Thesis for doctoral degree (Ph.D.)
2019

Epidemiology, prevention and control of Legionnaires' disease in Europe

Julien Beauté

Karolinska Institutet

From Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

EPIDEMIOLOGY, PREVENTION AND CONTROL OF LEGIONNAIRES' DISEASE IN EUROPE

Julien Beauté

Karolinska Institutet

Stockholm 2019

Published by Karolinska Institutet.
Printed by Arkitektkopia AB

ISBN 978-91-7831-603-8

Epidemiology, prevention and control of Legionnaires' disease in Europe

THESIS FOR DOCTORAL DEGREE (Ph.D.)

The dissertation will take place in the Lecture Hall Atrium, Nobels väg 12B, Campus Solna, Karolinska Institutet.

Wednesday, December 11, 2019 at 9:00

By

Julien Beauté

Principal Supervisor:

Prof. Pär Sparén
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Co-supervisor(s):

Prof. Johan Giesecke
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Sven Sandin
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Opponent:

Prof. Preben Aavitsland
University of Oslo

Examination Board:

Sofia Carlsson
Karolinska Institutet
Institute of Environmental Medicine

Hans Fredlund
Örebro University
Department of Laboratory Medicine

Nicola Orsini
Karolinska Institutet
Department of Public Health Sciences

“On n’explique qu’en comparant”.
– Emile Durkheim

ABSTRACT

Legionnaires' disease (LD) is a water-borne infection cause by Gram-negative bacteria *Legionella* spp. with virtually no person-to-person transmission. The clinical presentation is a severe pneumonia with a case fatality of approximately 10%. Known risk factors include increasing age, chronic lung disease and various conditions associated with immunodeficiency. Most cases are community-acquired and sporadic. LD is notifiable in the European Union (EU) and European Economic Area (EEA). LD incidence is thought to be increasing in Europe and the USA for reasons not fully understood, including climate change, changing demographics and improved surveillance. The overarching aim of this thesis was to explore various aspects of LD epidemiology, prevention and control using surveillance data.

In study I, we retrieved travel-associated Legionnaire's disease (TALD) surveillance data for 2009 from the European Surveillance System, and tourism denominator data from the Statistical Office of the European Union. We estimated the risk for TALD in several European countries and highlighted potential under-ascertainment of LD in some countries.

To confirm and generalize findings of studies performed at regional or national level, we investigated the effect of temperature, rainfall, and atmospheric pressure on short-term variations in LD notification rate in Denmark, Germany, Italy, and the Netherlands in Study II. We fitted Poisson regression models to estimate the association between meteorological variables and the weekly number of community-acquired LD cases. We found that the higher risk was associated with simultaneous increase in temperature and rainfall. These findings contributed to the growing evidence supporting a possible impact of climate change on LD incidence.

In Study III, we investigate the actors associated with LD recurrence in hotels. We conducted a retrospective cohort analysis and use survival analysis methods to estimate the association between hotels characteristics and the occurrence of a further case. We found that hotel size and previous association with multiple cases were predictors of the occurrence of a further case. This study also highlighted weaknesses in data collected in the surveillance scheme.

In Study IV, we used a large sample of LD over a 10-year period to look more closely at healthcare-associated (HCA) LD. We found that HCA LD cases are responsible for a major part of LD and differ from community-acquired cases in many aspects, including demographics, causative strains and outcome.

Taken together, the findings support the use of surveillance data for research purposes. They shed light on some epidemiological aspects of LD and inform the surveillance system for possible improvements.

LIST OF SCIENTIFIC PAPERS

I. Beauté J, Zucs P, de Jong B. Risk for travel-associated Legionnaires' disease, Europe, 2009. Emerg Infect Dis. 2012 Nov;18(11):1811-6.

II. Beauté J, Sandin S, Uldum SA, Rota MC, Brandsema P, Giesecke J, et al. Short-term effects of atmospheric pressure, temperature, and rainfall on notification rate of community-acquired Legionnaires' disease in four European countries. Epidemiol Infect. 2016 Aug 30:1-11.

III. Beauté J, Sandin S, de Jong B, Hallström LP, Robesyn E, Giesecke J, et al. Factors associated with Legionnaires' disease recurrence in hotel and holiday rental accommodation sites. Euro Surveill. 2019;24(20).

IV. Beauté J, Plachouras D, Sandin S, Giesecke J, Sparén P. Healthcare-associated Legionnaires' disease in Europe, 2008-2017 (Submitted)

RELATED PUBLICATIONS

Beauté J, Zucs P, de Jong B, European Legionnaires' Disease Surveillance Network. Legionnaires' disease in Europe, 2009-2010. Euro Surveill. 2013 Mar 07;18(10):20417.

Beauté J, Robesyn E, de Jong B. Legionnaires' disease in Europe: all quiet on the eastern front? Eur Respir J. 2013 Dec;42(6):1454-8.

Beauté J, European Legionnaires' Disease Surveillance Network. Legionnaires' disease in Europe, 2011 to 2015. Euro Surveill. 2017 Jul 06;22(27).

CONTENTS

LIST OF ABBREVIATIONS

AIC	Akaike's Information Criterion
ASR	Age-standardized rate
CI	Confidence interval
CRP	C-reactive protein
DALYs	Disability-adjusted life years
DFA	Direct fluorescent antibody
ECA	European Climate Assessment
ECDC	European Centre for Disease Prevention and Control
EEA	European Economic Area
ELDSNet	European Legionnaires' Disease Surveillance Network
EPIS	Epidemic Intelligence Information System
EU	European Union
GIS	Geographic information systems
HCA LD	Healthcare-associated Legionnaires' disease
HIV	Human immunodeficiency virus
HR	Hazard ratio
IATA	International Air Transport Association
ICU	Intensive care unit
LD	Legionnaires' disease
MAb	Monoclonal antibodies
NUTS	Nomenclature of territorial units for statistics
OR	Odds ratio
PCR	Polymerase chain reaction
RR	Relative risk
SIM	Subscriber identity module
TALD	Travel-associated Legionnaires' disease
TESSy	The European Surveillance System
UAT	Urinary antigen test
UK	United Kingdom
USA	United States of America
VIF	Variance inflation factor
WGS	Whole-genome sequencing
WHO	World health Organization
WSP	Water safety plan
YLL	Years of life lost

1 INTRODUCTION

Legionnaires' disease (LD) is a severe pneumonia caused by *Legionella* species (spp). These Gram-negative bacteria can be found in freshwater environments worldwide and often contaminate man-made water systems (1). The first description of the disease and its name came after a large outbreak among members of the American Legion in 1976 (2). People are infected by inhalation or less frequently by aspiration of aerosols containing *Legionella*, most commonly *L. pneumophila* serogroup 1 (1). Nonetheless, some other species may be involved as suggested by the reported association between handling potting soil and infection with *L. longbeachae* (3).

Although not common in outpatients, LD is one of the most common causative factor in community-acquired pneumonia admitted to intensive care units (ICU) (4). In Europe, 5000 to 7000 LD cases are reported each year, of which approximately 10% die. A limited number of countries account for most cases (5). Since 2011, the average notification rate increased from 0.97 to 1.2 LD cases per 100,000 population but masked important differences across countries. In many central and eastern European countries, notification rates were below 1 per million population, unlikely to reflect the local risk for LD. This could probably be explained by poor awareness among clinicians, limited diagnostic capability or capacity and low reporting (6). Approximately 70% of all reported cases are community-acquired, 20% travel-associated and 10% healthcare-related (5). Known risk factors were male sex, increasing age and various conditions associated with immunodeficiency (7).

Since the first description of the disease, most reported cases were sporadic but large outbreaks continued to occur. In 2001, 449 confirmed cases were reported in Murcia (Spain) in relation with a cooling tower (8). It is to date the largest outbreak ever reported. More recently, a large community outbreak occurred in Vila Franca de Xira near Lisbon, Portugal in 2014 (9). With nearly 400 cases, it was one of the largest outbreaks ever observed in Europe. The investigation identified industrial wet cooling systems to be the probable source of infection.

Outside of Europe, epidemiological information is also mostly provided by surveillance data. Surveillance schemes for LD are in place in North America (Canada and United States of America (USA)) and in other developed countries such as Australia or Japan but limited data are available from other parts of the world (7). In countries with available data, the main demographics of LD are similar to those observed in Europe (10).

For years, much of the attention focused on travel-associated cases (TALD) clusters and large community-acquired outbreaks because their detection prompted immediate control measures. Nonetheless, sources of the infection were seldom ascertained and outbreak investigation remains very challenging. In addition, relatively little is known on sporadic community-acquired or healthcare-associated cases. Environmental risk factors, especially of sporadic community-acquired LD remain poorly understood.

2 BACKGROUND

2.1 Microbiology

Legionellae are aerobic, Gram-negative, non-spore-forming gammaproteobacteria (11). Most human infection are caused by *L. pneumophila* serogroup 1 but numerous other species have been isolated in the environment and many of them are pathogenic in humans (12). There are 58 species in the genus *Legionella*, of which approximately 30 can cause infection in humans (13). However, it seems that most infections caused by *Legionella* species other than *L. pneumophila* occur in immunocompromised patients.

There is evidence that *Legionella* can virtually be found in any water environment whether natural or altered (1). Conversely, *Legionella* does not survive in dry environments. Water temperature plays an important role in in *Legionella* bacterial development. Katz et al showed that *L. pneumophila* multiplies at temperatures between 25 and 42°C (14). Under certain conditions, it may even be possible for mutant strains to grow below 20°C (15). At lower temperatures, *Legionella* will survive without multiplying (16). The alteration of aquatic environments by temperature could modify the balance between protozoa and bacteria, favoring the growth of *Legionella* (1).

In the environment, *Legionella* can be associated with complex biofilms or other microorganisms such as amoebae (1, 11). These associations can provide protection against extreme conditions, such as high or low temperature or the presence of chemical agents active against *Legionella* (e.g. chlorine).

Previous studies have suggested an impact of environmental conditions on LD incidence. Contributions of temperature, humidity or precipitation have been reported in studies with different methodologies and settings (17-24). Conclusions were at times divergent and the real impact of climate on LD incidence remains to be validated. Theoretically, any weather condition favoring the growth *Legionella* spp. or its presence in aerosols could potentially be associated with a higher LD incidence. Since climate change is expected to bring both an increase in heavy rainfall and higher temperatures, it is important to better understand the impact of weather on LD incidence (25).

2.2 Transmission

People are infected by inhaling aerosols contaminated by *Legionella*. Person-to-person transmission has been described only once (26). In most cases, contaminated aerosols contain *L. pneumophila* serogroup 1 (1) . Other species such as *L. longbeachae* are thought to infect people through other routes although yet not fully understood. Exposure to potting soils or compost, poor hand-washing after gardening activities may be associated with LD caused by *L. longbeachae* (27).

Several potential sources of infection have been identified, including cooling towers, hot and cold-water systems, spa and thermal pools, springs, humidifiers, domestic plumbing, sewage, potting mixes, and compost (16). A study has even suggested that the use of windscreen wiper fluid without added screenwash in motor vehicles could be a risk factor for LD (28). Other unusual transmission routes have also been reported. A humidifier filled with tap water caused the infection in an infant aged below 6 months (29).

2.3 Clinical presentation

Two forms of infection with *Legionella* are classically described under the term of Legionellosis. The pneumonic form is the Legionnaires' disease whilst Pontiac fever is a milder form of the infection without pneumonia, usually described as a self-limited influenza-like illness. Both incubation and duration of the disease are shorter for Pontiac fever (7). Pontiac fever is not notifiable in Europe and most cases are diagnosed during outbreaks when mild or asymptomatic cases are investigated (30).

The incubation period of LD is thought to be 2-10 days with a median of 7 days (7). However, shorter and longer incubations have been reported in outbreak reports. Thus, during the large outbreak that affected visitors of a flower show in the Netherlands in 1999, incubation periods ranged from 2 to 19 days (31).

LD is a severe pneumonia and its clinical and radiographic presentations are very difficult to distinguish from pneumonia caused by more common pathogens such as *Streptococcus pneumonia* (32, 33). It usually starts with a prodromal illness that may include unspecific symptoms such as headache, myalgia, asthenia, and anorexia. Respiratory symptoms may include cough, dyspnea, and chest pain. Cough does not systematically produce purulent sputum, which is a useful material for laboratory confirmation of the infection. Gastrointestinal and neurological symptoms are not uncommon (13). Other systemic disorders are common such as impaired renal and liver functions. Atypical presentations have been reported including cases with complete absence of respiratory symptoms (34).

2.4 Diagnosis and treatment

2.4.1 Diagnosis

The identification of the causative agent of LD – *L. pneumophila* – during the historical outbreak that stroke members of the American Legion was done by detecting specific antibodies in the serum of patients (35). Alongside culture, serology has been the main laboratory test used for diagnosis of LD in the early years of its history (11). In the past two decades, the urinary antigen test (UAT) has become the most used diagnostic test for LD. In Europe, it is approximately 80% of LD cases that are diagnosed with UAT (5). Some large reporting countries

such as Italy or Spain rely almost exclusively on UAT with 90-95% of their LD cases reported with UAT. The main limitation of UAT is that it only captures *L. pneumophila* serogroup 1. It means that infections caused by other species may remain underdiagnosed. However, the large diffusion of a convenient test such as UAT can also be beneficial. Thus, since its introduction in 1996 in Catalonia, Spain, community outbreaks were detected earlier and the case fatality decreased (36).

