Trends in **Microbiology**

CellPress

Scientific Life

LegionellaDB – A Database on *Legionella* Outbreaks

Inês G. Gonçalves,^{1,2} Henrique S. Fernandes,³ Ana Melo,¹ Sérgio F. Sousa,^{3,*} Lúcia C. Simões,⁴ and Manuel Simões ^{5,*}

LegionellaDB is the first database on *Legionella* outbreaks; it is based on a metadata analysis of peerreviewed manuscripts from PubMed and SCOPUS. LegionellaDB is dynamic and extensible, allowing users to search for specific outbreaks, suggest additional information to be included after curation, visualize statistical representations on specific outbreaks, and download selected data. The database is maintained online.

Legionella – The Source

Legionella is a bacterium, commonly found in water systems, that has been largely associated with pneumonia outbreaks since its discovery, when it was identified as the causative agent of an outbreak in a hotel in Philadelphia, in 1976, which affected 182 people and resulted in 29 deaths [1]. In its natural environment, Legionella does not present a high risk of causing infections to humans. However, when aerosolized and inhaled, the bacteria can cause Legionnaire's disease (LD) and Pontiac fever [2]. Currently, there are reports of more than 61 Legionella species and diverse serogroups among the species; Legionella pneumophila serogroup 1 is recognized to be the leading cause of LD, accounting for 90% of the total cases [3], although L. micdadei, L. bozemanii, L. dumoffii, and L. longbeachae are also known to be associated with LD [4].

The legionellae are Gram-negative rods commonly found in lakes, rivers, and other natural water systems as well as in soils [5]. Moreover, engineered structures, mainly hot-water systems, cooling towers, and air-conditioning systems, are also recognized as sources of Legionella. These provide excellent conditions for Legionella development [6], with temperatures between 20°C and 50°C, pH between 5.5 and 9.2, and the presence of iron salts. Additionally, dead legs and other surfaces favoring the presence of biofilms present optimal environments for Legionella species [6,7]. Protozoan grazing is another phenomenon relevant for Legionella growth and proliferation in stagnant waters suffering from inadequate disinfection [7]. Therefore, effective disinfection plans are fundamental to control Legionella in the bulk water, internalized in protozoa and in biofilms.

Despite being discovered after an outbreak in 1976, subsequent studies have shown that Legionella caused other pneumonia outbreaks in the past, dating back to 1949 [8]. Among these, there was a Legionella outbreak in 1974 in the same location as the 1976 event. It is now recognized that Legionella outbreaks may be under-reported and more prevalent than expected [9]. Nowadays, health practitioners and authorities have become more informed about LD, with a consequent increase in the number of cases reported [10]. In Europe, between 2011 and 2015, more than 30 000 cases were reported, of which 70% correspond to four countries: France, Germany, Italy, and Spain [11]. This increase in the number of outbreaks and cases is being recognized as related to climate change eliciting Legionella growth, and population aging [12]. Moreover, new techniques, such as sequence-based typing [13] and monoclonal antibody detection [14], have been essential to establish a link between clinical and environmental strains, enabling accurate identification of the source of the outbreak.

In recent years, the amount of data on Legionella outbreaks has been increasing significantly and the relatively extensive peer-reviewed literature on the topic motivated the development of LegionellaDB. This database provides the academic community, professionals in sectors affected by this problem (e.g., the industry and health sectors, and regulatory entities), and the public in general with a platform to search for outbreaks reported in peer-reviewed manuscripts and to collect specific information on a number of data (i.e., number of infections and deaths, place of occurrence, season, source, the analysis of Legionella in biofilms, the treatment applied prior to the outbreak, the action in response to the outbreak, in addition to microbiological and molecular information on the bacteria). The database further allows users to download data and suggest the inclusion of new outbreaks.

Organization of the Database

LegionellaDBⁱ provides geographic, environmental, biological, process, and public health data from *Legionella* outbreaks reported in peer-reviewed manuscripts. The database aims to be dynamic by allowing users to suggest new information that will be included after curation, as long as a valid reference from a peer-reviewed manuscript is associated with it.

