Applied Epidemiology in Victoria

Thesis submitted for the degree of
Masters of Philosophy (Applied Epidemiology)
Australian National University

Zoe Cutcher

BVSc(hons) MPH

December 2015

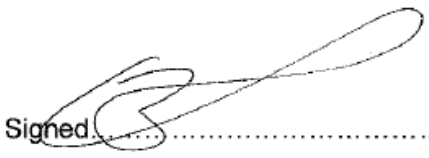
Academic Supervisor
Associate Professor Martyn Kirk

Field Supervisors
Ms Nicola Stephens & Mr Noel Cleaves

This page is intentionally blank

Originality statement

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at ANU or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at the Victorian Government Department of Health and Human Services or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation or linguistic expression is acknowledged'.



Signed.....

Date 08/12/1979

This page is intentionally blank

Acknowledgements

I wish to acknowledge my field supervisors Nicola Stephens and Noel Cleaves who provided excellent practical and strategic assistance during my field placement. Thank you both for your keen interest in my projects and unwavering flexibility and support. I'm also grateful to my academic supervisor Martyn Kirk who provided rigorous academic input, arranged my deployment to Vanuatu, and shared valuable time musing over my direction in career and life.

It has been a great privilege to be welcomed as one of the Communicable Disease Epidemiology and Surveillance team members. Thanks to my colleagues for friendship and sharing your professional expertise throughout my placement. In particular, I wish to thank Lucinda Franklin for sharing her role as *Legionella* epidemiologist and generously donating large amounts of time to brainstorm during my evaluation of the legionellosis surveillance system. I also wish to acknowledge Marion Easton and Nicola Stephens for making the *Salmonella* whole genome sequencing project such a fun journey of discovery. I wish to thank Joy Gregory whose ability to cut to the core of an issue made her an excellent sounding board.

I am indebted to the *Legionella* team, in particular Stuart Adcock for welcoming me to the team and teaching me about environmental control of *Legionella*. I was touched that the team displayed such great enthusiasm for the planned "*Legionella* Hot Water Survey" (HoWS study), and shared my disappointment when the study was abandoned. I'd also like to acknowledge the staff in Communicable Disease Prevention and Control for making me feel welcome, in particular Fran Tiplady who was a fantastic source of insight during the legionellosis surveillance system evaluation.

There are many others I wish to acknowledge, including Ben Howden, Glen Carter, Dieter Bulach and Tim Stinear from the Doherty Centre for Applied Microbial Genomics; Yvan Soares and Jojo Merilles from Secretariat of the Pacific Community; and Hellenson Bani and Markson Tabi from Lolowai Provincial Hospital, Vanuatu.

Above all, I wish to thank Jacqueline Davis for her patience and encouragement, for embracing weeks of single parenting, and for the gift of "thesis boot-camp" weekends. Jacs, I dedicate this volume to you.

Table of Contents

Abstract

The Health Protection Branch of the Victorian Government Department of Health and Human Services monitors and responds to incidents that could adversely affect the health of Victorians. During 2014-2015, I completed a field placement with the branch, assisting with numerous public health investigations and responses. In doing so I fulfilled the requirements of the Master of Philosophy in Applied Epidemiology (MAE). The skills I gained are demonstrated in this thesis.

Evaluation of a public health surveillance system is a core requirement for the MAE program. I evaluated Victoria's surveillance and response to legionellosis, which includes both disease surveillance and environmental surveillance and response arms. I found little evidence to support the current practice of sampling and disinfecting cooling towers around the home and workplace for sporadic cases. Improved co-ordination between databases and strategic use of spatial software could help develop more targeted and useful approaches in the future.

I embarked on two epidemiological projects. I designed a cross sectional study examining the prevalence of *Legionella* in domestic potable water and developed participant resources including letters to explain results, meeting the MAE requirement to communicate findings to a non-scientific audience. The study was not completed due to legal considerations; however the proposal and relevant participant resources are included as an appendix.

I completed an epidemiological project estimating the number of notified sporadic *Salmonella* Typhimurium 9 Phage type 9 cases likely to be associated with a recurrent outbreak source during a five year period. I examined 301 clinical *Salmonella* isolates, including sporadic and outbreak isolates from a series of linked outbreaks, and used multi-locus variable number tandem repeat analysis and whole genome sequence results to estimate the number of isolates genetically linked to the outbreak strain. Outbreak cases accounted for just one third of all isolates estimated to be closely related to the main outbreak clade. This project inspired my lesson from the field, in which I taught MAE colleagues how to analyse MLVA data.

I investigated an outbreak of *Salmonella* Typhimurium phage type 44 at a school function. I conducted a cohort study and interviewed twenty-nine out of thirty guests, of which ten were affected. Roast beef appetiser was the most likely food vehicle for

Salmonella infection. Cross-contamination from raw eggs during preparation was a possible source.

I analysed a public health dataset to assist a public health investigation into suspected antimony exposure in a rural mining town in Victoria. Residents were concerned about potential health effects from exposure to antimony dust from a local mine. Many sought urinary antimony testing to quantify exposure, with numerous elevated results. I used multivariate regression to examine risk factors for elevated urinary antimony and demonstrated residential proximity to the mine was not associated with urinary antimony results. Overwhelmingly, the largest risk factor for elevated results was the month of testing, consistent with false positive laboratory reports.

This thesis documents my experience and capabilities gained during the MAE program, and demonstrates my contribution to protecting the public health of Victorians.

Chapter 1. Summary of Field Experience

Contents

Introduction

My field placement for the Master of Philosophy in Applied Epidemiology (MAE) was in the Health Protection Branch of the Regulation, Health Protection and Emergency Management division of the Victorian Government Department of Health and Human Services. Within the branch the placement was divided between Communicable Disease Epidemiology and Surveillance and the Legionella Team, which sits within Environmental Health Regulation and Compliance. My experience included a three week deployment to Vanuatu to assist with disease surveillance following Tropical Cyclone Pam in March 2015.

Communicable Disease Epidemiology and Surveillance conduct surveillance for sixty five notifiable conditions in Victoria. Their work is focussed on data management, analysis and reporting for notifiable infectious diseases. They work closely with Communicable Disease Prevention and Control who collect data during public health follow up, and implement control measures to prevent or control outbreaks of infectious disease. The Legionella Team aims to reduce the burden of legionellosis on the Victorian community. Their work is divided between routine cooling tower inspections to ensure compliance with regulations, and investigation of potential environmental sources relating to notified cases of legionellosis in Victoria.

Field activities

The opportunities I encountered during my field placement enabled me to fulfil the requirements of the MAE. Beyond this, they provided me with unique insights into the workings of government, the challenges and opportunities presented by new technologies, and the central role epidemiologists play in a myriad of disease control activities. Through my placement with the Legionella team I developed significant knowledge and experience regarding the epidemiology, surveillance and control of *Legionella.* I was also involved in several other activities, including investigating relatedness between historical *Salmonella* isolates using genomic data*,* a field deployment to Vanuatu to support the establishment of an EWARN surveillance system following Tropical Cyclone Pam, investigating a variety of acute public health incidents, and assisting with Ebola preparedness. The skills I gained are demonstrated in this thesis.

During my placement I played an active role in Victoria's legionellosis disease surveillance system. This included coordinating information between the disease surveillance and environmental investigation systems, presenting surveillance reports at weekly surveillance meetings, troubleshooting database and mapping problems, assisting with outbreak investigations, and participating in the *Legionella* Set of National Guidelines (SONG) working group. I worked with a geographical information systems specialist to develop an automated geospatial workflow to rapidly detect legionellosis clusters and outbreaks and associated cooling towers. With the same GIS expert, I also explored methods to improve the use of GIS technology to strategically select cooling towers for inspection. These projects developed my expertise in the Geocortex (Esri™) mapping system, and throughout my placement I provided advice and training on its use to colleagues. I learnt that geospatial technologies are a powerful but underutilised tool for epidemiological analysis.

I evaluated Victoria's surveillance system for legionellosis. The evaluation, outlined in Chapter 2, led to several recommendations for improvement. These included developing an integrated database for disease and environmental data, and establishing a working group to improve laboratory methods for detection and typing of *Legionella*. As part of this work I collaborated with researchers from the Doherty Centre for Applied Microbial Analysis to investigate the utility of whole genome sequencing data for *Legionella* surveillance. The results are currently being prepared for publication. I also shared findings from my surveillance system evaluation with colleagues from the CDC and New York City Health Department to assist their response to a large legionellosis outbreak in New York City[1].

Through my work with the Legionella team I designed a cross sectional survey to investigate the prevalence of *Legionella* bacteria in residential hot water systems (the *Legionella* HoWS Survey). The project was developed to address concerns that non-cooling tower sources may account for a larger proportion of legionellosis cases than expected. The study was intended to inform sample size calculations for a future case control study investigating the association between *Legionella* contamination in residential hot water and illness with legionellosis. As part of this process, I developed a library of participant resources to assist with participant recruitment and result reporting, satisfying the MAE requirement to communicate results of a study to a lay audience.

[1]http://www.nbcnewyork.com/news/local/Legionnaires-Disease-New-York-City-Bronx-Cooling-Tower-Death-Sick-Outbreak-Source-Health-Department-322407892.html

Unfortunately, the HoWS survey was not pursued following receipt of departmental legal advice. Where *Legionella* was identified in a participant home, the study team had planned to advise participants they could choose whether or not to address it as the risk to health is likely negligible unless a person is in a high risk group for infection. However, we were advised landowners have a non-delegable legal duty of care to resolve the issue due to potential risk to people visiting the residence. In practice, resolving the issue is not always straightforward. It was decided the study findings may place an excessive burden on landowners, particularly because identifying landowners to gain consent for study participation may be problematic. For these reasons the study was not pursued. However the study protocol and results letters are included as an appendix to this volume (Appendix A). Note that the results letters reflect the team's planned approach to result reporting, where participants could use a 'risk based' approach to choose whether or not to address the issue. If the study is pursued in the future these resources need to be updated to reflect the landowner's non-delegable duty of care.

My epidemiological project appears in Chapter 3. The project has been prepared for publication and the late draft of an article for peer-reviewed publication makes up the body of the chapter. This project provided me a unique opportunity to engage with microbial genomics, a new frontier in communicable disease surveillance. I used multilocus variable-number tandem repeat analysis (MLVA) and whole genome sequence (WGS) data to investigate the number of notified salmonellosis cases potentially attributable to recurrent outbreak sources, with some interesting findings. Large scale WGS provides incredible discriminatory power and is likely to change the nature of public health surveillance over the next few years. Like any new technology it presents a variety of challenges. Investment of time and skill is required to identify the best use for the extensive data it generates. This project was one of Victoria's first forays into the use of this new technology for surveillance purposes, and provided a fantastic learning opportunity for me. The project inspired my lesson from the field, in which I taught MAE colleagues how to analyse MLVA data. This lesson is included as an appendix to this volume (Appendix C).

During my field placement I assisted with investigations into a variety of acute public health incidents. These included a number of legionellosis outbreaks, a suspected Zika virus outbreak in Vanuatu, a foodborne *Salmonella* outbreak, and suspected environmental exposure to excessive levels of antimony for residents in a regional

Victorian town. The salmonellosis outbreak and the antimony incident appear as Chapters 4 (outbreak investigation) and Chapter 5 (analysis of a public health data set) respectively. The salmonellosis outbreak has been prepared for publication, and the body of Chapter 4 is made up of the late draft of an article for peer reviewed publication. I presented the findings from the salmonellosis outbreak at the 2015 Communicable Disease Control conference in Brisbane; the abstract and slides are included as an appendix to Chapter 4.

The devastating West African Ebola epidemic during 2014-2015 led to a massive mobilisation of global public health effort, both for response at the source (if somewhat delayed) and local preparedness. I assisted with Victoria's preparedness activities, including assisting with the development of Victoria's case and contact questionnaires, preparing Victoria's disease surveillance database to capture this information, and evaluating a whole of Victorian Government Preparedness tabletop exercise. The experience provided some interesting insights into the challenges in developing a co-ordinated and timely national response in a system of federated states where each jurisdiction works within separate legislative frameworks.

In May 2015, I spent three weeks in Vanuatu on secondment with the Secretariat of the Pacific Community, where I provided technical assistance following Tropical Cyclone Pam's impact. I spent the majority of my deployment in the remote Penama Province where I provided capacity building for the establishment of an *Early Warning Alert and Response Network* (EWARN) Surveillance System and the EpiNet team, a local outbreak response team. Vanuatu had no established system for outbreak response prior to Cyclone Pam. The Penama EpiNet team was established as a pilot program with the hope of developing similar outbreak response teams in other provinces. During my deployment I supported the EpiNet team through its first outbreak investigation and response, a three-day field trip to an adjacent island to investigate a suspected Zika virus outbreak. The trip highlighted the challenges in developing sustainable surveillance mechanisms in resource poor settings. My final mission report is included as an appendix to this volume (Appendix B).

Summary of core activity requirements

Core requirement	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Appendix A	Appendix B	Appendix C
Response to an acute public health problem or threat			✓			✓	
Analysis of a public health dataset	✓	✓	✓	✓			
Evaluate or establish a surveillance or other health information system	✓					✓	
Design and conduct an epidemiological study		✓			✓		
Literature review	✓	✓			✓		
Report to a non-scientific audience					✓		
Advanced draft of a paper for peer-reviewed publication		✓	✓				
Oral presentation at national scientific conference			✓				
Teaching activities							✓

This page is intentionally blank

Chapter 2. An evaluation of the integrated environmental and infectious disease surveillance system for legionellosis in Victoria, Australia

Contents

Preface

My role

Victoria's unique approach to legionellosis surveillance that combines environmental and infectious disease surveillance has been operational since 2000 but never formally evaluated. I was tasked with designing and conducting an evaluation with a focus on the integration between the disease surveillance and environmental investigation systems.

I developed the evaluation plan in consultation with stakeholders and generated evidence through interviews, participant observation, document review and data analysis. I drafted the final report and developed recommendations for improvement. Throughout the evaluation stakeholders provided valuable feedback and advice.

During my field placement I also provided epidemiological support to the legionellosis surveillance system. This included signing off notifications, reviewing records for completeness, providing verbal reports at our weekly surveillance meetings, and assisting with outbreak investigations. I also developed expertise in the mapping software and assisted with the development of new mapping tools, as well as providing advice and troubleshooting when required.

Lessons Learnt

I learnt a great deal during this evaluation, not least of all the value in selecting the most appropriate framework for the task at hand. At times I found it difficult to address the aims of the evaluation using the surveillance evaluation framework. If I were to conduct an evaluation of this type again, I would select a program evaluation framework instead.

I gained technical knowledge relating to *Legionella,* including the challenges involved in pinpointing a source. I learnt that for complex public health problems a multidisciplinary team is essential. However I also learnt that such teams need to have clear leadership because the more complex a system is, the greater the risk of problems 'slipping through the cracks'.

I developed skills in the use of mapping software, and learnt that spatial analysis is underutilised in infectious disease epidemiology. I also observed that powerful

information technology can be underutilised if the right questions aren't asked, and if skilled operators are not available to ask them.

Public Health Implications

I found no evidence the system has reduced the burden of legionellosis in Victoria, although it provides excellent surge capacity for investigating significant outbreaks. I made several recommendations to improve both the integration and usefulness of the system.

I recommend reviewing the protocol for investigating environmental sources for sporadic cases, and instead focus on environmental investigations for outbreaks and clusters. This includes using more sophisticated spatial algorithms and improved molecular typing to identify areas that should be targeted for inspection.

Stakeholders have agreed in principal to explore changes to data management which should lead to significantly improved record keeping, which in turn may improve the capacity to review results of investigations to inform program planning.

Project Outputs

I produced an evaluation report for the Department of Health and Human Services, and presented my recommendations to the major stakeholders in the system. In addition, the evaluation is included here as a chapter in my thesis submission for the Master of Philosophy in Applied Epidemiology.

Acknowledgments

I gratefully acknowledge the contributions from the following:

- Lucinda Franklin, Legionella Epidemiologist, for helping to scope the project and assisting with technical quesitons throughout. Thanks for being such an excellent sounding board!

- Noel Cleaves, Manager of Environmental Health Regulation and Compliance and Nicola Stephens, Manager of Communicable Disease Epidemiology and Surveillance for great strategic advice

- Stuart Adcock, Team Leader Legionella Team, for educating me about *Legionella* ecology and control, and being ever-willing to provide data and documents to assist with the evaluation

- Clare Brazenor, Geospatial Administrator, for her passion and responsiveness, her interest in the system, and for her generosity in introducing me to technical geospatial concepts

- Jay Healy, Frances Tiplady, Jess Encena, and Shaun Coutts for sharing their thoughts and providing a refreshing practical viewpoint

- Caillean Michael, Lachlan Chapman, Laura Bales, Zina Maaliki, Prabhaka Gaware, Catherine Donoghue and Awnit Kumar for welcoming me into the team

- Kyra Chua for whole gneome sequence work

- Norbert Ryan from VIDRL, and John Coventry from MDU for answering my lab related questions

- Martyn Kirk for his endurance in producing detailed, thoughtful and specific feedback.

Abstract

Introduction: Victoria has a unique approach to legionellosis surveillance that combines infectious disease and environmental surveillance. The system aims to rapidly identify outbreaks or cases of public health significance and interrupt transmission from environmental sources to prevent further disease. This evaluation assessed the integration between the infectious disease and environmental surveillance systems and examined the usefulness of this combined approach for reducing transmission in Victoria.

Methods: The evaluation was guided by an evaluation framework. Evidence regarding system operation, simplicity, integration, stability and flexibility was gathered through participant observation and stakeholder interviews. Data quality and timeliness was assessed through an analysis of historical stored data and an audit of fifteen recent cases. Sensitivity was examined through a combination of the above methods. The system's usefulness for interrupting transmission from environmental sources was assessed by examining changes in legionellosis incidence over time including changes in frequency and size of clusters and outbreaks, and investigating the risk of *Legionella* contamination in cooling towers targeted for case investigations.

Results: The evaluation resulted in over thirty recommendations. The system is complex and not well integrated. There is excellent information technology but system stability is threatened by regular unstructured staff handovers within the disease surveillance system. Case confirmation and subsequent public health intervention is timely, except for cases notified on serology which are subject to significant delay. Data quality in the disease surveillance system is generally excellent; however geocoding accuracy and cluster/outbreak records are poor. Data quality in the environmental investigation system is hampered by the use of a database not designed to store case-related data. Sensitivity for detecting clusters or outbreaks is excellent. Sensitivity for confirming an environmental source is low, although this may be partially due to poor availability of molecular typing information. There is little evidence the system is useful for interrupting transmission from environmental sources for routine investigations, however it provides excellent surge capacity for large outbreaks.

Discussion: The system would benefit from improved integration, including the use of a shared database for storing case-related information and regular meetings between key stakeholders. There is no evidence that routine environmental investigations are

useful for interrupting transmission; however this assessment is necessarily coarse due to limitations in available data. System usefulness may be improved by transferring routine environmental investigative efforts away from sporadic cases, and redirecting efforts toward clusters. The identification of clusters could be made more specific through development and utilisation of culture-free molecular typing techniques. The geographical information system should also be leveraged to develop a spatial clustering workflow to detect statistical spatial clusters that fall outside the protocol definition, and these clusters should be used to guide selection of cooling towers for targeted inspection.

1. Introduction

In this evaluation I review Victoria's unique approach to the surveillance of legionellosis which combines disease surveillance and environmental investigation systems. I have followed the evaluation framework provided in the CDC guidelines for the evaluation of surveillance systems (1), with adjustment to suit the internal and programmatic focus of the evaluation.

In this report I begin by explaining the public health significance of legionellosis, then document the operation and internal stakeholders of the combined Victorian Legionellosis Surveillance System. Next, I describe the methods used to evaluate the system and present evaluation findings for relevant system attributes. Finally, I provide conclusions and recommendations for improvement.

1.1 Public Health Significance of Legionellosis

The CDC guidelines for evaluating public health surveillance systems (1) describe several factors influencing the public health importance of a given disease. These include the number of people affected, the resources required to control or treat disease, clustering of disease in time and space, public concerns, previous control efforts, and the level of preventability.

Legionnaires' disease is significant from a public health perspective due to high morbidity and occasional mortality amongst cases, a tendency to cluster in time and space, and capacity to generate acute public concern. The disease is at least partially preventable, which elevates its public health importance.

Historical Context

Legionella bacteria were first identified following a large outbreak of atypical pneumonia at a convention of American Legion ex-serviceman at a Philadelphia hotel in July 1976. One hundred and eighty cases were affected with 147 hospitalisations and 29 fatalities (2). Investigators suspected an airborne infectious agent and demonstrated an association between illness and spending time in the hotel lobby (2). However they could not identify any known pathogens. The outbreak resulted in wild public speculation and eventual closure of the hotel (3).

The mystery was solved the following year when a previously unidentified gram negative bacterium was isolated from the lungs of four affected patients (5).

Retrospective analysis of stored serum samples found at least two other unsolved outbreaks were likely caused by the newly identified organism (5). Investigators of one such outbreak had previously suspected the air-conditioning system as the most likely source of infection and exposed guinea pigs to water from the evaporative condenser. The Legionnaires' disease bacterium were subsequently identified in stored lung tissue harvested from these animals (6), establishing a link with the cooling system. The bacterium, later named *Legionella pneumophila,* was the first of over 50 *Legionella* species which have since been identified, approximately half of which have been associated with human disease (7).

Microbial Ecology

Legionella bacteria are widespread in both the natural and built environment (8, 9) and cause disease relatively infrequently given their ubiquitous nature (8). The bacteria occur naturally in aquatic environments including fresh and brackish water, mud and soil (10). They proliferate in biofilms of organic matter and microbes such as algae, protozoa, and other bacteria; multiplication occurs most rapidly between 30˚C and 43˚C (8). Cooling towers, hot water systems, warm water systems, drinking fountains, spas, car washes and other wet elements of the built environment are often colonised with *Legionella* due to the presence of warm water that facilitates bacterial growth (8).

