Epidemiol. Infect. (2014), **142**, 2360–2371. © Cambridge University Press 2014
doi:10.1017/S0950268813003476

Summer increase of Legionnaires' disease 2010 in The Netherlands associated with weather conditions and implications for source finding

P. S. BRANDSEMA[1]*, S. M. EUSER[2], I. KARAGIANNIS[3], J. W. DEN BOER[2] AND W. VAN DER HOEK[1]

[1] *Centre for Infectious Diseases, Epidemiology and Surveillance, National Institute for Public Health and the Environment, Bilthoven, The Netherlands*
[2] *Regional Public Health Laboratory Kennemerland, Haarlem, The Netherlands*
[3] *Department for Infectious Disease Epidemiology, Respiratory Infections Unit, Robert Koch Institute, Berlin, Germany*

Received 22 July 2013; Final revision 27 November 2013; Accepted 12 December 2013; first published online 23 January 2014

SUMMARY

During August and September 2010 an unexpected high number of domestic cases of Legionnaires' disease (LD) were reported in The Netherlands. To examine this increase, patient characteristics and results of source finding and environmental sampling during the summer peak were compared to other domestic cases in 2008–2011. This analysis did not provide an explanation for the rise in cases. A similar increase in LD cases in 2006 was shown to be associated with warm and wet weather conditions, using an extended Poisson regression model with adjustment for long-term trends. This model was optimized with the new data from 2008 to 2011. The increase in 2010 was very accurately described by a model, which included temperature in the preceding 4 weeks, and precipitation in the preceding 2 weeks. These results confirm the strong association of LD incidence with weather conditions, but it remains unclear which environmental sources contributed to the 2010 summer increase.

Key words: Climate – impact of, infectious disease control, infectious disease epidemiology, Legionnaires' disease, *Legionella*.

INTRODUCTION

Legionnaires' disease (LD) is a pneumonia acquired by inhalation of aerosols containing *Legionella* bacteria. Most cases of diagnosed LD are hospitalized with a severe pneumonia, often requiring intensive care [1]. The incubation period is usually considered to be 2–10 days, but may exceed 10 days in 14% of cases, with a maximum of 19 days [2, 3]. Although many *Legionella* species and serogroups have been identified, *Legionella pneumophila* serogroup 1 is the causative agent in most cases of LD [4].

The *Legionella* bacterium is an environmental micro-organism that can be found in soil and water. In the aquatic environment, especially in artificial water systems, *Legionella* may multiply at temperatures in the range of 25–45 °C and may survive temperatures up to 50 °C [5]. Recognized sources of LD outbreaks include wet cooling towers and water systems such as spa pools and showers [6]. However, in most sporadic

* Author for correspondence: Mrs P. S. Brandsema, Centre for Infectious Diseases, Epidemiology and Surveillance, National Institute for Public Health and the Environment (RIVM), A. van Leeuwenhoeklaan 9, PO Box 13720 BA Bilthoven, The Netherlands.
(Email: petra.brandsema@rivm.nl)

(non-outbreak) cases the source of infection remains unknown. In The Netherlands, during 9 years (2002–2010) of environmental investigations 1317 potential sources linked to LD patients have been sampled and annually only around four genotypic matches between a patient's isolate and an environmental strain have been found [7].

In The Netherlands an average of 345 LD cases per year were notified between 2008 and 2011 of which 62% were domestic, while 38% of cases had travelled abroad during their incubation period [8]. There is a seasonal pattern with most cases reported between May and October with a peak during the summer months. Between 2003 and 2008 there was an increasing trend in domestic cases, which may be explained by the increasing use of the urine antigen test, resulting in a decrease of under-diagnoses (P. S. Brandsema, RIVM, unpublished data). In 2009, however, a markedly low incidence was observed that could not be attributed to a reduced number of diagnostic tests [9]. The epidemic curve shows a marked peak in LD incidence in 2006 and 2010 [8, 10, 11]. With an age-standardized rate of 2·8/100 000 inhabitants, The Netherlands had the highest rate of LD in Europe in 2010 [12]. There were no outbreaks identified or changes in the surveillance system that could explain these variations in incidence, and real fluctuations in LD incidence from year to year are likely.

Previous studies have suggested that weather conditions are associated with the incidence of LD [13, 14]. The peak in LD cases in the summer of 2006 in The Netherlands was explained by warm and humid weather conditions [15]. Like 2006, the increase in 2010 was only observed in August and September. The weather conditions in summer 2010 were very similar to those in summer 2006 with a record hot July followed by a record wet August [16]. We therefore analysed the LD cases that occurred in 2010 with the aim of assessing the role of weather conditions in the incidence rate of sporadic LD. To examine alternative explanations for this increase we also describe the characteristics of the domestic cases and the results of environmental investigations.

METHODS

Notification system and source finding

In The Netherlands LD has been a mandatory notifiable disease since 1987. Laboratory-confirmed cases are reported by clinicians and microbiologists to the Municipal Health Services (MHS). For each notified case a source-finding investigation is conducted by the MHS. The patient or a relative is interviewed using a standardized questionnaire to identify exposure to potential environmental sources during a 14-day incubation period. This questionnaire includes travel history, healthcare or hospital admission prior to onset, work-related exposure and leisure activities such as gardening, visiting a spa or swimming pool, wellness centre, or garden centre, showering at sporting facilities, exposure to a fountain, car wash or other location with possible aerosol exposure. Cases and their potential sources are recorded anonymously by the MHS in the national infectious diseases surveillance database (Osiris). All identified potential sources from the source-finding investigation are also registered by source type and postcode in a national database (National Legionella Outbreak Detection Programme; NLODP) [7, 17]. This database is designed to detect clusters linked to a potential source or geographical area.

Environmental sampling

Environmental investigations are performed by the National Reference laboratory (NRL) if one of the following criteria is met: (i) a potential source of infection is linked to more than one patient in a 2-year period (cluster), (ii) three or more patients live within 1 km radius within a 6-month period (geographical cluster), (iii) a case stayed in a healthcare setting during the incubation period, or (iv) a clinical isolate is available. In July 2009 criterion (iv) for solitary patients was restricted to those patients who reported at least one other potential source besides the patients' dwelling. Environmental sampling for clusters or healthcare-associated cases, included the cluster site or healthcare setting, but other potential sources of infection reported by the cases were also sampled.