The isolation of *Legionella* spp. from respiratory secretions or any normally sterile site (i.e. culture) remains the gold standard. Before the era of PCR, culture was the sole method that would allow for matching clinical and environmental isolates. In recent years, only 10% of cases reported in Europe were culture-confirmed (5). This overall proportion masked important differences across countries. In 2015, some countries did not report any culture confirmations while 41% of LD cases reported by Denmark were culture-confirmed (37).

In Europe, an increasing number of LD cases have been reported with a diagnosis made by polymerase chain reaction (PCR). In 2015, the proportion of PCR ascertained LD cases was over 75% in Denmark (37). A study in New Zealand has suggested that the routine use of PCR had improved the detection of LD cases caused by *Legionella* spp. (mainly *L. longbeachae*) (38).

Other laboratory tests used in Europe include the detection of *L. pneumophila* antigen in respiratory secretions or lung tissue and serological methods. The use of these methods is now declining and becoming increasingly marginal (5). Figure 1 summarizes the type of specimens and diagnostic tests that can be used for detecting *Legionella* infections.

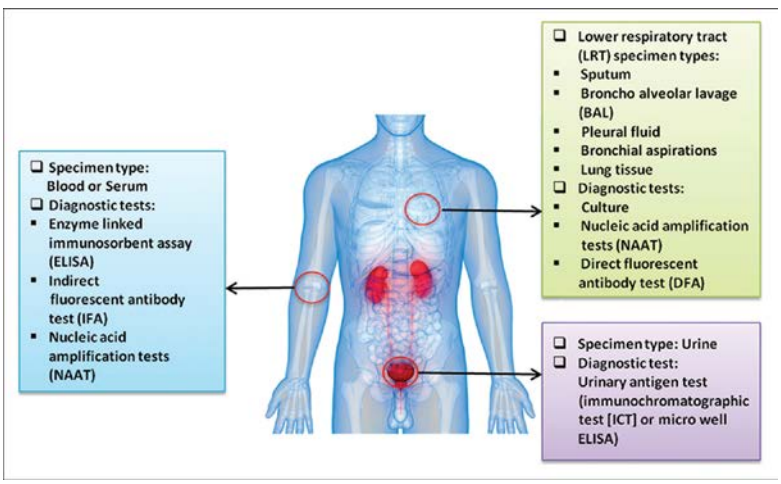

Figure 1. *Type of specimens and diagnostic tests for detecting Legionella infections. Source: Chaudhry (2018) (39)*

2.4.2 Treatment

Antibiotics with good intracellular action are effective against any form of legionellosis. Azithromycin (macrolide) and levofloxacin (quinolone) are the best first-line option (40, 41). β-lactam antibiotics are not active against *Legionella*. This is of importance since β-lactam antibiotics are usually the first-line treatment for community-acquired pneumonia. So far, very few *Legionella* strains have been reported with a reduced susceptibility to antibiotics with intracellular activity. One clinical isolate resistant to ciprofloxacin has been isolated in a patient with severe pneumonia (42). A recent meta-analysis comparing quinolones with macrolides suggested that patients receiving quinolones had a lower morality rate and shorter hospital stay (43).

2.5 Epidemiology

The exact incidence of LD is unknown. Most of the available epidemiological information on LD comes from surveillance data or outbreak investigations. It is estimated that approximately 5% of community-acquired pneumonia could be caused by *Legionella* (44). This proportion could even be higher in Europe. A European study looking at the impact of infectious diseases on population health using incidence-based disability-adjusted life years (DALYs) found that LD had the fifth highest burden after influenza, tuberculosis, human immunodeficiency virus (HIV) infection/AIDS and invasive pneumococcal disease (45) (Figure 2). Almost all DALYs associated with LD were due to years of life lost due to premature mortality (YLL).

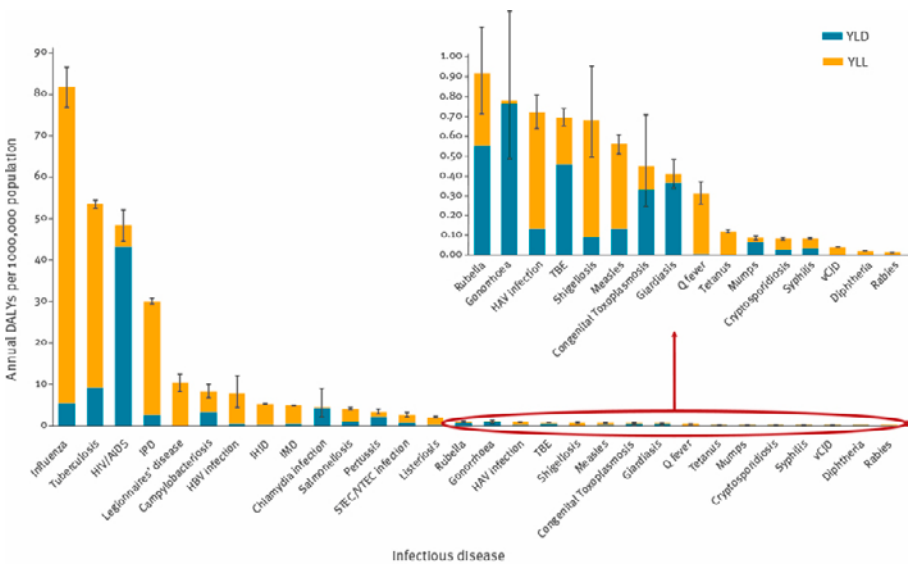

Figure 2. *Median annual DALYs per 100,000 population for selected infectious diseases, EU/EEA countries, 2009–2013. Source: Cassini, A. 2018 (45)*

EU/EEA: European Union/European Economic Area; HAV: Hepatitis A virus; HBV: Hepatitis B virus; HIV/AIDS: Human immunodeficiency virus infection; IHID: Invasive Haemophilus influenzae disease; IMD: Invasive meningococcal disease; IPD: Invasive pneumococcal disease; STEC/VTEC: Shiga toxin/verocytotoxin-producing Escherichia coli; TBE: Tick-borne encephalitis; vCJD: variant Creutzfeldt–Jakob disease; YLD: years lived with disability; YLL: years of life lost due to premature mortality. The error bars indicate the 95% uncertainty intervals.

Over the 2011–15 period, the age-standardized rate (ASR) of LD ranged 0.01 cases per 100,000 population in Bulgaria and Romania to 3.46 cases per 100,000 population in Slovenia (5). Most central and eastern European countries had ASR below 0.5 cases per 100,000 population (Figure 3).

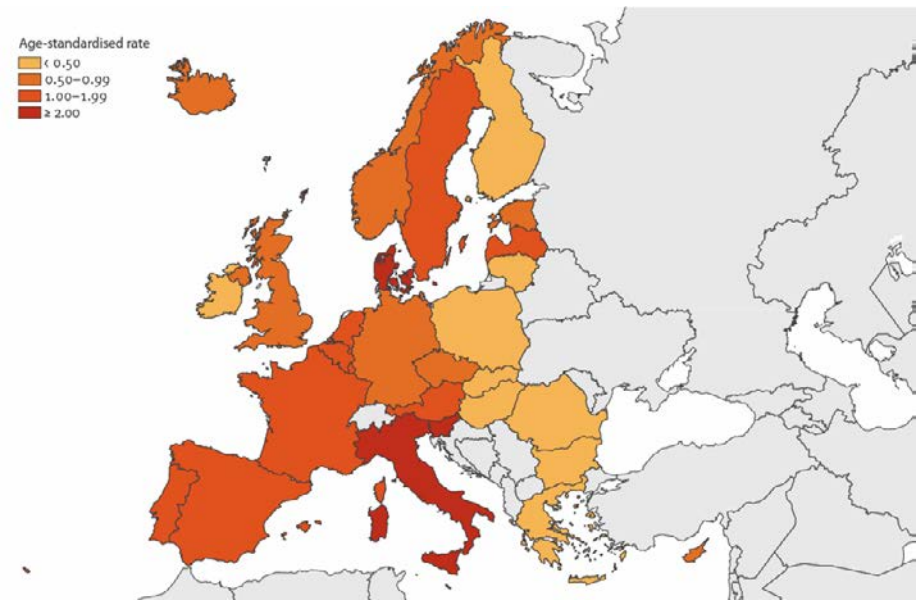

Figure 3. *Age-standardized rate of Legionnaires' disease per 100,000 population by country, European Union/European Economic Area, 2011–2015. Source: Beauté, J. 2017 (5)*

The overall notification rate for the EU/EEA continued to increase in the following years from 1.3 per 100 000 population in 2015 to 1.8 per 100 000 population in 2017 (46). The most recent data available suggest that the increase continued in 2018 (Figure 4). Comparable rates were reported in the USA (10, 47). However, a recent study based on hospitalization data carried out in Connecticut, USA suggested a substantial underdiagnosis with an estimated rate above 10 cases per 100 000 population (48).

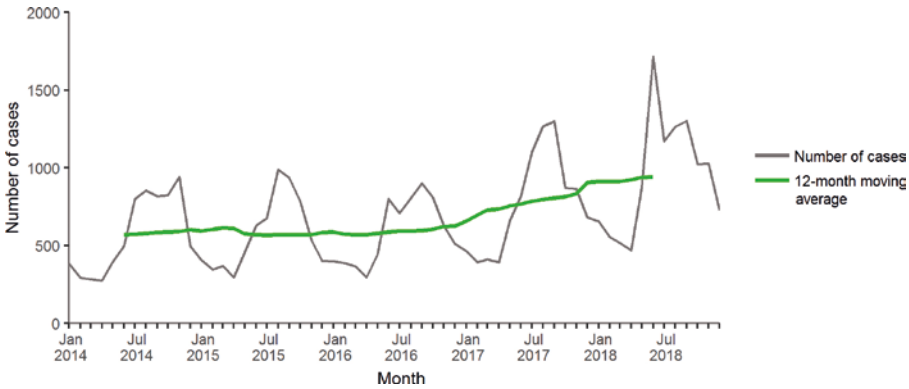

Figure 4. *Distribution of Legionnaires' disease cases by month, EU/EEA, 2014–2018*
Source: Country reports from Austria, Belgium, Bulgaria, the Czech Republic, Cyprus, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.

2.5.1 Demographics

Globally, demographics are quite similar across countries (7). The disease is rare in children and most cases occur in adults with a median age at date of onset between 60 and 65 years. Notifications rates increase with age and approximately 80% of reported cases occurred in people older than 50 years. LD is more common in males and the male-to-female ratio is approximately 2.5:1 (37).

2.5.2 Risk factors

Known risk factors for LD include increasing age, male sex, smoking, chronic lung disease, diabetes, and various conditions associated with immunodeficiency (49, 50). A recent population-based study carried out in the USA identified 12 clinical conditions associated with an increased risk of LD (51). In addition to previously known risk factors such as chronic lung disease, this study suggested that other clinical factors could play a role, including cardiovascular disease and neurological disease. that In addition, poverty and certain occupations may also be associated with a higher risk for LD (52). Thus, in a study carried out in the United States, working in transportation, repair, protective services, cleaning, or construction was associated with a higher risk for community-acquired LD.

2.5.3 Seasonality

In North America and in Europe, the monthly distribution of LD cases shows a clearly seasonality with most cases reported during the warm season (10, 37). For example, in 2015 approximately 60% of all cases reported in Europe had a date

of onset between June and October (7). In the USA the seasonality seemed to be less pronounced in states with mild climates (south and west census regions) (Figure 5). Data from Japan, South Korea and Taiwan suggested a similar pattern in other parts of the world (53, 54).

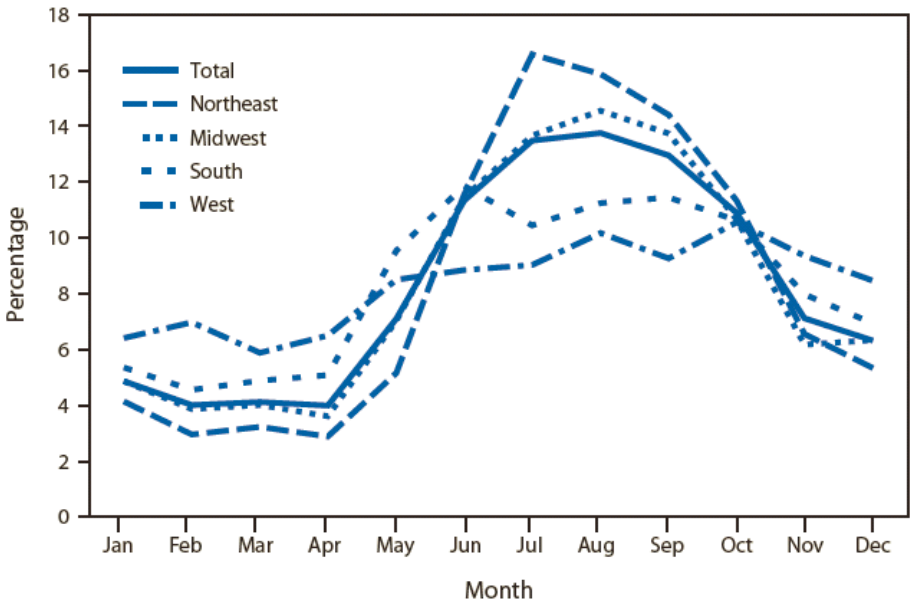

Figure 5. *Annual average percentage of legionellosis cases occurring annually, by month and U.S. Census region* – United States, 2000–2009. Source: Centers for Disease Control Prevention (10)*

** Northeast: Connecticut, Maine, Massachusetts, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania; Midwest: Indiana, Illinois, Michigan, Ohio, Iowa, Nebraska, Kansas, North Dakota, Minnesota, and Missouri; South: Delaware, District of Columbia, Florida, South Carolina, West Virginia, Kentucky, Louisiana, Oklahoma, and Texas; West: Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming, Alaska, California, Hawaii, Oregon, and Washington.*

2.5.4 Outcome

In Europe, the case fatality is approximately 10% but is usually higher in older age groups (5). When adjusting for age and sex, healthcare-associated cases were significantly associated with a higher risk for fatal outcome compared to other settings of infection. A study in France suggested that female sex, age, admission to ICU, renal failure, corticosteroid treatment and increased level of C-reactive protein (CRP) were associated with a higher mortality (55).