Database Implementation

PRISMA (preferred reporting items for systematic reviews and meta-analyses), an evidence-based minimum set of items for reporting in systematic reviews and metaanalyses, served as the guide for both data collection and analysis [15] using PubMed and SCOPUS as the source of information. Furthermore, search results were filtered using the keywords 'Legionella' OR 'legionnaires disease' AND 'outbreak', and by selecting reports published up to December 2020. Peer-reviewed manuscripts were evaluated for eligibility, initially based on the title, then on the abstract,

CellPress

and finally on the full text. The data were organized in a relational PostgreSQL database. The database was integrated in a web platform using PHP (version 7.0) to enhance the accessibility and analysis of the available data.

The database is composed of two backend tables that work together to provide the information in the web interface accessible to the user. The main table includes the year, place, city, country, season, number of cases, confirmed cases, and deaths, and the mean age of the patients. Furthermore, it showcases available data on the Legionella strains (clinical and environmental) identified in each outbreak, including their corresponding serogroups, sequencebased typing of strains, monoclonal antibody types, and whether there is a match between the clinical and environmental strains. Finally, the source of the outbreak, the maintenance plan, the disinfection treatment applied before and/or during the outbreak, and the source of information (DOI or PMID) are also listed in the database. Since some publications refer to multiple outbreaks, the bibliographic information was stored in a second backend table to avoid redundancy. Therefore, DOI or PMID were used to populate this second backend table, including the names of authors, title, journal, and year of each publication. When a page on a specific outbreak is opened, the code implementation collects data from both backend tables to present all the available information about the outbreak.

All the graphs displayed in the LegionellaDB website are automatically generated based on the information retrieved from the database using the data visualization JavaScript library Chart.js (version 2.5.0). Chart.js is available as an open-source library under the MIT license available at www.chartjs.org. This is also valid for the maps, which provide a general overview of all the outbreaks and an easy access to the available information. This

LegionellaDB						Home Explore	Map Statistics	More
	Search outbreaks					Search		
	The query returns 161 results.							
T Filter Results	Year	City	Country	Place	Cases	Source	Species/Serogroup	Mor
Since Year	2019	Quebec	Canada	Community	14	Unknown	Lp1	Q
Minimum Cases	2018	Toronto	Canada	Hospital	9	Unknown	Lp1	۹
City	2018	England	UK	Public Bathhouse	115	Water System	Unknown	۹
Country	2018	Modena	Italy	Hospital	3	Unknown	Lp1	۹
Select Country \$	2018	Bresso	Italy	Community	52	Water System	Lp1	Q
Place	2017	Geneve	Switzerland	Hospital	34	Unknown	Lp1	۹
Select Place \$	2017	Basel	Switzerland	Community	27	Cooling Tower	Lp1	Q
Source Select Source ¢	2017	Boxtel	Netherlands	Community	8	Biologic Wastewater	Lp1	۹
Species/Serogroup						Treatment Plant		
Select Species/: \$	2016	Boxtel	Netherlands	Community	6	Biologic Wastewater Treatment Plant	Lp1	۹
Search	2016	Sydney	Australia	Community	9	Unknown	Lp1	۹
Export Selection	0010	Sudnov	Australia	Community	e	Linknown	1.01	0



Trends in Microbiology

Figure 1. Search for Outbreaks. (A) The 'Explore' page enables users to explore the LegionellaDB database, showcasing a table with all cases collected in the platform. Additionally, it is possible to search for specific outbreaks using either the search bar, at the top, or the available filters. (B) Outbreak Data. Through the 'Explore' page, users can select a specific outbreak to get more detailed information that is not available in the general table. This new page includes the number of deaths, the season of occurrence, the mean age of the patients, the source of the infection, the type of treatment used before and upon the outbreak, and whether biofilms and protozoa were identified in the source of infection. In addition, the page contains a map with a marker for the city of occurrence as well as a radar chart that compares the numerical data of the outbreak with the average values for the outbreaks present in LegionellaDB. Moreover, users can also get information on the source manuscripts.