For each location sampling points were selected by trained NRL personnel in cooperation with the technical team of a facility (when available) to obtain a comprehensive collection of samples. Water samples (500 ml) and swab samples were obtained in accordance with national guidelines [18, 19]. Water samples were concentrated by filtration and filtered residues were resuspended in 1 ml sterile water. Of this suspension, 100 μl samples were cultured without dilution and after tenfold dilution on two media at 35 °C, as described previously [7]. Swab samples were dispersed by immersion in 1 ml sterile water and cultured as

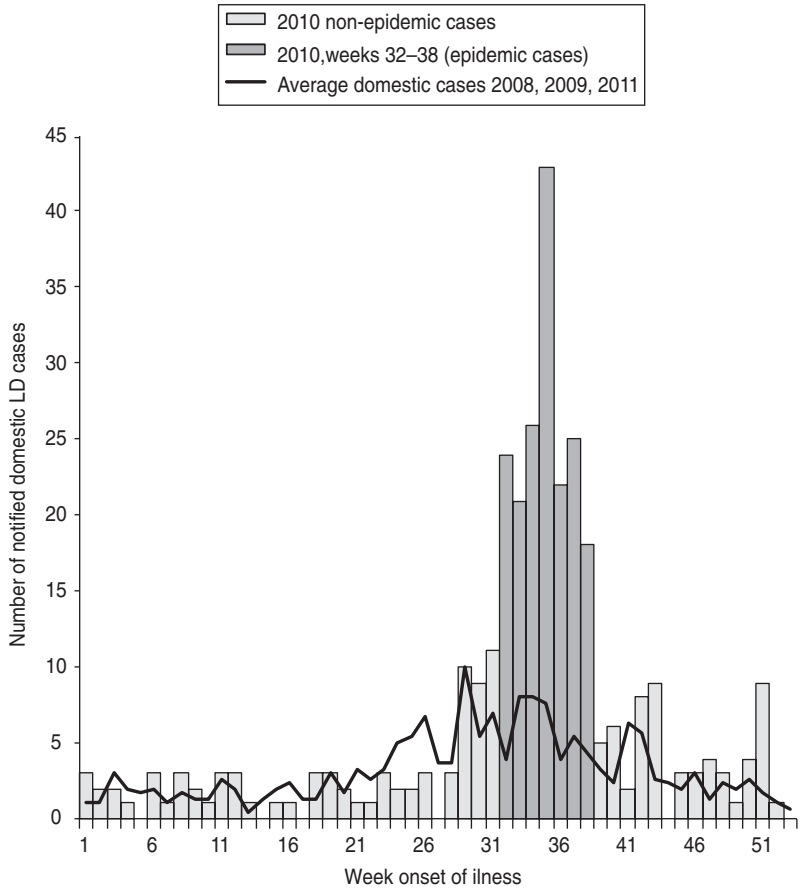

Fig. 1. Number of domestic Legionnaires' disease (LD) cases notified in The Netherlands with onset in 2010 compared to the average number of domestic cases in 2008, 2009 and 2011.

described above. Both patient and environmental *Legionella* isolates were serotyped using commercially available kits [7]. Sequence-based typing was performed, as recommended by the European Study Group for Legionella infections [20, 21].

Comparison of case characteristics and source-finding results

Epidemiological data on domestic cases were gathered from the Osiris database. Results of environmental sampling and typing of isolates were obtained from the NLODP database. All notified LD cases with pneumonia and onset of illness between 1 January 2008 and 31 December 2011 who had not travelled abroad during the 14-day incubation period (domestic cases), were included in the analysis.

The domestic LD cases during the peak in 2010 (onset weeks 32–38, 9 August–26 September 2010), further denoted as 'epidemic cases' (Fig. 1), were compared to the domestic cases from 2008 to 2011, excluding cases with illness onset in weeks 32–38 of 2010 ('non-epidemic cases'). Two-tailed χ^2 test or Fisher's exact test was used to test categorical variables for differences between groups. Continuous variables were compared using the Mann–Whitney U test. $P<0{\cdot}05$ was considered statistically significant.

Weather analysis

For the weather analysis, the dataset was extended to all notified domestic LD cases with onset of illness between 1 July 2003 and 31 December 2011. Only sporadic domestic cases were included in the model. For clusters only the index case was included. Meteorological data were obtained from the Royal Netherlands Meteorological Institute (KNMI) for the weather station in De Bilt, located in the middle of the country (see Fig. 2) [16].

To investigate the association of the incidence of domestic LD cases with weather variables, the extended Poisson regression model developed by