2.5.5 Outbreaks

LD was first described after an outbreak during a convention of the American Legion in 1976 (2). Although the source of the outbreak could never be confirmed (visiting the hotel lobby was a risk factor), the epidemiological curve is still presented in textbooks as a typical example of a point-source outbreak (56).

In Europe, the vast majority of cases is thought to be sporadic (>90%) (5). However, in the absence of sound cluster definition for most settings of infection, it is difficult to estimate the exact proportion of cases associated with the same probable source of infection. Clusters of travel-associated cases and healthcare-associated cases are probably easier to identify although the source of infection is seldom identified. Large outbreaks may be associated with specific morbidity. Thus, impaired health-related quality of life and posttraumatic stress disorder have been reported among survivors of LD outbreaks (57). Outbreaks attract a lot of media attention and can trigger changes in health policy (58).

During LD outbreaks, the localization and removal of the source of infection is essential to prevent further cases. In some outbreaks, the probable source of infection is easily identified because most cases stayed or visited in the same location. In large outbreaks, cooling towers are often identified as the source of infection (Table 1).

Table 1. Selection of large outbreaks of LD (>100 cases), 1976–2018. Adapted from Phin (2014) (7)

Place	Year	Number of cases	Case fatality	Source	Ref.
Philadelphia, USA	1976	182	16%	Not confirmed	(2)
Los Angeles, USA	1977–82	>200	-	Potable water	(59)
Bovenkaspel, Netherlands	1999	188	11%	Whirlpool spa	(31)
Melbourne, Australia	2000	125	3%	Cooling tower	(60)
Murcia, Spain	2001	449	1%	Cooling tower	(8)
Barrow-in-Furness, UK	2002	197	4%	Cooling tower	(61)
Miyazaki, Japan	2002	295*	2%	Public bathhouse	(62)
Sarpsborg, Norway	2005	103	10%	Industrial air scrubber	(63)
Pamplona, Spain	2006	146	0%	Cooling tower	(64)
Quebec, Canada	2012	182	8%	Cooling tower	(65)
Vila Franca de Xira, Portugal	2014	334	4%	Cooling tower	(9)
New York, USA	2015	138	12%	Cooling tower	(66)

* including suspected cases.

However, the source of an outbreak, especially a community outbreak, remains sometimes undetected. Classical approaches for outbreak investigation (case-control studies) are often unfruitful. The analysis usually focuses on the interaction of cases with their environment. Although the place of residence is easily obtained, the collection of detailed data on case movements during the potential exposure period can be challenging. In such cases, the investigation can benefit from other tools such as geographic information systems (GIS) (67). GIS are tools that collect, analyze and display data with any geographical component. The use of GIS during an LD outbreak may help identify spatial patterns in relation with a common source of infection (Figure 6). The analysis can include not only case data but also potential sources locations (e.g. cooling towers), demographic data, and meteorological data. Thus, there are example of successful investigations that identified the source by simulating the dispersion of aerosols emitted from a number of potential sources of infection (68). If basic spatial information for cases is usually easily retrieved (home location, place of work, places visited etc.), the collection of detailed travel routes or places of shorter stay during the potential exposure period can be very challenging with traditional questionnaire techniques. Data of subscriber identity module (SIM) cards from mobile phones have recently be used with success during a cholera outbreak to track population movements (69). A similar methodology could yield promising results in LD outbreak investigation.

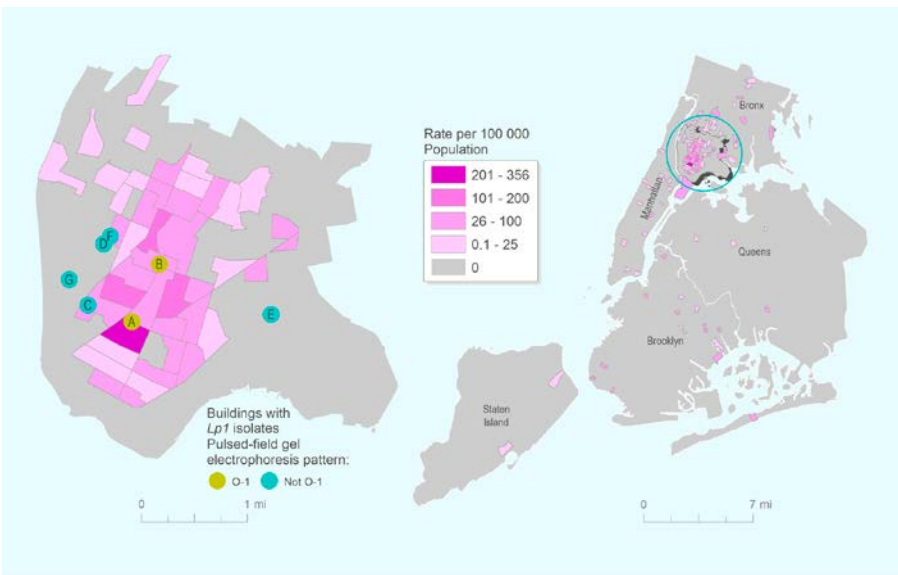

Figure 6. *Crude attack rates of Legionnaires' disease by census tract and cooling towers testing positive for Legionella pneumophila serogroup 1 (Lp1), Bronx, New York City, July 2 to August 3, 2015. Source: Weiss (2017) (71)*

Whole-genome sequencing (WGS) is increasing used in outbreak investigation. Theoretically, WGS could help matching isolates from clinical and environmental samples but often only highlight a wide genetic diversity across clinical and water isolates (70). LD outbreaks may be associated with multiple LD strains.

2.5.6 Setting of infection

For surveillance purposes, LD cases are usually classified by probable setting of infection. These settings are associated with some specific characteristics.

2.5.6.1 Community-acquired cases

Most LD cases are community-acquired cases and sporadic. In Europe, approximately 70% of all cases reported in the recent years were community-acquired, of which 5% were reported as part of a cluster. Countries are asked to report cases as having formed part of a cluster if one case was exposed to the same source as at least one other case with their dates of onset within a plausible time period (37).

2.5.6.2 Travel-associated cases

In Europe, cases are reported as travel-associated if they stayed at an accommodation site away from home during their incubation period. Cases who stayed in accommodation used for commercial purposes (such as hotels) should also be reported in dedicated surveillance scheme (see below).

Overall, approximately 20% of European LD cases are travel-associated (TALD), of which half travelled in their country of residence (37). TALD tend to be younger, especially those with a travel history abroad and have a lower case fatality. A study using European data estimated the overall risk associated with travel abroad at 0.3 cases/million nights. An increasing trend in risk from north-western to south-eastern Europe was observed with Greece having the highest risk (1.7 cases/million nights) (72).

Although TALD cases are more frequently associated with stays in hotels (73), *Legionella* in the water system was detected more frequently in ferries than in hotels (74). In hotels, cooling towers and/or potable water systems were the most frequent incriminated source while hot tubs were most commonly associated with cases occurring in ships.

A study suggested that the probability of successive LD cases to occur in European hotels depended on the country and the size of the hotel (75). The size of the hotel was also associated with reoffending accommodation, i.e. associated with further LD cases after a first investigation following a cluster notification (76).

2.5.6.3 Healthcare-associated cases

Legionnaires' disease is known to be a significant cause of nosocomial pneumonia leading to important costs both in treatment and prevention (77). In addition to inhalation, aspiration is thought to be another mode of transmission of HCA LD (78). Studies have suggested that a substantial proportion of hospitals may have their water systems colonized by *Legionella* but the percent positivity that should prompt action remains controversial (79). More worrying, studies performed in the USA suggested that most HCA LD cases were linked to contamination of the potable water systems (80). Hospitalized people are at higher risk for LD because they tend to be older and more likely to have chronic disease compare to the general population. Therefore, outbreaks of LD in hospitals are not uncommon (81). Indeed, a review of LD outbreaks suggested that approximately 25% of LD outbreaks occurred in healthcare settings (82).

Since nosocomial infection is more likely to occur in immunocompromised people, a higher proportion of non-*L. pneumophila* serogroup 1 would be expected. Thus, it has been shown that less than 50% of nosocomial cases can be diagnosed by urinary antigen detection (83). Case fatality is usually higher in nosocomial cases (≈30%) (37).

It is likely that healthcare-associated cases are both poorly diagnosed and reported throughout Europe. Of the 470 healthcare-associated cases reported in 2015, 343 (73%) were reported by France, Italy and Spain (37).

2.6 Surveillance

2.6.1 The European Legionnaires' disease Surveillance Network

Since 2010, the surveillance of LD in Europe has been carried out by the European Legionnaires' Disease Surveillance Network (ELDSNet) and coordinated by the European Centre for Disease Prevention and Control (ECDC). ELDSNet involves 28 EU Member States, Iceland and Norway (37).

It is mandatory to notify all cases of Legionnaires' disease in Europe. All cases meeting the EU case definition for LD should be reported to the European Surveillance System (TESSy), a database hosted by ECDC. According to the type of laboratory test used to ascertain the case, cases are classified as confirmed or probable (84).

LD is thought to be underreported for two main reasons. Firstly, it is underdiagnosed by clinicians. Especially when treating milder forms of chest infection, patients are not tested for LD before empirically prescribing broad-spectrum antibiotics that are likely to cover *Legionella* spp. Secondly, health professionals may fail to notify cases to health authorities due to the added administrative burden (1).

2.6.2 Indicator-based surveillance

Indicator-based surveillance refers to the collection of structured data relying on established routine surveillance systems. It is usually opposed to 'Event-based surveillance' in which unstructured data are collected through the screening of various sources (85). LD surveillance is indicator-based and relies on two different schemes: one covering all cases (comprehensive notifications) reported from European Union (EU) Member States, Iceland and Norway, the other covering all travel-associated cases of Legionnaires' disease (TALD), including reports from countries outside the EU/EEA.

The aims of these two schemes differ. The main objectives of collecting data on all nationally reported cases of LD are:

- to monitor trends over time and to compare them across Member States;
- to provide evidence-based data for public health decisions and actions at EU and/or Member State level;
- to monitor and evaluate prevention and control programs targeting LD at national and European level;
- to identify population groups at risk and in need of targeted preventive measures (37).

The surveillance of TALD aims primarily at identifying clusters of cases that may not otherwise have been detected at the national level, and enabling timely investigation and control measures at the implicated accommodation sites in order to prevent further infections.

2.6.2.1 Comprehensive notifications

Each year, all EU/EEA countries are invited to submit the LD data of the previous year to ECDC. All LD cases meeting the European case definition are included (Box) (84). This case definition was amended in August 2012 and it is no longer possible to report probable cases with an epidemiological link only.

European Union case definition for Legionnaires' disease

Clinical criteria:

Any person with pneumonia.

Laboratory criteria for case confirmation:

At least one of the following three:

- Isolation of Legionella spp. from respiratory secretions or any normally sterile site;
- Detection of Legionella pneumophila antigen in urine;
- Significant rise in specific antibody level to Legionella pneumophila serogroup 1 in paired serum samples.

Laboratory criteria for a probable case:

At least one of the following four:

- Detection of Legionella pneumophila antigen in respiratory secretions or lung tissue e.g. by DFA staining using monoclonal-antibody-derived reagents;
- Detection of Legionella spp. nucleic acid in respiratory secretions, lung tissue or any normally sterile site;
- Significant rise in specific antibody level to Legionella pneumophila other than serogroup 1 or other Legionella spp. in paired serum samples;
- Single high level of specific antibody to Legionella pneumophila serogroup 1 in serum.

Case classification

Probable case

Any person meeting the clinical criteria AND at least one positive laboratory test for a probable case.

Confirmed case

Any person meeting the clinical AND the laboratory criteria for case confirmation.

2.6.2.2 Travel-associated Legionnaires' disease

A travel-associated Legionnaires' disease (TALD) surveillance system at the European Union (EU) level has been in place since 1987 (86). Since 2010, ELDSNet members report TALD cases to ECDC on a daily basis through the web-based European Surveillance System (TESSy). TALD cases need to fulfil the official EU case definition for LD and to have a history of travel, i.e. at least one night spent in commercial accommodation away from home within the incubation period of LD.

ELDSNet defines a cluster of TALD as two or more cases who stayed at or visited the same commercial accommodation site in the two to ten days before onset of illness and whose onset is within the same two-year period. Interestingly, a study challenged the current definition for TALD cluster suggesting that a more flexible definition would allow the detection of more sites (87). In addition, ELDSNet defines a rapidly evolving cluster as at least three cases with dates of onset within a three-month period. The detection of a cluster in the Member States will prompt action in the accommodation and follow-up by health authorities of the measures taken. No action is required for accommodations associated with single cases. When ELDSNet detects a cluster, an investigation by public health authorities is required at the accommodation site. To be able to effectively prevent further cases, all notifications done through this scheme should be timely, i.e. shortly after occurrence of a case. In 2016, the median time from date of onset to reporting to ELDSNet was 19 days (range 6–47 days). ELDSNet subsequently notified the country where the accommodation site associated with the TALD cases was located within days (mostly the same or following day of reporting to ELDSNet) (88).