Transmission to humans usually occurs through inhalation of aerosols (11), though the 'infectious dose' required to cause disease remains controversial (10, 12). Cooling towers are the most commonly implicated source of *Legionella* infection, but contaminated water in hospital warm water systems, spas, humidifiers, decorative fountains and domestic showers has also been associated with disease (8, 13).

Appropriate design, management and regular disinfection of high risk water systems can help minimise *Legionella* colonisation (8). An exception is *L. longbeachae,* which thrives in warm soil. *L. longbeachae* infection is frequently associated with exposure to potting mix (14).

Clinical Spectrum

Infection with *Legionella* bacteria can result in outcomes ranging from subclinical infection or a mild illness known as Pontiac fever to a severe and sometimes fatal pneumonic illness referred to as Legionnaires' disease.

Legionnaires' disease affects a small proportion of people exposed to *Legionella* bacteria, however because of its severity the illness is of significant public health concern. It predominantly affects people predisposed to lung disease due to pre-existing lung pathology, co-morbidities, immunosuppression, older age or smoking (11). Males are also disproportionately affected. Following exposure to the bacteria, the illness has an incubation period of 2-10 days followed by headache and non-specific malaise, then development of fever and a non-productive cough. The illness progresses to atypical pneumonia with interstitial pulmonary infiltrates. Widespread consolidation can ultimately develop (2). Extra-pulmonary disease can also occur with significant consequences (7).

Morbidity and Mortality

The case fatality rate of the Legionnaires' convention outbreak was 16% (2), demonstrating the high mortality for Legionnaires' disease in the absence of appropriate treatment. However, appropriate antibiosis vastly improves recovery and recent case fatality estimates are closer to 5% (15, 16).

Although the case fatality rate has improved with the discovery of appropriate antibiosis, the illness still results in significant morbidity. Legionnaires' disease has been estimated to cause around 15% of community-acquired pneumonias (17, 18) and is the second most common cause of community-acquired pneumonia requiring intensive care (19). It is also a significant contributor to the nosocomial pneumonia burden due to colonisation of complex warm water systems in health-care facilities (20, 21). Hospitalisation often extends up to 8-10 days (15) and treatment including mechanical ventilation in intensive care is not uncommon. Severe cases can require extra-corporeal life support when mechanical ventilation fails (22).

Incidence of Disease in Victoria

Legionellosis has been a notifiable disease in Victoria since 1979 and notifiable nationally through the National Notifiable Disease Surveillance System since 1991. However notified cases underestimate the true burden of disease in the community because legionellosis is difficult to diagnose and can be treated without being confirmed as the specific etiologic agent. Diagnosis of sporadic Legionnaires' disease can be challenging as the clinical presentation is non-specific and no single diagnostic technique is both sensitive and specific for all species of *Legionella* (23, 24). In

addition, effective antibiotic treatment is available so clinicians can treat Legionnaires' disease empirically even in the absence of a diagnosis (15, 24).

Males and older adults are disproportionately affected by both *L. pneumophila* and *L. longbeachae*, as demonstrated in Figure 1a and 1b, although *L. longbeachae* affects a narrower age range. During the period 1991-2013, 89% of all notified legionellosis cases in Victoria were aged 40 or older (median age 60, range 15-99 years), and 70% of all cases were male. The majority (80%) of cases were not known to be associated with any outbreak or cluster and were assumed to be due to sporadic illness. *L. pneumophila* was the most common causative organism and was responsible for 80% of all cases. *L. longbeachae* accounted for 14% cases, and the remaining 6% were due to other species or were not further specified. Figure 2 shows the frequency of legionellosis cases over time according to causative organism, along with the crude rate of disease in the Victorian population between1991-2013.

Figure 1a. Age-sex chart for legionellosis cases due to *L. pneumophila* in Victoria during 2013

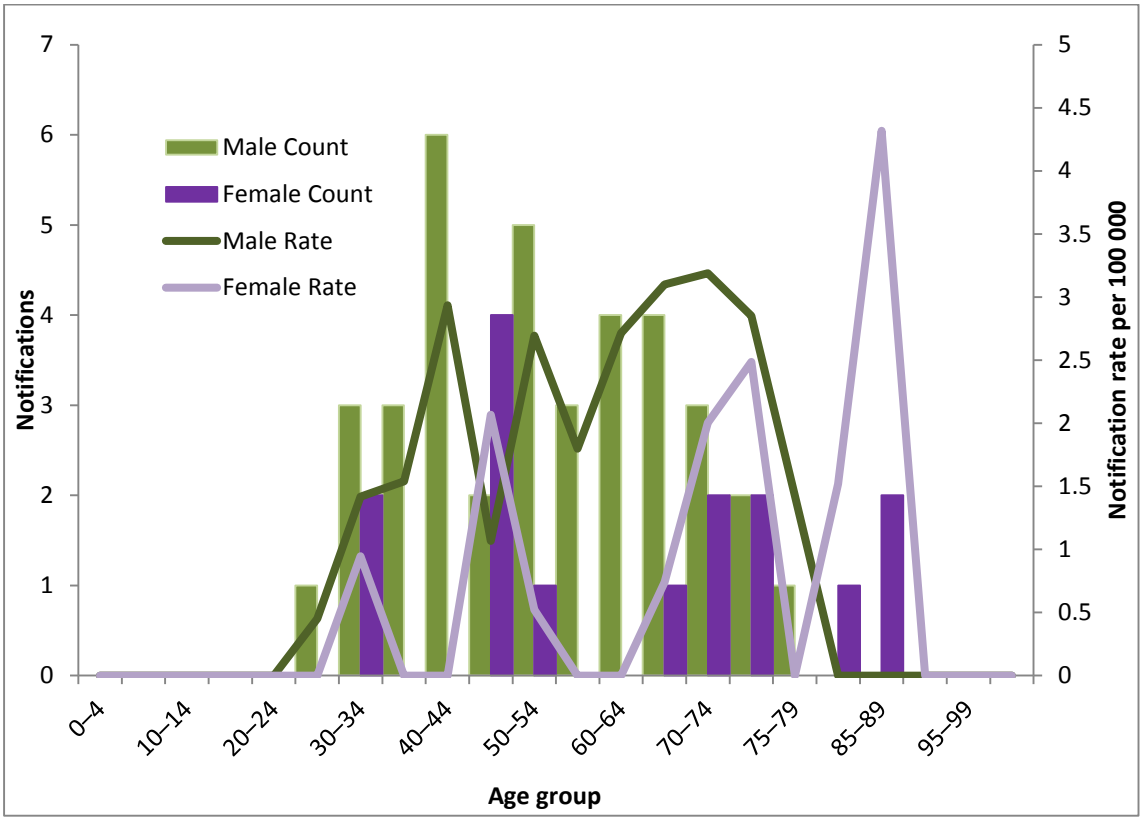


Figure 1b. Age-sex chart for legionellosis cases due to *L. longbeachae* in Victoria during 2013

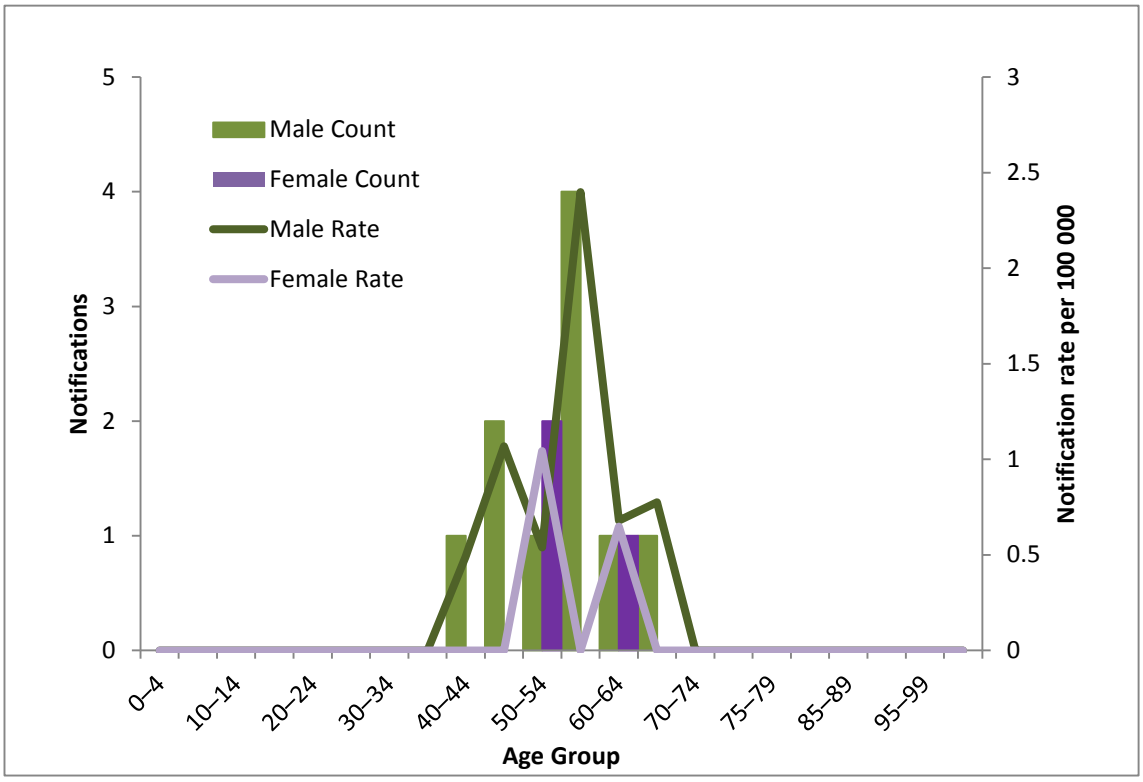


Figure 2. Victorian notified cases of *Legionella* by causative organism over time, Jan 1991 to 31 Dec 2013

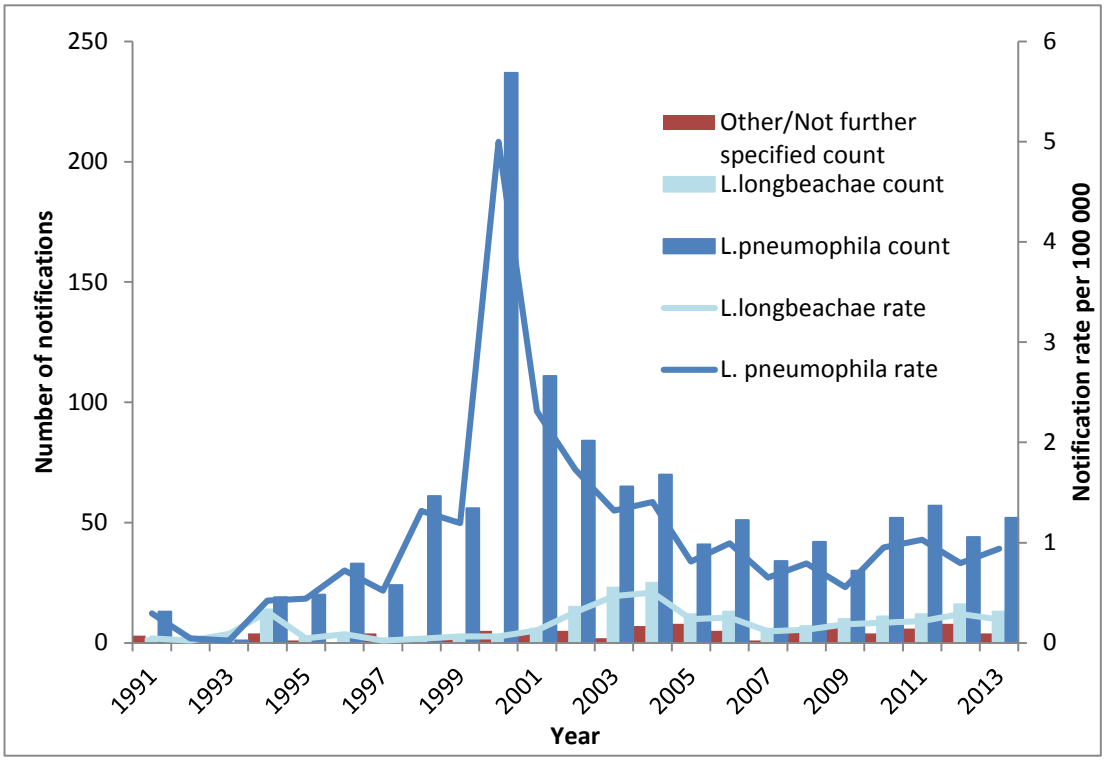


Figure 2 demonstrates the year with the largest annual number of notifications between 1991 and 2013 occurred in 2000, coinciding with a large outbreak at the Melbourne Aquarium. This year accounts for 245 or 16% of all *L. pneumophila* cases seen during the 22-year period. The Aquarium outbreak contributed 125 of these cases, but a further 120 notified cases were not associated with the outbreak. This is approximately twice the number of 'sporadic' (i.e., non outbreak-associated) cases recorded in any previous year. This phenomenon likely reflects increased awareness and testing due to the Aquarium outbreak in 2000 and suggests a substantial burden of cases previously went undetected. The increase in case detection and reporting persisted briefly then declined over time, and rates have stabilised since 2005.

The distribution of legionellosis is strongly influenced by geography. Victoria and South Australia have higher rates of illness due to *L. pneumophila* compared to the rest of Australia, where the mean rate for *L. pneumophila* during the period 2005-2013 was 0.69 per 100,000 persons per year compared to a mean of 0.84 per 100,000 persons per year in Victoria and South Australia (data retrieved through data request from NNDSS). However these rates are lower than that reported in Europe (26). In contrast, legionellosis due to *L. longbeachae* is relatively rare in Victoria, where the mean rate was 0.21 cases per 100,000 persons per year between 2005-2013 compared to 0.91 cases per 100,000 persons nationally.

Outbreaks in Victoria

The potential to cause outbreaks is one of the key reasons legionellosis is under surveillance. These are of public health significance both due to the clustering of disease in time and space (1) and due to the significant public and political concern generated by these events.

Between 1991 and 2013, 52 Legionnaires' disease clusters/outbreaks were recorded in Victoria, ranging in size from 2 to 125 cases and affecting at least 360 cases. *Legionella* were identified during environmental sampling in 20 of these. These originated from a variety of sites including industrial areas, shopping centres, a car wash, and a large tourist attraction.

Victoria developed a specific interest in Legionnaires' disease following an outbreak at the Melbourne Aquarium in 2000. The outbreak was Australia's largest with 125 cases and 4 deaths. The source was identified as a poorly disinfected cooling tower in the

large, recently opened international tourist attraction (28). The outbreak attracted significant international attention and highlighted the need for stringent regulation to ensure regular testing and treatment of cooling towers for *Legionella* (29). Following the outbreak detailed legislation was introduced in Victoria to ensure cooling towers are appropriately maintained and monitored according to a risk-based model (30).

Preventability

Legionnaires' disease can be prevented, or at least limited, through careful design and maintenance of the built environment (8). In addition, prompt identification of outbreaks allows rapid investigation to find the source of infection and public health intervention to limit the size of the outbreak.

Summary of Public Health Significance

Legionnaires' disease is a significant cause of both community-acquired and nosocomial pneumonia, especially pneumonia requiring intensive care. The bacteria are capable of causing sizable outbreaks resulting in significant public anxiety. Appropriate management of the built environment can reduce the risk of disease, and early detection of outbreaks provides the opportunity to limit further disease through public health intervention. Effective surveillance is essential to monitor trends, detect outbreaks, and guide public health interventions.

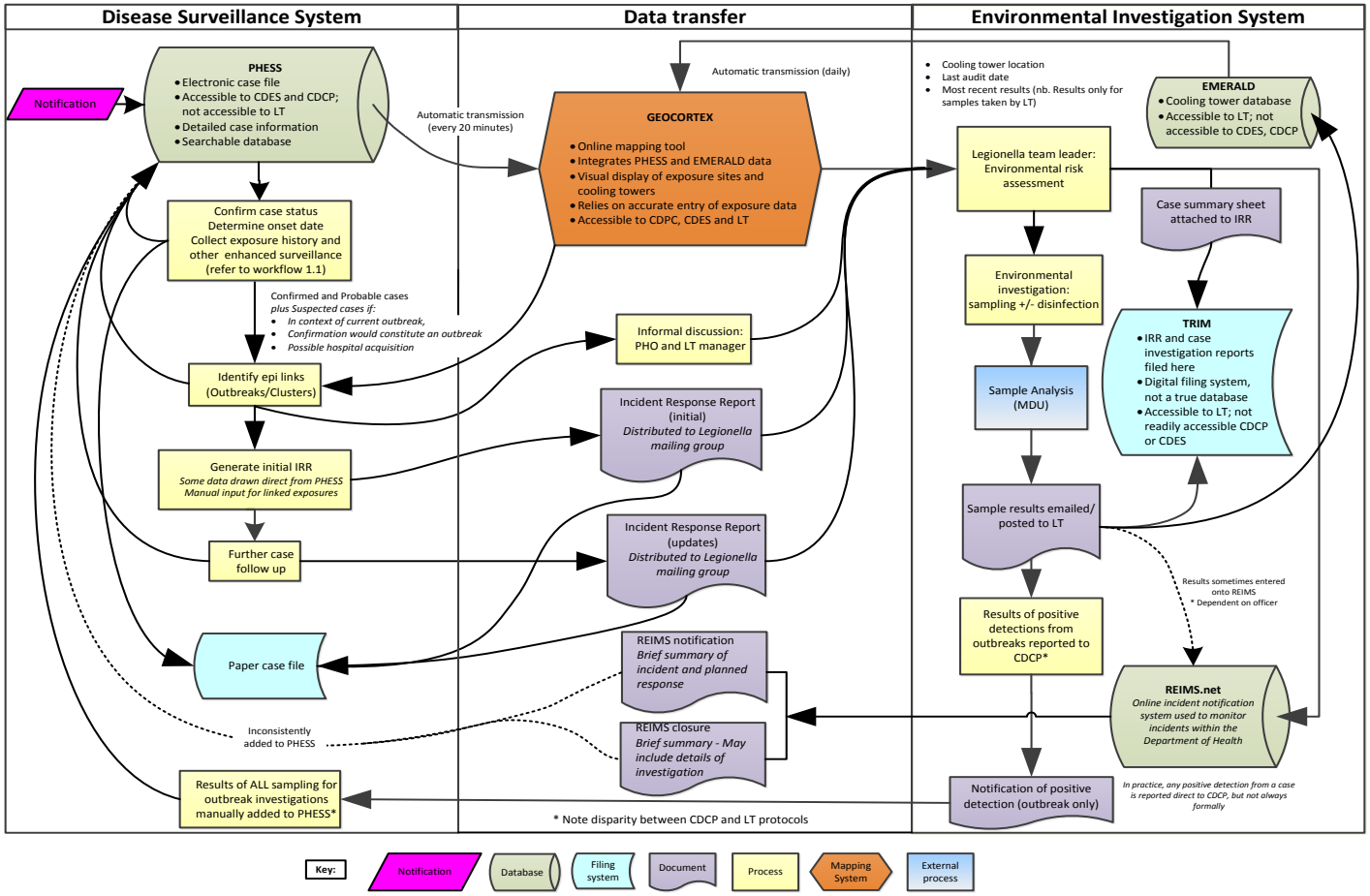
1.2 System Operation

The Legionellosis Surveillance System in Victoria operates from within the Health Protection Branch at the Department of Health and Human Services. The current structure of the Health Protection Branch is shown in Appendix 1.

The Disease Surveillance and Environmental Investigation components of the system operate separately. The disease surveillance system is shared between Communicable Disease Prevention and Control (CDPC) and Communicable Disease Epidemiology and Surveillance (CDES) sections. The former group undertakes case follow-up for notified cases of Legionellosis, while the latter is responsible for analysis and reporting on disease surveillance data gathered during public health follow-up by CDPC.

Environmental investigations for confirmed and some probable cases are conducted by the Legionella Team (LT) which sits within the Environmental Health Regulation and Compliance section. For ease of discussion in this report the activities of CDPC and CDES are referred to as the Disease Surveillance System, the activities of LT are referred to as the Environmental Investigation System, and the activities of the overall system are referred to as the combined Legionellosis Surveillance System. An overview of the two systems and the information flow between them is provided in Figure 3. The component systems are described below.

Figure 3 Victorian Legionellosis Surveillance System Information Storage and Transmission



1.2.a Disease Surveillance System

The activities of the Disease Surveillance System are outlined in the legionellosis disease case investigation protocol (31). The Surveillance System's documented objectives are as follows:

- Guide immediate action for cases of public health importance to prevent further transmission
- Detect and guide immediate action for outbreaks to prevent further transmission
- Monitor trends in Legionnaires' disease with respect to time, population groups, geography and other risk factors
- Guide the planning and implementation of policy, service provision, prevention strategies and other public health interventions
- Provide a basis for epidemiological research
- Monitor and evaluate the impact of interventions such as cooling tower regulation.

Legionellosis is classified as a 'Group A condition' in Victoria, reflecting its high public health priority. This designation means both clinicians and laboratories must notify the Department by phone or fax immediately upon diagnoses (confirmed or presumptive). The system uses the nationally agreed upon case definitions as shown in Box 1.