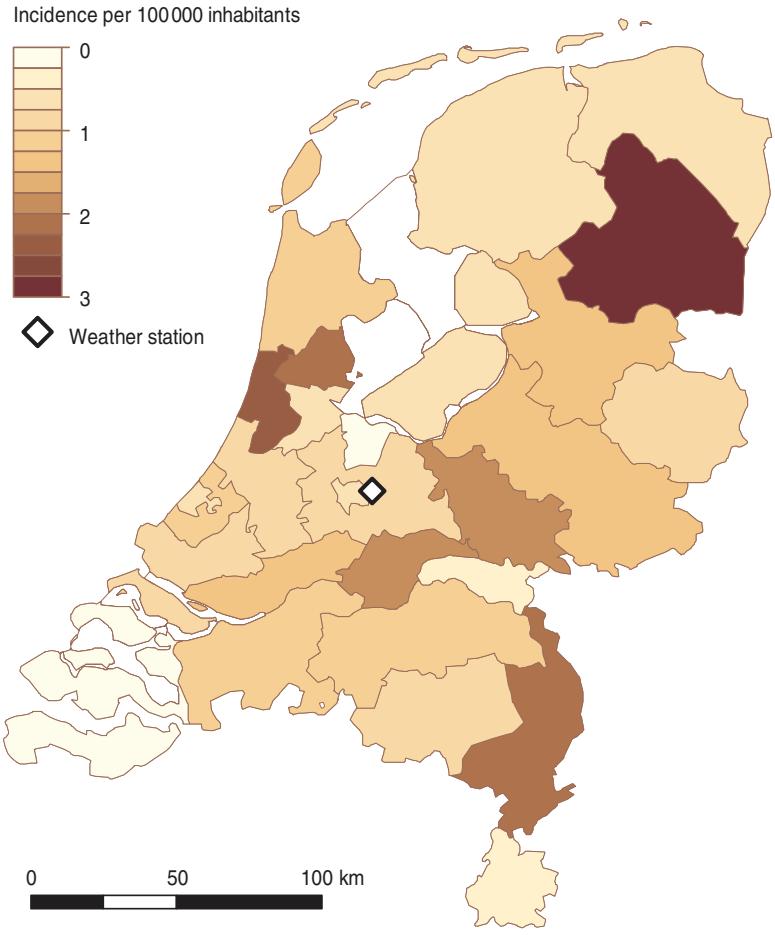

Fig. 2 [*colour online*]. Incidence rate of domestic Legionnaire's disease cases per municipal health region in weeks 32–38 (2010) and the location of the De Bilt weather station.

Karagiannis *et al*. [15] was used. This model includes relative humidity, average weekly temperature, precipitation intensity, and loess smoothing. First, the data from 2003 to 2007 that were used to develop this model, was updated with the recent data (2008–2011). Cases were attributed to an infection date by redistributing cases backwards in time according to a Gamma-distributed incubation period, for which the maximum-likelihood estimate according to Egan *et al*. [22] (1–15 days) was used. The infection dates were then aggregated to weekly data and case numbers were rounded to digits.

To control for long-term seasonality, a loess smoothing function for the LD count data was included in the model [23]. Loess is a flexible nonparametric regression method, also known as locally weighted polynomial regression. This method will fit a smoothed curve through a set of data points, using weighted quadratic least squares regression, giving more weight to variables near the point whose response is being estimated. Including a loess curve will accommodate over-dispersion in the data, as the model will be fitted to segments of the data. The smoothing window, with minimal partial autocorrelation in the residuals, was calculated based on the new data and set to 21 weeks. The loess function of the week preceding the week of infection was included in the adjusted model. The period included in the model was extended from weeks 18–39 in the original model to weeks 16–44 (mid-April to end of October) in the adjusted model.

To optimize the model for the new dataset all variables with a significant association in the original analysis were included: the mean weekly value for temperature, relative humidity, cloudiness, precipitation, and rainfall intensity (calculated as the weekly sum of precipitation divided by total duration of rainfall). In addition, the weekly mean atmospheric pressure and rainfall duration were included, as well as a second rainfall intensity variable (calculated as

a daily ratio of precipitation by rainfall duration). In order to examine the influence of a longer period of warm weather, the mean temperature for 2, 4 and 6 weeks preceding the infection date were also included as variables in the analysis. Furthermore, a 2-week variable was included for the rainfall variables (precipitation, duration, intensity). To account for possible nonlinear associations with LD incidence, all weather variables were transformed into categorical variables with cut-off points at every 10th percentile, except for temperature and cloudiness, which were categorized into six categories (percentiles <10, 10–25, 25–50, 50–75, 75–90, >90).

The variables were examined by univariate Poisson regression and significant variables were included in a multivariable Poisson regression analysis with backwards variable selection. For the multivariable analysis it was decided that only one long-term temperature variable should be included in the model simultaneously. Nested models were compared with the likelihood ratio (LR) test. For competitive models the model with the smallest Akaike's Information Criterion (AIC) value was chosen [24]. The fit of the model was assessed with Pearson's goodness-of-fit χ^2 test.

Statistical analysis of case characteristics was performed with IBM SPSS Statistics, version 19 (IBM Corp., USA) and Poisson regression was performed using Stata release 12 (StataCorp, USA).

RESULTS

Patients' characteristics and source-finding results

A total of 1368 LD cases with onset of illness between 1 January 2008 and 31 December 2011 were notified, of which 539 cases were excluded from the analysis because they had travelled abroad during the incubation period. This left a total of 829 domestic cases in the analysis of which 195 had an onset of illness in 2008, 142 cases in 2009, 317 cases in 2010 and 175 cases in 2011. In total 179 cases had an onset of illness from weeks 32–38 in 2010, compared to an average of 41 cases in this period from 2008 to 2011 (Fig. 1). The LD cases in the epidemic group were reported in 26 of the 28 MHS regions. The increased LD incidence was observed in 24 of the 28 MHS regions, although the rise in cases was most pronounced in four non-adjacent regions (Fig. 2).

The epidemic and non-epidemic groups were comparable with respect to patients' characteristics and proportion of community-acquired cases (Table 1). In the non-epidemic group, serology was used more often as diagnostic method compared to the epidemic group. The 32 patients' isolates that were available from the epidemic group consisted of 14 different *Legionella* strains, of which no particular *Legionella* type or strain stood out. *L. pneumophila* serogroup 1 was isolated most often (91%), and the sequence type (ST) type found most frequently was ST47 (AFLP type 004 Lyon). This is a common pattern found in cases in The Netherlands [25].

The source-finding investigations suggest that cases in the epidemic group may have been more mobile than the non-epidemic group, as 61·5% *vs.* 53·5% ($P=0·06$) of cases did not stay within their place of residence during their incubation period, they also visited other places. There were six cases in the epidemic group that could be allocated to four different small clusters. Although slightly more environmental sampling was conducted for the epidemic cases, this did not result in a higher success rate of *Legionella* detection. In the epidemic group *Legionella* was found in 15/47 (31·9%) of the sampled cases, which was 8·4% of the total number of 179 epidemic cases. In the non-epidemic group 66/142 (46·5%) of sampled cases had a positive result in environmental investigations, which was 10·2% of the total number of 650 non-epidemic cases. This difference was not significant ($P=0·08$).

