorporated into logistic regression models. Modification of effect was considered to have occurred if \( P \) was <.05 for the interaction term coefficient [28]. All analyses were performed using SAS (version 8.01; SAS Institute) or Stata (version 8.0; Stata Corporation).

**RESULTS**

**Epidemiological profile of legionellosis in the GPMA.** There were 240 cases of legionellosis reported in the GPMA between 1 January 1995 and 31 December 2003. A male:female predominance was seen in the occurrence of cases, and the incidence of legionellosis increased with increasing age. Most cases were sporadic: only 14 cases (6%) occurred in association with recognized outbreaks. Crude and age- and sex-standardized rates of legionellosis are presented in table 1.

Two hundred four case patients (85%) had a documented date of onset of symptoms. There were no significant differences in sex, immunocompromised status, occurrence as part...
of an outbreak, or likelihood of survival between individuals with a documented date of onset of symptoms and those without a documented date of onset of symptoms. Individuals with missing dates of onset of symptoms were significantly younger (mean age, 54.5 vs. 60.6 years; \( P = .04 \)), less likely to be documented smokers (14% vs. 36%; \( P = .02 \)), and more commonly from Philadelphia and Montgomery Counties (92% vs. 57%; \( P < .001 \)), compared with those with a known date of onset of symptoms.

**Time trends, seasonality, and weather effects, using aggregate monthly case counts.** We created Poisson regression models to assess seasonality and time trends, using monthly aggregate case counts. By use of a model incorporating terms for calendar year and calendar month, we found strong evidence both for seasonality of legionellosis (\( P < .001 \)) and for increasing case counts over time (\( P < .001 \)). The overall goodness of fit of the time trend model was excellent (\( \chi^2 = 714.9 \) [537 df]; \( P < .001 \)). Actual and predicted case counts are presented in figure 2.

In univariable and multivariable models, risk of legionellosis increased with increasing average monthly temperature and average relative humidity and was inversely related to mean monthly barometric pressure. A weak but statistically significant increase in risk of occurrence of legionellosis was seen with increasing total monthly precipitation in univariable models, but it did not persist after other meteorological exposures were controlled for (table 2).

**Case-crossover analysis of acute weather effects.** Increased occurrence of legionellosis was associated with precipitation and increased relative humidity in the period that likely preceded infection (i.e., 6–10 days before occurrence of cases). The IRR for legionellosis 6–10 days after rainfall was \( 2.48 \) (95% CI, 1.30–3.12), whereas the IRR for legionellosis for each 1% increase in relative humidity was \( 1.076 \) (95% CI, 1.048–1.106). Legionellosis was not associated with precipitation or humidity during the likely incubation period (i.e., within 5 days of onset of disease) or >10 days before occurrence of cases.

An inverse association was seen between average wind speed 6–10 days before occurrence of cases and the incidence rate for legionellosis, for each kilometer per hour increase in wind speed (IRR, 0.986 [95% CI, 0.978–0.993]). No association between wind speed and occurrence of cases was seen using other effect periods. Occurrence of cases of legionellosis was associated with high atmospheric pressure >10 days before occurrence of cases and low pressure within 5 days of occurrence of cases, a finding that may reflect the movement of weather fronts resulting in rainfall in the intervening period [29] (figure 3). There was no relationship between daily or multiday maximum or minimum temperature readings and the incidence rate for legionellosis (data not shown).

The possibility of a dose-response relationship between rainfall, relative humidity, and wind speed was evaluated using a 6–10-day effect period. A strong, positive dose-response relationship was seen for both rainfall and relative humidity, whereas a negative dose-response relationship was seen for wind speed (table 3).

We assessed the possibility that meteorological exposures had different effects on the risk of occurrence of legionellosis among case patients with differing health and demographic characteristics. In case patients, no significant modification of the effects of meteorological exposures was seen for sex, age, smoking status, or immunocompromised status. Effects were also unchanged when outbreak-associated cases were excluded from the analysis and when fatal and nonfatal cases were analyzed separately.

**DISCUSSION**

Legionella species are an important cause of severe pneumonia in the United States [4, 24, 30, 31]. We analyzed data collected by public-health authorities in the GPMA during a 9-year period, using both traditional methods for time series data and a case-crossover methodology. We found that legionellosis occurred predominantly during hot months but was associated acutely with rainy, humid periods during those months. This finding is consistent with both the thermophilic nature of Legionella species
and the aquatic ecological profile of this organism [1, 32]. Our findings also provide novel insights into the importance of the physical environment in determining the occurrence and seasonality of bacterial diseases in the developed world.