2.7 Prevention and control

LD is a preventable disease and key to prevent LD is to ensure the proper maintenance of water systems. The plan for water risk management developed by the World health Organization (WHO) provides a framework applicable to *Legionella*-related issues (16). This so-called water safety plan (WSP) consist of three main components: (a) system assessment; (b) monitoring; (c) management and communication.

There are several existing regulations and guidelines for *Legionella* control. Most of them share three common principles (89). First, they highlight the importance of avoiding and monitoring spots that favor the growth of *Legionella*. Second, they propose measures to limit water stagnation, which is propitious to *Legionella* proliferation. Last, they require sufficiently high temperature to prevent the growth of *Legionella*. For instance, there is evidence suggesting that hot water temperature and frequent running showers could reduce *Legionella* contamination of domestic household (90).

Unfortunately, it seems that many LD outbreaks could be associated with deficiencies in environmental control as suggested by a study carried out in the USA (91).

2.8 Research priorities

Among the main research priorities identified by Phin et al., there are several aspects that are worth mentioning, including LD epidemiology, outbreak investigation, diagnostics tests, and ecology (7).

2.8.1 Epidemiology

First, in the absence of reliable estimates of the disease incidence, it is difficult to evaluate the real burden of LD, in Europe and in the rest of the World. Therefore, it is important to provide better estimates of LD incidence and to quantify associated morbidity and mortality. Since LD remains a relatively rare disease, a better understanding of host factors (e.g. genetic or immunologic) associated with a higher susceptibility to *Legionella* would help target people at higher risk of infection.

This thesis investigates LD epidemiology in two specific settings, travel (Study I) and healthcare (Study IV). Study I paid extra attention to the somehow neglected LD associated with domestic travel.

2.8.2 Outbreak investigation

Recent years have seen an increasing use of new tools in outbreak investigation. For LD outbreaks, GIS tools could be promising.

Study III took advantage of the data generated by the longtime functioning scheme of TALD surveillance in Europe, in which information on control measures are systematically collected.

2.8.3 Diagnostic tests

The accuracy of LD diagnostic tests should be improved. The landscape of LD diagnostic tests is currently dominated by UAT with their limitations, especially the incapacity to detect species other than *L. pneumophila* serogroup 1. The increasing use of PCR may change this landscape if standardized methods are defined and applied.

Focusing on LD cases in healthcare settings, Study IV explored various LD strains and the characteristics of the cases that they caused.

2.8.4 Ecology

To better control and potentially eradicate *Legionella* from water systems, a better understanding of its ecology is needed. Risk associated with certain concentrations in different setting should be better estimated. The development of environmental surveillance could help map the risk of LD. This would be helped by a better understanding of the environmental drivers of LD.

Using data from four European countries at subnational level, Study II tried to quantify the role of several environmental drivers on the incidence of community-acquired LD cases.

Study III used the information collected in the near-real-time surveillance of TALD, which include the results of environmental investigation following the detection of a cluster of TALD cases.

3 AIMS

The overarching aim of this thesis was to explore various aspects of LD epidemiology, prevention and control using surveillance data.

The specific aims of the studies are as follows:

Study I: To assess the risks for TALD in European countries on the basis of travel patterns and to provide an estimate of the extent of under-ascertainment by country of destination;

Study II: To test and investigate the effect of temperature, rainfall, and atmospheric pressure on short-term variations in LD notification rate;

Study III: To identify factors associated with the occurrence of further cases after implementation of control measures to improve prevention and control of LD in travelers.

Study IV: To describe the epidemiology of HCA LD using EU-level surveillance data and to determine how it differs from the epidemiology of community-acquired LD in terms of seasonality, demographics, causative pathogens and outcome.

4 DATA

4.1 Legionnaires' disease data

In both Studies II and IV we used LD surveillance data collected through the annual scheme (cf. 2.6.2.1). Each year, EU/EEA countries report all LD cases meeting the EU case definition to ECDC. Although most surveillance systems are similar, there are some differences across countries (92). With the exception of Belgium, all countries had surveillance systems with national coverage. Through this scheme, some of the variables collected are common to most diseases under EU/EEA surveillance. This set of basic epidemiological variables includes age, sex, date of onset, date of diagnosis, place of residence, and outcome. In addition to these basic variables, there are some disease-specific epidemiological variables, which include laboratory method, importation status, probable country of infection, cluster status, pathogen information (LD strain, monoclonal subtype, and sequence type), probable setting of infection, and results of possible environmental investigations. Data completeness is high for most basic variables (>90%). However, completeness was poor for some disease-specific variables. For example, less than 5% of reported cases had information on the sequence type of the causative strain during 2010–2015 (37).

In Study II we extracted a subset of these data, including all community-acquired LD cases reported by Denmark, Germany, Italy, and The Netherlands with onset date in 2007–2012. We aggregated cases by onset week and region of residence (NUTS 2). Community-acquired is a diagnosis of exclusion. A case is community-acquired if there is no history of travel or admission to a hospital in the 2 to 10 days prior to disease onset.

In Study IV we included all locally-acquired cases reported during the years from 2008 to 2017. We defined a locally-acquired case as any case not reported as travel-associated. We used the following variables for the analysis: age, sex, date of disease onset, probable setting of infection, cluster status, laboratory method used for diagnosis, pathogen and clinical outcome (dead or alive).

4.2 Travel-associated Legionnaires' disease data

The variables collected through the near-real-time surveillance scheme of TALD are very similar to those collected for comprehensive notifications. In addition, this scheme collects information on travel history. Travel history includes accommodation type (e.g. hotel), arrival and departure dates, and the location of the accommodation.

Study I was based on TALD surveillance data. We aggregated TALD cases by reporting country and destination country for the year 2009. We restricted the analysis to European residents travelling in EU/EEA countries. The use of surveillance data is considered a valid source to estimate risk in travelers (93).

4.3 Epidemic Intelligence Information System data

The Epidemic Intelligence Information System for ELDSNet (EPIS-ELDSNet) is a web-based communication platform used by nominated public health experts to detect and follow-up travel-associated clusters of LD. ECDC staff investigates TALD data on a daily basis and identifies cluster of TALD (cf. 2.5.6.2). Notification of cluster to the member states and follow-up of control measures are both carried out in the EPIS-ELDSNet platform.

For Study III we used EPIS-ELDSNet data and included all hotel and holiday rental accommodation sites in the EU/EEA that were associated with a cluster of TALD cases notified between 1 June 2011 and 31 December 2016.

4.4 Tourism denominator data

For Study I, travel denominator data were obtained from the Statistical Office of the European Union (Eurostat) (94). We used the total number of nights spent, by destination country. This includes all nights spent in a collective accommodation establishment or in private tourist accommodation for personal or professional purposes by EU/EEA residents, aged 15 or older. Most countries collected such information through household surveys. To ensure maximum data quality countries are required to follow the instructions described in the *Methodological manual for tourism statistics* (95).

4.5 Meteorological data

For Study II, meteorological data were extracted from the European Climate Assessment & Dataset project (ECA&D). ECA provides access to homogenized high-quality datasets based on daily station series maintained by national meteorological institutes. Previous studies have demonstrated the quality of ECA datasets (96). For the purpose of study II, meteorological variables of interest were aggregated by week at regional level (NUTS2) for the period 1 January 2007−31 December 2012 (Figure 7).

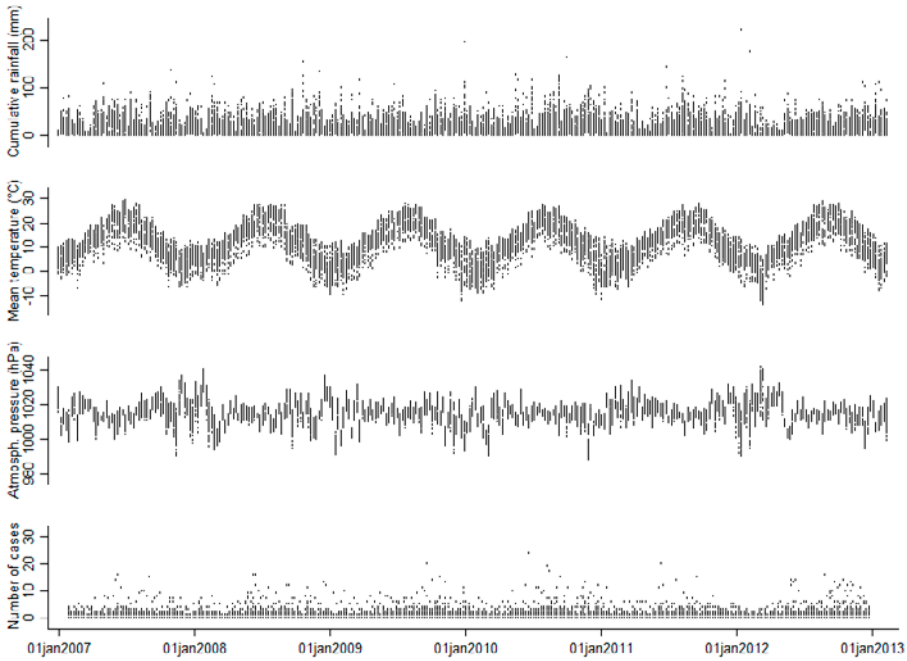

Figure 7. *Weekly average temperature, cumulative rainfall, average atmospheric pressure and number of Legionnaires' disease cases at NUTS2 level, Denmark, Germany, Italy, and the Netherlands, 2007−2012*

4.6 Accommodation size data

In Study III, we looked at the risk of occurrence of a further TALD case after implementation of control measures in accommodation sites. Since this risk is likely to be associated with the number of guests visiting the accommodation, it was necessary to control for this. Information on the annual number of guests was unfortunately not available. Therefore, we decided to use the number of rooms as a proxy. Since surveillance data did not capture this information, we searched the number of rooms for each accommodation in two of the most popular travel website companies (Booking.com and TripAdvisor). This was a tedious work with manual investigation of nearly 400 accommodation sites.

4.7 Ethical considerations

All studies relied on surveillance data routinely collected by ECDC. These data are submitted by EU/EEA Member States in compliance with the EU regulations, especially Decision 1082/2013 and its Implementing Decision (84). LD is part of the 56 communicable diseases for which ECDC coordinates surveillance

activities as stated by the Article 3 of its founding regulation. These surveillance data are anonymized and processed for public interest in the area of public health. Therefore, informed consent was not required or subject to national policies. Yet, most of these data are case-based and contain personal information.

Both Study I and II were based on aggregate subset of LD surveillance data. For Study I, data were aggregated at national level, by destination country for the year 2009. For study II, data were aggregated by week of onset at regional level (NUTS2). Aggregate data do not fall under the law of ethical review for research in Sweden.

Study III was based on accommodation data. Since it was not possible to identify any of the accommodations included on the analysis, there was no risk of undermining commercial interests.

Data used for study IV were anonymized. This means that no individual could be identified. The variables used for the purpose of this analysis were age, sex, date of disease onset, probable setting of infection, cluster status, laboratory method used for diagnosis, pathogen and clinical outcome (dead or alive).

5 STATISTICAL METHODS

5.1 Poisson regression

Study II aimed to estimate the association between meteorological variables and the weekly number of community-acquired LD cases. The outcome measure was the weekly number of cases for each geographical area (NUTS2 region) given its population (included as an offset). Poisson regression is a method used to analyze rates or counts of rare events (97). It allows the comparison of different exposure groups, estimating and controlling for effects that change over time. We used Poisson regression to estimate relative risks (RR) from rate ratios and their 95% confidence intervals (CI).

Since the median incubation duration for LD is approximately one week, we assumed that a time lag of one week between exposure to weather conditions and disease onset to be the most likely. Yet, the weather conditions observed in previous weeks could also play a role. Therefore, we allowed for delayed exposure effects up to four weeks before date of onset. To compare the goodness-of-fit of our models, we used Akaike's Information Criterion (AIC) (98). We selected models with the lowest AIC because they are thought to minimize the information loss.

The three exposure variables considered (cumulative rainfall, mean temperature and mean atmospheric pressure) might share some collinearity. For example, low atmospheric pressure is likely to be associated with rainfall. to address potential problems related to multicollinearity between continuous covariates, we calculated the variance inflation factor (VIF) (99). VIF is an indicator quantifying the severity of multicollinearity. If there is collinearity among the variables, VIF is expected to increase sharply.

5.2 Modelling seasonality and long-term trends

Since LD incidence is known to have a pronounced seasonality, it is necessary to control for the seasonal patterns in the regression model. Otherwise, it would not be possible to distinguish the seasonal patterns from the short-term associations between weather conditions and LD incidence (Study II). Bhaskaran et al. proposed three alternative for modelling seasonal and long-term patterns (100) (Figure 8):

a) The first and simplest approach is to split the study period into short intervals and to include an indicator variable for each interval in the model (time-stratified model). We discredited this approach because it would have generated too many parameters. In addition, the differences observed between two adjacent intervals may be difficult to interpret.

b) The second approach is to model long-term patterns by fitting Fourier terms in the model. Although such cyclic regression would model smoothly seasonality, we rejected it because it would force the timing of each peak to be identical for each cycle.

c) The third option proposed by Bhaskaran is to fit a spline function of time. Spline functions are polynomial functions joined by knots. This was our preferred option because spline functions allow seasonal patterns to change over time and can also capture non-seasonal long-term trends. The only drawback is their mathematical complexity. For the purpose of study II, we selected restricted cubic spline functions with 3 degrees of freedom (DF) for knots and the spline basis centered on the median value of the exposure (default setting in Stata).