The patients' dwelling (private residence), which overall was the most frequently sampled source, was sampled less frequently for the epidemic cases (Table 2). However, if the dwelling was sampled, *Legionella* was found less often in the dwelling of epidemic cases, although this difference was not significant. Commercial accommodations (hotels, campsites, holiday homes) were sampled more frequently, but this was a seasonal effect since the included period for the epidemic group was the holiday season, while the whole year was included for the non-epidemic group. For cases with illness onset in August and September, there was no difference between the groups for spending the night elsewhere ($P=0·12$). As garden centres were overrepresented in the sampled sources for the epidemic group, we also looked at outdoor gardening activities reported by cases in August and September. There was no difference between the groups. In the epidemic group 29 (16·2%) cases reported gardening activities during incubation time, compared to 28 (17·3%) of the 162 cases with onset in August or September from

Table 1. *Characteristics of Legionnaires' disease cases and source-finding results in the epidemic group (weeks 32–38 in 2010) compared to domestic cases in 2008–2011 (excluding weeks 32–38 in 2010). [Source: National Infectious Diseases Surveillance database (Osiris)]*

	Epidemic group*		Non-epidemic group*		*P* value
Patients' characteristics	*N* = 179	(100%)	*N* = 650	(100%)	
Male gender	124	(69·1%)	463	(71·2%)	0·61
Age >62 years (=median age)	92	(51·4%)	311	(47·8%)	0·87
Underlying disease†	63	(36·2%)	253	(39·8%)	0·39
Smoking‡	106	(60·9%)	347	(54·6%)	0·14
Hospital admission	177	(98·9%)	628	(96·7%)	0·19
Death	6	(3·4%)	43	(6·6%)	0·16
Setting					
Community acquired	157	(87·7%)	573	(88·2%)	0·87
Nosocomial/healthcare setting	5	(2·8%)	19	(2·9%)	0·93
Domestic travel	17	(9·5%)	53	(8·2%)	0·57
Unknown	0	—	4	(0·6%)	0·58
Diagnostic methods					
Culture confirmed	36	(20·1%)	140	(21·5%)	0·68
Urine antigen test	158	(88·3%)	560	(86·2%)	0·46
Serology	5	(2·8%)	58	(8·9%)	<0·01
Polymerase chain reaction	26	(14·5%)	74	(11·3%)	0·25
Patient isolates§	*N* = 32	(100%)	*N* = 132	(100%)	—
L. pneumophila sg 1	29	(90·6%)	121	(91·6%)	0·74
L. pneumophila sg 2–14	2	(6·3%)	7	(5·3%)	0·69
Most frequent sequence types					
ST47	13	(40·6%)	49	(37·1%)	0·71
ST62	2	(6·3%)	14	(10·6%)	0·46
ST46	3	(9·4%)	8	(6·1%)	0·45
ST45	3	(9·4%)	5	(3·8%)	0·19
Non-*L. pneumophila*	1	(3·1%)	4	(3·0%)	1·0
Source finding					
Case mobility (range)					
Travel outside MHS region	43	(25·4%)	159	(26·0%)	0·88
Travel outside place of residence	104	(61·5%)	327	(53·5%)	0·06
Range unknown	10	(5·6%)	39	(6·0%)	0·84
Clustered cases	6	(3·4%)	28	(4·3%)	0·57
Cases with environmental sampling	47	(26·3%)	142	(21·8%)	0·21
Cases with positive sampling	15	(31·9%)	66	(46·5%)	0·08

MHS, Municipal Health Service.

* Epidemic group: domestic cases with onset disease in weeks 32–38 (2010). Non-epidemic group: domestic cases with onset of disease in 2008, 2009, 2010 (excluding weeks 32–38) and 2011.

† Underlying disease was unknown in 2·8% and 2·2% of cases, respectively.

‡ Smoking was unknown for 2·8% and 2·0% of cases, respectively.

§ Source: Database National Reference Laboratory for Legionella in Haarlem, The Netherlands.

the non-epidemic group. More wellness centres were sampled in the non-epidemic group, as there had been a number of clusters in wellness centres during that period [26].

The environmental sampling of potential sources in Table 2 showed no significant differences in the sampling results between the groups. *Legionella* was detected in 21·5% of sampled sources linked to epidemic cases, and slightly more often in sources linked to non-epidemic cases (27·6% of sampled sources). However, this comparison may be biased, because garden centres were overrepresented in the epidemic group, and wellness centres were overrepresented in the non-epidemic group. As can be seen in Table 2, garden centres were often negative for *Legionella* by environmental sampling, while wellness

Table 2. *Results of environmental investigation of potential sources linked to LD patients. (Source: Database National Legionella Outbreak Detection Programme, National Reference Laboratory for Legionella in Haarlem, The Netherlands)*

Type of source sampled	Number of sources sampled					Positive sampled sources				
	Epidemic group*		Non-epidemic group*		*P* value	Epidemic group		Non-epidemic group		*P* value
Private house	24	(30·4%)	105	(38·2%)	0·20	5	(20·8%)	28	(26·7%)	0·55
Hospital or healthcare setting	8	(10·1%)	22	(8·0%)	0·55	4	(50%)	12	(54·5%)	0·83
Garden centre	14	(17·7%)	27	(9·8%)	0·05	1	(7·1%)	2	(7·4%)	1·00
Wellness centre	1	(1·3%)	16	(5·8%)	0·14	1	(100%)	13	(81·3%)	1·00
Accommodation†	9	(11·4%)	14	(5·1%)	0·05	1	(11·1%)	2	(14·3%)	1·00
Cooling tower	3	(3·8%)	17	(6·2%)	0·58	1	(33·3%)	9	(52·9%)	1·00
Car wash	4	(5·1%)	18	(6·5%)	0·79	0	—	2	(11·1%)	1·00
Place of work	2	(2·5%)	14	(5·1%)	0·54	1	(50%)	2	(14·3%)	0·35
Other	14	(17·7%)	42	(15·3%)	0·60	3	(21·4%)	6	(14·3%)	0·40
Total no. of sources sampled	79	(100%)	275	(100%)	—	17	(21·5%)	76	(27·6%)	0·28

* Epidemic group: domestic cases with onset disease in weeks 32–38 (2010). Non-epidemic group: domestic cases with onset of disease in 2008, 2009, and 2010, excluding weeks 32–38 (2011).
† Accommodation: hotel, campsite or holiday home.

centres were often positive for *Legionella* by environmental sampling. If we exclude these types of sources from the comparison, *Legionella* was found in 23·4% of sources linked to epidemic cases, and 26·3% in sources linked to non-epidemic cases.