Trends in Microbiology



representation is accomplished using the free version of a visualization JavaScript library, amCharts (version 4, copyrighted by @amcharts). Both amCharts and Chart.js were implemented in LegionellaDB as standalone libraries, following the respective license agreements, avoiding external dependencies and possible downtimes caused by the unavailability of external servers.

Database Functionalities Outbreak Search

The 'Explore' page provides the user with several possibilities to search for outbreaks through the 'Filter Results' functionality (Figure 1A). The user can choose from a variety of filters (year, city, country, place of occurrence, source, and strain) to narrow the results according to their aims. Each row of the table corresponds to a single outbreak where information on the year, city, country, place of occurrence, source, and strain is displayed. Nonetheless, by selecting one of the rows, the user is taken to a new page which comprises all the information on that specific outbreak. The reference(s) and link(s) for the source manuscript(s) are also provided (Figure 1B).

Visualization of Spatial Information Map Visualization

The 'Map' page showcases spatial information on the outbreaks registered by representing the geographic location of occurrence (Figure 2A). The locations are marked with circles whose radius changes according to the number of cases capturing the magnitude of each outbreak.

Statistics

In the 'Statistics' page (Figure 2B), the user can visualize information regarding the outbreaks registered in the database, accompanied by general information on Legionella and its main risk factors. In particular, this page presents charts elucidative on the evolution of Legionella outbreaks,





Trends in Microbiology

Figure 2. Data Visualization. (A) The 'Map' page allows users to assess where in the world Legionella outbreaks are more frequent and reported in peer-reviewed manuscripts. Users can interact with the markers to get more details on the outbreak (year and city of occurrence and number of cases) and can choose between a map and a globe view. (B) The 'Statistics' page offers multiple data representations, providing the users with trends on Legionella outbreaks, their effects, and their main causes.

their impact in terms of number of cases lead to Legionella outbreaks, and the and deaths, the places where outbreaks sources that lead to more severe tend to occur, the main risk factors that

outbreaks.

Trends in Microbiology

CellPress

User Interaction

One of the unique features of LegionellaDB is the 'Suggest and Download Data' functionality available in the 'More' page, wherein users can complete a form to suggest new data to be included in the database. The user is required to indicate the DOI of the peer-reviewed manuscript(s) from which the data were collected and to provide a valid email address. This step critically relies on the user's interest to collaborate. Finally, users are offered the opportunity to download the entire database as a CSV file.

Critical Analysis

An overall data analysis from the 'Statistics' page proposes that the number of outbreaks has increased over the years. The numbers of related infections and deaths are lower than those observed in the first years that were reported. This is due, in part, to improvements in diagnostics and the treatment of Legionella infections. It is further evident that L. pneumophila has the highest prevalence among all Legionella species, in particular strains of serogroup 1 (Lp1), with a higher incidence in age groups between 50 and 80, and in seasons of the year where the temperature is higher. Although natural water systems are thought to be critical sources of Legionella species, the number of outbreaks reported with an origin from that source is significantly lower than from cooling towers and air-treatment systems. This is attributed to a lack of, or complete absence, of management and disinfection plans for the process and the equipment, as highlighted for diverse outbreaks listed in the database. Such conditions are critical for the planktonic proliferation of Legionella in the bulk water and/or internalized within a protozoan host as well as for biofilm formation and harboring of Legionella. LegionellaDB shows that microbiological sampling of biofilms from surfaces, and protozoa analysis, are typically neglected when searching for the source of an outbreak - even though the presence of biofilms on surfaces, and the existence of protozoa, are recognized by the scientific community for their role in the ability of *Legionella* to thrive in environmental and industrial systems [6,7].

Concluding Remarks

LegionellaDB is the first database on Legionella outbreaks. It provides geographic, environmental, biological, process, and public health data from over 160 outbreaks reported in peer-reviewed manuscripts. This aims to be a continually updated resource, allowing users to visualize selected data on each outbreak - both online and as a downloadable spreadsheet. While the database provides the academic community, and professionals working on Legionella, with information from a range of peer-reviewed reports on the outbreaks, it also depends on the users to make this web platform a dynamic and extensible resource by suggesting new data for inclusion following curation.