Box 1. Case Definitions for Legionellosis in Victoria*

Confirmed Case
A confirmed case requires laboratory definitive evidence AND clinical evidence

Laboratory definitive evidence
- Isolation of *Legionella* OR
- Detection of *Legionella* urinary antigen OR
- Sero-conversion or a significant increase in antibody level or a fourfold or greater rise in titre to *Legionella*

Clinical Evidence
Fever OR Cough OR Pneumonia

Probable Case
A probable case requires laboratory suggestive evidence AND clinical evidence

Laboratory suggestive evidence
- Single high titre to *Legionella*
- Detection of *Legionella* by nucleic acid testing OR
- Detection of *Legionella* by fluorescence assay

Clinical Evidence
Fever AND Cough
OR
Pneumonia

**Extract from the Department of Health Victoria's legionellosis (Legionnaires' Disease) case investigation protocol (31)*

The surveillance system uses locally unique definitions for clusters and outbreaks, which are shown in Box 2. These differ from the outbreak definition provided in the Set of National Guidelines for Legionellosis (SoNG).The SoNG does not define clusters and outbreaks separately, instead referring to them interchangeably, and defines a cluster/outbreak as 2 cases sharing an exposure site within 100m during a 3 month period (page 5) (32). Note that neither the SoNG nor the Victorian legionellosis case investigation protocol provide a definition for outbreaks or clusters of *L. longbeachae*, which is commonly associated with potting mix rather than aerosol generation. Outbreaks associated with *L. longbeachae* are rarely reported.

Box 2. Outbreak and cluster definitions for the Victorian Legionellosis Surveillance System*.

Cluster
A cluster is defined as two or more cases who have a common exposure (ie they are linked to within 500 metres of the same geographical location in their incubation period) and their dates of onset of illness are within 3 months of each other.

Outbreak
A cluster is defined as two or more cases who have a common exposure (ie they are linked to within 500 metres of the same geographical location in their incubation period) and their dates of onset of illness are within 14 days of each other.

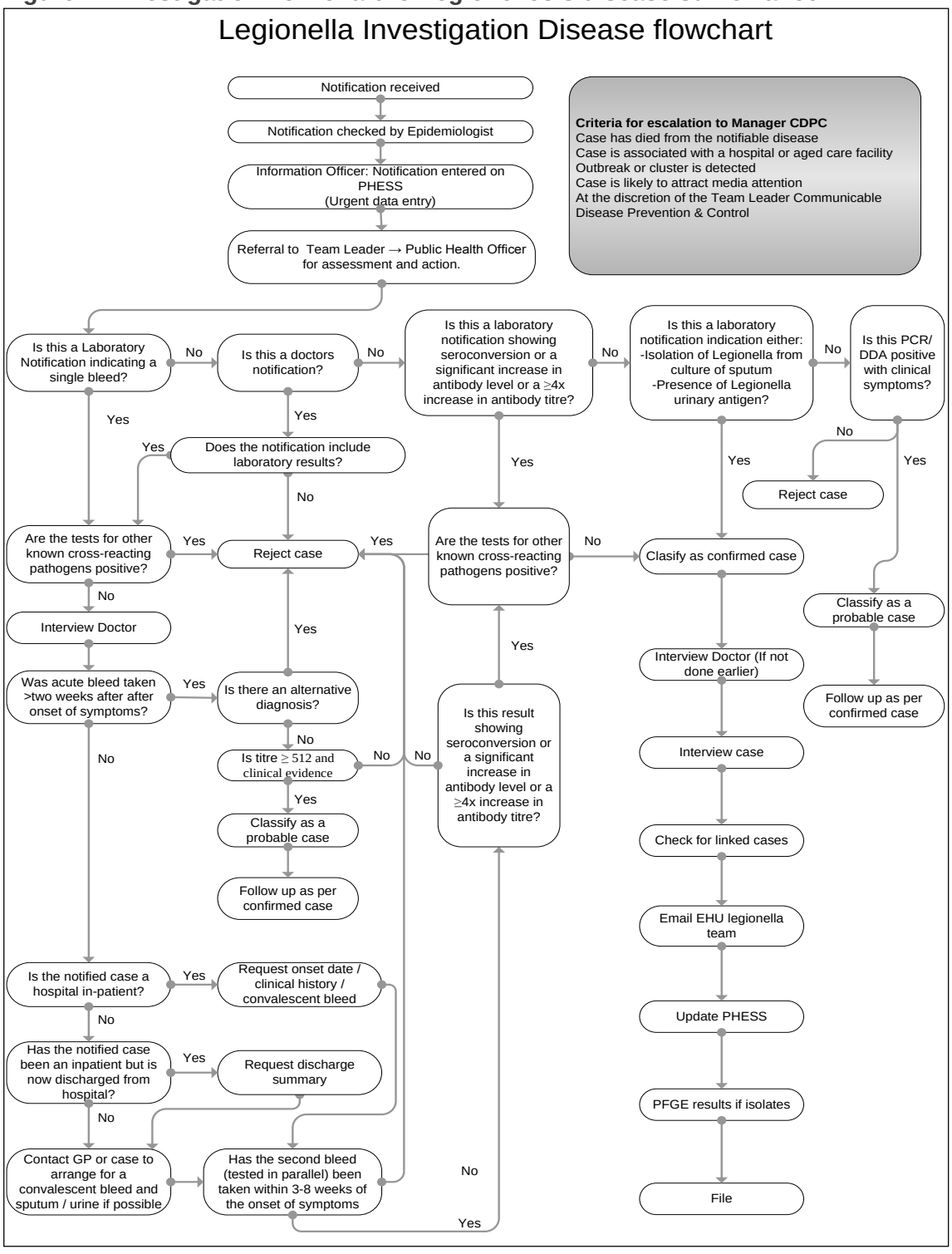
**Extract from the Department of Health Victoria's legionellosis (Legionnaires' Disease) case investigation protocol (31)*

Case Follow Up

An Epidemiologist from CDES reviews all legionellosis notifications and forwards them to an Information Officer at CDPC who enters available data into PHESS within an hour of receipt. A designated *Legionella* Public Health Officer (PHO) commences case follow-up within 24 hours of notification. Public Health Officers perform disease follow up on a rotating basis. One officer is solely responsible for all legionellosis follow up for six to nine months; after this time a 'disease rotation' occurs and a new PHO commences legionellosis follow up.

As outlined in Figure 4, case follow-up entails requesting follow up serology and/or confirmatory testing of urine or respiratory specimen at the Victorian Infectious Diseases Reference Laboratory (VIDRL); applying case definitions; collecting enhanced surveillance data for all confirmed and probable cases; and identifying clusters and outbreaks of disease.

Figure 4. Investigation flow chart for legionellosis disease surveillance*



**Extract from the Department of Health Victoria's legionellosis (Legionnaires' Disease) case investigation protocol (31)*

Enhanced surveillance data includes demographics, clinical details, illness onset date and an inventory of exposure sites, including exposure to potting mix. Exposure sites are any sites cases recall attending during the acquisition period (2-10 days prior to

illness onset) and represent the potential sources of a case's exposure to the *Legionella* bacteria.

The PHO geocodes exposure sites after entering the address into the database, which enables them to be mapped. Geocoding involves manually prompting PHESS to reference a mapping database, which provides latitude and longitude that matches the address. Importantly PHESS does not currently provide any immediate feedback to the user regarding geocoding accuracy, nor does it allow manual override of geocoding when a poor geocode match is returned.

The PHO next uses a spatial software package called Geocortex (Esri™) to identify whether exposure sites are shared by multiple cases, which signifies an outbreak or cluster. At the time of writing (December 2015), some PHOs also conducted a manual 'postcode search' within PHESS to identify shared exposure sites. This method was used prior to the introduction of Geocortex in February 2014 and the requirement to continue or discontinue it has not been clarified in the case investigation protocol (31).

Information Transfer

PHESS transmits exposure site data (and limited case data) to the Geocortex mapping system every twenty minutes. PHESS also automatically transmits core disease surveillance data to the National Notifiable Disease Surveillance System daily. The *Legionella* Epidemiologist provides a verbal report summarising recent legionellosis notifications and follow-up activities during weekly surveillance meetings.

The PHO communicates a summary of case information to the Environmental Investigation System using an Incident Response Report (IRR). This is a reporting template generated through PHESS that summarises the clinical details, exposure period and exposure sites for the case of interest as well as highlighting any links with other cases. IRRs are circulated by email to representatives from CDPC, LT, CDES, and the Office of the Chief Health Officer. They are generated for all confirmed and probable cases, and for suspect cases if (31):

- The case is associated with a possible hospital source
- In the context of an outbreak
- Confirmation of the suspect case would constitute an outbreak

Note that IRRs are generated for each individual case, but not for clusters or outbreaks. An example IRR is provided in Appendix 2. PHO's update and recirculate IRRs as new information is received.

Outbreaks

When an outbreak is identified an outbreak record is opened in PHESS, and thus the outbreak is provided with its own PHESS case number. Cases associated with the outbreak are "linked" to the outbreak record. The requirement to create an outbreak record for clusters is not articulated clearly in the *Legionella* case-follow-up protocol. As such, cluster records are inconsistently created. Outbreak and cluster records are differentiated through naming conventions. Outbreaks are named "year-month location legionellosis outbreak", while clusters are named "year-month location legionellosis cluster".

For outbreaks (but not clusters) a case conference is convened between the LT, CDPC and Office of the Chief Health Officer to plan the response. There are no formal written reporting mechanisms for outbreaks or clusters.

Resources

Legionellosis is one of sixty-five notifiable conditions under surveillance in Victoria. Many of the resources required to operate the Legionellosis Disease Surveillance System are shared across other notifiable diseases. CDPC resources include a manager, 16 Public Health Officers, 3 senior Public Health Officers, 3 senior officers with roles in planning, relationships and projects, and 8 Information Officers. CDES resources include five Epidemiologists, a manager, and two IT staff. Approximate resources required to operate the Disease Surveillance System from within this pool include one full time PHO, access to the PHESS database and associated IT support, access to Geocortex and associated support, an Information Officer (0.4 full time equivalent), an Epidemiologist (0.2 full time equivalent), as well as IT, communications, and laboratory expenses.

1.2.b Environmental Investigation System

The LT conduct an environmental risk assessment for each notified legionellosis case, and conduct an environmental investigation and disinfection if indicated. The LT is also responsible for regulation and compliance of cooling towers in Victoria. The team

administers Division 1 of Part 7 of the Public Health & Wellbeing Act 2008 (33) relating to cooling tower systems (CTS) and the Public Health & Wellbeing Regulations 2009 (30) relating to *Legionella* control in water delivery systems. Their documented objectives are:

- To minimise the impact of Legionnaires' disease on the Victorian community.
- To identify, and render safe, potential sources of the disease

Regulatory Activities

The regulatory work of the LT team is underpinned by legislation introduced in response to a large Legionnaires' disease outbreak at the Melbourne Aquarium (28) which highlighted the need for regulation of CTS to mitigate risk due to *Legionella* colonisation (29).

Regulation was initially introduced as the Building (*Legionella*) Act 2000 (34), and is now incorporated under the Public Health & Wellbeing Act 2008 (33) and the Public Health and Wellbeing Regulations 2009 (30). The regulations require landowners of sites housing CTS to register all CTS with the Department of Health and Human Services; to prepare and follow CTS risk management plans; to conduct regular maintenance and testing for *Legionella*; and to undergo an annual third party audit to ensure compliance. *Legionella* Team Environmental Health Officers inspect and sample CTS from sites that fail to complete annual audits or who receive a non-compliant audit. The team also conducts random compliance inspections.

The mandatory registration of CTS is managed using Emerald, a central database of all registered cooling towers in Victoria. The Emerald cooling tower database contains information on all registered and decommissioned cooling towers in Victoria including location, registration status, results of annual audits and results of LT inspections. This centralised database provides a unique advantage to investigation of CTSs as potential sources of legionellosis in Victoria, as it enables case exposure sites to be mapped against cooling tower locations to identify potential sources of infection.

Case Investigation

The LT conduct a risk assessment and public health action if indicated in response to confirmed legionellosis cases, probable cases diagnosed by PCR, and probable cases in a residential care facility or hospital.

The LT commence case investigations within 24 hours of receiving formal notification of a new case through an Incident Response Report (IRR) from CDPC. The LT do not have access to PHESS. Case information provided through the IRR is supplemented by the Geocortex mapping system. Geocortex displays data from both PHESS (case exposure data) and EMERALD (CTS data) concurrently, providing a rapid inventory of potential cooling tower sources for each case.

The Standard Operating Procedures for Environmental Investigation of *Legionella* provides guidelines for sampling and disinfection of cooling towers in response to notified cases (35). It is standard procedure to sample and disinfect any CTS within 500m of case homes, and those at the workplace, plus any CTS within 500m of an outbreak. Cooling towers within 500m of cluster locations are noted for targeting during the routine cooling tower inspection program. Figure 5 describes the protocol for case investigation followed by LT.

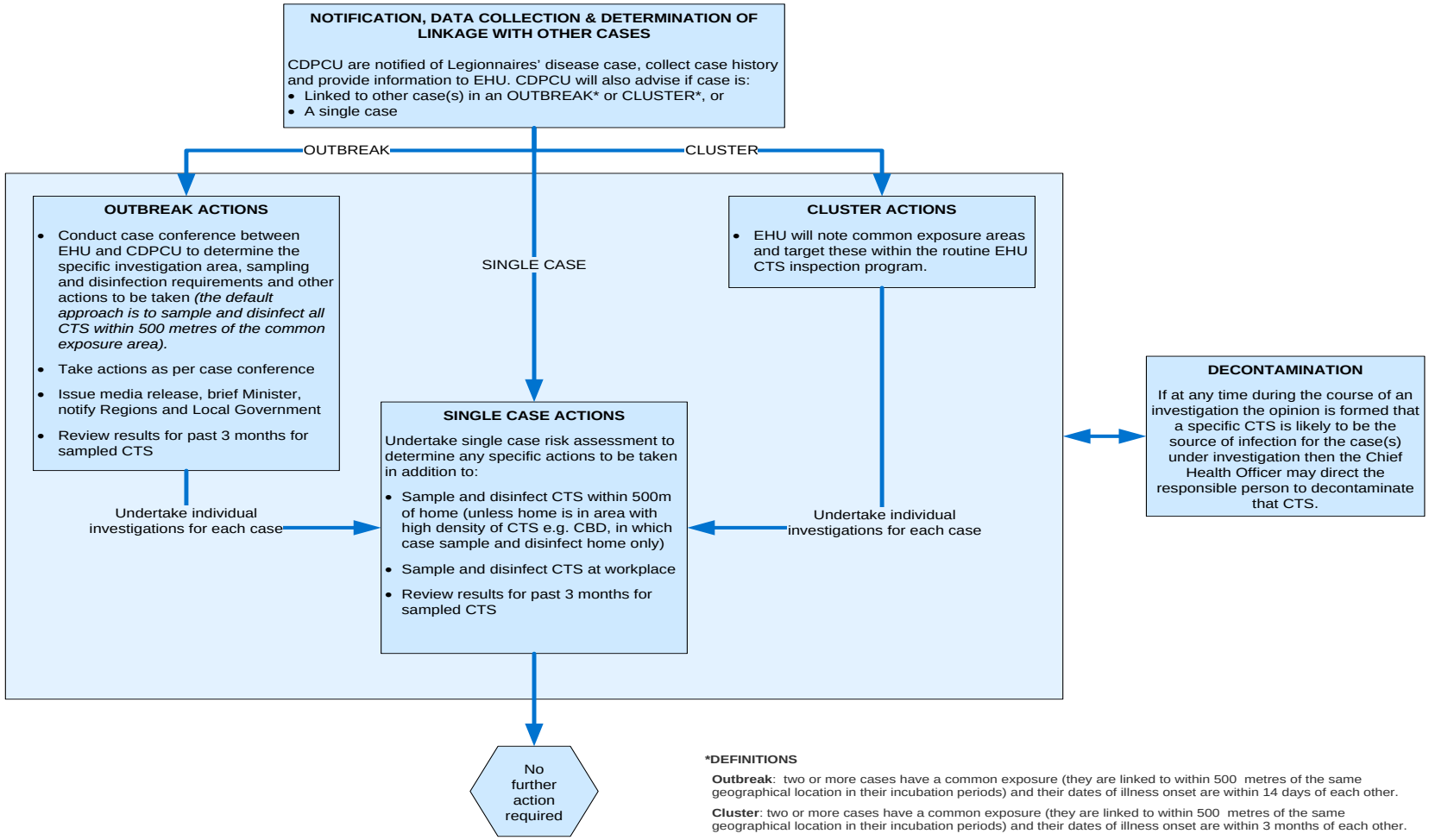
Inspectors collect a water sample from each cooling tower for *Legionella* and Heterotrophic Colony Count (HCC) testing. HCC measures the number of colony-forming bacterial units present in a system and is an indicator of system maintenance. Samples are analysed including culture and identification at Victoria's *Legionella* pubic health reference laboratory, the Microbiological Diagnostic Institute (MDU). Isolates of *L. pneumophila* serogroup 1 are further typed using Pulse-Field Gel Electrophoresis (PFGE).

Any cooling tower with an elevated HCC (>200,000 cfu/ml) or a *Legionella* detection undergoes a disinfection (if not already performed at the time of sample collection) and is resampled at weekly intervals until the problem is resolved.

Non-cooling tower sources may be sampled and possibly disinfected if warranted by a risk assessment. A high index of suspicion is maintained for hospitals and aged care facilities as potential sources of *Legionella* because of the significance of an outbreak in those facilities. Where cases have spent considerable time at such facilities during their acquisition period, the Environmental Investigation System will sample and require

disinfection of the warm water system and CTS (if present) at the site. Other sources that may be investigated include CTS beyond 500m from the home or in the vicinity of the workplace, warm water systems, warm water car washes, spas or fountains and ice machines.

Figure 5. Environmental investigation workflow for cases of legionellosis in Victoria, 2015 (*extract from Environmental Investigation Standard Operating Procedures) (35)*



RIEMS Incident Notification

Upon receipt of an IRR from CDPC the LT open an incident in a web-based incident management system called RIEMS. RIEMS automatically sends an email notification to previously nominated recipients whenever a new incident is opened, when the incident details are updated, or when an incident is closed. RIEMS recipients for *Legionella* events include the following:

- All recipients of IRR (CDPC *Legionella* interest group)
- Environmental Health Team
- Legionella Team
- Water Program
- Radiation Team
- Regional Public Health Managers
- Regional Environmental Health Officers
- Health Protection Branch Executive
- Health Protection Emergency Management

An example RIEMS notification is shown in Appendix 3.

The RIEMS web-based record can be accessed to review all updates relating to a particular incident. Historically RIEMS was updated when any further information was provided and following completion of a case to provide LT with a record of case progress. However the system has not kept pace with changing technology and is no longer used consistently and these updates are now rarely completed. RIEMS has been used historically across parts of the department to manage and track incidents. RIEMS has never been adopted by CDPC or CDES as the teams manage communicable disease incidents using purpose built software (currently PHESS and prior to that a legacy system called NIDS).

Case Summary Sheets

The LT summarise each case investigation on a Case Summary Sheet. This is a word document appended to each IRR. An example Case Summary Sheet is provided in Appendix 4.

The LT files each IRR along with its completed Case Summary Sheet in TRIM, the department's digital filing system. When updated IRRs are received, information from the previous Case Summary Sheet must be copied over to the new Case Summary Sheet and updated to reflect new information. The new IRR and appended Case Summary Sheet are filed in TRIM alongside the previous documents for the case. These documents can be retrieved from the TRIM filing system using a "title" search for the case ID number or by browsing within the case investigation folder for the relevant year. A recent TRIM innovation also searches within the text of documents, eg searching for "humidifier" will bring up all Case Summary Sheets that list humidifiers as an exposure, along with any other electronic documents filed within the Department of Health and Human Services that include the word "humidifier".

The TRIM digital filing system has not been rolled out to CDPC or CDES, so these teams cannot readily access files in TRIM. The LT Standard Operating Procedures also stipulate the Case Summary Sheet should be attached to RIEMS when a case is closed in the RIEMS system, however in practice this is rarely done.

Case Summary Sheets are designed to record information regarding individual cases. There is no equivalent record keeping in place for outbreaks. Currently, records regarding investigations for outbreaks are included on the Case Summary Sheets for one or some of the cases involved in the outbreak but there is no systematic way of deciding which Case Summary Sheet should record the information. Consequently record keeping for outbreaks (and clusters) tends to be fragmented across multiple Case Summary Sheets.

Inspection Records

Findings of CTS inspections for case investigations are entered into the team's cooling tower database, Emerald. An example inspection report from Emerald is shown in Appendix 5.

Inspections relating to case investigations are marked by selecting “Case Investigation” as the reason for inspection from a drop down menu to differentiate them from inspections conducted for other purposes. There is no option for “Outbreak Investigation”. There is also no designated field to record the PHESS ID number for the case the investigation relates to. The PHESS case ID number is usually noted along with other remarks in a free-text comments field.

For outbreak related inspections, a PHESS case ID number for one of the outbreak cases, not the outbreak itself, is usually noted in the comments box. Note that the sample inspection record shown in Appendix 5 relates to an outbreak and the PHESS number in the comments box was the most recent case that prompted sampling; however, this is not evident from the record. Emerald does not capture whether the sample was collected due to a home or workplace exposure for the single case, or due to a shared exposure with another case.

No specific inspection record is made for non-cooling tower sources. The Emerald system can only accommodate cooling tower system records and so cannot be used to store information relating to non-cooling tower sources, such as hot water systems, fountains, spa pools, etc. Notes about inspection findings for non-cooling tower sources are occasionally made in the Case Summary Sheets, but this is inconsistent.

Sample Results

Laboratory results are received within 10 days of sample collection, and include the case PHESS ID number. Results are transcribed onto the relevant Case Summary Sheet. For CTS results, the *Legionella* result (*Legionella* species, concentration) is also manually entered into Emerald, and the result notification is filed in TRIM in a folder for the CTS.