Weather analyses

A total of 1589 domestic cases were notified with onset of illness occurring between 1 July 2003 and 31 December 2011. After exclusion of 75 clustered cases, 1514 (95·3%) cases were left in the dataset, of which 1156 had been infected during the warmer period of the year (weeks 16–44). The model developed by Karagiannis *et al.* [15] was applied to the data for domestic LD cases from week 27 (2003) to week 44 (2008) (2003–2008 model, $n=676$) and a prediction was made for the cases for 2009–2011 ($n=480$). Although this model resulted in an adequate prediction of the low LD incidence in 2009 and the high incidence in 2010, the variance (R^2) explained by the model was reduced from 43% in the original model to 32% in the 2003–2008 model. Running the model with the extended dataset up to 2011 (2003–2011 model, $n=1156$) improved the R^2 to 39%, but markedly decreased the fit of the model (Pearson's goodness-of-fit χ^2 $P=0·10$).

Replacing the original rain intensity variable (based on the week sum) with the alternative value of rain intensity based on the daily ratio (2003–2011a model) improved the R^2 of the model to 42% resulting in a better fit (Pearson's goodness-of-fit χ^2 $P=0·68$). This alternative variable of rain intensity was used for further analysis.

To allow for long-term associations, the 2-, 4- and 6-week mean temperatures and 2-week variables for rainfall were introduced in the analysis. In the univariate models (without loess correction) the variance was best explained by the 6-week temperature ($R^2=19·9\%$), the 2-week rain intensity ($R^2=17·2\%$), 4-week temperature ($R^2=16·0\%$) and the model with 2-week precipitation ($R^2=14·3\%$). These variables also had the smallest AIC values.

Figure 3 shows the incidence rate ratios (IRR) for 6-week mean temperature, and 2-week variables for rainfall intensity, duration, and precipitation. In the univariate analysis these long-term variables had a higher IRR than the temperature and rainfall variables that enclosed a shorter period. The IRR of all variables are presented in Table 3.

In the multivariable analysis, the variables for long-term mean temperature and precipitation were consistently found to be significant predictors in different competitive models. However, since the precipitation

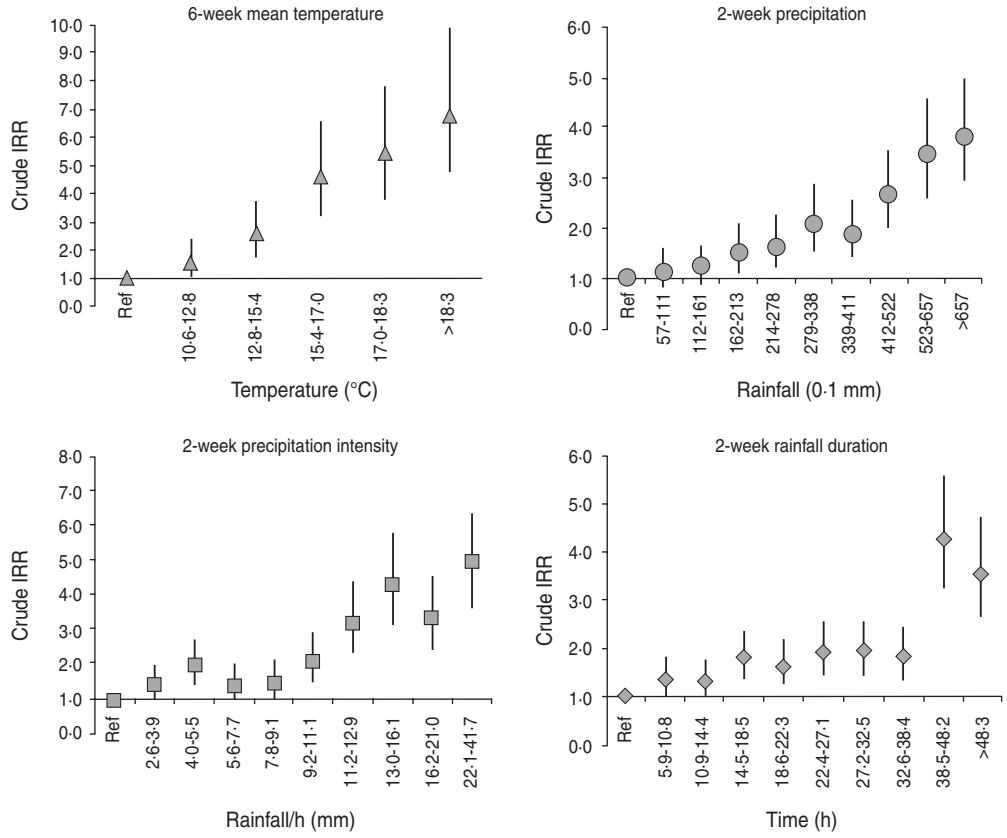

Fig. 3. Univariate Poisson regression analysis of 6-week mean temperature, 2-week rainfall intensity, 2-week precipitation, and 2-week rainfall duration. Incidence rate ratio (IRR; symbol) and 95% confidence interval (line) per category.

variables (1- or 2-week period for precipitation sum, duration, and intensity) are closely associated, different combinations of these variables resulted in near comparable results.

The final model included the 4-week mean temperature, 2-week rainfall duration and the 2-week rainfall intensity. After a 4-week period of warm weather (but not the hottest weather) the incidence is 2·2 times higher than after a colder period (4-week mean temperatures below 10·5 °C). Long-lasting and intense rainfall further contributes to an increased LD incidence. This model (2003–2011b model) was able to explain 46·5% of the variance in the epidemiological data. The AIC of the model was 967·3 and it had an excellent goodness-of-fit, based on Pearson's residuals ($P = 0{\cdot}99$). This is an improvement to the 2003–2011a model based on the mean weekly temperature, relative humidity, and precipitation intensity (AIC 1064·0 and Pearson's goodness-of-fit, $P = 0{\cdot}68$). In Figure 4 the prediction of the final model for the cases in 2010–2011 is shown if the model is fitted on data up to 2009 (2003–2009b model).

