

Warm, wet weather associated with increased Legionnaires' disease incidence in The Netherlands

I. KARAGIANNIS^{1,2*}, P. BRANDSEMA¹ AND M. VAN DER SANDE¹

¹ Rijksinstituut voor Volksgezondheid en Milieu (National Institute for Public Health and the Environment – RIVM), Bilthoven, The Netherlands

² European Programme for Intervention Epidemiology Training (EPIET), Stockholm, Sweden

(Accepted 10 June 2008; first published online 17 July 2008)

SUMMARY

It has been suggested that warm and humid weather is related to a high incidence of Legionnaires' disease (LD), but no data on this association existed in The Netherlands. The objective of this study was to investigate the short-term effects of the weather on LD in The Netherlands. National LD surveillance and meteorological data were obtained. We analysed the data using Poisson regression, adjusting for long-term trends, and using principal components analysis. The highest weekly incidence of LD occurred when the mean weekly temperature was +17.5 °C. Mean weekly relative humidity, temperature and precipitation intensity were associated with LD incidence in the multivariable model. Warm, humid and showery summer weather was found to be associated with higher incidence of LD in The Netherlands. These results may be used to predict an increase in the number of cases of LD in The Netherlands during the summer.

Key words: Cluster analysis, Legionnaires' disease, weather.

INTRODUCTION

Legionellosis refers to a number of diseases caused by infection with bacteria of the Legionellaceae family. Although more than 50 *Legionella* spp. have been identified, *Legionella pneumophila* is the cause of most Legionnaires' disease (LD) cases [1]. The disease onset is usually 2–14 days after infection with the bacteria [2] and may cause serious morbidity with pneumonia, fever, nausea, abdominal pain, diarrhoea, vomit and bradycardia. Neurological symptoms, although rare, include headache, lethargy and encephalopathy.

The first described outbreak of LD took place

in 1965 in patients of a psychiatric hospital in Washington, DC, and had a documented case fatality of 17.3%. However, it was not until 1976, after a fourth outbreak of LD had occurred in the United States, that the bacterium was isolated and named [3]. In The Netherlands, the first documented large LD outbreak took place in 1999 after common exposure to a contaminated spa pool at a flower exhibition. There were 188 confirmed and probable *L. pneumophila* pneumonia cases with 21 deaths (case fatality 11.2%) [4]. In 2006, a second outbreak of LD involving 31 cases with three deaths was notified in Amsterdam, after airborne transmission of the bacteria from a contaminated cooling tower [5, 6].

Following this located point source outbreak, there was a considerable rise in sporadic LD cases in late summer that could not be attributed to a common

* Author for correspondence: Mr I. Karagiannis, RIVM – EPI, Postbus 1/pb 75, 3720 BA, Bilthoven, The Netherlands.
(Email: ioannis.karagiannis@rivm.nl)

source or changes in the reporting system. This increase in sporadic cases followed a record hot July and a record wet August. A similar generalized outbreak was observed in the United Kingdom [7, 8], but not in neighbouring Belgium and Germany [9]. Nevertheless, it was hypothesized that ‘something to do with the weather’ could be responsible for this increase in LD cases. This study attempts to investigate short-term effects of the weather on the incidence of sporadic LD cases in The Netherlands, which might contribute to our understanding of the determinants of this unexpected increase in notified cases in 2006. Furthermore, this study may prove useful to predict similar increases in the future and could, therefore, support early warning for the risk of future outbreaks.

MATERIALS AND METHODS

The data used in the present analysis are from the national surveillance database (OSIRIS) and include any person that was notified with LD between 1 July 2003 and 30 September 2007. The infection date was calculated from the date of symptoms onset by deducting 5 days, the median incubation period of LD. All dates were then transformed to weeks and the analysis was based on weekly observations in order to avoid dealing with zero counts and reduce random error. Patients who had travelled outside The Netherlands during the incubation period were not included in this study. The analysis was restricted to the warmest period of the year (April–September) because the fewer LD cases in the cold period could influence the results disproportionately. This period is often quoted as the warmest part of the year in temperate climates [10] and roughly coincides with weeks 18–39. Moreover, the two seasons are characterized by different atmospheric circulation patterns, which is also in favour of a separate analysis; we were interested in understanding what determines the different epidemiological pattern in different summers, rather than increase our understanding on the seasonality of LD. Cases that were identified as being part of a known cluster of LD cases were also excluded from this study; only the first cases in time (‘index cases’) were left in the analysis.