Sample results for non-cooling tower sites are filed as word documents or PDFs in TRIM. Historically, results were filed in a “warm water incident” TRIM folder for a given year. As a new folder was started each year, results for a given facility could be scattered across multiple folders. Notes relating to each investigation/disinfection were noted in the relevant Case Summary Sheets, further dispersing the information. In 2015 a new system has commenced which partially addresses this problem. A folder is generated for each facility or home address sampled. This folder will store all sample results for that facility over time in a single location.

PFGE results for environmental samples are received intermittently, at intervals of up to 6 months. Result notifications are filed in the TRIM digital filing system under the relevant CTS system or non-CTS site folder. PFGE results for environmental samples are not added to Emerald or to the Case Summary Sheet. They are inconsistently communicated to the disease surveillance system and thus inconsistently added to the case records in PHESS.

Results of environmental investigations are not routinely fed back to CDPC. However if *Legionella* is detected in any CTS, CDPC are notified using a *Legionella* detection notification form to enable them to check for cases in the area. This notification does not include a field for PHESS case ID, and does not specify whether the inspection is case related.

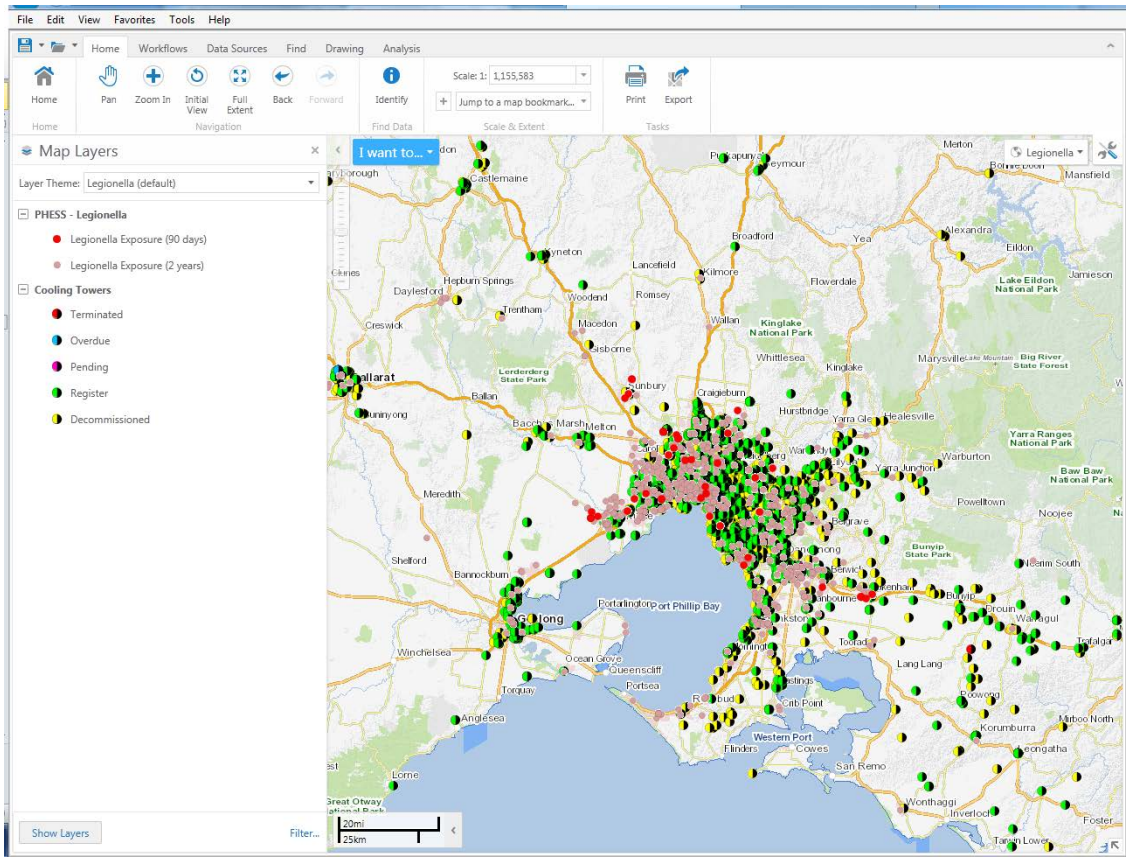
Resources

The LT team includes a Team Leader, two and a half full time equivalent Inspection Officers and one full time student intern. It shares a number of resources with other teams in the Environmental Health Regulation & Compliance Section (one systems & database manager; one trainee, as well as a shared registration and licensing team). It operates on an annual budget of approximately $560,000. The Team recovers approximately $420,000 annually in registration related income.

1.2.c Geocortex

Geocortex is a web-based digital mapping system that integrates case exposure data from PHESS and cooling tower data from Emerald by displaying multiple data tables in a simultaneous spatial array. The system is managed by a Geographic Information Systems Administrator from Modelling, GIS and Planning Products in the System Intelligence and Analytics Branch. Separate Geocortex windows are available to PHESS and Emerald users. Figure 6 shows the Geocortex mapping interface. Clicking on a case exposure site or cooling tower opens a pop-up window that describes the site attributes.

Figure 6. Geocortex mapping interface



Case Exposure Data

PHESS exports two case exposure data tables to Geocortex. One table includes exposure and summary data for confirmed and probable cases notified during the most recent 90 days; the other provides data for cases during the most recent 2 years. Summary data includes the following attributes:

- PHESS case ID
- Organism cause
- Illness onset date
- Event classification eg. confirmed or probable
- Exposure site details eg address, latitude, longitude, LGA, health region

Case exposure sites recorded in PHESS are only displayed on Geocortex if certain criteria are met. These criteria are:

- Case classification is confirmed or probable
- Symptom onset date must be recorded
- Exposure sites must be geocoded
- Geocoded sites that have had their address details edited must be re-geocoded, because editing an eddress deletes the previous geocode.
- A twenty minute lag is observed between exposure site entry into PHESS and transmission of data to Geocortex. This means if PHOs check for outbreaks or clusters immediately after updating details, these will not yet be included on the map.

Cooling Tower Data

Cooling tower data are exported to Geocortex from Emerald every 24 hours. Cooling tower layers include registered cooling towers, decommissioned cooling towers, and layers for systems with overdue or pending registration. Cooling tower attributes include:

- Cooling tower system identifier (CTSID)
- Cooling tower site identifier (SIDID)
- Registration status
- Legionella result at last sample (detected/not detected)
- HCC concentration at last sample
- Results obtained from latest sample (yes/no)
- Risk management plan reviewed/implemented/compliant at last inspection (yes/no for each)
- Site details eg address, latitude, longitude, LGA, health region

Functions

Geocortex allows basic exploration and analysis of spatial data displayed on the maps. Most functions are easy to use after basic training but are not readily accessible to those without prior mapping experience or advanced IT skills. Functions include searching the maps to identify addresses, applying filters to display features of interest from a particular layer (eg exposure sites for a particular case, or series of cases) and buffering localities to identify features within the chosen buffer distance. There is also a distance measure tool that is useful for measuring the distance between exposure sites and potential sources. The system also has the capacity to select a subset of features for display in an attribute table which can then be exported to comma-separated values or excel files. Where required, new data layers can be imported in comma-separated value or excel format. Users can add notes or drawings to the maps to mark important features, and projects can be saved for later reference.

Cluster and Outbreak Detection

Geocortex features a customised function that streamlines the detection of clusters and outbreaks. This workflow automatically identifies exposure sites for a case of interest within the “90 day” data layer and applies a buffer of a specified distance to each site. The workflow then identifies any exposure sites within the buffered areas and returns an attribute table listing all sites for the case of interest as well as sites for any other cases within the area of interest. Figures 7a and 7b shows the workflow input and output display. Note that the output shown in Figure 7b includes both a cluster and an outbreak associated with case 320152304931, but these are not specifically noted by the output table and could be missed by an untrained operator.

Figure 7a. Geocortex workflow for cluster/outbreak detection: Workflow input

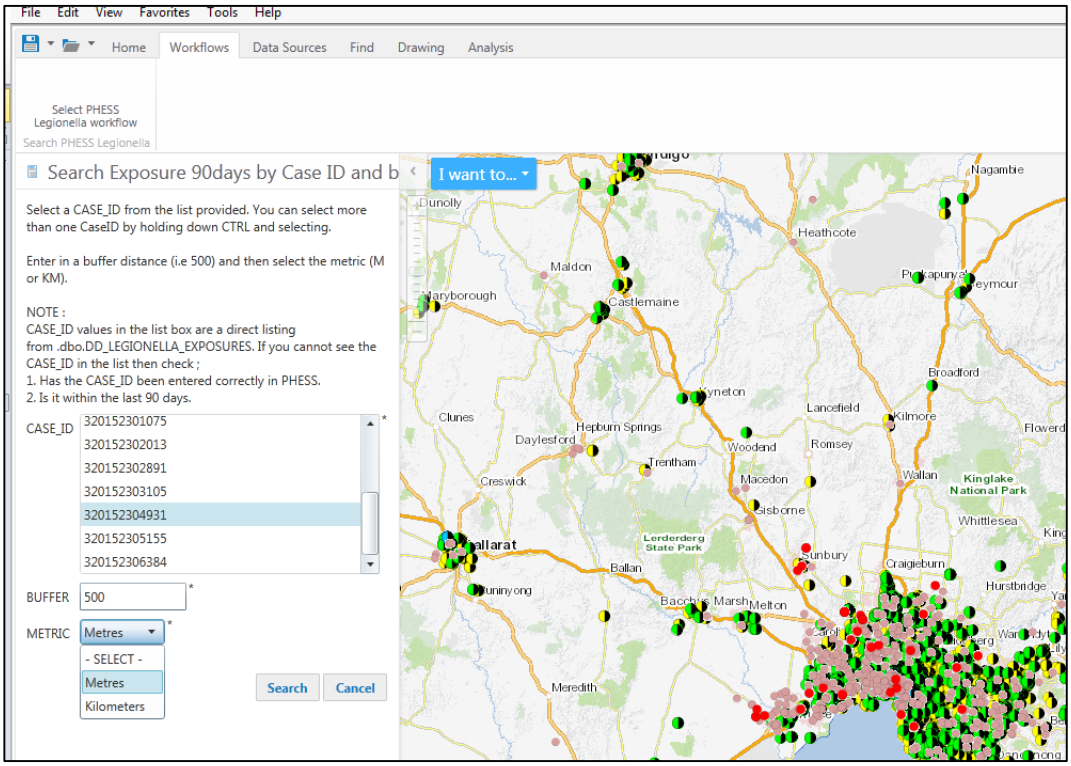
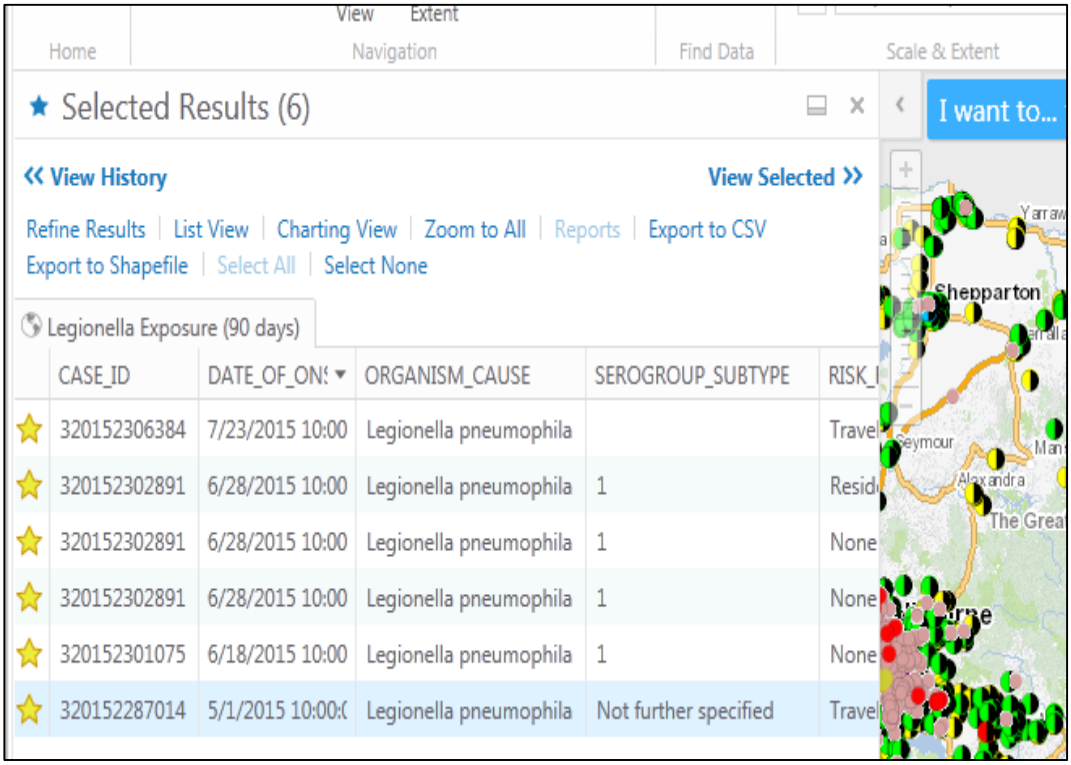


Figure 7b. Geocortex workflow for cluster/outbreak detection: workflow output



1.3 Stakeholders

The Legionellosis Disease Surveillance and Environmental Investigation Systems in Victoria have a number of internal stakeholders. Table 1 lists internal stakeholders and their interest in the surveillance system.

This evaluation was guided by consultation with managers from LT, CDPC, and CDES. These stakeholders collaborated to identify the intended purpose of the evaluation, focus the design, and provide feedback on the draft evaluation plan. Internal stakeholders continued to be a rich source of information during the evaluation through qualitative methods.

Due to the internal focus of this evaluation, consultation with external stakeholders was considered out of scope.

Table 1. Internal stakeholders of the Legionellosis Surveillance System, Victoria

Stakeholders	Interest in the system
Communicable Disease Prevention and Control (CDPC)	
Information Officers	Enter disease notifications and laboratory results into PHESS database Important role in data quality
Public Health Officers	Case follow up and update PHESS database with case clinical, demographic and exposure data Ensure exposure sites map properly Identification of clusters/outbreaks Refer case information to LT for environmental investigations Communicate clusters and outbreaks to CDPC Senior Public Health Officers, CDES, Office of the Chief Health Officer and LT
Senior Public Health Officers	Provide support and advice to PHO's Initial escalation point for clusters/outbreaks Provides strategic advice regarding clusters and outbreaks or unusual cases
Manager, CDPC	Manages activities of Information Officers and Public Health Officers Escalation point for important cases eg. If cases die, are associated with hospitals or aged care facilities, if an outbreak or cluster is detected, or media attention is anticipated

Table 1. continued

Communicable Diseases Epidemiology and Surveillance (CDES)	
Legionellosis Surveillance Manager	Epidemiologist responsible for analysing, interpreting and reporting on legionellosis disease surveillance data Manages the disease surveillance system, including ensuring case definition is consistently applied and data is complete Provides epidemiological support during outbreaks/complex clusters
Database Manager, CDES	Manages the Public Health Event Surveillance System (PHESS) database, which is the repository for legionellosis and other notifiable disease surveillance data Facilitates data sharing with Geocortex mapping service
Manager, CDES	Oversees the work of the surveillance manager and database manager Provides support, advice and strategic guidance around surveillance system and outbreaks as necessary
Legionella Team (LT)	
Legionella Environmental Health Officers	Monitor cooling tower risk management compliance; includes random and targeted site visits, sampling for *Legionella*/HCC and reviewing maintenance records Risk assessment and environmental investigations for cases and outbreaks Implements public health interventions as directed by the manager, *LT*
Team Leader, LT	Oversees activities of environmental health officers Co-ordinates environmental investigation and response to notified cases Provides advice on environmental risk management for health care facilities Key role in risk assessment
Manager, Environmental Health Regulation and Compliance	Oversees the work of the LT Provides support, advice and strategic guidance around regulation and compliance and response to outbreaks as necessary

Table 1. continued

Database manager, LT	Maintains database of cooling tower inspection data Facilitates data sharing with Geocortex mapping service
Modelling, GIS and Planning Products	
GIS Administrator	Development and maintenance of Geocortex mapping software (interface between cooling tower and disease surveillance data)
Office of the Chief Health Officer	
Senior Medical Advisers	Provide advice to CDPC Provide a conduit between CDPC/CDES/LT and Chief Health Officer Require succinct and rapid updates integrating information from both CDPC and LT during outbreaks
Chief Health Officer	Statutory functions under the Public Health and Wellbeing Act (2008) Provides advice to the Minister, Secretary, and healthcare providers on matters relating to public health including legionellosis outbreaks, clusters or cases of public health importance The Government's media spokesperson on matters relating to control of disease, including legionellosis

2. Methods

The evaluation purpose was developed in consultation with key stakeholders. The aim was to assess the integration between the Disease Surveillance and Environmental Investigation components of the system and generate recommendations for improvement.

The evaluation framework was adapted from the Centres for Disease Control and Prevention (CDC) guidelines for evaluation of surveillance systems (1), with some adjustment to account for the internal focus of the evaluation.

In this evaluation I assessed system performance against relevant attributes identified in the CDC framework, plus an additional attribute unique to this system called “Integration”. Thus, I examined performance in terms of the following attributes:

a. Simplicity
b. Integration
c. Stability
d. Timeliness
e. Data quality
f. Sensitivity
g. Flexiblity
h. Acceptability

I considered each attribute as it relates to the co-ordination between the Disease Surveillance and Environmental Investigation arms of the system. Note the CDC evaluation framework includes the attributes Predictive Value Positive and Representativeness, which I excluded from the current evaluation framework as they were not relevant to the evaluation focus.

I examined the usefulness of the system in terms of whether it had reduced the burden of disease in Victoria. This included examining changes in legionellosis notifications in Victoria over time, assessing whether the system reduced the number and size of clusters and outbreaks over time, and reviewing the risk of *Legionella* detection in cooling towers targeted for case investigations relative to randomly sampled cooling towers.

I generated evidence for the evaluation using mixed methods, including participant observation, stakeholder interviews, document review, and extraction and analysis of data from relevant databases (PHESS, Emerald, TRIM, and the Laboratory Information Management System (LIMS) database at MDU).

Stakeholder interviews were guided by semi-structured questionnaires. All stakeholders were asked to comment on system simplicity, flexibility, acceptability and the Geocortex mapping software. They were also invited to comment on specific areas relevant to their work. Interview notes were recorded manually, then entered into Excel (Microsoft, 2011) and coded into themes relating to each attribute.

I reviewed documents including departmental protocols, hard copy case records, and digital case records from PHESS, TRIM, Emerald, and RIEMS. I examined data extracted from PHESS, TRIM and Emerald to assess timeliness and completeness, and compared against LIMS data and relevant hard copy documents for accuracy.

I audited fifteen cases to document the flow and completeness of data through the system. Cases included a convenience sample of seven consecutive cases notified between June-July in 2014, plus a random selection of eight cases notified during the period August 2014-March 2015. Additional examples are provided where relevant.

3. Results

3.1 System Attributes

3.1.a Simplicity

Simplicity reflects both the structure and the ease with which a surveillance system operates and should be maximised as much as possible while still meeting surveillance objectives (1). In this evaluation I considered the simplicity of operations within each system where they influence, or are influenced by, the co-ordination between the two systems. I assessed simplicity through stakeholder interviews and participant observation. Stakeholders who contributed through interviews are listed in Appendix 6. Integration between the Disease Surveillance and Environmental Investigation Systems is a structural factor with a major impact on the simplicity of the overall system and I examined this as a separate attribute.

Disease Surveillance System

The Legionellosis Disease Surveillance System is not a simple system. Surveillance is complicated by the difficulties associated with diagnosing legionellosis by serology, both when confirming diagnoses and determining the causative species. Geocortex simplifies identification of clusters and outbreaks; however complexities around the use of Geocortex could be pitfalls. Identifying suspect cases associated with outbreaks is challenging but important for timely response. Defining clusters and outbreaks using the definitions provided in the legionellosis protocol (31) is not straightforward where multiple linked cases occur over a prolonged time period, and record keeping for these events can be complex.

Classification of Cases Diagnosed on Serology

Diagnosing cases based on serology is complicated because antibody titres in response to legionellosis are slow to rise, and cross reactions occur both with non-*Legionella* pathogens and with other *Legionella* species (36). Further, the additional information required when classifying such cases is not clearly articulated in the protocol or captured by the case definition which make classifying some cases seem unnecessarily complicated.

The case definition for confirmed and probable legionellosis cases notified on serology was provided in Box 1. However, in practice the classification is more complex than

this. According to the case investigation workflow shown in Figure 4, additional considerations for confirmed cases include:

- The second serology test must be tested in parallel with the first. This means the first blood specimen must be retained and retested in the same test run as the second test, to account for poor test repeatability
- The second blood specimen must be collected within 3-8 weeks of illness onset
- Tests for other cross-reacting pathogens must be negative.

For probable cases, the additional considerations are:

- The single high titre must be collected at least 2 weeks after symptom onset
- The single high titre must be ≥ 512
- Tests for other cross-reacting pathogens must be negative.

These additional requirements lead to confusion as they are not succinctly summarised in the protocol. Currently, these requirements are only presented in the workflow shown in Figure 4, but do not appear elsewhere in the document. Case classification for serology cases would be simplified by clearly outlining these requirements within the text of the protocol, or ideally, capturing them within the case definition.