DISCUSSION

This study confirms that the record high LD incidence in summer 2010 was related to a long period of warm weather which was followed by intense rainfall. The Poisson regression model, with 4-week mean temperature, 2-week rainfall intensity, and 2-week rainfall duration described the association of the weather conditions with LD incidence remarkably well. The underlying mechanism is biologically plausible with increased growth of *Legionella* bacteria in the environment during the warm weeks and dissemination of the bacteria during the wet weeks.

An overview of the meteorological conditions shows a marked resemblance in the weather conditions in the summers of 2006 and 2010, the two years with the unexplained increased LD incidence. In 2006 July was the warmest for that month in The Netherlands since 1901, while in 2010 July was ranked fifth with a mean temperature of 19·9 °C compared to a normal value of 17·4 °C [16]. The wettest August month in more than a century occurred in 2006,

Table 3. *Incidence rate ratios (IRR) and 95% confidence intervals (CI) obtained by univariate Poisson regression (crude IRR) and IRR from in the final multivariable model after multivariable Poisson regression analysis with loess correction (adjusted IRR)*

		Univariate analysis			Multivariable model with loess correction		
Weather variable	Reference category*	Category with highest IRR	Crude IRR†	95% CI†	Category with highest IRR	Adjusted IRR†	95% CI†
1-week mean temperature	3·3–10·2 °C	4–5 (15·4–19·3 °C)	2·6	1·9–3·4		—	—
2-week mean temperature	4·5–10·6 °C	5 (17·2–18·6 °C).	4·2	3·1–5·6		—	—
4-week mean temperature	6·1–10·5 °C	6 (18·5–22·3 °C)	5·8	4·1–8·3	5 (17·2–18·5 °C)	2·2	1·5–3·2
6-week mean temperature	6·1–10·6 °C	6 (18·3–20·9 °C)	6·8	4·7–9·8		—	—
Weekly precipitation	0–0·3 mm	10 (377–1338 mm)	2·5‡	1·9–3·1		—	—
2-week precipitation	0–0·6 mm	10 (657–1708 mm)	3·8	2·9–5·0		—	—
Weekly rainfall duration	0–0·4 h	9 (21·3–26·7 h)	2·3‡	1·8–3·1	9 (21·3–26·7 h)	1·6	1·2–2·2
2-week rainfall duration	0–0·5 h	9 (38·5–42·2 h)	4·3	3·2–5·6		—	—
1-week rainfall intensity	0−0·35 mm/h	10 (12·6–32·2 mm/h)	2·4	1·8–3·1		—	—
2-week rainfall intensity	0–2·5 mm/h	10 (22·5–41·7 mm/h)	5·0	3·7–6·7	10 (22·5–41·7 mm/h)	2·3	1·6–3·2
1-week mean relative humidity	50·4–68·6%	7 (81·6–83·1%)	3·0	2·1–4·1		—	—
1-week mean atmospheric pressure	997–1009 hPa	2 (1009–1011 hPa)	1·5§	1·2–1·9		—	—
1-week mean cloudiness	0·5–2·7 oktas	3 (3·7–5·1 oktas)	3·1	2·4–4·0		—	—
1-week mean sunshine	226–618 J/cm	6 (1466–1592 J/cm)	2·6‖	2·0–3·5		—	—
Intercept multivariable model			—	—		0·5	0·3–0·8

* The variables for temperature and cloudiness were transformed to six categories, all other variables to 10 categories.

† The IRR and 95% CI for the category with the highest IRR is presented.

‡ A significant decrease was found for weekly precipitation, category 4 (54–77 mm, IRR = 0·7) and for weekly rainfall duration, category 2 (0·5–2·7 h, IRR = 0·7).

§ An increase in IRR for atmospheric pressure, categories 2–4, was followed by a significant decrease for categories 7–10, with the smallest IRR at category 8 (1018–1020 hPa).

‖ A similar result (IRR 2·6, 95% CI 2·0–3·5) was found for sunshine, category 3 (866–1150 J/cm).

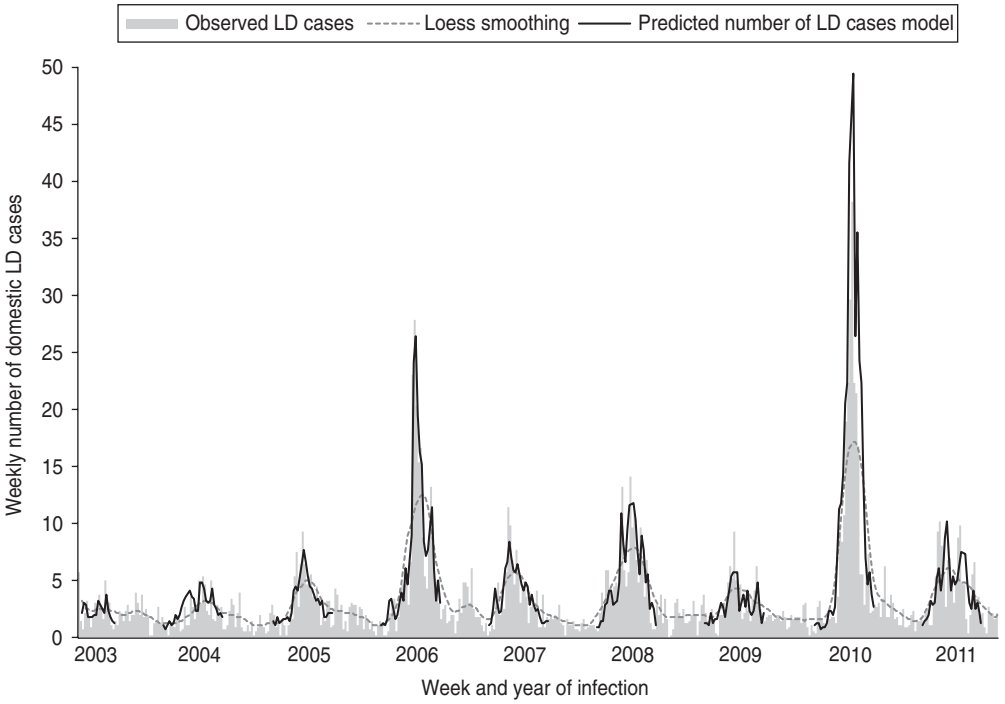

Fig. 4. Actual number of notified domestic, sporadic Legionnaires' disease (LD) cases by year and week of infection, 2003–2011 (grey bars) and predicted number of LD cases for 2003–2011 with the adjusted Poisson regression model, based on the 4-week mean temperature, the 2-week rain intensity, 1-week rain duration (line), and long-term correction (loess, dotted line). For this graph the model was fitted on data from weeks 16–44 (2003–2009) (191 weeks, $n = 769$ cases) and run to predict LD cases in 2010–2011.