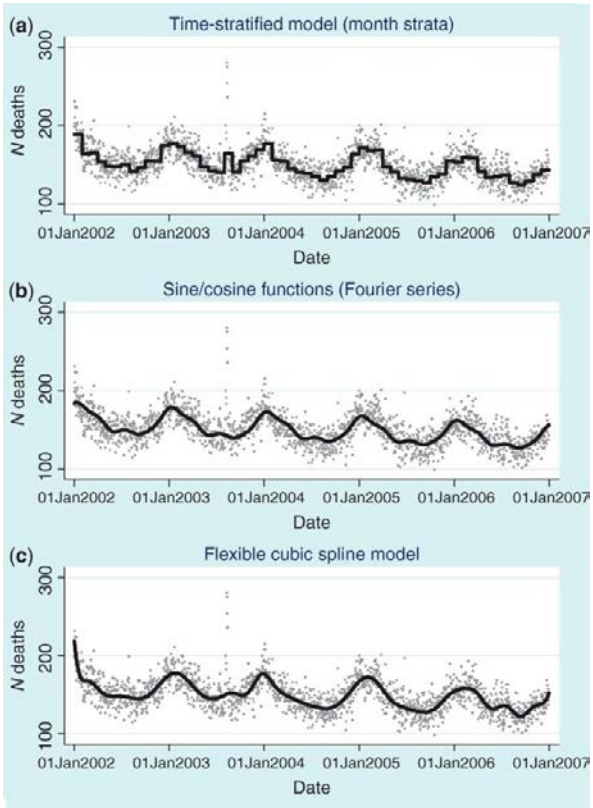

Figure 8. *Three alternative ways of modelling long-term patterns in the data (seasonality and trends). Source: Bhaskaran (2013) (100)*

5.3 Survival analysis

In Study III we examined the occurrence of further LD cases after implementation of control measures in accommodation sites associated with a cluster of TALD. The outcome of interest was the time to a further LD case. Survival analysis deals with time-to-event outcomes. In survival analysis there is no need to assume that the risk of occurrence of the event is constant over time. In our study, it would be reasonable to assume that the risk of occurrence of a further LD case may be lower immediately after implementation of control measures but could increase later on if the measures effects wane over time.

Survival analysis relies on a survivor function S(t) and a hazard function h(t). The hazard function represents the instantaneous rate at time t. In Study III, h(t) corresponds to the rate of occurrence of further cases after the report on control measure (number of cases per 100 accommodation-years). The survival function is the probability that an accommodation will not experience the event of interest (i.e. occurrence of a further case) up to and including time. Since the exact failure time is known (date of occurrence of a further case) it is possible to estimate the exact failure and censoring times by the Kaplan-Meier estimate (97). In study III we reported cumulative incidence of accommodations sites associated with a further TALD case (i.e. inverse of survival function using Kaplan-Meier estimate) and compared different groups using the log-rank test. Accommodations sites for which no further LD cases was reported were censored on 31 December 2016, which was the end of the study period (right censoring).

To quantify the differences in survival across groups, we fitted Cox proportional hazards models. The Cox regression has several advantages, one that the baseline risk ("hazard") does not have to be modeled explicitly. Instead, the Cox models assuming that the ratio of the hazards between the groups of interest is constant over time. There are various methods to assess the validity of this assumption, including plotting Schoenfeld residuals (101).

5.4 Logistic regression

In Study IV, we compared binary outcomes (e.g. dead or alive) between two exposure groups, which were two probable settings of infection (community or healthcare). For such analysis, logistic regression is a commonly used method (97). Logistic regression models the association between exposure and binary outcome variables in terms of odds ratio (OR). It is then possible to derive confidence intervals (CI) by using the standard error of the log OR to calculate a CI for the log OR.

In study IV, we fitted two main models. The first one estimated the OR of HCA LD compared to community-acquired LD. The second estimated the OR of fatal LD compared to non-fatal LD. For both models, we included a small number of variables (first model: age, sex, reporting year, and reporting country; second model: age, sex, reporting year, reporting country, and probable setting of infection).

6 MAIN RESULTS

6.1 Risk for travel-associated Legionnaires' disease (Study I)

Of the 607 TALD cases reported among European residents travelling in EU/EEA countries in 2009, 363 (60%) were related to domestic travel, i.e. travel in their country of residence. The top three travel destinations (France, Italy, and Spain) accounted for 72% of all TALD cases. TALD cases were associated with stays in hotels (70%), campsites (8%), private accommodations rented for commercial purposes (6%), apartments (5%), cruise ships (<1%), and other accommodations (10%). In 2009, EU/EEA residents spent two billions nights in Europe, of which 66% were in their country of residence. France, Italy, and Spain accounted for 46% of all nights spent.

In 2009, the average risk for TALD in Europe in 2009 was 0.30 cases/1 million nights (95% CI 0.27–0.32). The highest for domestic travel was in Italy (0.66 cases/1 million nights) and for non-domestic travelers in Greece (0.88 cases/1 million nights). Using the best reporting countries as reference (the UK, the Netherlands, France, and Denmark), we estimated a pooled overall risk of 0.55 cases/million nights and a pooled risk of 1.68 cases/1 million nights when traveling to Greece (Table 2). We observed the highest level of under-ascertainment in Greece, Portugal, and Austria (Germany did not report cases in domestic travelers until 2012.

Table 2. Expected risk for Legionnaires' disease in European travelers to non-domestic destinations in Europe, based on reference data reported by the United Kingdom, the Netherlands, France, and Denmark, Europe, 2009

Destination	Risk in travellers (cases/million nights)	Incidence ratio (95% conf. interval)	Total cases (n)	
			Reported	Estimated
Greece	1.68	7.2 (4.2-12.2)	34	98
Italy	1.40	6.0 (3.9-9.2)	209	463
Germany	1.19	5.1 (2.9-8.7)	22	353
Portugal	1.06	4.6 (2.1-9.0)	20	44
Austria	1.01	4.4 (2.2-8.2)	20	95
Spain	0.57	2.5 (1.6-3.8)	98	188
France	0.53	2.3 (1.6-3.3)	137	145
Netherlands	0.33	1.4 (0.8-2.4)	21	26
UK	0.23	1.0 (ref.)	45	53
Other countries	0.90	3.9 (2.3-6.4)	42	282
Total	0.55	-	607*	1 127

* A case may have a travel history involving more than one country

6.2 Short-term effects of meterological conditions on incidence of Legionnaires' disease (Study II)

Four countries (Denmark, Germany, Italy and the Netherlands) accepted to participate in the study by providing data at regional level (NUTS2). This represented 77 NUTS regions with a corresponding population of 164 million inhabitants. Of the 8,708 LD cases reported during 2007−2012, 8,093 (93%) had available information on both onset date and place of residence. We excluded cases with onset in the first four weeks of 2007 for which we had no or partial exposure data (the time series for meteorological data started on 1 January 2007). Finally, we included 7,961 cases in the analysis.

We found a positive association between weekly cumulative rainfall and an increased risk of LD. We observed the association with the highest risk and the lowest AIC with a lagged effect of 1 week (RR 1.13 for every 10-mm increase, 95% CI 1.12–1.14). We found a positive association between weekly mean temperature and an increased risk of LD. We observed the association with the highest risk and the lowest AIC with a lagged effect of 3 weeks (RR 1.05 for every 2°C increase, 95% CI 1.03–1.07).

We kept in the adjusted model meteorological variables with the lag associated with the highest RR and lowest AIC. There was no indication of multicollinearity between these variables according to calculated VIC (<10). With no weekly rainfall as a reference, the estimated adjusted RR of LD for weekly cumulative rainfall >40 mm with a lagged effect of 1 week was 2.14 (95% CI 1.90–2.42; rate 182 vs. 62 LD cases/10 million population). With weekly mean temperature <10°C as a reference, the estimated adjusted RR of LD for weekly mean temperature of 15–19°C with a lagged effect of 3 weeks was 2.00 (95% CI 1.75–2.28, rate 120 vs. 54/10 million population). Interestingly, the effect of temperature plateaued above 20°C.

We found positive interactions between increasing weekly cumulative rainfall (1 week lag) and increasing weekly mean temperature (3 weeks lag) (Table 3).

Table 3. Estimated relative* risk and 95% CI of community-acquired Legionnaires' disease for an interaction between weekly cumulative rainfall (one week lag) and weekly mean temperature (three weeks lag), Denmark, Germany, Italy and the Netherlands, 2007−2012

Weekly cumulative rainfall (one week lag)	Weekly mean temperature (three weeks lag)			
	<10°C	10 to 14°C	15 to 19°C	≥20°C
<10 mm	1 (ref.)	1.20 (1.04-1.37)	1.66 (1.42-1.95)	1.50 (1.24-1.81)
10 to 19 mm	1.13 (1.01-1.28)	1.53 (1.31-1.79)	2.00 (1.69-2.37)	1.70 (1.38-2.11)
20 to 29 mm	1.31 (1.15-1.50)	1.82 (1.54-2.17)	2.77 (2.34-3.27)	2.66 (2.14-3.32)
≥30 mm	1.37 (1.21-1.55)	2.28 (2.00-2.61)	3.50 (3.00-4.08)	2.90 (2.38-3.54)

* Relative risks from Poisson regression including covariates year (2007−2012), NUTS2 (one intercept for each region), population, weekly cumulative rainfall (one week lag), weekly mean temperature (three weeks lag), weekly mean atmospheric pressure (one week lag), adjusted for season using a cubic spline function with five knots, and an interaction term.

We found the highest RR for weekly mean temperature of 15–19 °C and cumulative rainfall >30 mm compared to temperature <10 °C and rainfall <10 mm (RR 3.50, 95% CI 3.00–4.08).

6.3 Factors associated with Legionnaires' disease recurrence in hotels (Study III)

During 1 June 2011−31 December 2016, 395 accommodation sites in the EU/EEA were notified with a cluster of TALD cases. Of these, 357 (90%) had information on both follow-up of control measures and number of rooms. Of these 357 accommodations, 90 (25%) were associated with at least one further case after the report on measures taken (12.4/100 accommodation-years). We observed higher cumulative incidences for accommodation sites associated with a previous case compared with those that were never associated with any case before the cluster (Figure 9). After 3 years of follow-up, 50% of the accommodations previously reported with two cases or more were associated with a further case.

Accommodation sites with 36 rooms or more had a higher risk of a further case compared to those with less than 36 rooms (HR>2). Accommodations previously associated with two cases or more had a HR of 2.26 (95%CI: 1.40–3.64). We found no association between the detection of *Legionella* in the water system nor the type of disinfection and the risk of a further case.

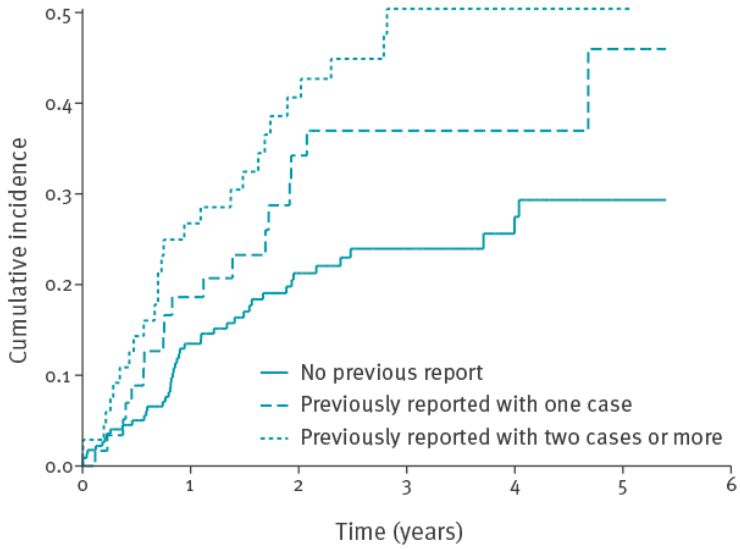

Figure 9. *Cumulative incidence of hotel and holiday rental accommodations sites associated with a further TALD case after control measures, by previous report status, EU/EEA, 1 June 2011–31 December 2016. Source: Beauté, J. 2019 (102)*

6.4 Healthcare-associated Legionnaires' disease (Study IV)

Over the 2008−2017 period, 30 countries reported 64,409 LD cases, of which 57,175 (88.8%) had the information available for inclusion. Of these, 40,411 (70.7%) were reported as community-acquired, 11,512 (20.1%) as travel-associated, 4,315 (7.6%) as healthcare-related and 937 (1.6%) as associated with other settings. Finally, we included 44,726 LD cases in the analysis reported by 29 countries, of which 40,411 (90.4%) were community-acquired and 4,315 (9.6%) HCA LD. Of the 4,315 HCA LD cases, 2,937 (68.1%) were nosocomial cases and 1,378 (31.9%) linked to other healthcare facilities.

The proportion of HCA LD cases was higher in female compared with male cases (14.3% vs. 7.8%; $p<0.01$). The male-to-female ratio was lower in younger and older age groups (0.9:1 below 20 years and at 80 years and over), peaking at 2.2:1 for those 40–49 years of age. When adjusting for age, sex, year and reporting country, females were more likely to have acquired their infection in a hospital compared with males (OR: 1.60, 95%CI: 1.49-1.71). Compared with those aged 50-59 years, people younger than 20 years were twice as likely to be reported as HCA (OR: 2.04, 95%CI: 1.25-3.33). At 60 years of age and over, the risk of being reported as HCA increased with age peaking in those aged 80 years and over (OR: 4.58, 95%CI: 4.11-5.12).