Determining Causative Species

Another complexity in the disease surveillance system is that serological determination of the species of *Legionella* responsible for *Legionella* infection is unreliable (37, 38, 36). The Public Health Laboratory Network of Australia advises “serological tests for *Legionella* species other than *L*. *pneumophila* serogroup 1 are predictive of legionellosis caused by the *Legionella* species antigen used or a related *Legionella* species” (36). This means for example, a case of legionellosis diagnosed as *L. longbeachae* on serology may be caused by *L. pneumophila* serogroup 1, or any other species of *Legionella*. This complexity is addressed in the *Legionella* disease case investigation protocol, for example page 22 states “There should be a high degree of suspicion for any case notified that may have a link to a cluster or outbreak. This would include cases with a low titre (e.g. 256) or notified cases of a different species (e.g. *L. longbeachae* and *L. pneumophila*)”. However, this

is not well known amongst PHOs. For example, when asked whether they would check for shared exposure sites for *Legionella* cases caused by two different species of *Legionella* one PHO replied “No, I don’t see why you would”. This reflects a generally poor understanding within the Disease Surveillance System of the unreliability of speciation of *Legionella* infection based on serology.

Identifying Suspect Cases Involved in Outbreaks

Identifying involvement of suspect cases notified based on serology during outbreaks is not simple. Cases notified with an initial low titre are classified as suspect cases until further results allow either confirmation or rejection of the case. There can be between twenty and thirty suspect cases on the system at any one time. Ideally suspect cases would be considered when identifying outbreaks and clusters to maximise timeliness of the system. The disease follow-up protocol advises that for suspect cases, an IRR should be sent “in the context of a possible outbreak, or where confirmation of a case would constitute an outbreak” (page 25 (31)). It also advises that during an outbreak a high index of suspicion should be maintained for potentially linked cases with a low titre. These requirements are challenging because:

- Only confirmed and probable cases are transmitted to the Geocortex mapping system which is used to identify clusters and outbreaks
- There is no mechanism in place to identify where “confirmation of a case would constitute an outbreak”, as spatial relationships are not routinely examined for suspect cases
- Where an outbreak has been identified, the PHO must manually review postcodes for all suspect cases to identify those that may potentially be involved in the outbreak. If they identify a case whose residential postcode matches the outbreak they manually locate the residential address in Geocortex to assess proximity to the outbreak. In practice this process is time consuming and only undertaken during large outbreaks.

Transmitting data for suspect cases to Geocortex would strengthen the system by enabling these cases to be included in the usual automated workflow to readily identify suspected outbreaks and clusters.

Defining Clusters and Outbreaks

The combined Legionellosis Surveillance System employs rigid outbreak and cluster definitions. These aim to simplify response to linked cases by automating decision-making, however in practice clusters and outbreaks do not always align neatly with these strict definitions. Where multiple cases occur in a small geographical area over weeks to months it is common to see temporally overlapping clusters and/or outbreaks. Defining and generating accurate records for such events can be complex. Data quality and consistency of outbreak records is discussed further under “Data Quality”.

The possibility that two or more cases may attend popular sites during their incubation period through chance alone is an added complexity and contributes to the problem of overlapping clusters and outbreaks. Molecular typing information could be used to differentiate between such cases to improve the specificity of outbreak detection. However the system uses molecular typing infrequently because Victoria’s current typing approach (PFGE) relies on sputum collection and culture, both of which are challenging to achieve in practice (42, 43).

Environmental Investigation System

The basic operations of the Environmental Investigation System are simple. However data storage and retrieval is complex because there is no designated database for case related data, and because outbreak and cluster investigations are not specifically documented.

Basic Operations

Investigating environmental sources of *Legionella* is operationally simple. The Standard Operating Procedures (35) clearly outline which sites should be sampled, including cooling tower and non-cooling tower sources. Environmental sampling is a straightforward process involving collection of a water sample and submission for culture at MDU. Interpretation of results can be challenging which is discussed further under sensitivity.

Data Storage and Retrieval

Data storage and retrieval is a major complexity within the Environmental Investigation System. The problem arises because the main database used by the LT is not

designed to store case-related data; thus most case-related data is electronically filed in word documents and PDFs. The problem is compounded during outbreaks or cluster as there are multiple linked cases and it is not always clear where information will/should be stored.

Case investigation information stored in Emerald cannot be easily queried as Emerald does not have a field for the PHESS case identifier. This means it is not possible to generate simple statistics using Emerald data such as average number of sites inspected per case or total number of cases with CTS inspections each year. Similarly, because non-cooling tower case investigations are not captured in any database it is difficult to retrieve and analyse data including sampling and detection frequency for non-cooling tower exposures.

Clusters and Outbreaks

Data management for clusters and outbreaks is a particular challenge for the Environmental Investigation System. The system does not create a dedicated record or utilise a unique identifier for cluster or outbreak related investigations. This means there is no specific record of environmental investigations conducted in response to each outbreaks/cluster. Environmental samples collected as part of outbreak investigations are allocated a PHESS ID from one of the cases within the outbreak (generally the most recent case at the time of sampling). However, cases within outbreaks frequently also have non-outbreak related samples collected. This situation makes it difficult to determine retrospectively which samples relate specifically to outbreak investigations and which relate to routine case investigations.

The lack of specific record keeping for clusters/outbreaks also makes it challenging to compile a complete record of sites sampled during an outbreak investigation. Stakeholders reported it is difficult to keep track of which sites have been inspected and when results were due for sampled systems during large or complex investigations. This is compounded by a delay in updating case inspection information in the Emerald database. Inspection details are usually entered into Emerald within 24-48 hours but during large outbreaks this may be significantly delayed because the officers do not have access to EMERALD in the field. The team relies on memory to recall which towers have been inspected until inspection details are entered. This creates a risk that some sites may be missed, or that sites may be sampled (and disinfection ordered) twice by different inspectors particularly during larger outbreaks.

The key complexities identified for sporadic and outbreak associated cases in the Disease Surveillance and Environmental Investigation Systems are summarised in Table 2.

Table 2. Key complexities for sporadic and outbreak associated cases in Disease Surveillance and Environmental Investigation Systems

Event Type	Disease Surveillance System	Environmental Investigation System
Sporadic cases	Classifying serology cases Determining causative species	Data retrieval, particularly for large numbers of cases or exposure types
Outbreaks and clusters	Identifying involvement of suspect cases Defining and recording overlapping clusters and outbreaks	No designated record for each cluster/outbreak Keeping track of inspections Data retrieval for cluster/outbreak investigations

Geocortex

Geocortex streamlines the information interface between the Disease Surveillance and Environmental Investigation Systems. The customised work-flow for cluster and outbreak detection and cooling tower identification simplifies activities within both systems. However, geocoding of exposure sites and ensuring complete and accurate transfer of data to the Geocortex mapping system are complex tasks, and there are some intricacies to Geocortex which are not well understood. Many stakeholders reported a lack of confidence using the system due to inadequate training.

The workflow for identification of clusters and outbreaks ensures all exposure sites for cases are easily taken into account (assuming all exposure sites are properly mapped), thus simplifying the task of identifying clusters and outbreaks. The workflow also rapidly identifies CTS near case exposure sites for one or multiple cases, simplifying the risk assessment process for the LT.

While Geocortex simplifies identification of clusters, outbreaks, and potential cooling tower exposures, it has some complexities. If exposure sites are not accurately geocoded and transmitted to Geocortex, the mapping system cannot account for them. The ‘black box’ nature of the Geocortex workflow means there is a strong potential for

omissions to go undetected. The evaluation identified poor knowledge among PHOs regarding the complexities of Geocortex including limited awareness of pre-requisites for transmission of data to the maps, and limited understanding of the need and method for quality checking to ensure mapping was accurate and complete.

Geocortex is simple to use but requires appropriate training and practice. However no practical training was provided when Geocortex was introduced. During evaluation interviews over half of all stakeholders expressed a lack of confidence in navigating the system for basic tasks, and some staff within the Environmental Investigation System could not use the system at all, instead relying on other staff to conduct risk analysis on their behalf.

3.1.b Integration

Integration refers to how well the Disease Surveillance and Environmental Investigation Systems coordinate and is a major structural factor affecting the ease of operation for the combined Legionellosis Surveillance System. Integration is strongly influenced by organisational structure and leadership, and data storage and access.

Structure and Leadership

Roles in the combined Legionellosis Surveillance System are distributed across four teams and two branches in the Department. Operational activities are undertaken by three teams within the Health Protection Branch, while Geocortex is administered from within the System Intelligence and Analytics Branch. The three teams within the Health Protection Branch fall under different management streams and are characterised by distinct professional cultures. Meetings between stakeholders from the three key teams only occur in response to outbreaks. The evaluation interviews highlighted that no individual within the system understood how the entire system operated.

The distribution of roles across multiple teams and a flattened organisational structure means the system operates without clear leadership. Below the level of the Branch Director no single person has engagement across the whole system, and consequently the component systems are scantily integrated. This lack of role definition impacts poorly on efficiency and effectiveness, as demonstrated in the following examples:

- **PHESS identifiers for clusters and outbreaks:** When outbreaks (and ideally clusters) are identfied, the Disease Surveillance System records the

incident and allocates it a unique PHESS identification number; however this number is not communicated to the Environmental Investigation System. The Environmental Investigation System struggles to effectively keep track of outbreak (and cluster) investigations as they do not assign outbreaks their own unique record number. The outbreak/cluster PHESS number would be ideal for this purpose if it were shared by the Disease Surveillance System in a systematic way.

- **Dual notifications following a case of Legionellosis**: CDPC circulates an IRR email to alert stakeholders when a notified case of Legionellosis is confirmed. Shortly afterwards LT also provides a RIEMS email notification. A single alert would be adequate in the majority of cases, however, the systems have not co-ordinated this.

- **Data accessability**: Recently CDPC decided to convert all IRRs to PDF format before circulating them. However, this format interferes with LT record keeping system, as it does not allow them to update the attached Case Summary Sheet. As the LT was not aware of the planned change, and the Disease Surveillance System was not aware of the LT requirements for record storage the problem was not identified until after the IRRs had been altered and circulated in PDF form.

- **Communication of *Legionella* detections**: The Environmental Investigation System notifies the Disease Surveillance System when *Legionella* is detected in an environmental sample. The notification form does not specify whether the detection relates to a case investigation, and does not include a field for PHESS number. During the evaluation one PHO was in the *Legionella* role for 4 months before being added to the email list to receive these notifications.

Data Storage and Access

The Disease Surveillance and Environmental Investigation Systems use separate databases, with database access restricted to the teams involved with each system. This inhibits integration within the system and leads to duplication of effort in an attempt to maintain the required information within each part of the system. It also means data relating to case investigations is fragmented across multiple systems.

Table 3 lists the databases used by the system, their functions and users within the system.

Table 3. Databases in the combined Legionellosis Surveillance System

Database	Type of database	Function	Access
PHESS	Data repository (web-based)	Disease surveillance data	CDPC CDES → selected data exported to Geocortex
TRIM	Electronic document management system	Stores documents relating to environmental investigation for cases including Case Summary Sheets and laboratory results for environmental sampling	LT *PHOs can gain access by completing a TRIM training course, arranging for TRIM installation on their PC, and gaining security access through the LT leader. However the 6 monthly 'disease rotation' system makes this impractical.*
EMERALD	Data repository (desktop)	Cooling tower regulation and compliance data	LT → selected data exported to Geocortex
RIEMS	Incident management system	Incident notification Incident management (historical function)	LT *Has not been adopted by CDPC. Could be accessed with appropriate training/protocol adjustments, but is no longer updated consistently*

Officers within the Environmental Investigation System are reliant on information communicated in IRRs and via Geocortex when making risk assessments regarding cases. Where information changes over time it is difficult to check/verify it's accuracy, and missing or inaccurate updates are not uncommon (see also "Data Quality"). In addition due to limited understanding of the functions in other parts of the system

PHOs are not always aware of which information updates would be helpful to the Environmental Investigation System.

3.1.c Stability

Stability is the reliability of the system including ability to be operational when needed (1). The main factors affecting stability for the combined Legionellosis Surveillance System are human resources and information technology.

Human Resource Stability

Human resource skills and training have an important impact on stability within the Disease Surveillance System due to the complexities of legionellosis follow-up. In CDPC, Public Health Officers rotate their disease responsibilities every 6-9 months. Because rotations occur simultaneously across the team the outgoing PHO trains the incumbent PHO while learning the protocol for a new disease. Similarly, the incumbent PHOs learn the legionellosis protocol while handing over previous duties to another PHO. Legionellosis handovers are unstructured and usually conducted in a single session. The legionellosis Epidemiologist and senior PHOs are not routinely involved, and handovers do not include any formal training on the use of Geocortex. This regular staff changeover risks gradual degradation of the skills and knowledge required for effective legionellosis surveillance. For example during the evaluation, several weeks after a handover the new *Legionella* PHO:

- Did not know how to open or operate the Geocortex mapping system
- Was unaware of the need or method for ensuring exposure sites were properly geocoded
- Did not know how to interrogate the maps to check for outbreaks or clusters.

Further examples of loss of knowledge with each handover are included under the discussion on “Data Quality”.

Information Technology

Information technology stability in the system is excellent. The three key systems PHESS, EMERALD and Geocortex all experience minimal unscheduled outages. PHESS reported four unscheduled outages during the 2014 calendar year. Only one

was greater than fifteen minutes in duration, and this lasted less than an hour. Emerald reported one unscheduled outage during the same period. This lasted half a day and occurred following a planned system upgrade. Geocortex reported three outages with a maximum duration of two hours. Each system is supported by dedicated IT staff that are available to resolve unexpected outages during working hours.

3.1.d Timeliness

Timeliness is the amount of time between steps in a surveillance system (1). The combined Legionellosis Surveillance System aims to implement immediate response to outbreaks or cases of public health significance and timeliness is therefore a critical attribute. I assessed timeliness by calculating the time between key milestones within the system, with a focus on detection and response. Table 4 outlines these milestones along with the data sources and methods used to assess dates. I excluded *L. longbeachae* cases diagnosed by serology because timeliness is less of a priority for these cases, given these infections are less prone to outbreaks than other Legionella species.

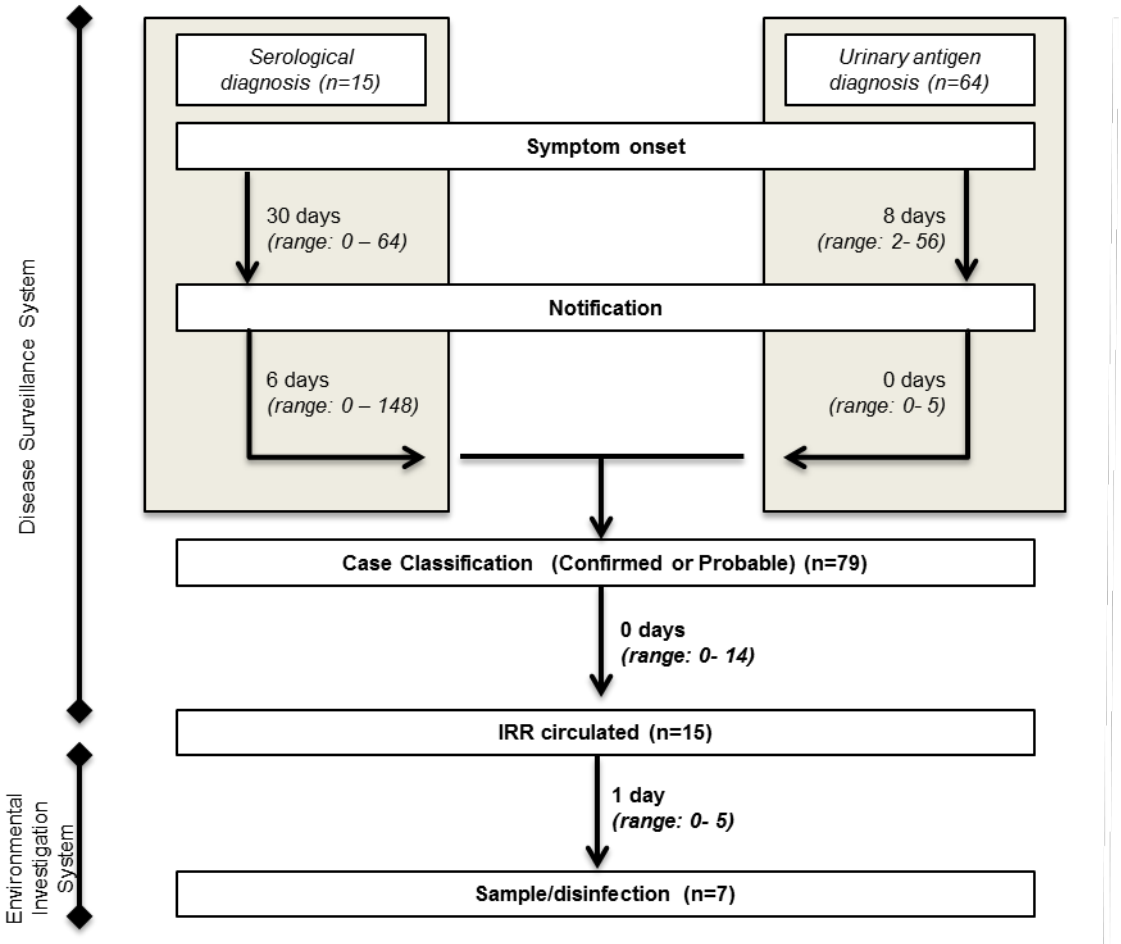
Table 4. Key timeliness milestones, records reviewed and how dates were retrieved from the system.

Milestone	Sample size	Sample description	How dates retrieved
Symptom onset date	79 cases	All confirmed and probable cases, 2014*	Date field extracted from PHESS
Notification date	79 cases	All confirmed and probable cases, 2014*	Date field extracted from PHESS for the first notification received
Confirmation date (confirmed or probable)	79 cases	All confirmed and probable cases, 2014*	Date field extracted from PHESS
Date Incident Response Report (IRR) circulated	15 cases	Audit cases – convenience sample	Manual review of email records
Date source sampled/ disinfected	7 cases (28 inspections)	All audit cases where an environmental investigation was conducted	Date of first inspection manually retrieved from Case Summary Sheet

**Excluding L. longbeachae cases diagnosed on serology*

Timeliness within the disease surveillance system is excellent, although timeliness of case confirmation could be improved for cases diagnosed by serology. Timeliness within the Environmental Investigation System is also excellent in the majority of cases. Figure 8 outlines the median (and range) number of days between each key milestone within the system. Timeliness is presented separately for cases diagnosed by serology and urinary antigen detection because the time between symptom onset date, notification date and case classification date differs substantially with method of diagnosis.

Figure 8. Timeline of key milestones in the Legionella surveillance system, Victoria 2014*



* Figure shows median number of days (and range) between key milestones, excluding *L. longbeachae* cases notified by serological diagnosis.

Onset to Notification Date

Symptom onset date and notification date are recorded for all cases in PHESS. Data was extracted and the interval assessed for confirmed and probable cases notified

during 2014 (excluding *L longbeachae* cases notified based on serology). The majority of cases diagnosed by urinary antigen detection were notified promptly after symptom onset. Half were notified within 8 days of symptom onset. There were a small number of outliers. The longest interval extended up to 56 days in a patient who described an 8 week history of intermittent fever and cough prior to diagnosis. This variability is an important consideration as it reduces the reliability of exposure information due to loss of recall, and reduces the potential for timely identification of clusters and public health intervention.

Cases diagnosed by serology had a longer delay to notification. After excluding cases due to *L. longbeachae,* there were 15 serologically diagnosed cases including seven due to *L. pneumophila* (one due to serotype 7-14, and the others not able to be serotyped); seven due to *Legionella* not further specified; and one due to *L. micdadei.* There was a median interval of 30 days between symptom onset and notification date, and a range of up to 64 days. This delay reflects, in part, the time interval required to collect a follow up blood sample. It may also be influenced by patient factors including illness severity. Patients with a less acute illness are more likely to delay presentation to a health facility, and clinicians may be more likely to select a broad approach to diagnosis such as an atypical pneumonia serological screen, in preference to a specific *Legionella* urinary antigen test.

Notification to Case Classification Date

Case classification is when case status is altered from “suspect” to “confirmed” or “probable”, triggering public health intervention. The date of classification was extracted directly from PHESS and the interval between notification and case classification calculated for confirmed and probable cases during 2014 (excluding *L longbeachae* cases notified based on serology). Case classification was timely for cases diagnosed by urinary antigen detection. Over half of all cases were confirmed the same day the notification was received. The longest interval between notification and case classification was five days. This delay reflects the time it takes PHO’s to contact doctors to determine whether the patient meets the clinical requirements of the case definition. This can be delayed where cases are diagnosed in an emergency department and there is a large turn over in medical staff, making it difficult to contact the treating clinician.

In a small number of cases delay in case classification of 1-3 days occurred as PHOs awaited confirmation of urinary antigen results at VIDRL before confirming a case. This occurs because primary laboratories sometimes notify the department by phone of positive results as “preliminary findings”, then send urine to VIDRL for confirmation before issuing a final laboratory result. Confirmation of results by a reference laboratory is not required by the case definition.

The interval between notification and case classification was substantially longer for cases diagnosed by serology. For these cases the median delay between notification and case classification was six days. However there was substantial variation with the delay extending to over one hundred days for two cases (134 and 148 days, respectively). When added to the delay between symptom onset and notification, this delay is sometimes substantial, for example one patient had a total delay of 212 days between symptom onset and confirmation.

Part of this delay is accounted for by 3-6 week lag required for collection of follow up serology for those cases notified after a single bleed. However a significant portion of the delay occurs due to workflow because there is no system for managing serology case follow-up, making it challenging to keep track of which cases require action. In addition, some serology cases could be classified as probable based on the first bleed. However a portion of these might later be rejected based on a stable titre in the second bleed. Rather than update the case classification to probable at the earliest possible time, PHOs sometimes wait for the second serology result to avoid the possibility of conducting public health follow up only to reject the case later. For cases that aren’t rejected this results in a significant lag before case interview, delaying the potential identification of outbreaks and increasing the likelihood of poor recall of exposure sites.