followed by August 2010. Extreme rainfall resulting in the flooding of roads occurred on 26 August 2010, and the infection date calculated for the cases in the epidemic group shows a peak that coincides with this day. Summer 2011 also had abundant rainfall, but in that year, no increase in LD cases was observed. However, with an average temperature of 15·9 °C in July and 16·9 °C in August the temperature in summer 2011 remained well below the normal climatic values. Although the weather observations explain a part of the LD incidence, the model still required a loess smoother to accurately predict the incidence. This indicates that there is still variation in the incidence that is not yet accounted for. This is a topic for further research.

A limitation of this study is that the meteorological data was derived from only one weather station, while local differences in weather conditions, particularly the amount of precipitation, can be substantial, even in a small country such as The Netherlands. To develop a more advanced model, the use of data from local weather stations linked to the individual cases could be investigated. We found that the 2-week rainfall variables were better predictors than the 1-week variables. This is possibly due to the design of the model, in which cases are attributed to an infection date 1 or 2 weeks before onset of illness. To achieve an optimal fit for a model based on weather variables, without a loess smoother, a model based on daily weather data may be required. Such a model with daily local weather data should also take into account the complete incubation period for each individual case.

The results of our model are consistent with previous findings. Hicks *et al.* [27] also described a sharp rise in LD cases in the Mid-Atlantic region of the USA in 2003, which coincided with a period of record-breaking rainfall. Using negative binomial regression, Hicks *et al.* found that increased rainfall and higher mean monthly temperature were independently associated with an increase of the legionellosis incidence. The IRR of 1·07 found by Fisman *et al.* [13] for the average monthly temperature is practically identical to the IRR found for in our model; when the variables are included in our model without transformation to categories, the IRR for 4-week temperature is 1·08. Fisman *et al.* also identified an association with increased rainfall (odds ratio 2·48) in the 6–10 days before onset of disease. In that study the LD incidence was not associated with precipitation $\leqslant 5$ days

or >10 days before the occurrence of cases. A Spanish study which compared LD cases with non-*Legionella* pneumonia, also identified increased rainfall as a significant risk factor for sporadic cases of LD [28]. This is relevant information for physicians, who should consider diagnostic testing for *Legionella* pneumonia in patients presenting with community-acquired pneumonia after a period of warm weather followed by extensive rainfall.

While the present study confirms a strong association of LD incidence with weather conditions it remains unclear which environmental sources may be responsible for the increase of LD cases in summer 2010. A comparative analysis of the case characteristics and source-finding results failed to provide an explanation for the increase in the LD incidence during summer 2010. There were no apparent changes in diagnostic methods that may account for increased numbers of cases. Furthermore, the patient population appeared to be the regular patient population for domestic LD cases in The Netherlands, and there were no changes made to the case definition for legionellosis. The finding that 74·6% of patients had not left their MHS region to travel to other regions of the country excluded the possibility of a point-source outbreak to account for the rise in the number of cases. This was confirmed by the clinical isolates which consisted of 14 different strains of *Legionella*. Source-finding investigations also failed to identify multiple common sources that could account for a number of smaller outbreaks. Environmental sampling was done for only 26% of cases, since most cases were sporadic and geographically scattered. *Legionella* was detected for only a third of epidemic cases for which environmental sampling was performed. These source-finding results suggest that we may not be sampling the true or most important sources that are contributing to the summer increase of LD cases.

The current method of source finding is focusing on known water-related sources, such as spa pools and plumbing installations in buildings. This method is usually effective in identifying the source of clusters and point-source outbreaks. However, the success rate for sporadic LD cases is low. If there is an increased growth of *Legionella* in the natural environment, such as in soil or surface water, or an increased exposure to *Legionella* which is present in the natural environment, this will not be detected. Recently, however, *L. pneumophila* has been found in rainwater from puddles on roads and flooded sewers [29, 30].

It is plausible that the warm weather in 2006 and 2010 would favour growth of *Legionella* in the natural environment. The intense rainfall and flooding that followed, may have led to aerosol formation in the environment, causing increased exposure to *Legionella*. Further studies are needed to determine if *Legionella* from the natural environment may actually cause transmission resulting in *Legionella* infections, and which environmental sources contribute to the LD incidence.

Wet cooling towers are established sources of infection for LD outbreaks, and they have also been described as source for sporadic cases of LD [31, 32]. However, their role in a weather-associated increase of LD incidence needs further investigation. During the peak in 2010 only three cooling towers were investigated. Although there has been a legal requirement for registration of wet cooling towers in The Netherlands since January 2010, the current registration is still incomplete, and it is likely that cooling towers are often overlooked in source-finding investigations by municipal health services.

The traditional way of source finding will not help to identify currently unknown environmental sources. However, without a clear understanding of the true sources, we will be unable to prevent these weather-related LD epidemics. Therefore, in addition to the traditional way of source finding, we need to think 'out of the box' and explore other possibilities for source finding, using tools such as spatial analysis and new sampling techniques such as air-sampling.