Of the 4,859 culture-confirmed cases reported with a causative pathogen, 4,739 (97.5%) were due to *Legionella pneumophila*. This proportion was similar in community-acquired and HCA LD cases (97.4% vs. 98.1%; p=0.31). Of the 4,533 laboratory-confirmed cases due to *L. pneumophila* reported with a serogroup, 4,137 (91.3%) were due to due to *L. pneumophila* serogroup 1. This proportion was higher in community-acquired cases compared with HCA LD cases (92.3% vs. 85.1%; p<0.01). Of the 107 community-acquired cases with culture confirmation due to other *Legionella* species, 48 (44.9%) were due to *L. longbeachae*. No HCA LD case was reported with *L. longbeachae*.

Of the 856 culture-confirmed cases due to *L. pneumophila* serogroup 1 isolates that were subtyped using monoclonal antibodies (MAb), 679 (79.3%) were MAb 3/1 positive. This proportion was higher in community-acquired cases compared with HCA LD cases (83.6% vs. 43.3%; p<0.01).

Of the 32,379 cases with known outcome, 3,448 (10.7%) died. When adjusting for age, sex, year and reporting country, HCA LD cases were associated with a higher risk for fatal outcome compared with community-acquired cases (OR: 3.02, 95%CI: 2.75-3.32).

7 DISCUSSION

7.1 Main findings

7.1.1 Risk of TALD and under-ascertainment

In Study I, we estimated risk for TALD in most touristic destination countries in Europe. These findings are valuable because most similar studies rely on travel clinics, which are likely to miss LD cases (103). Steffen et al. provided health risk among travelers using logarithmic scales (104). Based on European data collected in 2003, he estimated the TALD risk to be between 0.0001% and 0.001% per month of stay in developing countries. We found the highest risk associated with travel to Greece with 1.68 cases per 1 million nights. This corresponds to 0.005% per month of stay, which is five times higher than Steffen's estimate. We found the lowest risk estimate associated with travel to the UK (0.001% per month of stay). These results are compatible since travelers visiting developing countries are likely to be younger than those staying in Europe.

Study I also revealed high levels of under-ascertainment in Austria, Germany, Greece, and Portugal. In all these countries, LD notification rate increased at a faster pace than the EU/EEA average during 2009–18. The EU/EEA average notification doubled during this 10-year period from 1.1 to 2.2 case per 100,000 population (105). Over the same period, LD notification rate was multiplied by 2.5 in both Austria and Portugal, by 3 in Germany, and by 4.3 in Greece. In 2012, Germany started reporting TALD cases to ELDSNet. Our findings did not trigger the increase observed in these countries but it confirmed that our results were sound.

7.1.2 Community-acquired Legionnaires' disease and weather conditions

In Study III, we confirmed the role of both rainfall and temperature as key environmental factors associated with LD incidence. Previous studies in Europe and northern America reported similar findings (17, 19, 20, 23). The sequence of warm weather followed by heavy rainfall as conditions associated with high LD incidence was also reported by a study carried out in the Netherlands (18). We found that the association between LD incidence and cumulative rainfall to have a linear shape. However, it seemed that unusually high temperatures were not associated with further increase in LD incidence.

7.1.3 Recurrence of TALD in hotels

In Study III, our results suggested that approximately 30% of hotels associated with TALD cluster were reported with at least one further case within two years. This proportion was comparable but higher than previous estimates made by a study

carried out during 1993–2000 (75). There may be two main reasons explain the difference. First, the ELDSNet database is more mature with more accommodation sites and therefore more likely to detect further cases. Second, surveillance of LD improved over the last decade and LD cases are now more likely to be reported to ELDSNet than they were 20 years ago.

Our findings suggested that the risk of recurrence was independent of the measures taken after the detection of a TALD cluster.

7.1.4 Healthcare-associated Legionnaires' disease

HCA LD disproportionally affects older people in Europe. Yet, HCA LD also concerns younger age groups. LD cases in children below 20 years of age were twice more likely to be HCA cases compared to those people that are 50–59 year-olds. HCA LD were associated with less virulent strains of *Legionella* as suggested by earlier work (106).

HCA LD cases are more severe than community-acquired cases. Countries with less performant surveillance systems seemed to capture HCA LD more easily than cases associated with other settings.

7.2 Strengths

7.2.1 Legionnaires' disease surveillance data

The interpretation of any analysis based on surveillance data will depend on the validity of these notification data (56). Among the main factors that may create biases, most are unlikely to have a large impart of LD surveillance data. First, LD is a severe condition. It means that factors influencing health-seeking behaviors (e.g. distance to healthcare facilities or cost) will probably not change the probability for a case to seek medical care. Second, the case definition of LD includes clinical criteria (i.e. presence of pneumonia). Therefore, issues related to the possibility to diagnose asymptomatic cases of a disease (e.g. screening of chlamydia) are here irrelevant. Last, the vast majority (>90%) of LD cases in these surveillance data were classified as confirmed according to the EU case definition. This means that the laboratory test used to ascertain the diagnosis was highly specific.

In both Study I and III we focused on TALD data, which are unique. Studies looking at travel-associated infection in European travelers tend to focus on overseas destinations and exotic pathogens (107). Infections like LD, which are unlikely to be diagnosed in travel clinics and can be associated with domestic travel are usually overlooked (103).

7.2.2 Pooling data from different countries

The other main strength of European LD surveillance data is its international component. All 30 EU/EEA Member States report data to the comprehensive notifications scheme and approximately 25 countries report to the near-real-time surveillance of TALD (22 EU/EEA Member States and three non-EU/EEA countries in 2015 (37)). All reporting countries use the same reporting protocol. These schemes allowed us to have large sample size for our studies. Although only four countries participated in study II, we were able to include approximately 8,000 community-acquired cases, allowing adjustment at the regional level and analysis by month (108). We could also identify similarities and possible difference across countries in terms of LD environmental drivers. In study IV, we included more than 4,000 HCA LD cases, which allowed us to look at strains less commonly associated with LD.

Another advantage of pooling data from several countries is to allow comparison and benchmarking. In Study I, we took advantage of the data from Denmark, France, the Netherlands, and the UK to highlight under-ascertainment in other countries.

Last, running analyses at the supranational level may help confirm hypotheses made at lower geographical level. In study II, we explored the contribution of different environmental factors, of which some had already been studied at national or regional level. However, it was difficult to understand whether the differences observed between studies were real or explained by somehow different methodologies.

7.3 Limitations

7.3.1 Surveillance data

For all studies, we used surveillance data, i.e. data generated for – at least partly – a purpose different from the research objectives. The use of secondary data can be problematic if the data are inadequate to answer the research question (109).

One of the main limitations of surveillance data is that there is always a risk of under-ascertainment of cases. First, physicians may fail to suspect LD and/or prescribe the laboratory test that will confirm the diagnosis. Reports from central and eastern European countries suggest that lack of clinician awareness could explain the low notification rates observed (110, 111). Second, cases may not be reported to the surveillance system. Many European countries carried out capture-recapture studies to assess the magnitude of LD underreporting (112-114). If the factors influencing under-ascertainment are associated with either the exposure or the outcome of our studies, there may be an ascertainment bias. Although we thought this was unlikely, we adjusted our analyses by country.

There are other issues with data quality of surveillance data, including completeness and validity (115). Completeness refers to the proportion of variables for which there are missing and/or unknown fields. Overall, data completeness was high for the variables included in our analyses. However, only a small proportion of cases were culture-confirmed (<15%). Since culture confirmation is the best method to characterize further LD strains, we could not fully describe the causative strains in Study IV. The validity of the data is the ability to represent the reality. In terms of validity, the main limitation of our studies is the possible misclassification of cases for the probable setting of infection. Although there is a European definition for TALD, there is currently no consensus for HCA LD. Community-acquired LD is a diagnosis of exclusion (i.e. non-HCA and non-travel). It is therefore for the epidemiologist registering cases to classify them according the most probable setting. Since investigating every sporadic case would be both costly and likely to yield negative results, there is no environmental investigation for most cases. Therefore, it is impossible to assess the level of misclassification. Yet, there are reasons to think this phenomenon is limited. . Most of all, LD is a rare event so any changes taking place before the incubation is likely to play a role in the infection (e.g. admission to hospital).

Last, if pooling data from different countries present many advantages, it also poses a number of analytical and procedural problems related to heterogeneity (116). We can distinguish three main types of heterogeneity:

a) Heterogeneity of surveillance systems refers to differences in the design and operation of surveillance (e.g. if some countries have a surveillance system covering only a few regions of their territory). Although LD surveillance systems differ across EU/EEA countries, there are important similarities. All countries but Belgium have a system with comprehensive coverage and LD is notifiable in all countries (92). The main problem is under-ascertainment of cases, especially in central and eastern European countries (see above). Yet, we do not think under-ascertainment in some countries introduced substantial biases in our analyses. There were no major changes during the periods considered in our studies and analyses of risk factors were adjusted (116).

b) Heterogeneity in disease determinants refers to real epidemiological differences across countries (e.g. different exposure to risk factors). Exposure to weather conditions favourable to the growth of *Legionella* are likely to vary across country and even within countries. Thus, the overall risk is probably higher in southern Europe compared to nortern countries. Howver, the situation may be more complex as suggested by the increasing trend form west to east reported in France (117). Northern Italy and southern Switzerland seem to be region particularly favourable to high LD incidence (20, 108). Such heterogeneity is deemed problematic when the effectiveness of intervention differs across countries. In study III, we found no significant differences across countries in terms of effectiveness of control messures.

c) Heterogeneity of data quality refers to variation in terms of quality across countries (e.g. completeness is higher in country A compared to country B). For study II, we only included the four countries with sufficient data quality for the variable place of residence. For study IV, we excluded Sweden because its cases lacked inormation on the probable setting of infection.

7.3.2 Travel data

In study I, we used travel data provided by Eurostat (94). These data have limitations. First, most countries collect these data through household surveys, which are at risk of recall bias. Second, the number of nights spent by destination countries were not available by age group. It is possible that EU/EEA residents aged over 50 years (i.e. people at increased risk for LD) have different travel patterns than younger residents. This could have biased our findings. Eurostat data publishes travel data with delay. We ran our analysis for study I in 2011 but travel data were only available until 2009.

There are alternative data sources for travel data but none was fully satisfactory at the time. Different methodologies used to capture travel patterns result in different estimations of traveler numbers (118). Recent studies used different data source to estimate travel patterns such as the International Air Transport Association (IATA) data (119). IATA data seem reliable but may miss a large part of European travelers who stay in Europe or in their residence country.

7.3.3 Ecological fallacy

In study II, we looked at the association between environmental condition and the incidence of community-acquired LD. Data were aggregated at regional level (NUTS2). Since the study unit was a group of people rather than the individual, it can be qualified an ecological study (120). Such study is at risk of ecological fallacy, i.e. assuming that the results found at group level are valid at individual level. To avoid such risk, we should have collected exposure data for all cases. In this case, it would have meant collecting individual exposure to meteorological conditions for the nearly 8,000 cases included in the analysis, which was not possible.

7.3.4 Censoring in survival analyses

In study III, accommodations were right-censored, i.e. censored because of study termination. Theoretically, this should pose no problem to the analysis if we can assume that the risk of occurrence of a further case was independent of the time of entry to the study (report of control measures after a cluster of TALD). If this is true, we could assume independent censoring (or non-informative censoring) (121). Yet, the number of TALD cases increased during 2011–15 for reasons that are not fully understood (37). It means that it is possible that the risk of occurrence of a further case increased during the study period.

7.3.5 Confounding

When the association between an exposure variable and the outcome is influenced by other factors or variables, there is risk to draw wrong conclusions because of confounding bias. Confounding occurs when a variable is associated with both exposure and outcome (97). However, a confounding variable is not part of the causal chain between exposure and outcome. In Study IV, age was a possible confounder. Old age is associated with both severe outcome (e.g. death) and probability to be HCA. We controlled for confounding by stratifying age in eight age groups.

7.4 Future perspectives

LD incidence has been increasing in both Europe and the USA. Individuals can risk will get LD even with low level of *Legionella* in the causative water system (122). In addition to known risk factors (e.g. age, male sex), some behavioral factors may be associated with LD infection. Such factors may be occupational (e.g. plumbing) or linked to daily activities such as showering or gardening. Futures studies should investigate further the role of these behavioral factors.

Outbreak investigations often yielded disappointing results. The use of WGS suggested that LD outbreaks could be caused by several strains (122). The inclusion of molecular data in LD epidemiological studies could improve our understanding of LD transmission.

Systematic collection of information on preventive and control measures, especially in surveillance schemes such as the ELDSNet surveillance of TALD could help determine the best options to prevent LD cases.

8 CONCLUSIONS

Using data from the year 2009, Study I revealed high levels of under-ascertainment in Austria, Germany, Greece, and Portugal. It also provided fresh estimates of TALD risk among European travelers, which need regular updates (123).

Study II confirmed the impact of rainfall and temperature on LD incidence. This is of importance in a context in which there is increasing concern that climate change may lead to a higher incidence of LD (25, 124, 125). Yet, it is unclear whether behavioral factors could be associated with community-acquired cases and weather conditions.

In Study III, we identified possible risk factors associated with TALD recurrence in hotels. These findings may help ELDSNet prevent TALD cases after detection of a cluster. In addition, we proposed several improvements to the surveillance scheme, especially regarding data collection. Information on *Legionella* strains and control measures should be more detailed and harmonized in order to understand whether some control measures are effective than others.