Meteorological data were collected from the weather station of the Royal Dutch Meteorological Institute (KNMI) in De Bilt. The meteorological variables that were included in this analysis were mean weekly temperature, weekly sunshine duration,

mean weekly cloudiness, mean weekly wind speed, weekly precipitation, mean weekly rainfall intensity and mean weekly relative humidity (RH).

Analysis with weather variables

We investigated the associations between the weather and LD incidence by extending Poisson regression generalized linear models (GLMs) with loess smoothers [11, 12]. The extended model allows the inclusion of non-parametric smooth functions to model the potential nonlinear dependence of morbidity from LD on weather variables and the season. The assumption of this model is that

$$\log[E(Y)] = \beta_0 + S(X) + \beta_1(X_1) + \dots + \beta_p(X_p),$$

where Y is the weekly number of LD cases, $E(Y)$ is the expected value of this count, X_i ($i = 1, 2, \dots, p$) are the covariates and S is a smoothing function. To control for long-term seasonality, we used loess smoothing for the LD count data [13]. The equivalent smoothing window was chosen with the help of partial autocorrelation plots and plots of residuals. We had decided in advance that the smoothing window should not be less than 2 months in order to avoid eliminating underlying short-term patterns, as has been suggested by other authors [14, 15]. The smoothing window that minimized partial autocorrelation in the residuals after loess smoothing was 28 weeks.

To allow for the assessment of nonlinear associations between weather variables and LD incidence, we also performed quadratic transformation of the weather variables. The contribution of all variables in the Poisson regression models was assessed with the use of Wald tests.

Analysis with weather types

Principal components analysis (PCA), a factor analysis technique that rewrites the original data matrix into a new set of components that are linearly independent and ordered by the amount of variance they explain was used in order to limit the number of the used variables [16]. To do so, component loadings were calculated, which expressed the correlation between the original variables and the newly formed components. Each week was then expressed by its particular set of ‘component scores’, which are weighted summed values, the magnitudes of which depended on the weather observations for each week and the principal component loading. Thus, weeks

Table 1. Comparison with use of Student's *t* test between 'warm' and 'cold' periods in mean weekly temperature, relative humidity, precipitation, sunshine duration, precipitation intensity, atmospheric pressure at sea level and number of cases with Legionnaires' disease (LD) infected in The Netherlands. Time period: 1 July 2003 to 30 September 2007 ($n=222$ weeks)

	Weeks 18–39	Rest of the year	<i>P</i> value
Mean temperature (°C)	16.3	7.0	<0.0001
Mean relative humidity (%)	77.8	84.6	<0.0001
Mean precipitation (mm)	17.6	15.5	0.9799
Mean sunshine duration (h)	45.5	23.7	<0.0001
Mean precipitation intensity (mm/h)	1.56	0.98	<0.0001
Mean sea-level pressure (hPa)	1015.8	1015.6	0.7025
Mean weekly cases of LD (<i>n</i>)	4.4	2.3	0.0002

Sources: The Royal Dutch Meteorological Institute (KNMI); the national surveillance database (OSIRIS).

with similar meteorological conditions tend to exhibit proximate component scores. At the same time, the use of too many independent variables is avoided; the new components reflect the synergy of the initial variables. We then used a clustering procedure to group weeks with similar component scores into the same categories of weeks that are meteorologically more homogeneous. The method used here was the average linkage method, which is considered the most efficient method in clustering meteorological variables [17]. The mean weekly LD incidence was then calculated for every synoptic (weather) category, along with its standard deviation, to ascertain the distribution of LD incidence by synoptic category.

Missing values management

For six cases, the date of symptom onset was unknown. To estimate that date for those patients, we calculated the median lag between date of symptom onset and date of report in OSIRIS for the rest of the patients. That lag, namely 5 days, was then deducted from the report date of the six patients with unknown symptom onset date. Then, as for the rest of the cases, an extra 5 days were deducted from their estimated date of symptoms onset to calculate the most probable date of exposure and infection.