The understanding of the epidemiological profile of legionellosis has been derived largely from investigations of outbreaks [3, 24]. However, although aerosol-related legionellosis outbreaks may be spectacular [33, 34], most cases of legionellosis in the United States occur sporadically [24, 30]. The epidemiological profile of sporadic legionellosis remains poorly understood, though it has been suggested that infection occurs most commonly when susceptible hosts aspirate contaminated drinking water or inhale aerosols produced by showerheads [35, 36].

Rainfall acutely increases microbial contamination and sediments in drinking water [37]. Organic sediments and the presence of other microorganisms increase *Legionella* replication, whereas sediments diminish the effectiveness of chlorination. [1, 38–40]. Thus, our finding of a 2.5-fold increase in occurrence of cases with precipitation just before the known incubation period of this organism is consistent with an increase in environmental *Legionella* species as the mechanism of increased risk. An increase in *Legionella* burden could occur because of a direct increase in *Legionella* species in source waters or water distribution systems or because of increased replication of the organism in colonized plumbing systems [41–43].

Such a mechanism is quite plausible: in tropical areas, the incidence of infectious diseases with water-related ecology (e.g.,

---

**Figure 2.** Trends in legionellosis in the 5-county greater Philadelphia metropolitan area. Bars represent actual numbers, and the black curve depicts numbers of cases predicted by a multivariable Poisson regression model including sine and cosine terms. Occurrence of cases is seasonal (summertime predominant) with a gradual increase in total cases over time (incidence rate ratio for each successive year, 1.16 [95% confidence interval, 1.02–1.32]).

---

**Table 2.** Monthly weather patterns and incidence of legionellosis in the greater Philadelphia metropolitan area.

<table>
<thead>
<tr>
<th>Meteorological element</th>
<th>Univariable models</th>
<th>Multivariable model&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR (95% CI)</td>
<td></td>
</tr>
<tr>
<td>Relative humidity, mean, %</td>
<td>1.10 (1.07–1.13)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Wind speed, mean, km/h</td>
<td>0.75 (0.71–0.80)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Maximum temperature, mean, °C</td>
<td>1.08 (1.06–1.10)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Minimum temperature, mean, °C</td>
<td>1.09 (1.07–1.11)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Precipitation, mm</td>
<td>1.0005 (1.0001–1.01)</td>
<td>.02</td>
</tr>
<tr>
<td>Atmospheric pressure, mean, kPa</td>
<td>0.47 (0.25–0.86)</td>
<td>.01</td>
</tr>
</tbody>
</table>

<sup>a</sup> The model was also adjusted for calendar year. Pressure and precipitation were excluded from multivariable models because of lack of statistical significance, whereas minimum temperature was excluded because of collinearity with maximum temperature.

---

NOTE. Incidence rate ratios (IRRs) reflect change in legionellosis risk per unit change in the meteorological variable in question. CI, confidence interval.
Legionella and Precipitation

Figure 3. Daily weather and risk of legionellosis in the greater Philadelphia metropolitan area. Incidence rate ratios and 95% confidence intervals for occurrence of cases are plotted on the Y-axis, whereas lags before occurrence of cases are plotted on the X-axis. An increase in the risk of occurrence of cases is seen with precipitation and increasing humidity using a 6–10 day lag (A and B), whereas increased wind speed during the same interval is associated with decreased risk (C). Dashed vertical lines separate the likely incubation period (at extreme right) from the period before the likely incubation period (at extreme left). Note that plots use log scales and that scales of Y-axes differ between panels.

Regardless of the mechanism, our findings provide epidemiological evidence for a direct role of the physical environment in occurrence of infectious disease in an urban-suburban setting in the developed world. The use of a case-crossover study design allowed distinction of acute weather effects from long-term seasonal weather patterns. Because case-crossover design is characterized by self-matching, our findings are unlikely to be due to seasonal changes in population behavior or to seasonal exposure to sources of Legionella species (e.g., with increased fre-

Table 3. Dose-response relationship between meteorological exposures and risk of legionellosis in the greater Philadelphia metropolitan area.