Case Classification to Incident Report Circulation

Circulation of the Incident Response Report is a key milestone within the system as it transfers essential information from the Disease Surveillance System to the Environmental Investigation System, triggering public health intervention. The date of circulation is not recorded in the disease surveillance database, thus the parameter was assessed by examining the fifteen cases selected for audit purposes. For each case a manual search within the departmental email system was conducted to identify the date the first IRR was circulated. This date was compared against the case classification date extracted from PHESS. Incident Response Report circulation was

timely for the majority of cases. Twelve of fifteen (80%) IRRs were sent on the day of case classification. One was sent the following day, one three days after confirmation, and one was circulated 14 days after case confirmation due to challenges in contacting the case for interview.

The time between classifying cases and circulating the IRR can reflect the time it takes PHO's to make contact with the case, conduct the case interview, ascertain exposure sites, and conduct a search for linked cases. This is in line with the legionellosis disease investigation protocol that states "an internal e-mail should be prepared and sent to key department staff as soon as possible after case status is confirmed, the exposure history taken, and a search for epidemiological links has been completed" (section 10.1) (31). However, timeliness of environmental investigations could be improved through the provision of an initial IRR containing residential and work address (if available), followed by an update containing additional exposure sites when available. This is because the majority of environmental investigations are conducted around the residential and workplace addresses, which are frequently available before the complete exposure history is determined.

Incident Response Report to Sampling/Disinfection

The interval between IRR circulation and sampling/disinfection of exposure sites was determined for the fifteen audit cases. Where more than one IRR was sent the interval between the latest IRR and sample collection was used. The sampling/disinfection date of the first site sampled was manually retrieved from the Case Summary Sheets filed in TRIM. Seven of the fifteen audited cases had an environmental investigation conducted. Four of these investigations occurred within one business day of receiving the IRR. Two investigations were conducted more than one business day after receiving the IRR, and one was conducted five days later.

3.1.e Data Quality

Data quality refers to both completeness and validity of the data recorded by a surveillance system (1). In this evaluation I assessed the quality of data separately for the Disease Surveillance and Environmental Investigation Systems, with a focus on data that is required by and shared between the two systems.

Disease Surveillance System

Case follow up data

I examined completeness and validity of data required for case follow-up by reviewing the progress of fifteen cases through the system. Cases reviewed for the audit are presented in Appendix 7. As shown in Table 5, data completeness and validity varied across fields. Completeness was excellent for the true onset date and acquisition period, but reduced to around 80% for geocoding exposure addresses and for recording whether an environmental investigation was conducted. Clusters were poorly recorded; there were three clusters and one outbreak associated cases in the audit and only the outbreak was recorded in PHESS. Data was valid (over 90% correct) for case classification and whether an environmental investigation was conducted, but slightly less reliable for organism type. Validity of dates and exposure site information provided to the Environmental Investigation System in IRRs was low, indicating updated information had not been communicated to the Environmental Investigation System.

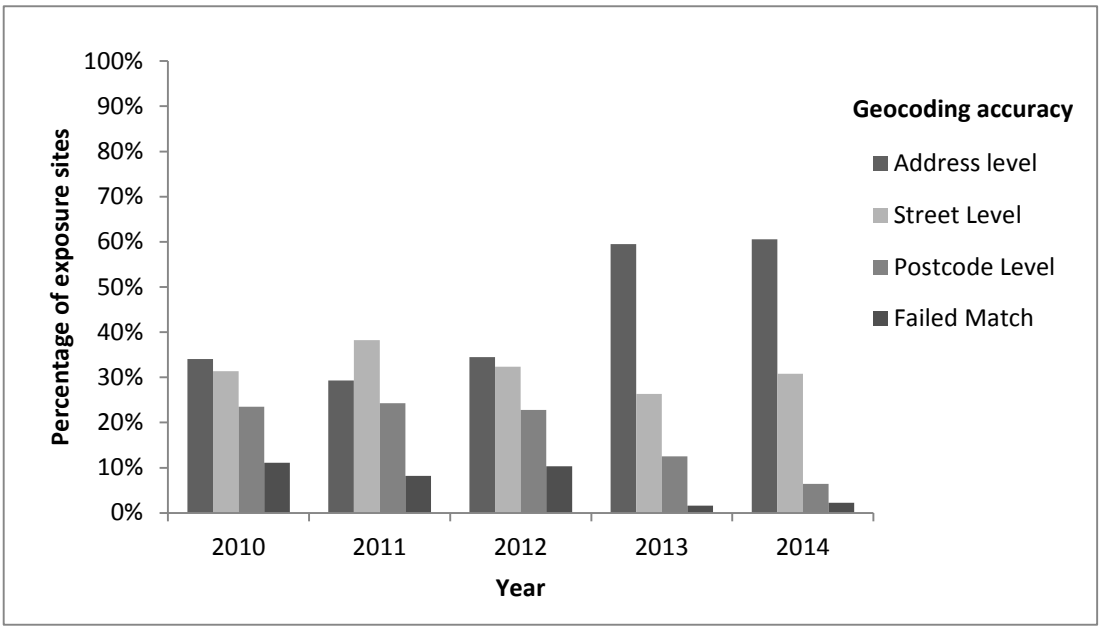
Table 5. Completeness and validity of case follow up data

Data field	Records reviewed	Complete/Valid	
Completeness	**n**	**no**	**%**
Outbreak/cluster record created	4	1	25
"Was environmental investigation conducted?"	15	12	80
Exposure Sites geocoded	80	72	89
True onset date complete	15	14	93
Acquisition period complete	15	15	100
Validity			
IRR matches onset date and exposure sites recorded in PHESS	15	11	73
Organism/Cause correctly applied	15	13	87
"Was environmental investigation conducted?" field correct	12	11	92
Case correctly classified	15	15	100

Geocoding Accuracy

To assess the geocoding accuracy of exposure site addresses stored in PHESS I extracted all addresses for exposure sites recorded during 2010-2014 and geocoded them using ArcGISv10(Esri™), referencing the department's internal geocoding database. I then plotted the geocoding accuracy (as reported by ArcGIS). This approach was necessary because PHESS does not currently report the accuracy of geocoding. The validity of address data stored in PHESS is presented in Figure 9. During 2013 and 2014 around 60% of exposure site addresses stored in PHESS geocoded to address level accuracy, meaning they could be mapped to a specific address. Around 30% geocoded to street level, meaning they could only be mapped to the centre of the street. For short streets this may be a reasonable approximation, but for longer streets or roads this may be several 100m from the actual exposure site. The remainder were only geocoded to postcode level (mapped to the centre of the postcode) or could not be matched at all. An increase in geocoding accuracy for address data in 2013 coincided with the introduction of the PHESS database which was accompanied by intensive efforts to improve address standardisation.

Figure 9. Geocoding accuracy for exposure addresses recorded in PHESS (n=2043)



The poor validity of geocoding can be attributed to limitations in the current geocoding interface in PHESS. The geo-coder is sensitive to spelling, formatting and other address variations, and does not provide immediate feedback to the user regarding the

accuracy of the match. Further, staff members from CDPC were unaware that the geocoding 'confidence' level reported on IRRs refers to the confidence of the geocode regardless of whether the match is successful at address, street or postcode level. Thus a site geocoded with 100% confidence may be 100% accurately matched only to postcode level.

PHESS Records of Environmental Detection

I assessed completeness of PHESS records for environmental detection of *Legionella* in case investigations by examining the number of *Legionella* environmental detections recorded in PHESS over the period 2006-2014, and comparing these against those recorded in MDU's LIMS database. Maintenance of complete records of environmental *Legionella* detections in the PHESS database is essential as PHESS is the only database in the system that stores case-related environmental investigation data in a searchable format. However recording of environmental detections in PHESS is poor. The *Legionella* Quick Entry Guide (39) outlines how to enter environmental *Legionella* results into PHESS' laboratory package to create a searchable record. Table 6 shows that environmental detections were consistently recorded in the appropriate PHESS field historically, but no detections have been recorded since 2011.

Table 6. PHESS records of environmental *Legionella* detections, 2006 – 2014

Year	Number of cases	Environmental detection recorded in PHESS (laboratory package)	
		no.	%
2006	69	1	1.45
2007	41	3	7.32
2008	54	3	5.56
2009	48	2	4.17
2010	67	4	5.97
2011	75	0	0
2012	68	0	0
2013	70	0	0
2014	91	0	0
Total	**586**	**13**	**2%**

Results from MDU were available for comparison from January 2010 onwards. Between January 2010 and Dec 2015 MDU reported 22 case-related environmental detections of *Legionella*. This apparent discrepancy occurs because since 2010 results for environmental detections have been noted in the free text 'comments' box rather than the PHESS laboratory package. Results in this field cannot be readily extracted. This change is likely to have arisen through an omission of detail in a PHO handover, demonstrating the importance of complete, thorough and structured handovers, well documented databases and standard operating procedures to ensure stability in surveillance practice.

Outbreak and Cluster Records

Incomplete recording of clusters was documented through the case audit and has been previously discussed. To assess completeness for existing records, I extracted outbreak and cluster records created during the period 2010-2014 from PHESS and reviewed key data fields. To assess validity of the cluster versus outbreak designation for each record I reviewed the onset dates and exposure sites for cases linked to each record and noted any combination of cases that met the formal cluster or outbreak definitions, then compared these against the recorded incident type, as derived from the incident name.

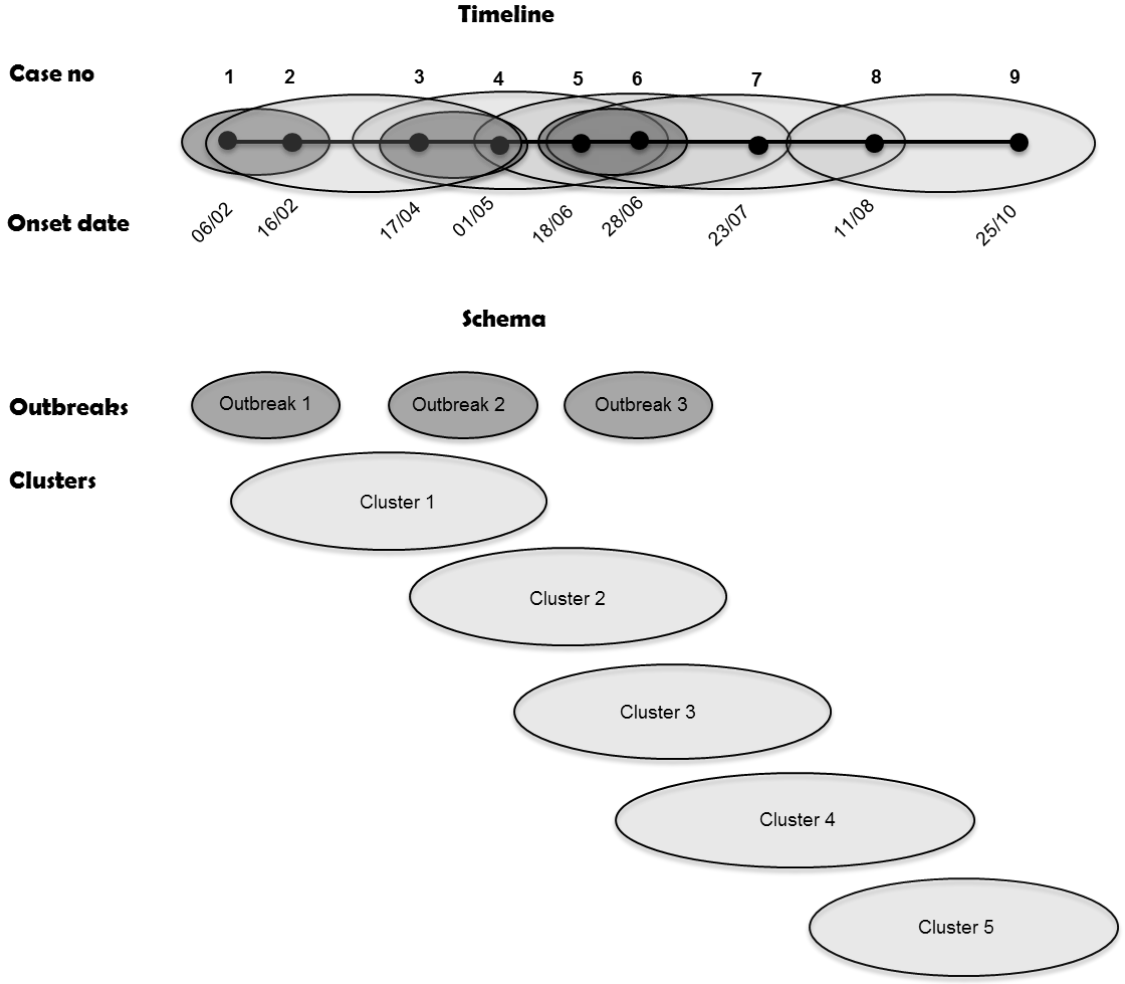
Outbreak records provide important epidemiological information, and are of particular interest in the Legionellosis Surveillance System because the system is geared toward identifying and resolving outbreaks. Completeness of key data fields in cluster and outbreak records was poor. As presented in Table 7 only organism/cause and outbreak reference fields were above 90% complete. The number of cases (laboratory confirmed) was the next most complete field with just over half complete, while completeness for all other assessed fields was below 50%. Poor completeness for outbreak and cluster records may arise because the outbreak record structure is modelled around gastrointestinal outbreaks and there is no Quick Entry Guide to guide record keeping for legionellosis outbreaks/clusters.

Table 7. Completeness of outbreak and cluster records, 2010-2014 (n=17)

Data field	Records reviewed	Complete	
	n	no.	%
Organism/cause	17	17	100
Outbreak reference ID	17	16	94
Number of cases (lab confirmed)	17	9	53
Date of onset of first case	17	8	47
Date of onset of last case	17	7	41
Number of hospitalisations	17	3	18
Number of deaths	17	3	18
Was an organism positively identified	17	3	18
Were water samples obtained	17	0	0

Designation of incident type for clusters and outbreaks was inconsistent due to a tendency for these events to overlap and a lack of clear guidance on naming conventions. Figure 10 demonstrates overlapping clusters and outbreaks at a Melbourne venue during 2015. Each case is represented by a dot on the timeline; the case number appears above the timeline and the onset date for the case is denoted below the timeline. Ovals on the timeline represent clusters (lighter ovals) or outbreaks (darker ovals) defined according to the protocol definitions. Clusters and outbreaks have been individually depicted in the schema below the timeline to demonstrate timing of individual events. However in reality the events overlap on the timeline, producing a single large cluster that extends over 8 months. Note that the size of each circle is proportional to duration of event, not number of cases.

Figure 10. Timeline showing multiple cases at a busy Melbourne venue during 2015, resulting in numerous overlapping clusters and outbreaks.



Designation as cluster or outbreak is recorded using naming conventions as outlined in the description of the system. Seventeen cluster and outbreak records were created between 2010-2014, including seven incidents recorded as clusters, eight recorded as outbreaks and two unspecified incident types. Over half (4/7) of all records designated as "clusters" included outbreaks within the linked cases, ie included linked cases with onset dates less than 14 days apart. Conversely, close to 40% (3/8) of records designated as "outbreaks" included clusters within the linked cases, ie included linked cases with onset dates up to 90 days apart. Overall, 46% of events designated as either outbreak or cluster included linked cases fitting the alternate definition.

Record keeping for clusters or outbreaks that fall outside the protocol definitions can be challenging. One atypical incident included multiple cluster/outbreaks in a small geographical area over a five-month period. Separate records were created for each smaller incident, as well as a larger record that included all cases that had occurred

across the area. This assisted with management of the incident but may cause confusion when it comes to counting the number of incidents or the number of cases associated with outbreaks over time. In another example, *Legionella* was isolated from the cooling towers at a case's workplace and this site was considered to be the most likely source of exposure. Another case worked at the same workplace as the case, but a cluster record was not created for the two cases as the onset dates were just over 3 months apart, even though epidemiologically it is likely the cases had the same source. However the second case had returned from an overseas holiday and was linked to a cluster in the holiday destination, as well as a cluster at the airport in spite of having a primary source identified at the workplace.

Environmental Investigation System

The Environmental Investigation System does not generate case-related data, defined as "a collection of items of information" (40). Instead case records for individual cases are maintained in Case Summary Sheets (word documents) and inspection records in the Emerald database where they are not readily linked to a specific case. Information relating to cases must be manually extracted and assembled from Case Summary Sheets, the emerald database and TRIM records if required for analysis. This makes it difficult to interrogate information generated by the system.

I manually reviewed the completeness of records for the fifteen audit cases in Case Summary Sheets and the Emerald database to determine how completely environmental investigations were documented. An example Case Summary Sheet is provided in Appendix 4, and an Emerald inspection record is shown in Appendix 5.

I assessed data completeness for the fields presented in Table 8 by reviewing the percentage of fields 'not missing', apart from the fields 'cooling towers that have been sampled and disinfected' and 'disinfection dates'. To validate the completeness of records for 'cooling towers that have been sampled and disinfected', I compared the number of cooling towers listed on the case summary sheet against the number of cooling tower samples recorded under the PHESS ID for each audit case in MDU's LIMS database. Sites found to be decommissioned on attendance (as noted in Emerald) were excluded from the calculation. Assessing completeness of disinfection dates in Case Summary Sheets was complicated by the fact that disinfection is indicated for some inspections but not all, and the sheet does not include a field specifically capturing whether disinfection occurred (yes/no). Rather, it includes only a

field for ‘disinfection date’. To differentiate between true missing ‘disinfection date’ and circumstances where disinfection was not indicated (and therefore, where the field was intentionally left blank), I compared the number of disinfection dates recorded on Case Summary Sheets with disinfections noted in Emerald for each system listed on the Case Summary Sheet. This approach assumes the Emerald record is accurate and does not account for disinfections not captured in Emerald. Sites inspected for a case investigation but not recorded on the Case Summary Sheet (ie sites identified only through the LIMS data extract) were excluded from the calculation. Results are presented in Table 8.

Table 8. Completeness of Environmental Investigation System case records

Data Field	Records reviewed	Complete	
Case Summary Sheets	**n**	**no.**	**%**
Cooling towers within 500m of the home Y/N	15	15	100
Results recorded (CTS near home)*	2	0	0
Cooling towers at workplace Y/N	15	9	60
Results recorded (CTS at work place)*	4	0	0
Any other sites inspected Y/N	15	8	53
Results recorded (other sites)*	1	0	0
Cooling towers sampled and disinfected	25	20	80
Date disinfected	12	8	75
Links to other cases noted	4	0	0
Link noted on other case's Case Summary Sheet	4	0	0
Emerald cooling tower database	**n**	**no.**	**%**
Entered as case investigation	25	23	92
Legionella test results entered	21	20	95
Case ID noted in comments box	25	25	100
HCC results entered	21	21	100

** Completeness for sample results was only assessed if an inspection was conducted; ie where the response to “Cooling tower within 500m of the home”, “Cooling tower at workplace” or “Any other sites inspected” was “Yes”.*

Because cases had a variety of exposures, and the number of towers on any given cooling tower property varies, the number of records available to review changed with each data field and this is reflected in the column “Records reviewed”. The fifteen audit cases included three clusters and one outbreak. Twenty five cooling towers were inspected in response to these cases, including sites near the home for two cases, near the workplace for four cases, and at “other” (not further specified) for one case. Twelve towers were disinfected. Three non-cooling tower samples were collected, one from a humidifier, one from a car wash, and one from a spa. Two samples tested positive for *Legionella*, one from a cooling tower (associated with in outbreak) and one from a domestic humidifier.

Completeness of Case Summary Sheets was excellent for recording RIEMS notification number and whether there were cooling towers within 500m of the home. The high level of completeness for the field “were there any cooling towers within 500m of the home” is likely to reflect the question’s location at the beginning of the Case Summary Sheet. There was poor completeness for the similar questions “were there any cooling towers at the workplace” and “were any additional sites inspected”.

Disinfection date was entered in the Case Summary Sheet for only 75% (8/12) of sites that had cooling towers disinfected. As there is no binary (yes/no) field capturing whether disinfection occurred, this means for 25% of sites that were disinfected there is no readily accessible record of disinfection in association with the case.

Links to other cases were not noted on the Case Summary Sheet for any of the cases associated with outbreaks and clusters. As the Case Summary Sheet is stored with the Incident Response Report, it is possible to cross check the Incident Response Report for any linked cases. However noting it on the Case Summary Sheet provides evidence the link has been noted by the Environmental Investigation System and that an appropriate risk assessment has been conducted.

For linked cases, no record was made to distinguish which inspections related to the outbreak/cluster and which related to the single case investigation.

Only 80% (20/25) of cooling tower sites sampled for case investigations were recorded on the Case Summary Sheets. This makes it difficult to compile a complete list of sites inspected for a particular case or cluster of cases.

Results of environmental samples were not noted on the Case Summary Sheets for any cases. The fifteen audit cases included four cases with cooling tower investigations and two cases with non-cooling tower investigations. *Legionella* was isolated in two cases (one cooling tower, one humidifier) and these results were available in Emerald (for cooling tower samples) and TRIM files (for non-cooling tower samples). However none of the results (detection or non-detection) were recorded in the Case Summary Sheets. This means, to determine whether *Legionella* was detected for a given case one must review the Case Summary Sheet to identify which sites were investigated and cross reference Emerald records for each system in each site inspected, and also refer to TRIM for non-cooling tower samples to determine whether *Legionella* was detected during the investigation.