RESULTS

In total, 707 cases of LD with The Netherlands as the most probable country of infection were notified to RIVM through OSIRIS between 1 July 2003 and 30

September 2007. Of these cases, 432 (61.1%) had been infected during the warm period of the year (weeks 18–39). Table 1 presents the basic characteristics of the weather during the two distinct periods at the KNMI weather station in De Bilt, as well as the weekly number of LD cases in The Netherlands. The warm period is drier and 9.3 °C warmer than the cold period and, even though total precipitation does not differ statistically significantly between the two periods, it tends to fall more intensely in the warm period ($P<0.0001$).

Weather variables

Univariate analysis

The weekly incidence of LD appears to maximize when the mean weekly temperature is +17.5 °C ($P<0.001$). Higher temperatures in the 2 weeks preceding exposure further contribute to a higher LD incidence ($P=0.002$ and $P=0.004$ respectively). An increase in weekly sunshine by 1 h results in an increased LD incidence of 1.8% (95% CI 1.2–2.4). Similarly, weeks with increased cloud cover are related with the highest LD incidence; more cases of LD occur when the cloudiness is 7 oktas, i.e. when the mean weekly cloud cover is seven-eighths. RH is also associated with LD incidence: a 1% increase in RH is associated with an increased LD incidence of 6.4% (95% CI 4.7–8.2). LD occurrence is highest when average weekly precipitation intensity is 3 mm/h ($P=0.004$). Last, LD incidence is highest when weekly precipitation is between 40 mm and 60 mm. Univariate analysis results are given in Table 2.

Table 2. Univariate (crude) and multivariable (adjusted) Poisson regression-obtained incidence rate ratios (IRRs) and respective 95% confidence intervals (CI), adjusted for long-term trends ($n=101$)

Variable	Crude IRR	95% CI	Adjusted IRR	95% CI
Mean relative humidity (%)	1.064	1.047–1.082	1.051	1.029–1.075
Mean sunshine (h)	0.982	0.977–0.988	n.a.	n.a.
Square of cloudiness [(oktas) ²]	1.024	1.017–1.031	n.a.	n.a.
Mean temperature* (°C)	2.302	1.600–3.310	1.647	1.125–2.410
Square of mean temperature [(°C) ²]	0.974	0.964–0.985	0.986	0.975–0.997
Mean weekly precipitation (mm)	1.037	1.025–1.049	n.a.	n.a.
Square of mean weekly precipitation [(mm) ²]	0.9997	0.9996–0.9999	n.a.	n.a.
Mean precipitation intensity (mm/h)	2.025	1.483–2.765	1.148	1.042–1.266
Square of mean precipitation intensity [(mm/h) ²]	0.895	0.831–0.965	n.a.	n.a.

n.a., Not available.

* Mean weekly temperature has a quadratic association with the weekly incidence of Legionnaires' disease in the multivariable model.

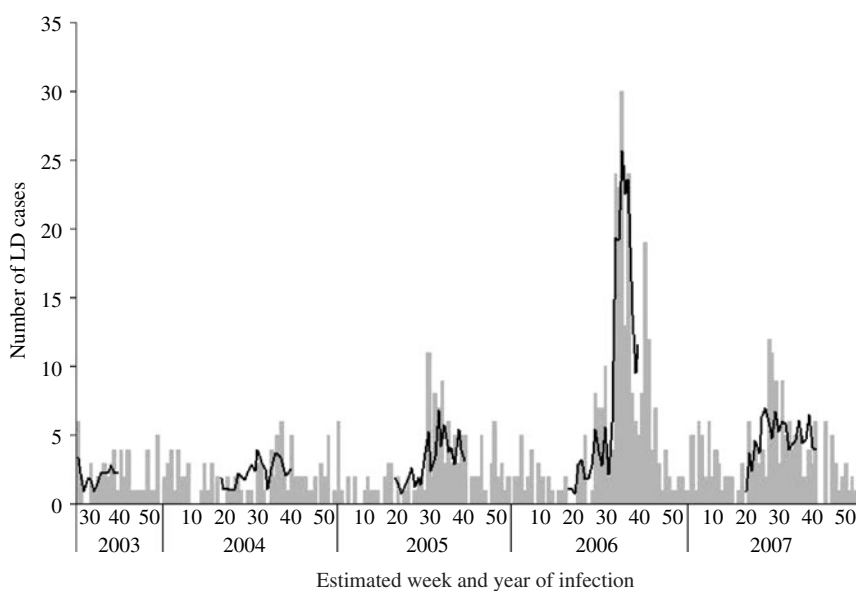


Fig. 1. Actual notified (■) and predicted (—) weekly number of domestic sporadic Legionnaire's disease (LD) cases in The Netherlands, 2003–2007.