<table>
<thead>
<tr>
<th>Quintile</th>
<th>Precipitation a</th>
<th>Relative humidity b</th>
<th>Average wind speed c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st (referent)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2nd</td>
<td>1.64 (0.94–2.88)</td>
<td>2.26 (1.19–4.26)</td>
<td>0.24 (0.096–0.58)</td>
</tr>
<tr>
<td>3rd</td>
<td>1.41 (0.84–2.36)</td>
<td>1.76 (0.98–3.16)</td>
<td>0.12 (0.05–0.27)</td>
</tr>
<tr>
<td>4th</td>
<td>1.65 (0.96–2.82)</td>
<td>3.93 (2.18–7.09)</td>
<td>0.02 (0.008–0.064)</td>
</tr>
<tr>
<td>5th</td>
<td>2.40 (1.43–4.02)</td>
<td>3.59 (2.06–6.28)</td>
<td>0.46 (0.20–1.05)</td>
</tr>
</tbody>
</table>

**NOTE.** A 6-10 day effect period was used for all exposures except atmospheric pressure, for which an 11-14 day effect period was used. CI, confidence interval; IRR, incidence rate ratio.

a Wald χ², 9.14 (1 df); P = .003, test for linear trend.
b Wald χ², 25.84 (1 df); P <.001, test for linear trend.
c Wald χ², 18.01 (1 df); P <.001, test for linear trend.

malaria, cholera, and severe melioidosis) is influenced by wet weather [44–46]. In temperate regions of North America, modest increases in the turbidity of treated water have been associated with diarrheal disease in urban centers [47, 48]. Although Legionella replication is diminished by chlorination, the use of less-effective chlorination systems in municipal water supplies has emerged as a risk factor for nosocomial legionellosis, and the Legionella species colonizing plumbing systems may be relatively chlorine resistant [49, 50].

Table 3. Dose-response relationship between meteorological exposures and risk of legionellosis in the greater Philadelphia metropolitan area.
frequency of showering or swimming). Self-matching also eliminates confounding that might occur as a result of differential exposure to *Legionella* species by individuals with different susceptibility to infection. Finally, because this design is useful when outcomes of interest are rare, it is ideally suited to analysis of public-health surveillance data [12, 13].

We noted distinct seasonality in the occurrence of cases of legionellosis in the GPMA. Although the seasonality of infectious diseases has been recognized since the time of Hippocrates [51] and is sufficiently striking to influence current disease surveillance efforts (e.g., influenza) [52], the seasonality of non-vectorborne infectious diseases remains poorly understood [53–56]. Better understanding of mechanisms underlying seasonality and weather in occurrence of infectious diseases is of more than academic interest: recognition of weather and seasonal patterns in occurrence of legionellosis may enhance the likelihood that clinicians recognize this underdiagnosed disease [4, 31], improve the accuracy of “syndromic” surveillance systems currently in use for identification of bioterrorist attacks, and inform public-health policy in the context of global climate change [53, 57, 58].

Although the increase in legionellosis cases in Philadelphia during the study period coincided with periods of heavy rainfall [10, 11], it is not possible to determine whether this increase was due to changing epidemiological profile, improved diagnostic technology [8, 9], or a combination of these factors. Nonetheless, our findings suggest that changes in rainfall patterns may have been responsible for at least some component of this increase.

The present study has several limitations. As in any study that makes use of environmental data, it is possible that some exposures may have been misclassified. However, random misclassification of exposure status would be likely to weaken the strength of observed associations, suggesting that our estimates of risk may be lower-bound estimates [59].

A second limitation of the present study is the circumscribed nature of case data available through public-health surveillance systems. In this instance, the lack of detail on individual cases was counterbalanced by our desire to include all known cases of legionellosis in the GPMA in our analysis. Nonetheless, it is likely that even these records represent only a subset of the cases of legionellosis that occurred in the GPMA during the study period, as a result of underdiagnosis and underreporting of disease. It should be noted that undercounting of cases of legionellosis would bias our analysis only if the likelihood of case reporting were correlated with environmental conditions (e.g., cases that occur during rainy periods are more likely to be diagnosed and reported). Finally, it might be argued that clustering of cases in association with outbreaks might have resulted in spurious identification of coincident weather patterns as risk factors for disease. However, exclusion of outbreak-associated cases did not change the results of our analysis.

In summary, we found that a recent increase in legionellosis in the GPMA was associated with marked summertime seasonality. However, the use of a case-crossover approach identified rainfall and humidity, rather than increased temperature, as predictors of occurrence of legionellosis. Although further study is needed, this finding is consistent with the contention that contamination of potable water is important in the genesis of sporadic legionellosis. The analytic approach employed here appears to be a valuable tool for identification of environmental influences on occurrence of acute communicable diseases.

**Acknowledgments**

We wish to thank Dr. Zekeria Berhane, Jennifer Breaux, and 2 anonymous reviewers, for their valuable suggestions related to this study.

**References**


18. Case definitions for infectious conditions under public health surveil-