In study IV, we analyzed one of the largest data set of HCA LD to date. This allowed a precise description of the LD strains involved in HCA LD, especially less common strains such as *L. pneumophila* non-serogroup 1.

9 ACKNOWLEDGEMENTS

To my supervisors, who accepted to guide this work knowing that my contribution to the life of the department would be limited with my full-time work at ECDC. I liked our discussion and our lunches at KI in which we would discuss anything but my PhD! To Pär, for your kindness and positive spirit. To Sven, for constantly challenging my positions and helping me explore further statistics. To Johan, for connecting ECDC and KI, your meticulous reviews and *joie de vivre*!

To Edoardo for introducing me to MEB and supporting me since day one!

To my co-authors for providing valuable input to these studies.

To administrators at MEB: Camilla Ahlqvist and Alessandra Nanni for kindly supporting me for administrative requests.

To my colleagues at ECDC for helping me embark on such endeavor. To Andrew, Bruno and Phillip for understanding that a PhD could be beneficial to both personal development and to the agency.

To my brothers in arms at ECDC: Alessandro, Ettore, Hakim, Jonathan, and Tarik.

To ELDSNet and all participating countries for providing data and input to discussions.

To my parents for your never-ending support.

To Leila, for our life together. I know I am not always easy…

To Zeno and Cesare, for projecting me in the future!

To all those I forgot to mention here!

10 REFERENCES

1. Fields BS, Benson RF, Besser RE. Legionella and Legionnaires' disease: 25 years of investigation. Clin Microbiol Rev. 2002;15(3):506-26.

2. Fraser DW, Tsai TR, Orenstein W, Parkin WE, Beecham HJ, Sharrar RG, et al. Legionnaires' disease: description of an epidemic of pneumonia. N Engl J Med. 1977;297(22):1189-97.

3. de Jong B, Zucs P. Legionella, springtime and potting soils. Euro Surveill. 2010;15(8):19497.

4. File TM. Community-acquired pneumonia. Lancet. 2003;362(9400):1991-2001.

5. Beaute J, European Legionnaires' Disease Surveillance Network. Legionnaires' disease in Europe, 2011 to 2015. Euro Surveill. 2017;22(27).

6. Beaute J, Robesyn E, de Jong B, European Legionnaires' Disease Surveillance Network. Legionnaires' disease in Europe: all quiet on the eastern front? Eur Respir J. 2013;42(6):1454-8.

7. Phin N, Parry-Ford F, Harrison T, Stagg HR, Zhang N, Kumar K, et al. Epidemiology and clinical management of Legionnaires' disease. Lancet Infect Dis. 2014;14(10):1011-21.

8. Garcia-Fulgueiras A, Navarro C, Fenoll D, Garcia J, Gonzalez-Diego P, Jimenez-Bunuales T, et al. Legionnaires' disease outbreak in Murcia, Spain. Emerg Infect Dis. 2003;9(8):915-21.

9. Shivaji T, Sousa Pinto C, San-Bento A, Oliveira Serra LA, Valente J, Machado J, et al. A large community outbreak of Legionnaires disease in Vila Franca de Xira, Portugal, October to November 2014. Euro Surveill. 2014;19(50):20991.

10. Centers for Disease Control Prevention. Legionellosis --- United States, 2000-2009. MMWR Morb Mortal Wkly Rep. 2011;60(32):1083-6.

11. Mercante JW, Winchell JM. Current and emerging Legionella diagnostics for laboratory and outbreak investigations. Clin Microbiol Rev. 2015;28(1):95-133.

12. Muder RR, Yu VL. Infection due to Legionella species other than L. pneumophila. Clin Infect Dis. 2002;35(8):990-8.

13. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. Lancet. 2016;387 (10016):376-85.

14. Katz SM, Hammel JM. The effect of drying, heat, and pH on the survival of Legionella pneumophila. Ann Clin Lab Sci. 1987;17(3):150-6.

15. Soderberg MA, Rossier O, Cianciotto NP. The type II protein secretion system of Legionella pneumophila promotes growth at low temperatures. J Bacteriol. 2004;186(12):3712-20.

16. Bartram J. Legionella and the prevention of legionellosis. Geneva: WHO; 2007.

17. Hicks LA, Rose CE, Jr., Fields BS, Drees ML, Engel JP, Jenkins PR, et al. Increased rainfall is associated with increased risk for legionellosis. Epidemiol Infect. 2007;135(5):811-7.

18. Karagiannis I, Brandsema P, M VDS. Warm, wet weather associated with increased Legionnaires' disease incidence in The Netherlands. Epidemiol Infect. 2009;137(2):181-7.

19. Ricketts KD, Charlett A, Gelb D, Lane C, Lee JV, Joseph CA. Weather patterns and Legionnaires' disease: a meteorological study. Epidemiol Infect. 2009;137(7):1003-12.

20. Conza L, Casati S, Limoni C, Gaia V. Meteorological factors and risk of community-acquired Legionnaires' disease in Switzerland: an epidemiological study. BMJ Open. 2013;3(3).

21. Dunn CE, Rowlingson B, Bhopal RS, Diggle P. Meteorological conditions and incidence of Legionnaires' disease in Glasgow, Scotland: application of statistical modelling. Epidemiol Infect. 2013;141(4):687-96.

22. Halsby KD, Joseph CA, Lee JV, Wilkinson P. The relationship between meteorological variables and sporadic cases of Legionnaires' disease in residents of England and Wales. Epidemiol Infect. 2014;142(11):2352-9.

23. Garcia-Vidal C, Labori M, Viasus D, Simonetti A, Garcia-Somoza D, Dorca J, et al. Rainfall is a risk factor for sporadic cases of Legionella pneumophila pneumonia. PLoS One. 2013;8(4):e61036.

24. Brandsema PS, Euser SM, Karagiannis I, Den Boer JW, Van Der Hoek W. Summer increase of Legionnaires' disease 2010 in The Netherlands associated with weather conditions and implications for source finding. Epidemiol Infect. 2014;142(11):2360-71.

25. Sakamoto R. Legionnaire's disease, weather and climate. Bull World Health Organ. 2015;93(6):435-6.

26. Correia AM, Ferreira JS, Borges V, Nunes A, Gomes B, Capucho R, et al. Probable Person-to-Person Transmission of Legionnaires' Disease. N Engl J Med. 2016;374(5):497-8.

27. Amodeo MR, Murdoch DR, Pithie AD. Legionnaires' disease caused by Legionella longbeachae and Legionella pneumophila: comparison of clinical features, host-related risk factors, and outcomes. Clin Microbiol Infect. 2010;16(9):1405-7.

28. Wallensten A, Oliver I, Ricketts K, Kafatos G, Stuart JM, Joseph C. Windscreen wiper fluid without added screenwash in motor vehicles: a newly identified risk factor for Legionnaires' disease. Eur J Epidemiol. 2010;25(9):661-5.

29. Moran-Gilad J, Lazarovitch T, Mentasti M, Harrison T, Weinberger M, Mordish Y, et al. Humidifier-associated paediatric Legionnaires' disease, Israel, February 2012. Euro Surveill. 2012;17(41):20293.

30. Tossa P, Deloge-Abarkan M, Zmirou-Navier D, Hartemann P, Mathieu L. Pontiac fever: an operational definition for epidemiological studies. BMC Public Health. 2006;6:112.

31. Den Boer JW, Yzerman EP, Schellekens J, Lettinga KD, Boshuizen HC, Van Steenbergen JE, et al. A large outbreak of Legionnaires' disease at a flower show, the Netherlands, 1999. Emerg Infect Dis. 2002;8(1):37-43.

32. Mulazimoglu L, Yu VL. Can Legionnaires disease be diagnosed by clinical criteria? A critical review. Chest. 2001;120(4):1049-53.

33. Granados A, Podzamczer D, Gudiol F, Manresa F. Pneumonia due to Legionella pneumophila and pneumococcal pneumonia: similarities and differences on presentation. Eur Respir J. 1989;2(2):130-4.

34. Delicata M, Banerjee A. A rare presentation of Legionnaires' disease. BMJ Case Rep. 2015;2015.

35. McDade JE, Shepard CC, Fraser DW, Tsai TR, Redus MA, Dowdle WR. Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. N Engl J Med. 1977;297(22):1197-203.

36. Alvarez J, Dominguez A, Sabria M, Ruiz L, Torner N, Cayla J, et al. Impact of the Legionella urinary antigen test on epidemiological trends in community outbreaks of legionellosis in Catalonia, Spain, 1990-2004. Int J Infect Dis. 2009;13(6):e365-70.

37. European Centre for Disease Prevention and Control. Legionnaires' disease in Europe, 2015. Stockholm: ECDC; 2017. Available from: http://dx.doi.org/10.2900/692621.

38. Murdoch DR, Podmore RG, Anderson TP, Barratt K, Maze MJ, French KE, et al. Impact of routine systematic polymerase chain reaction testing on case finding for Legionnaires' disease: a pre-post comparison study. Clin Infect Dis. 2013;57(9):1275-81.

39. Chaudhry R, Sreenath K, Agrawal SK, Valavane A. Legionella and Legionnaires' disease: Time to explore in India. Indian journal of medical microbiology. 2018;36(3):324-33.

40. Woodhead M, Blasi F, Ewig S, Garau J, Huchon G, Ieven M, et al. Guidelines for the management of adult lower respiratory tract infections--full version. Clin Microbiol Infect. 2011;17 Suppl 6:E1-59.

41. Pedro-Botet ML, Yu VL. Treatment strategies for Legionella infection. Expert Opin Pharmacother. 2009;10(7):1109-21.

42. Bruin JP, Koshkolda T, EP IJ, Luck C, Diederen BM, Den Boer JW, et al. Isolation of ciprofloxacin-resistant Legionella pneumophila in a patient with severe pneumonia. J Antimicrob Chemother. 2014;69(10):2869-71.

43. Burdet C, Lepeule R, Duval X, Caseris M, Rioux C, Lucet JC, et al. Quinolones versus macrolides in the treatment of legionellosis: a systematic review and meta-analysis. J Antimicrob Chemother. 2014;69(9):2354-60.

44. Arnold FW, Summersgill JT, Lajoie AS, Peyrani P, Marrie TJ, Rossi P, et al. A worldwide perspective of atypical pathogens in community-acquired pneumonia. American journal of respiratory and critical care medicine. 2007;175(10):1086-93.

45. Cassini A, Colzani E, Pini A, Mangen MJ, Plass D, McDonald SA, et al. Impact of infectious diseases on population health using incidence-based disability-adjusted life years (DALYs): results from the Burden of Communicable Diseases in Europe study, European Union and European Economic Area countries, 2009 to 2013. Euro Surveill. 2018;23(16).

46. European Centre for Disease Prevention and Control. Annual epidemiological report for 2017 – Legionnaires' disease. Available from: https://www.ecdc.europa.eu/sites/portal/files/documents/AER_for_2017-Legionnaires-disease_1.pdf.

47. Dooling KL, Toews KA, Hicks LA, Garrison LE, Bachaus B, Zansky S, et al. Active Bacterial Core Surveillance for Legionellosis – United States, 2011-2013. MMWR Morb Mortal Wkly Rep. 2015;64(42):1190-3.

48. Cassell K, Gacek P, Rabatsky-Ehr T, Petit S, Cartter M, Weinberger DM. Estimating the True Burden of Legionnaires' Disease. Am J Epidemiol. 2019;188(9):1686-94.

49. Marston BJ, Lipman HB, Breiman RF. Surveillance for Legionnaires' disease. Risk factors for morbidity and mortality. Arch Intern Med. 1994;154(21):2417-22.

50. Ginevra C, Duclos A, Vanhems P, Campese C, Forey F, Lina G, et al. Host-related risk factors and clinical features of community-acquired legionnaires disease due to the Paris and Lorraine endemic strains, 1998-2007, France. Clin Infect Dis. 2009;49(2):184-91.

51. Cooley LA, Pondo T, Francois Watkins LK, Shah P, Schrag S, Active Bacterial Core Surveillance Program of the Emerging Infections Program N. Population-Based Assessment of Clinical Risk Factors for Legionnaires' Disease. Clin Infect Dis. 2019.

52. Farnham A, Alleyne L, Cimini D, Balter S. Legionnaires' disease incidence and risk factors, New York, New York, USA, 2002-2011. Emerg Infect Dis. 2014;20(11):1795-802.

53. Chen NT, Chen MJ, Guo CY, Chen KT, Su HJ. Precipitation increases the occurrence of sporadic legionnaires' disease in Taiwan. PLoS One. 2014;9(12):e114337.

54. Han BS, Lee MJ, Kwon YH, Lee WC. A Comparative Study of the Epidemiological Aspects of Legionnaires' Disease: Outbreaks in Korea and Japan, 2010-2014. Journal of clinical medicine research. 2017;9(1):67-70.

55. Chidiac C, Che D, Pires-Cronenberger S, Jarraud S, Campese C, Bissery A, et al. Factors associated with hospital mortality in community-acquired legionellosis in France. Eur Respir J. 2012;39(4):963-70.

56. Giesecke J. Modern infectious disease epidemiology. 3rd ed. ed. Boca Raton, FL: CRC Press; 2017.

57. Lettinga KD, Verbon A, Nieuwkerk PT, Jonkers RE, Gersons BP, Prins JM, et al. Health-related quality of life and posttraumatic stress disorder among survivors of an outbreak of Legionnaires disease. Clin Infect Dis. 2002;35(1):11-7.