Multivariable analysis

The multivariable model suggests that, adjusting for long-term trends, mean weekly precipitation intensity, mean weekly temperature and mean RH contribute to the explanation of LD incidence in The Netherlands. Mean weekly temperature has a quadratic association with LD incidence. An increase in the mean weekly precipitation intensity by 1 mm/h results in an increased LD incidence of 14.8% (95% CI 4.2–26.6), while higher weekly values of mean RH are

also associated with higher LD incidence; a 1% increase in mean weekly RH is associated with a higher LD incidence of 5.1% (CI 95% 2.9–7.5). There was no statistically significant interaction between the variables that were included in the final multivariable analysis model.

The multivariable analysis model, based purely on meteorological data and long-term trends of LD incidence, was able to explain 43.3% of the variance in the epidemiological data. Actual and predicted weekly case counts are presented in Figure 1. The variables

Table 3. Average meteorological characteristics and average weekly number of Legionnaires' disease (LD) cases for the 10 revealed synoptic categories (n = 98 weeks*)

Synoptic category	Average temperature (°C)	Relative humidity (%)	Sunshine duration (h)	Cloudiness (oktas)	Weekly precipitation (mm)	No. of LD cases	No. of weeks per weather category
1	18.8	71.4	58.2	3.74	4.90	2.69	16
2	23.2	68.0	64.9	3.00	2.23	3.50	4
3	16.9	82.2	32.8	6.13	41.4	8.84	19
4	16.5	80.6	48.0	3.85	9.96	4.14	28
5	14.1	78.7	37.5	5.56	21.1	1.50	10
6	11.7	74.5	47.6	5.19	8.78	1.40	10
7	15.6	58.2	79.4	2.14	0	0	2
8	13.3	83.0	22.7	7.05	57.1	4.00	3
9†	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	1
10	14.3	88.0	24.9	5.69	24.3	10.0	5

n.a., Not available.

* Three weeks not categorized due to sparse data.

† Only 1 day was classified in category 9.

contributing to the multivariable analysis model are presented in Table 2.

Weather types

Factor analysis limited the number of used variables to three components that could still explain 76.4% of the variability existing among all weather variables initially used in the analysis. Ten major synoptic weather classifications were identified using the clustering procedure (Table 3).

The synoptic category with the highest LD incidence is category 10; this category includes weeks with a mean temperature close to the overall mean for weeks 18–39 at the De Bilt station. The average RH in that category is 88.0% and the sun shines for 24.9 h/week (an average of 3.6 h/day). On the contrary, weather categories 5 and 6 are linked with the fewest LD cases per week; the weather is then sunnier with 37.5 h and 47.6 h of sunshine per week respectively. Category 7 has an average of zero LD cases per week, but this category was encountered only twice in the sample. The hottest summer weeks, represented in category 2 (mean weekly temperature 23.2 °C) do not coincide with the highest LD incidence in the country.

DISCUSSION

Two major outbreaks of LD have been described in The Netherlands since 1999, increasing interest in this disease. In 2006, a record number of LD cases were notified throughout the country; most LD cases

reported that year were sporadic and could not be assigned to any clearly defined outbreaks.

The results published in this study show a possibly direct role of the natural environment in the epidemiology of *Legionella*. Our time-series methodology provides evidence that LD incidence in The Netherlands is highest when the weather is warm and wet during the summer. Very hot days, although they do not occur too often, do not coincide with the highest incidence of LD throughout the year. Even though disease occurrence can be confounded by factors such as different population behaviour depending on the weather, our study attempts to control for long-term and seasonal patterns, focusing on short-term effects of the weather on LD incidence only.