Case-related Emerald fields were all over 90% complete. Note that while 92% of the case related cooling tower inspections were entered as ‘case investigations’, 52% of inspections (13 cooling towers) were outbreak or cluster inspections. These were all entered into Emerald as “case investigation” as there is no option for “outbreak investigation” in Emerald. PHESS case number was consistently noted in the comments box along with other details. However, there was inconsistency in the way it was recorded with some officers including spaces between the digits, and others not.

3.1.f Sensitivity

According to the CDC guidelines sensitivity can relate to both the ability to detect individual cases and the ability to detect outbreaks (1). In this evaluation I focused on the sensitivity to detect clusters and outbreaks using the Geocortex interface, and the sensitivity for finding a source during environmental investigations.

Sensitivity for Detecting Clusters and Outbreaks

The sensitivity for detecting spatial clusters and outbreaks using the Geocortex mapping system is potentially excellent. The automated workflow provides a comprehensive check of all exposure sites within 500m of exposures for the case of interest. However in practice the sensitivity of the workflow is reduced by the quality and range of available data, the output layout, and by the distance selected for the buffer search.

Quality of Available Data

Many of the data quality issues affecting cluster and outbreak detection have been discussed previously. Briefly, the sensitivity of cluster and outbreak detection is limited by the following data quality issues:

Geocoding accuracy: Accurate geocoding is essential because the workflow identifies exposures sites within 500m of each other. As discussed under “Data Quality” only 60% of exposures site addresses recorded for cases during 2013-2014 could be geocoded to address level, reducing the sensitivity of the outbreak detection tool.

Completeness of geocoding: As shown previously in Table 5, only 89% of 80 exposure sites for the 15 audit cases were geocoded. To ensure detection of all clusters and outbreaks ideally the target would be closer to 100%. Sites that are not geocoded are excluded from the sampling frame for outbreak detection.

Non-notifiable cases: Non-Victorian residents with exposures in Victoria are omitted from the Geocortex data tables as they are classified as “Not Notifiable’ in PHESS, and therefore cannot be assessed by the outbreak detection workflow.

Recall of exposure sites: Cases may have imperfect recall of sites visited during their acquisition period, which means the reported sites may not represent all the sites a case attended. This may be particularly true for cases interviewed while still symptomatic (as confusion is a common symptom of Legionnaires’ disease) and for cases with a serological diagnosis who may be interviewed weeks or months after their illness.

Available Data Range

The sensitivity of the Geocortex workflow for cluster detection is limited by its execution in the “*Legionella* 90 days” data table which includes exposure sites for cases notified within the past 90 days. However, cluster detection requires a review of cases with *onset dates* within 90 days of the case’ *symptom onset date*, rather than *notification date* within 90 days of the *date of case classification*.

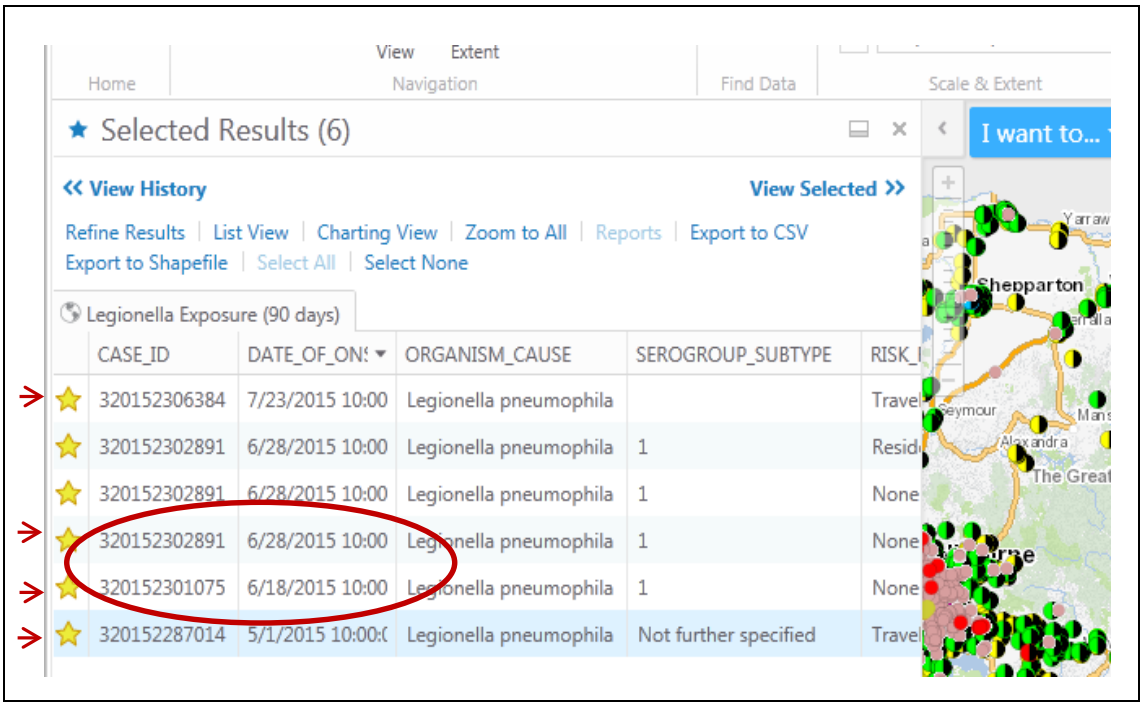
Consider a serology case confirmed 44 days after symptom onset (the median interval for serology cases during 2014, excluding *L.longbeachae* cases). The Geocortex workflow would be conducted following confirmation and case interview, ie 44 or more

days after symptom onset. To detect all possible clusters with the current case, exposure sites for cases with *onset dates* up to 134 days prior to the current date (44 days + 90 days) need to be examined. The sampling frame limits the sensitivity of the workflow as the 90 day data table only includes cases with a notification date within the previous 90 days. Thus some clusters will be missed by the workflow cluster checks.

Output Layout

The sensitivity of the workflow for detecting clusters and outbreaks may be affected by the need for careful interpretation of the workflow output. The output is presented as a table of exposures, including those for the case of interest as well as exposures from other cases within 500m of these notified during the preceding 90 days. The tables do not highlight cases that are different to the case of interest and therefore need to be carefully checked to identify potential outbreaks or clusters. The dates of onset and exposure sites for additional cases detected within the buffer ranges must also be examined to decide whether the shared exposure constitutes an outbreak or a cluster and whether these represent single or multiple sites. Figure 11 shows a workflow output in which both a cluster (marked with arrows) and an outbreak have been detected, but these are not specifically noted by the output table. For descriptive purposes the outbreak has been circled and the cluster is marked with arrows; however no such alert appears in practice.

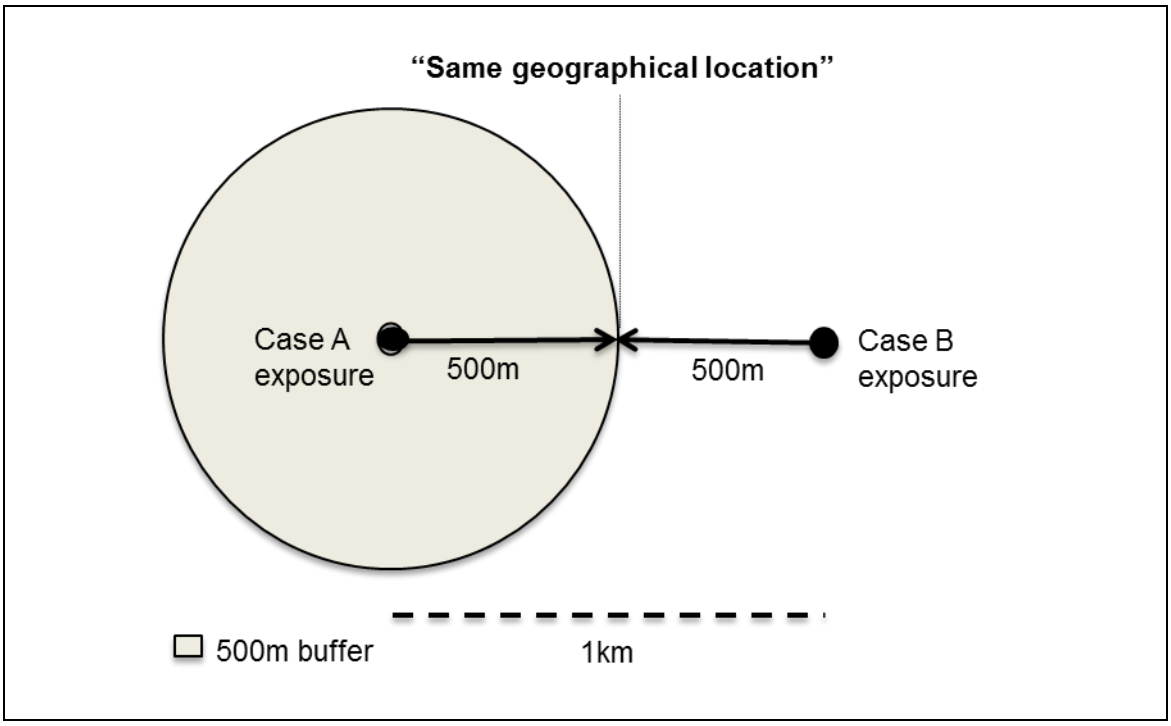
Figure 11. Geocortex workflow output table demonstrating an outbreak and a cluster



Buffer Search Distance

Sensitivity of the workflow for detecting clusters and outbreaks can be adjusted by altering the size of the buffer applied to search for other case exposures. The standard practice is to apply a 500m buffer to identify any exposure sites within 500m of the exposure sites reported by the current case. However the case definition for clusters or outbreaks includes cases that “are linked to within 500 metres of the same geographical location in their incubation period”. The wording in this definition leaves the size of the required buffer open to interpretation, as demonstrated in Figure 12. The scenario shows the exposure sites for two cases, case A and case B. Each exposure site lies within 500m of the ‘same geographical location’. However the exposure site for case B lies 1km from case A and therefore is not detected using a 500m buffer from case A. The ‘same geographical location’ is not captured by the workflow as the system only searches for case exposure sites.

Figure 12. Two cases with exposure sites within 500m of a shared geographical location and failure of a 500m buffer to detect both exposure sites



Sensitivity for Identifying a Source

I examined the ability of the Environmental Investigation System to detect the environmental source of infection during case investigations between 2010-2014. I examined sensitivity for both implicating and confirming sources according to the definitions used by the system (39), whereby "implicated" means *Legionella* was isolated from an epidemiologically linked environmental source, regardless of whether the species was identical to the case species, while "confirmed" means the environmental detection was matched by PFGE to a clinical isolate from the case.

I requested environmental sampling results and associated PHESS case numbers from MDU's LIMS database for the period 2010-2014, and calculated sensitivity using the formula:

$$\text{Sensitivity} = \frac{\text{number of cases with a source implicated } \textit{(or, confirmed)}}{\text{number of cases with environmental investigation}}$$

The Environmental Investigation System conducted environmental investigations for 138 cases during the period 2010-2014. This included cooling tower investigations for 111 cases and non-cooling tower investigations for 68 cases. A number of cases had both sources investigated. Record keeping practices meant it was not possible to

differentiate between single case and outbreak investigations. As presented in Table 9, the sensitivity for implicating a cooling tower source was 14%, while non-cooling tower investigations implicated a source for 19% of investigations. For non-cooling tower sources, the species from the implicated source was often difficult to compare to the clinical species, as five of these cases were diagnosed as "*Legionella* not further specified" based on urinary antigen result (Enzyme immunoassay (EIA) positive and immunochromatography (ICT) negative). Interestingly all of these cases had *L. anisa* detected from environmental shower samples.

Table 9. Sensitivity for implication of environmental sources for cooling tower and non-cooling tower investigations

Source	Investigations	Sources inspected	Source implicated		Source implicated with species match	
	n.	n.	n	%	n	%
Cooling tower	111	658	16	14	15	14
Non-cooling tower	68	336	13	19	6	9
All	138	994	28*	20	15	11

* Note one case had both a cooling tower and a non-cooling tower source implicated

The 138 environmental investigations included inspection and sample collection from 658 cooling towers. *Legionella* bacteria were isolated from 2.8% (19/658) cooling tower samples (some sites had multiple positive systems). This detection rate is comparable to the overall positivity rate for randomly collected CTS samples during the same period (3.9%), reflecting the fact that *Legionella* are ubiquitous organisms that can be cultured from many different environmental sources in the absence of associated disease (41). Thus, at least some of the sources 'implicated' by the detection of *Legionella* during environmental investigation will reflect the background contamination rate in cooling tower (and presumably, also non-cooling tower) systems. This highlights the necessity for further typing of both case and environmental specimen in order to confirm environmental sources as the cause of disease.

Sensitivity for *confirming* sources (ie demonstrating matching PFGEs for clinical and environmental isolates) relies on the availability of both environmental PFGE results and clinical isolates from cases. However, PFGE is only performed if environmental isolates are *Legionella pneumophila* 1, and thus PFGE results were available for around half of all environmental *Legionella* isolates during the period. Amongst the 28 cases that had a *Legionella* environmental detection, only seven also had a clinical

isolate available. Three of these had a PFGE result available for the epidemiologically linked environmental *Legionella* isolate. Two of these isolates matched on PFGE, one each for cooling tower and non-cooling tower investigations. Thus the sensitivity for confirming sources was 0.9% and 1.5% for cooling tower and non-cooling tower investigations respectively.

Availability and Typing of Clinical Isolates

The exceptionally low sensitivity for confirming sources is due, in part, to the small number of clinical isolates available for comparison with environmental samples. Implicating an environmental source of infection requires availability of clinical isolates as well as environmental isolates. Between 2010-2014, sputum sample results were available for only 120 of 365 confirmed and probable cases. Clinical isolates can be challenging to obtain as legionellosis is often characterised by a dry cough, making respiratory specimen problematic to collect (42). Respiratory specimen may be more likely to be collected in critically ill patients where interventions such as intubation provide an opportunity for sample collection. Note that the PHESS field recording whether or not sputum had been collected was left blank in 57% of cases during 2010-2014, raising the question of whether the PHO's had discussed the importance of sputum collection with treating clinicians.

Even once a sputum sample is collected, isolation of *Legionella* bacteria can be challenging, particularly if the sample is collected after commencement of antibiotics. Of the 120 confirmed and probable cases where a sputum sample was collected during the period 2009-2014, only 56% (67/120) were culture positive. A recent study found *Legionella* could be isolated from only around two-thirds of hospitalised cases confirmed on urinary antigen, but this improved to 80% where samples were collected within 48 hours of hospital admission (43).

A method to provide clinical typing information in the absence of a *Legionella* clinical isolate has recently been evaluated (43). The approach begins with qPCR detection of *Legionella* in respiratory specimen, followed by direct amplification of *Legionella* DNA and the use of sequence based typing (SBT) to analyse 7 genes that can be used to differentiate between strains of *Legionella*. While the technique provides an excellent solution to typing in the absence of a clinical isolate, the clonal nature of *Legionella* can limit the capacity of SBT to differentiate between common strains in some areas (44). However in the absence of a clinical isolate this approach may add significant value to

epidemiologic investigations. Whole genome sequence analysis is emerging as a viable surveillance technology and provides the most discriminatory method for analysing epidemiologically linked *Legionella* specimen, but currently relies on obtaining both clinical and environmental isolates (44).

3.1.g Flexibility

Flexibility allows the surveillance system to adapt to changing circumstances with minimal additional resources (1). The combined Legionellosis Surveillance System is characterised by excellent flexibility for response. Both the Disease Surveillance and Environmental Investigation Systems have excellent surge capacity for response to outbreaks or other incidents. The PHESS database is a flexible system with the capacity to accommodate cooling tower and non-cooling tower environmental investigation data which could help overcome current challenges with the storage and retrieval of case related environmental data.

Flexibility to Respond to Outbreaks

The Legionellosis Surveillance System has excellent flexibility for outbreak response. PHOs rotate duties every six-nine months. This ensures there are always a number of officers in the team who understand the protocol for legionellosis case-follow up and can be mobilised to provide additional support if required during outbreaks. The Environmental Investigation System also has excellent surge capacity. The LT is predominantly involved in regulatory activities; indeed 70% of cooling tower inspections conducted by the team are regulatory rather than case related. The team is able to rapidly re-prioritise and respond when a case requires environmental investigation.

Flexibility to Integrate Data

The information technology available to the system is very flexible and provides a means to integrate the data into the future. The PHESS database can readily be extended to capture cooling tower and non-cooling tower environmental investigation results, and an automated reporting system is available that can transmit this data from PHESS to Emerald to prevent double entering results.

3.1.h Acceptability

I explored the willingness of staff to be involved with the system through stakeholder interviews. The system was considered acceptable by most staff, although in general

acceptability would increase with improved coordination between the Disease Surveillance and Environmental Investigation Systems.

Acceptability for Disease Surveillance System Staff

Staff within the Disease Surveillance System felt the system was acceptable overall. Most staff commented that acceptability would improve with improved communication and co-ordination between the Disease Surveillance and Environmental Investigation Systems. In particular staff members were eager to receive feedback summarising results from environmental investigations, consistent with the culture of feedback within clinical professions. PHOs expect to receive a summary of environmental investigations before 'closing' or completing each case. When they don't receive these results they find it difficult to decide when to close a case which has implications for their workflow. Staff felt feedback of results would improve work satisfaction and also provide a more complete epidemiological record. Interestingly the protocols for the two systems conflict on this point. Section 6.9 of the Legionellosis disease investigation protocol advises "The CDPC PHO should obtain copies of any environmental sampling conducted and enter these results" (31), while the protocol for the Environmental Investigation System advises CDPC need only be advised of results that test positive for *Legionella* during outbreak investigations (35).

Many stakeholders within the disease surveillance system commented that the *Legionella* PHO would benefit from an induction with the LT, including a field trip to inspect and sample a cooling tower. This would improve collaboration between teams and increase co-ordination by developing the PHO's understanding of the activities within the Environmental Investigation System.

Acceptability for Environmental Investigation Staff

Staff within the Environmental Investigation System found the system acceptable although acceptability was reduced by the lack of transparency between systems. Most commented that acceptability was improved when there was regular communication from the PHO, particularly when the PHO made personal contact. Many had little confidence in the accuracy of information gathered through case interview, and reported it was not uncommon to encounter additional exposures during field investigations that were not reported at interview. Some suggested that ideally a second interview should be completed when cases were recovered to ensure exposure

information was complete, however it was acknowledged this could be resource intensive and the benefits gained through interview may need to be weighed against the resources involved. Acceptability was also reduced by inaccuracies in geocoding which was compounded by the inability for LT staff to access PHESS to check or correct geocoding when required to correct an exposure site on the maps.

Acceptability of Geocortex

Geocortex was widely accepted amongst interviewees. However many staff commented it is underutilised, and felt the current system is "just scratching the surface" of what can be achieved. Most staff felt the system had a great deal of potential but they had not been adequately trained in using the system, and acceptability would improve with training.

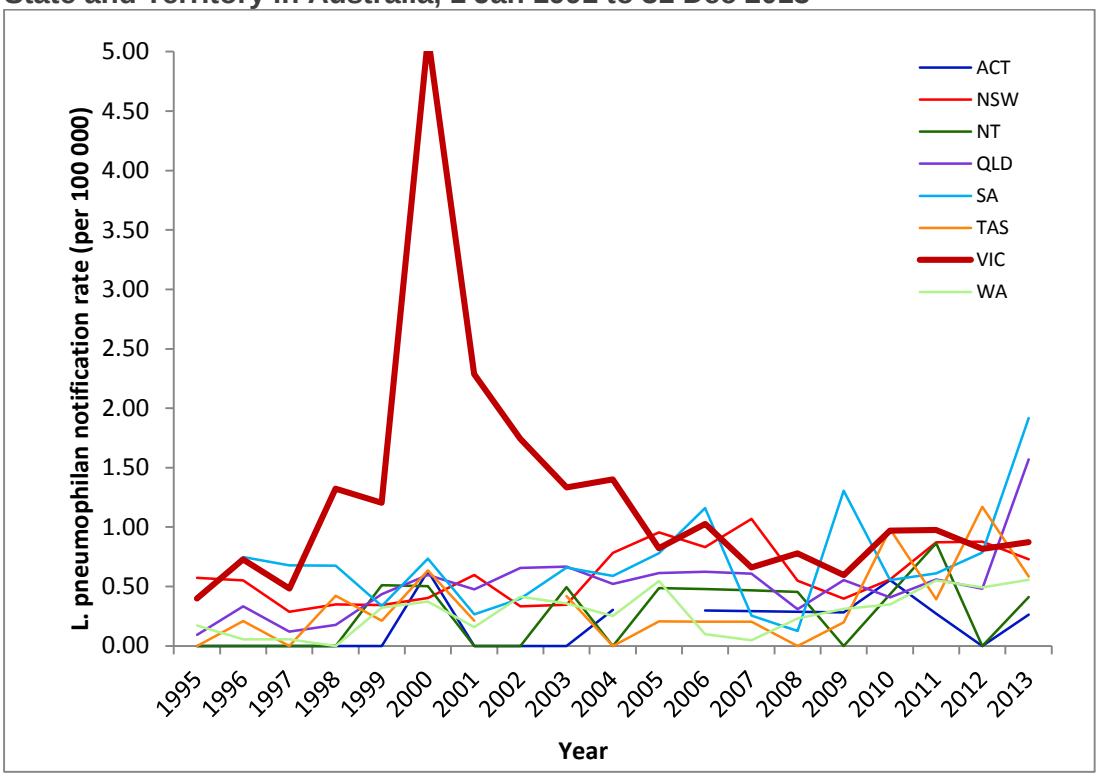
3.2. System Usefulness

I assessed the usefulness of the system for reducing transmission from environmental sources. I assessed this by reviewing its impact on the overall incidence of legionellosis and outbreaks of legionellosis in Victoria, and by examining the risk of *Legionella* detection in cooling tower systems targeted during case investigations. The system is unique to Victoria and focuses on reducing clusters/outbreaks which are predominantly caused by *L. pneumophila*. Therefore I compared rates of legionellosis due to *L. pneumophila* for the periods before and after the introduction of the system, and compared Victorian rates with those observed in other jurisdictions. Next, I examined the frequency and size of clusters/outbreaks before and after introduction of the system. Finally, I assessed the usefulness of environmental investigations by comparing the risk of *Legionella* detection in cooling towers targeted for case investigations with the risk of detection in randomly selected cooling tower samples.