ACKNOWLEDGEMENTS

The authors thank the staff of the municipal health services for providing epidemiological data on individual cases and source-finding results. We thank Jacqueline Brouwer-De Vries, Jacob Bruin and Wim Houtenbos from the Regional Public Health Laboratory Kennemerland, Haarlem, The Netherlands, for their support with the source-finding database and genotyping. We thank Jan van de Kassteele (Department of Statistics, Mathematical Modelling and Data Logistics of the National Institute for Public Health and the Environment) for his valuable suggestions during the analysis of the weather data.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Vergis EN, Akbas E, Yu VL.** Legionella as a cause of severe pneumonia. *Seminars in Respiratory and Critical Care Medicine* 2000; **21**: 295–304.
2. **Den Boer JW, *et al.*** A large outbreak of Legionnaires' disease at a flower show, the Netherlands, 1999. *Emerging Infectious Diseases* 2002; **8**: 37–43.
3. **Greig JE, *et al.*** An outbreak of Legionnaires'disease at the Melbourne Aquarium, April 2000: Investigation and case-control studies. *Medical Journal of Australia* 2004; **180**: 566–572.
4. **Yu VL, *et al.*** Distribution of Legionella species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: an international collaborative survey. *Journal of Infectious Diseases* 2002; **186**: 127–128.
5. **Bartram J, *et al.*** *Legionella and the Prevention of Legionellosis*. Geneva: World Health Organization, 2007.
6. **Fields BS, Benson RF, Besser RE.** Legionella and Legionnaires' disease: 25 years of investigation. *Clinical Microbiology Reviews* 2002; **15**: 506–26.
7. **Euser SM, *et al.*** Wellness centres: an important but overlooked source of Legionnaires disease. Eight years of source investigation in the Netherlands, 1 August 2002 to 1 August 2010. *Eurosurveillance* 2012; **17**(8) : pii=20097.
8. **Brandsema PS, *et al.*** Annual report surveillance respiratory infectious diseases [in Dutch]. RIVM Report 210231008 (http://www.rivm.nl/en/Library/Scientific/Reports/2011/oktober/Annual_report_surveillance_respiratory_infectious_diseases). Accessed 30 May 2013.
9. **Euser SM, *et al.*** Diagnostic testing for Legionnaires' disease in the Netherlands between 2007 and 2009: a possible cause for the decline in reported Legionnaires' disease patients. *European Journal of Clinical Microbiology & Infectious Diseases* 2012; **31**: 1969–1974.
10. **Swaan CM, *et al.*** National increase of legionellosis notifications in the months August–September 2006 [in Dutch]. *Infectieziekten Bulletin* 2007; **18**: 98–101.
11. **Joseph CA, van der Sande M.** Unexplained summer increase in non-travel-related legionellosis in the UK and Netherlands. *Eurosurveillance* 2006; **11**(42): pii=3065.
12. **Beauté J, Zucs P, de Jong B, on behalf of the European Legionnaires' Disease Surveillance Network.** Legionnaires' disease in Europe, 2009–2010. *Eurosurveillance* 2013; **18**(10): pii=20417.
13. **Fisman DN, *et al.*** It's not the heat, it's the humidity: wet weather increases legionellosis risk in the greater Philadelphia metropolitan area. *Journal of Infectious Diseases* 2005; **192**: 2066–2073.
14. **Ricketts KD, *et al.*** Weather patterns and Legionnaires' disease: a meteorological study. *Epidemiology and Infection* 2009; **137**: 1003–1012.
15. **Karagiannis I, Brandsema P, Van der Sande M.** Warm, wet weather associated with increased Legionnaires' disease incidence in The Netherlands. *Epidemiology and Infection* 2009; **137**: 181–187.
16. **KNMI.** Koninklijk Nederlands Meteorologisch Institute, maand en seizoenoverzichten. (http://www.knmi.nl/klimatologie/maand_en_seizoensoverzichten/). Accessed 22 January 2013.
17. **Den Boer JW, *et al.*** Outbreak detection and secondary prevention of Legionnaires' disease: A national approach. *International Journal of Hygiene and Environmental Health* 2007; **210**: 1–7.
18. **The Netherlands Standardization Institute (NEN).** NPR 6569: 1992 – Bacteriological examination of water – explanatory comments on sampling and preservation according to NEN 6559. Delft: NEN, 1992.
19. **The Netherlands Standardization Institute (NEN).** NEN-EN-ISO 19458: 2007 – Water quality – sampling for microbiological analysis. Delft: NEN, 2007.
20. **Gaia V, *et al.*** Consensus sequence-based scheme for epidemiological typing of clinical and environmental isolates of Legionella pneumophila. *Journal of Clinical Microbiology* 2005; **43**: 2047–2052.
21. **Ratzow S, *et al.*** Addition of neuA, the gene encoding N-acylneuraminate cytidylyl transferase, increases the discriminatory ability of the consensus sequence-based scheme for typing Legionella pneumophila serogroup 1 strains. *Journal of Clinical Microbiology* 2007; **45**: 1965–1968.
22. **Egan JR, *et al.*** Modeling Legionnaires' disease outbreaks: estimating the timing of an aerosolized release using symptom-onset dates. *Epidemiology* 2011; **22**: 188–198.
23. **Cleveland WS, Devlin SJ.** Robust locally-weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 1988; **74**: 829–836.
24. **Akaike H.** A new look at the statistical model identification. System identification and time-series analysis. *IEEE Transactions on Automatic Control* 1974; **19**: 716–723.
25. **Den Boer JW, *et al.*** Genotypic comparison of clinical Legionella isolates and patient-related environmental isolates in The Netherlands, 2002–2006. *Clinical Microbiology and Infection* 2008; **14**: 459–66.
26. **Euser SM, *et al.*** Legionella source detection in the Netherlands 2009–2010 [in Dutch]. *Infectieziekten Bulletin* 2011; **22**, 160–163.
27. **Hicks LA, *et al.*** Increased rainfall is associated with increased risk for legionellosis. *Epidemiology and Infection* 2007; **135**: 811–817.
28. **Garcia-Vidal C, *et al.*** Rainfall is a risk factor for sporadic cases of Legionella pneumophila pneumonia. *PLoS One* 2013; **8**(4): e61036.
29. **Sakamoto R, *et al.*** Legionella pneumophila in rainwater on roads. *Emerging Infectious Diseases* 2009; **15**: 1295–1297.
30 **Schalk JA, *et al.*** Isolation of Legionella pneumophila from pluvial floods by amoebal coculture. *Applied and Environmental Microbiology* 2012; **78**: 4519–4521.
31. **Bhopa RS, *et al.*** Proximity of the home to a cooling tower and risk of non-outbreak legionnaires' disease. *British Medical Journal* 1991; **302**: 378–383.
32. **Ricketts KD, *et al.*** Wet cooling systems as a source of sporadic Legionnaires' disease: a geographical analysis of data for England and Wales, 1996–2006. *Journal of Epidemiology & Community Health* 2012; **7**: 618–623.