The two methods used to analyse the data gave comparable results. In the first analysis, where GLMs were used, gloomy and wet summer weather was shown to be associated with the weeks with the highest LD morbidity in The Netherlands. Extensive cloud cover, low sunshine, mean temperatures around 17.5 °C, high RH readings and intense precipitation were independently found to be linked to the highest LD incidences. In the multivariable model, temperature, precipitation intensity and RH could explain 43.3% of total LD incidence variability. The clustering technique, on the other hand, showed that very humid and wet conditions with little sunshine and close to average temperatures are associated with the highest LD incidence.

The use of the PCA analysis, which resulted in several distinct weather types, can prove easier to

interpret compared to the results of the Poisson regression analysis. Mean weekly temperature has, for instance, a quadratic association with the incidence of LD in The Netherlands. Even though this finding allowed us to see that LD incidence is maximized for mean weekly temperatures of 17.5 °C and that the effect of the latter is not linear, incidence rate ratios (IRRs) cannot be interpreted easily. On the other hand, the categorization of periods of time into well-defined weather types – such as, indeed, ‘warm, wet weather’ – can facilitate the understanding of the effect of weather conditions on the incidence of LD.

Some of the weather variables that appear in the univariate analysis results fail to be statistically significant in the multivariable model. This can be explained through collinearity between some of the weather variables; low cloud cover is, for instance, associated with little or no precipitation and high sunshine totals. Rainy days are, on the other hand, more humid than dry ones. Hence, the selection of some variables in the model automatically excludes other variables from being included in the model.

Regarding precipitation and RH, our results show that warm and wet weather patterns, but not the hottest ones, are associated with more LD cases. This finding is consistent with the results found by Fisman *et al.* [16], although the subject was approached with a different methodology. Our findings also correspond to the ecological profile of *Legionella* [18], a principally aquatic microorganism.

The present study has some limitations. Data on the most likely country of infection rely wholly on the answers given by patients, so misclassification is possible to some extent. Moreover, the definition of most likely country of infection used here was the same as in the national surveillance system; that means that patients who had been outside the country during part of the incubation period were classified as having acquired their infection abroad, which may be untrue for some of them. However, random misclassification of this exposure would lessen the strength of the associations we found.

A second limitation of this study is the source of meteorological data, which are derived from the De Bilt weather station. It may be preferable to select data from different stations around the country to allow for better estimation of the individual exposure of LD cases to the weather conditions. However, De Bilt is in the middle of the country and shares some of the maritime climate features

of the west and north of the country and some of the more continental elements of the climate in the south and east of the country; for these reasons, it can be considered representative of The Netherlands [19].

Even though disease occurrence can be confounded by factors such as different population behaviour depending on the weather, our study attempts to control for long-term and seasonal patterns, focusing on short-term effects of the weather on LD incidence only.

We have deliberately chosen to exclude from our analysis all LD cases that were part of a cluster, except for the index cases, i.e. the cases with the earliest symptom onset date per cluster. Each cluster is seen as an independent incident of human exposure to *L. pneumophila*. However, the purpose of the present study was to explore the influence of the weather on the transmission of the pathogenic organism from the environment to the human population. The investigation of how weather influences the transmission dynamics within a cluster of LD cases was beyond our study objectives.

The results of the present study come to some agreement with the preliminary results of Ricketts *et al.*, who, through a case-crossover approach, suggest that there could be an association between the incidence of LD and RH [20]. Understanding the influence of weather on the incidence of LD could help clarify the underlying mechanism that resulted in such an increased number of sporadic cases of LD in The Netherlands in 2006 and could help predict the impact of new outbreaks. Blatny *et al.* suggested that the highest concentrations of *L. pneumophila* in air samples close to a known source in Norway were measured during cloudy and not very hot days [21].

Specific weather variables can be used to better understand the underlying association between the weather patterns and the incidence of LD. The possible transition towards a warmer climate with fewer days with precipitation but with more abrupt changes and heavier rainfall, as described in other science fields, could mean that the epidemiological profile of some diseases are affected by the weather changes [22]. However, the underlying biological mechanisms of the associations between weather and LD incidence still remain unexplained. Additional research in this field could help provide more evidence for the biological plausibility of ‘warm, wet weather’ being associated with more LD cases.