58. Sonder GJ, van den Hoek JA, Bovee LP, Aanhane FE, Worp J, Du Ry van Beest Holle M, et al. Changes in prevention and outbreak management of Legionnaires disease in the Netherlands between two large outbreaks in 1999 and 2006. Euro Surveill. 2008;13(38).

59. Shands KN, Ho JL, Meyer RD, Gorman GW, Edelstein PH, Mallison GF, et al. Potable water as a source of Legionnaires' disease. JAMA. 1985;253(10):1412-6.

60. Greig JE, Carnie JA, Tallis GF, Ryan NJ, Tan AG, Gordon IR, et al. An outbreak of Legionnaires' disease at the Melbourne Aquarium, April 2000: investigation and case-control studies. The Medical journal of Australia. 2004;180(11):566-72.

61. Bennett E, Ashton M, Calvert N, Chaloner J, Cheesbrough J, Egan J, et al. Barrow-in-Furness: a large community legionellosis outbreak in the UK. Epidemiol Infect. 2014;142(8):1763-77.

62. Okada M, Kawano K, Kura F, Amemura-Maekawa J, Watanabe H, Yagita K, et al. [The largest outbreak of legionellosis in Japan associated with spa baths: epidemic curve and environmental investigation]. Kansenshogaku zasshi The Journal of the Japanese Association for Infectious Diseases. 2005;79(6):365-74.

63. Simonsen O, Wedege E, Kanestrom A, Bolstad K, Aaberge IS, Ragnhildstveit E, et al. Characterization of the extent of a large outbreak of Legionnaires' disease by serological assays. BMC Infect Dis. 2015;15:163.

64. Castilla J, Barricarte A, Aldaz J, Garcia Cenoz M, Ferrer T, Pelaz C, et al. A large Legionnaires' disease outbreak in Pamplona, Spain: early detection, rapid control and no case fatality. Epidemiol Infect. 2008;136(6):823-32.

65. Levesque S, Plante PL, Mendis N, Cantin P, Marchand G, Charest H, et al. Genomic characterization of a large outbreak of Legionella pneumophila serogroup 1 strains in Quebec City, 2012. PLoS One. 2014;9(8):e103852.

66. Lapierre P, Nazarian E, Zhu Y, Wroblewski D, Saylors A, Passaretti T, et al. Legionnaires' Disease Outbreak Caused by Endemic Strain of Legionella pneumophila, New York, New York, USA, 2015. Emerg Infect Dis. 2017;23(11):1784-91.

67. Bull M, Hall IM, Leach S, Robesyn E. The application of geographic information systems and spatial data during Legionnaires disease outbreak responses. Euro Surveill. 2012;17(49).

68. Nygard K, Werner-Johansen O, Ronsen S, Caugant DA, Simonsen O, Kanestrom A, et al. An outbreak of legionnaires disease caused by long-distance spread from an industrial air scrubber in Sarpsborg, Norway. Clin Infect Dis. 2008;46(1):61-9.

69. Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. PLoS Med. 2011;8(8):e1001083.

70. Garner E, Brown CL, Schwake DO, Rhoads WJ, Arango-Argoty G, Zhang L, et al. Comparison of Whole-Genome Sequences of Legionella pneumophila in Tap Water and in Clinical Strains, Flint, Michigan, USA, 2016. Emerg Infect Dis. 2019;25(11):2013-20.

71. Weiss D, Boyd C, Rakeman JL, Greene SK, Fitzhenry R, McProud T, et al. A Large Community Outbreak of Legionnaires' Disease Associated With a Cooling Tower in New York City, 2015. Public health reports. 2017;132(2):241-50.

72. Beaute J, Zucs P, de Jong B. Risk for travel-associated legionnaires' disease, Europe, 2009. Emerg Infect Dis. 2012;18(11):1811-6.

73. Rota MC, Cano Portero R, Che D, Caporali MG, Hernando V, Campese C. Clusters of travel-associated Legionnaires disease in Italy, Spain and France, July 2002 – June 2006. Euro Surveill. 2007;12(11):E3-4.

74. Mouchtouri VA, Rudge JW. Legionnaires' Disease in Hotels and Passenger Ships: A Systematic Review of Evidence, Sources, and Contributing Factors. J Travel Med. 2015;22(5):325-37.

75. Ricketts KD, Slaymaker E, Verlander NQ, Joseph CA. What is the probability of successive cases of Legionnaires' disease occurring in European hotels? Int J Epidemiol. 2006;35(2):354-60.

76. Ricketts KD, Yadav R, Rota MC, Joseph CA, European Working Group for Legionella I. Characteristics of reoffending accommodation sites in Europe with clusters of Legionnaires disease, 2003-2007. Euro Surveill. 2010;15(40).

77. Sabria M, Yu VL. Hospital-acquired legionellosis: solutions for a preventable infection. Lancet Infect Dis. 2002;2(6):368-73.

78. Blatt SP, Parkinson MD, Pace E, Hoffman P, Dolan D, Lauderdale P, et al. Nosocomial Legionnaires' disease: aspiration as a primary mode of disease acquisition. Am J Med. 1993;95(1):16-22.

79. Allen JG, Myatt TA, Macintosh DL, Ludwig JF, Minegishi T, Stewart JH, et al. Assessing risk of health care-acquired Legionnaires' disease from environmental sampling: the limits of using a strict percent positivity approach. Am J Infect Control. 2012;40(10):917-21.

80. Skerrett SJ. Prevention of Health Care-Associated Legionnaires Disease. JAMA Netw Open. 2018;1(2):e180232.

81. Demirjian A, Lucas CE, Garrison LE, Kozak-Muiznieks NA, States S, Brown EW, et al. The importance of clinical surveillance in detecting legionnaires' disease outbreaks: a large outbreak in a hospital with a Legionella disinfection system-Pennsylvania, 2011-2012. Clin Infect Dis. 2015;60(11):1596-602.

82. Hamilton KA, Prussin AJ, 2nd, Ahmed W, Haas CN. Outbreaks of Legionnaires' Disease and Pontiac Fever 2006-2017. Curr Environ Health Rep. 2018.

83. Helbig JH, Uldum SA, Bernander S, Luck PC, Wewalka G, Abraham B, et al. Clinical utility of urinary antigen detection for diagnosis of community-acquired, travel-associated, and nosocomial legionnaires' disease. J Clin Microbiol. 2003;41(2):838-40.

84. Commission Implementing Decision 2012/506/EU of 8 August 2012 amending Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision No 2119/98/EC of the European Parliament and of the Council.

85. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Euro Surveill. 2006;11(12):212-4.

86. Joseph CA, Ricketts KD. From development to success: the European surveillance scheme for travel associated Legionnaires' disease. Eur J Public Health. 2007;17(6):652-6.

87. Rota MC, Bella A, Caporali MG, Nicolau A, Drasar V, Ricci ML, et al. Travel-associated Legionnaires' disease: would changing cluster definition lead to the prevention of a larger number of cases? Epidemiol Infect. 2018:1-6.

88. Robesyn E, Payne Hallstrom L, Young J, de Jong B. Timeliness of travel-associated Legionnaires' disease surveillance. Cent Eur J Public Health. 2018;26(2):154-5.

89. Van Kenhove E, Dinne K, Janssens A, Laverge J. Overview and comparison of Legionella regulations worldwide. Am J Infect Control. 2019.

90. Hayes-Phillips D, Bentham R, Ross K, Whiley H. Factors Influencing Legionella Contamination of Domestic Household Showers. Pathogens. 2019;8(1).

91. Garrison LE, Kunz JM, Cooley LA, Moore MR, Lucas C, Schrag S, et al. Vital Signs: Deficiencies in Environmental Control Identified in Outbreaks of Legionnaires' Disease – North America, 2000-2014. MMWR Morb Mortal Wkly Rep. 2016;65(22):576-84.

92. Surveillance systems overview [Internet]. 2018 [cited 8 March 2019]. Available from: https://ecdc.europa.eu/en/publications-data/surveillance-systems-overview-2017.

93. Leder K, Wilson ME, Freedman DO, Torresi J. A comparative analysis of methodological approaches used for estimating risk in travel medicine. J Travel Med. 2008;15(4):263-72.

94. Annual data on trips of EU residents [Internet]. [cited 21/03/2017]. Available from: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tour_dem_tnw&lang=en.

95. Eurostat. Methodological manual for tourism statistics – 2014 edition. Luxembourg: Publications Office of the European Union; 2015.

96. Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, et al. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. International Journal of Climatology. 2002;22(12):1441-53.

97. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. ed. Malden, Mass.: Blackwell Science; 2003.

98. Pawitan Y. In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford: Oxford University Press; 2013.

99. Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: Wiley; 1980.

100. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. Int J Epidemiol. 2013;42(4):1187-95.

101. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis Part III: multivariate data analysis -- choosing a model and assessing its adequacy and fit. Br J Cancer. 2003;89(4):605-11.

102. Beaute J, Sandin S, de Jong B, Hallstrom LP, Robesyn E, Giesecke J, et al. Factors associated with Legionnaires' disease recurrence in hotel and holiday rental accommodation sites. Euro Surveill. 2019;24(20).

103. Beaute J. Travel-associated infections in Europe. Lancet Infect Dis. 2015;15(8):878-9.

104. Steffen R, deBernardis C, Banos A. Travel epidemiology--a global perspective. Int J Antimicrob Agents. 2003;21(2):89-95.

105. European Centre for Disease P, Control. Surveillance Atlas of Infectious Diseases [Internet]. ECDC; [Available from: http://ecdc.europa.eu/en/data-tools/atlas/Pages/atlas.aspx.

106. Helbig JH, Bernander S, Castellani Pastoris M, Etienne J, Gaia V, Lauwers S, et al. Pan-European study on culture-proven Legionnaires' disease: distribution of Legionella pneumophila serogroups and monoclonal subgroups. Eur J Clin Microbiol Infect Dis. 2002;21(10):710-6.

107. Schlagenhauf P, Weld L, Goorhuis A, Gautret P, Weber R, von Sonnenburg F, et al. Travel-associated infection presenting in Europe (2008-12): an analysis of EuroTravNet longitudinal, surveillance data, and evaluation of the effect of the pre-travel consultation. Lancet Infect Dis. 2015;15(1):55-64.

108. Beaute J, Sandin S, Uldum SA, Rota MC, Brandsema P, Giesecke J, et al. Short-term effects of atmospheric pressure, temperature, and rainfall on notification rate of community-acquired Legionnaires' disease in four European countries. Epidemiol Infect. 2016:1-11.

109. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. ed. Philadelphia, Pa.: Lippincott Williams & Wilkins; 2008.

110. Stypulkowska-Misiurewicz H, Czerwinski M. Legionellosis in Poland in 2017. Przegl Epidemiol. 2019;73(2):151-5.

111. Tomova I, Marinov R, Maeva I. First cluster of travel-associated legionnaires' disease detected in Bulgarian citizens. Euro Surveill. 2007;12(4):E070412 2.

112. Van Hest NA, Hoebe CJ, Den Boer JW, Vermunt JK, Ijzerman EP, Boersma WG, et al. Incidence and completeness of notification of Legionnaires' disease in The Netherlands: covariate capture-recapture analysis acknowledging regional differences. Epidemiol Infect. 2008;136(4):540-50.

113. Rota MC, Cawthorne A, Bella A, Caporali MG, Filia A, D'Ancona F, et al. Capture-recapture estimation of underreporting of legionellosis cases to the National Legionellosis Register: Italy 2002. Epidemiol Infect. 2007;135(6):1030-6.

114. Nardone A, Decludt B, Jarraud S, Etienne J, Hubert B, Infuso A, et al. Repeat capture-recapture studies as part of the evaluation of the surveillance of Legionnaires' disease in France. Epidemiol Infect. 2003;131(1):647-54.

115. European Centre for Disease Prevention and Control. Data quality monitoring and surveillance system evaluation : a handbook of methods and applications. Stockholm: ECDC; 2014. Available from: http://dx.doi.org/10.2900/35329.

116. European Centre for Disease Prevention and Control. Managing heterogeneity when pooling data from different surveillance systems. Stockholm: ECDC; 2019. Available from: http://dx.doi.org/10.2900/83039.

117. Campese C, Descours G, Lepoutre A, Beraud L, Maine C, Che D, et al. Legionnaires' disease in France. Med Mal Infect. 2015;45(3):65-71.

118. Behrens RH, Carroll B. The challenges of disease risk ascertainment using accessible data sources for numbers of travelers. J Travel Med. 2013;20(5):296-302.

119. Tataryn J, Vrbova L, Drebot M, Wood H, Payne E, Connors S, et al. Travel-related Zika virus cases in Canada: October 2015-June 2017. Can Commun Dis Rep. 2018;44(1):18-26.

120. Sedgwick P. Ecological studies: advantages and disadvantages. BMJ. 2014;348:g2979.

121. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annual review of public health. 1997;18:83-104.

122. Herwaldt LA, Marra AR. Legionella: a reemerging pathogen. Curr Opin Infect Dis. 2018;31(4):325-33.

123. Steffen R, Amitirigala I, Mutsch M. Health risks among travelers--need for regular updates. J Travel Med. 2008;15(3):145-6.

124. Rubin R. Why Are Legionnaires Disease Diagnoses Becoming More Common in the United States? JAMA. 2018;319(17):1753-4.

125. Walker JT. The influence of climate change on waterborne disease and Legionella: a review. Perspect Public Health. 2018;138(5):282-6.

Karolinska Institutet

Arkitektkopia AB, Bromma, 2019