Acknowledgments

This work was financially supported by Base Funding -UIDB/00511/2020 of LEPABE and UIDB/00081/2020 of CIQUP funded by national funds through the FCT/ MCTES (PIDDAC); Project Biocide_for_Biofilm - PTDC/ BII-BTI/30219/2017 - POCI-01-0145-FEDER-030219, ABFISH - PTDC/ASP-PES/28397/2017 - POCI-01-0145-FEDER-028397 and ALGAVALOR - POCI-01-0247-FEDER-035234, funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) and by national funds (PIDDAC) through FCT/MCTES; Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UIDB/04469/2020 unit and BioTecNorte operation (NORTE-01-0145-FEDER-000004) funded by the European Regional Development Fund under the scope of Norte2020 - Programa Operacional Regional do Norte; Applied Molecular Biosciences Unit -UCIBIO, which is financed by national funds from FCT (UIDB/04378/2020). H.S.F. acknowledges FCT for his PhD grant SFRH/BD/115396/2016.

Declaration of Interests

There are no interests to declare.

Resources

ⁱhttps://legionelladb.biosim.pt/

¹Faculty of Engineering, Department of Chemical Engineering, University of Porto, Porto, Portugal

²M2BE, Aragon Institute of Engineering Research, Mechanical Engineering Department, University of Zaragoza, Zaragoza, Spain

³UCIBIO/REQUIMTE, BioSIM, Departamento de Biomedicina, Faculdade de Medicina da Universidade do Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal ⁴CEB, Centre of Biological Engineering, University of Minho,

"CEB, Centre of Biological Engineering, University of Minho, Campus de Gualtar, Braga, Portugal

⁵LEPABE, Faculty of Engineering, Department of Chemical Engineering, University of Porto, Porto, Portugal

*Correspondence:

sergiofsousa@med.up.pt (S.F. Sousa) and mvs@fe.up.pt (M. Simões).

https://doi.org/10.1016/j.tim.2021.01.015

© 2021 Elsevier Ltd. All rights reserved.

References

- 1. Fraser, D. et al. (1977) Legionnaires' disease. N. Engl. J. Med. 297, 1189–1197
- Diederen, B. (2008) Legionella spp. and Legionnaires' disease. J. Infect. 56, 1–12
- Fields, B. (1996) The molecular ecology of legionellae. Trends Microbiol. 4, 286–290
- Fliermans, C. (1996) Ecology of *Legionella*: from data to knowledge with a little wisdom. *Microb. Ecol.* 32, 203–228
- van Heijnsbergen, E. et al. (2015) Confirmed and potential sources of *Legionella* reviewed. *Environ. Sci. Technol.* 49, 4797–4815
- Dilger, T. et al. (2018) Legionella contamination in warm water systems: A species-level survey. Int. J. Hyg. Environ. Health 221, 199–210
- Taylor, M. et al. (2009) Legionella, protozoa, and biofilms: interactions within complex microbial systems. Microb. Ecol. 58, 538–547
- Armstrong, C. and Miller, G. (1985) A 1949 outbreak of Pontiac fever-like illness in steam condenser cleaners. *Arch. Environ. Health* 40, 26–29
- Cunha, B. *et al.* (2016) Legionnaires' disease. *Lancet* 387, 376–385
- Parr, A. et al. (2015) Legionellosis on the rise. J. Public Health Manag. Pract. 21, E17–E26
- 11. Beauté, J. (2017) Legionnaires' disease in Europe, 2011 to 2015. Euro Surveill. 22, 30566
- Walker, J. (2018) The influence of climate change on waterborne disease and *Legionella*: a review. *Perspect. Public Health* 138, 282–286
- Gaia, V. et al. (2005) Consensus sequence-based scheme for epidemiological typing of clinical and environmental isolates of Legionella pneumophila. J. Clin. Microbiol. 43, 2047–2052
- Helbig, J.H. et al. (1997) Antigenic lipopolysaccharide components of *Legionella pneumophila* recognized by monoclonal antibodies: possibilities and limitations for division of the species into serogroups. J. Clin. Microbiol. 35, 2841–2845
- Shamseer, L. *et al.* (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 349, g7647-g7647