ACKNOWLEDGEMENTS

The authors thank Barbara Schimmer and Susan Hahné, epidemiologists at the Department of Epidemiology and Surveillance of the National Institute for Public Health and the Environment (RIVM), and Esther Kissling, EPIET fellow at the Institute of Public Health in Brussels, Belgium, for their support and provision of useful comments throughout the study.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Diederens BMW**. *Legionella* spp. and legionnaires' disease. *Journal of Infection* 2008; **56**: 1–12.
2. **Coordinating Center for Infectious Diseases/Division of Bacterial and Mycotic Diseases, CDC**. Legionellosis: Legionnaires' disease (LD) and pontiac fever (http://www.cdc.gov/ncidod/dbmd/diseaseinfo/legionellosis_g.htm). Accessed 12 October 2005.
3. **Thackers SB, et al.** An outbreak in 1965 of severe respiratory illness caused by the Legionnaires' disease bacterium. *Journal of Infectious Diseases* 1978; **138**: 512–519.
4. **Boer JW den, et al.** A large outbreak of Legionnaires' disease at a flower show, the Netherlands, 1999. *Emerging Infectious Diseases* 2002; **8**: 37–43.
5. **v.d. Hoek JAR, Ijzerman EPF, Coutinho R**. Legionella outbreak in Amsterdam: cooling tower as source. *Nederlandse Tijdschrift voor Geneeskunde* 2006; **150**: 1808–1811.
6. **v.d. Sande M**. Cooling towers as a source of *Legionella pneumophila* infections in The Netherlands. *Infectieziektenbulletin* 2008; **19**: 13–17.
7. **HPA**. National increase in cases of legionnaires' disease. *Communicable Diseases Report Weekly* 2006; **16**: news (<http://www.hpa.org.uk/cdr/archives/2006/cdr3706.pdf>).
8. **Joseph C, v.d. Sande M**. Unexplained summer increase in non-travel related legionellosis in the UK and Netherlands. *Eurosurveillance* 2006; **11**: E061018.1 (<http://www.eurosurveillance.org/ew/2006/061018.asp#1>).
9. **Swaan CM, et al.** Increase of Legionnaires' disease notifications in The Netherlands, August–September 2006. *Infectieziektenbulletin* 2007; **18**: 98–101.
10. **Michelozzi P, et al.** Assessment and prevention of acute health effects of weather conditions in Europe, the PHEWE project: background, objectives, design. *Environmental Health* 2007; **6**: 12.
11. **Hastie T, Tibshirani R**. *Generalized Additive Models*, 1st edn. London: Chapman and Hall, 1990, pp. 29–31.
12. **Schwartz J**. Generalized additive models in epidemiology. In: International Biometric Society, Invited Papers. *17th International Biometric Conference 1994*. Hamilton, Ontario, Canada: International Biometrics Society, 1994, pp. 55–80.
13. **Cleveland WS, Devlin SJ**. Robust locally-weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 1988; **74**: 829–836.
14. **Kassomenos PA, et al.** Atmospheric circulation types and daily mortality in Athens, Greece. *Environmental Health Perspectives* 2001; **109**: 591–596.
15. **Kassomenos PA, Gryparis A, Katsouyanni K**. On the association between daily mortality and air mass types in Athens, Greece during winter and summer. *International Journal of Biometeorology* 2007; **51**: 315–322.
16. **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.
17. **Kalkstein LS, Tan G, Skindlov JA**. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Climatology and Applied Meteorology* 1987; **26**: 717–730.
18. **Stout JE, Yu VL, Best MG**. Ecology of *Legionella pneumophila* within water distribution systems. *Applied and Environmental Microbiology* 1985; **49**: 221–228.
19. **Verbeek K**. The state of climate in The Netherlands. KNMI, 2003 [in Dutch] (http://www.knmi.nl/kennis/centrum/de_toestand_van_het_klimaat_in_Nederland_2003).
20. **Ricketts K, et al.** Meteorological study of Legionnaires' Disease cases in England and Wales. *22nd Meeting of the European Working Group for Legionella Infections (EWGLI)*, 2–5 June 2007, Stockholm and Uppsala, Sweden.
21. **Blatny JM, et al.** Air Sampling near the source of a large outbreak in Norway. In: *Proceedings of the 4th National Legionella Symposium*. Haarlem: The Netherlands, 2007, pp. 8–9.
22. **Campbell-Lendrum D, Corvalán C, Neira M**. Global climate change: implications for international public health policy. *Bulletin of the World Health Organization* 2007; **85**: 235–237.