Impact on L. pneumophila Notifications Over Time

To assess the impact of notifications over time relative to other jurisdictions, I requested notification data according to species for each jurisdiction for each year between 1995-2013 and calculated rates per 100,000 using population data from the Bureau of Statistics (27). As shown in Figure 13, the Victorian notification rate is amongst the highest in Australia and does not appear substantially reduced following the introduction of the combined Legionellosis Surveillance System during 2000.

Figure 13. Notification Rate of Legionellosis due to *L. pneumophila* according to State and Territory in Australia, 1 Jan 1991 to 31 Dec 2013*



* Disease data from the National Notifiable Disease Surveillance System; population data from Australian Bureau of Statistics (27).

Next, I formally assessed whether Victorian incident rates differed during the period prior to the introduction of the system compared to the following period using negative binomial regression in STATA (45). The enzyme immunoassay urinary antigen test was introduced in 1995 (46) resulting in an increase in case detection during 1995-1997; therefore I selected 1998 and 1999 as baseline years and examined incident rates in the years following the introduction of the system (2001-2013) relative to these. Note that the year 2000 was excluded as it was an outbreak year. There was no significant difference in the incident rates of disease in the years after the introduction of the system (incident rate ratio for annual incidence during the years after the introduction of the system relative to the incidence during the years prior to the system 0.86, 95% confidence interval 0.51-1.44).

Impact on the Frequency and Size of Clusters/Outbreaks

The combined Legionellosis Surveillance System responds to legionellosis cases by identifying potential sources of transmission and rendering them safe to prevent further disease. If the system were useful for this purpose the frequency and size of clusters

and outbreaks should be reduced following the introduction of the system. To assess this, I examined the frequency and size of clusters and outbreaks recorded in the PHESS database over time, excluding those with overseas exposures. Due to inconsistencies in record keeping (as discussed under 'Data Quality') clusters and outbreaks could not be accurately stratified, and thus are presented together. Inconsistencies in record keeping also mean not all clusters (and potentially outbreaks) are recorded in the database. In addition, poor availability of molecular typing information (due to infrequent sputum collection and subsequent culture) means the outbreak and cluster definitions likely have poor specificity (ie events that are not truly clusters/outbreaks may be counted as clusters/outbreaks). These significant limitations mean this data needs to be interpreted with caution. Figure 14a presents the number of clusters/outbreaks each year, while Figure 14b presents the number of cases recorded for each incident between 1998 and 2014, including a line of best fit showing the trend in number of cases per incident over time. An unusually large number of clusters and outbreaks were recorded during 2003. The graphs demonstrate there was no reduction in the frequency or size of outbreaks/clusters following the introduction of the combined surveillance system in the year 2000.

Figure 14a. Annual number of legionellosis clusters/outbreaks in Victoria, 1998 – 2013.

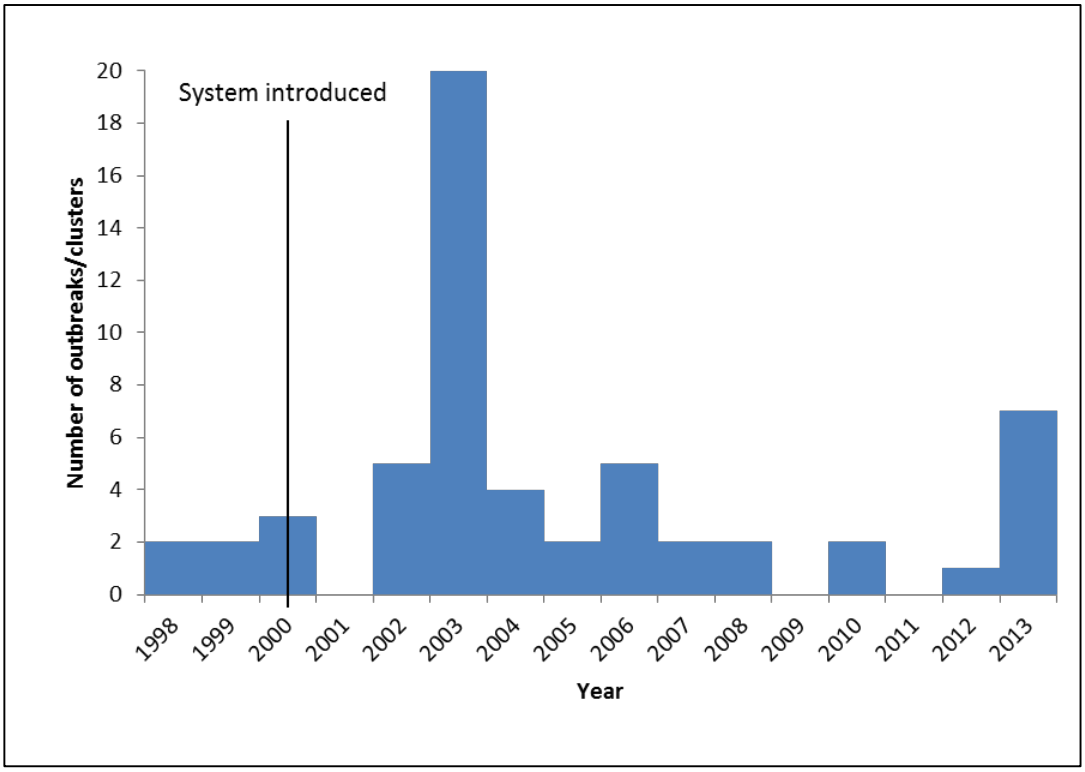
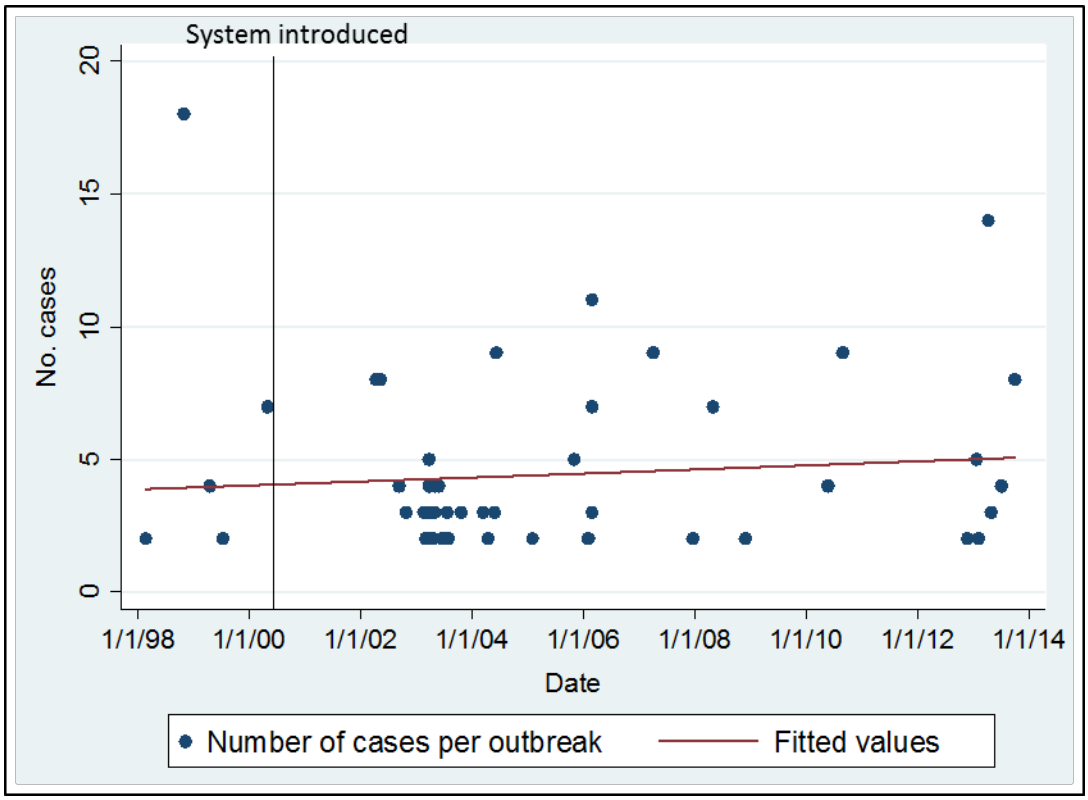


Figure 14b. Number of legionellosis cases per outbreak/cluster Victoria 1998-2014



Risk of Legionella in cooling towers targeted for case investigation

I assessed the usefulness of the current protocol for environmental investigations by comparing the risk of *Legionella* detection in case-related cooling tower inspections against the risk of detection in cooling tower samples collected for other purposes during the period 2010-2014.

I extracted cooling tower inspection data from Emerald and analysed it in STATA (45). There were 37 inspection types recorded during the period. I identified major inspection types in consultation with LT staff and collapsed inspection records into meaningful categories. Inspection records listed in Emerald as “other” ”survey” or “record only” were excluded from the analysis. I calculated the risk of Legionella detection for each inspection category over the five year period, and formally compared the risk of detection for case investigations against the risk of detection for randomly selected towers using chi square analysis. It was not possible to stratify the analysis to assess whether there was a higher risk of Legionella detection for outbreak versus cluster or single case investigations as these categories are not differentiated in the Emerald

cooling tower database. Table 10 presents the results of the analysis. The number and per cent of missing results for each category is also presented.

Table 10. Cooling tower Legionella detections according to inspection type, 2010 – 2014.

Inspection Type		*Legionella* Detected			Result Missing	
	n	No.	%	95% CI (%)*	No.	%
Case	987	27	2.74	(1.81, 3.96)	114	11.55
Random	482	19	3.94	(2.39, 6.09)	92	19.09
Compliance	3,112	81	2.60	(2.07, 3.2)	766	24.61
Follow up	1,228	66	5.12	(4.18, 6.79)	377	29.27
Total	*5,869*	*193*	*3.29*	*(2.84, 3.78)*	*1,349*	*22.99*

95%CI: 95% Confidence Interval

The risk of identifying *Legionella* in cooling towers sampled for case investigations was only 2.74%. This was lower but not significantly different to the risk of detecting *Legionella* in randomly sampled towers (chi square P value=0.119). The similarity between risk of *Legionella* detection in case-related versus randomly selected sites suggests the sites currently targeted for case investigation do not pose an increased risk to the public compared to other cooling towers.

At the time of the inspection, cooling tower operators are directed to disinfect sites linked to a case (ie. within 500m of home, at the workplace or within 500m of an outbreak) within 24 hours. The only exception is sites investigated for cluster investigations or cooling towers beyond 500m from shared exposure sites in an outbreak investigation. In these cases the need for disinfection is determined by a risk analysis and sample results. The financial burden of disinfection is borne by the cooling tower operator. The finding that sites inspected for case-related investigations have the same risk of *Legionella* detection as a random sample of sites makes this request for immediate disinfection difficult to justify.

Sample results were missing for over 20% of all inspections. This missing data may bias the reported risk of detection upwards because *Legionella* detections are likely to be more reliably recorded than non-detections. The proportion of missing sample results varied between inspection types. Those with a higher percentage of missing data are likely to be disproportionately affected by recording bias.

4. Discussion

The Victorian combined Legionellosis Surveillance System is a complex system that integrates infectious disease and environmental surveillance components. The use of the Geocortex mapping system simplifies the interface between the two systems, however staff require appropriate training to avoid some important pitfalls. Overall the integration between the two systems could be significantly improved through more regular, structured engagement between key stakeholders and access to a shared database.

The system has stable information technology but experiences some instability due to human resourcing. Regular staff handovers within the Disease Surveillance System result in loss of skills and knowledge in the complex system. These could be overcome through structured handovers including a handover checklist and a dedicated training session on Geocortex.

Case notification and confirmation is timely for cases notified by urinary antigen or respiratory specimen, but could be improved for serology cases by developing targets to guide timely follow-up. Communication of cases to the Environmental Investigation System and subsequent case investigation occurs promptly.

Data quality was good for the majority of data fields examined within the Disease Surveillance system. However the quality of geocoding and information updates to the Environmental Investigation System could be improved and *Legionella* environmental detections have not been entered into the database properly since 2010.

Data quality within the Environmental Investigation System was hampered by the use of multiple databases and reliance on Emerald, which is not designed to store case related data. Case Summary Sheets provide the only comprehensive record of case related investigations, however these were frequently incomplete which means collating information on case investigations is labour intensive and error-prone. Record keeping for outbreaks and clusters was poor across both the Disease Surveillance and Environmental Investigation Systems.

The system has excellent sensitivity for detecting clusters and outbreaks, although inaccurate geocoding reduces this sensitivity. The specificity of outbreak and cluster detection could be improved with increased availability of molecular typing to identify cases that were infected with the same or similar strain of *Legionella*. The system

implicates environmental sources moderately frequently, but this may reflect background contamination rates as a large number of sources can be sampled for a single investigation. The system has remarkably low sensitivity for confirming environmental sources through molecular typing. This is at least partially attributable to the low number of cases with clinical specimen available, and the fact that typing of environmental isolates is only undertaken for a portion of isolates *(L. pneumophila 1)*.

The system is flexible and provides excellent surge capacity for outbreak response, and the PHESS database offers flexibility that should enable integration of the two systems into a single database. Overall the system was considered acceptable by stakeholders, although acceptability would improve with improved integration between the two components.

Based on the available data, the system does not appear useful for reducing transmission from environmental sources. The introduction of the system had minimal impact on the notification rates of legionellosis due to *L. pneumophila*, and has not reduced the frequency or size of clusters and outbreaks observed. The lack of change in cluster and outbreaks may be a record keeping artefact due to inconsistency in record keeping over time. However, the finding that cooling tower sites currently targeted for case investigations are at no greater risk of contamination with *Legionella* than randomly selected cooling towers suggests the usefulness of targeting these sites for sporadic case investigations is limited. However, the Environmental Investigation System is almost cost neutral and operates within a regulatory framework. Disease investigation activities may be seen as an extension of regulatory activities.

The Environmental Investigation system samples and disinfects systems within 500m of case homes or at the workplace for all cases. This approach is consistent with the findings of a Glasgow study that found people living within 500m of a cooling tower had three fold the risk of illness compared to people living further than one kilometre from a cooling tower (47); however this study has not been validated in the Australia. Systems within 500m of other case exposure sites are tested in response to outbreaks (and occasionally clusters). However, aerosols spread from a contaminated cooling tower have been shown to cause illness in cases exposed up to 6km from the most likely source of an outbreak (48). This extensive aerosol spread makes it difficult to identify the environmental source for legionellosis even where cooling towers are the source. In Victoria, systems greater than 500m from case exposure sites may be tested in

response to larger outbreaks or clusters, but it is not practical or efficient to test all such towers for routine investigations.

The usefulness of the system may be improved through more strategic selection of cooling towers for investigation. Investigative efforts currently spent on sporadic case investigations may be better invested in response to clusters and outbreaks. Currently, very little environmental investigative effort is focussed on clusters. Strategic cooling tower selection may include increased use of molecular typing to improve the specificity of cluster and outbreak definitions, and more sophisticated use of mapping to target areas associated with spatial clustering of disease.

Although there is little evidence the system has had impact on the overall burden of legionellosis in Victoria, the system is undoubtedly useful for the rapid detection and prompt investigation of significant outbreaks. The structure of both CDPC and the *LT* provides excellent surge capacity when required for large outbreak investigations. Usefulness of the system may be improved with more strategic selection of cooling towers for sampling, and improved molecular typing to improve specificity of outbreak and cluster detection.

5. Recommendations

A number of recommendations to improve the operation and usefulness of the system are presented below. These include major and minor recommendations for the combined Legionellosis Surveillance System, the Disease Surveillance System, the Environmental Investigation System and for the Geocortex mapping system.

5.1 Combined Legionellosis Surveillance System

Major Recommendations for Combined System

5.1.1 Use a single database

The system would benefit from moving toward a single repository of information for case related data. PHESS has the potential to store detailed environmental investigation information for cases for both cooling tower and non-cooling tower exposures. This could include systems/sites investigated for each case, whether the investigation was for a sporadic case, cluster or outbreak, laboratory results for samples, and the reason for sampling (eg residence, workplace, outbreak, or other). This information could be linked to the case of interest or to relevant clusters/outbreaks resulting in a complete, easily accessible record of case information in a single database.

Integration would improve significantly if the two systems worked form a shared database. Adopting PHESS for storage of data relating to Environmental Investigations would overcome many of the case-related data storage challenges faced by the Environmental Investigation System as PHESS is specifically designed to store case-related information. Access to the PHESS database would reduce the need for email updates, enable the LT to freely check and verify exposure sites, onset dates, or other relevant information, and would also enable stakeholders from the Disease Surveillance System to access results of environmental investigations.

5.1.2 Review definition and protocol for clusters versus outbreaks

The distinction between clusters and outbreaks is unique to Victoria. The SoNG provides only a single definition (32). Clusters and outbreaks in Victoria frequently overlap and the definitions can cause confusion. Recorded outbreaks in the system often extend into clusters, and clusters often contain outbreaks. This may reflect sustained transmission from a single source, or coincidental movements in cases that

are unrelated but difficult to differentiate using current typing methods. Currently, little investigative effort is applied to clusters.

Consideration should be given to adopting a single definition and investigating all incidents with the same vigour. Ideally the definition would align with the guidelines provided in the SoNG (32) which considers a cluster/outbreak to be two or more cases within 100m of a shared exposure and onset dates within 90 days. However, other definitions are used, for example in New Zealand, outbreaks are 2 or more cases with a single exposure site and dates or onset within 6 months of each other (49) although all 19 outbreaks since 1990 involved cases separated by a month or less. In the Netherlands, outbreaks are any two cases with a shared exposure site and onset dates within 2 years (13).

5.1.3 Invest in geospatial analysis for identification of spatial clustering

To overcome limitations of current molecular typing methods for identifying clusters and outbreaks, the system should invest significantly in the development of geo-spatial methods for the detection of spatial clustering. See discussion under "Recommendations for Geocortex".

5.1.4 Engage through regular, structured meetings

System integration would benefit by having a clearly defined committee of responsible parties who meet regularly to provide leadership and review the components of the system. The aim of such a group would be to ensure key decision makers understand the complex system, activities remain co-ordinated, data flow is unimpeded and any changes to one component of the system are accounted for by the other. The committee could regularly review recent cases and investigation findings to ensure completeness and identify issues for resolution. Selecting representatives for this committee will be important. While Public Health Officers should be represented, it will be essential to also include a consistent representative from CDPC (such as a team leader) to ensure continuity due to the rapid changeover of Public Health Officers with each disease rotation. The *Legionella* Epidemiologist and LT leader should also be represented.

5.1.5 Establish a working group to improve molecular typing

The system would benefit from engaging with VIDRL and MDU to establish a working group to develop and evaluate a culture free qPCR sequence based typing technique, enabling typing of clinical samples in the absence of an isolate. Ideally, the departmental arm of the group would also work to incorporate molecular typing information into outbreak/cluster definitions to increase the specificity of these definitions. Sequence based typing of clinical and environmental isolates should also be evaluated as a potential means to link clinical and epidemiologically linked environmental samples in the absence of a clinical isolate.

Ultimately, the discriminatory power of whole genome sequencing should be assessed, including reviewing the diversity of whole genome sequences amongst Victorian environmental isolates. If the method has good discriminatory power, the working group should work towards development of meta-genomic methods to enable whole genome sequencing on clinical samples in the absence of an isolate. Once whole genome sequencing is available to type both clinical and environmental samples, consideration should be given to adopting whole genome sequencing in preference to PFGE or SBT to confirm environmental sources.

The departmental arm of the working group could also work toward increasing collection of sputum samples for typing by raising awareness amongst PHOs and clinicians of the importance of doing so.

5.1.6 Orient incumbent PHOs with the Legionella Team

Officers from both the Disease Surveillance and Environmental Investigation Systems suggested incumbent PHOs would benefit from an extended orientation session with the LT, ideally including a cooling tower inspection. Such an arrangement would be beneficial in building organisational networks as well as improving operational knowledge of the overall system.

Minor Recommendations for Combined System

5.1.7 Streamline incident notifications

During stakeholder interviews a number of senior stakeholders expressed a preference to receive fewer notifications regarding legionellosis. Currently many stakeholders

receive both the IRR and the RIEMS incident notification to notify them of a case of legionellosis. While the information received is relevant for some stakeholders, the recipient lists could be reduced. For example, IRRs could be routinely sent to the *LT* for single cases, and distributed more widely only when a cluster/outbreak is identified.

5.1.8 Improve investigation of cases with a serological diagnosis of L. longbeachae

This includes raising awareness amongst PHOs and the LT that *Legionella* is difficult to speciate based on serology, and thus cases diagnosed with *L. longbeachae* on serology could be due to any species. As such, consideration should be given to conducting a full investigation including collecting a complete exposure site history and Environmental Investigation (if indicated) as for any other case of legionellosis.

5.1.9 Develop a protocol for identifying clusters/outbreaks due to L. longbeachae.

Currently there is no systematic way of identifying clusters/outbreaks due to *L. longbeachae* with shared potting mix/compost exposure. The system would benefit from a protocol for cross checking potting mix/compost exposures for *L. longbeachae* cases to identify shared non-geographical exposures (eg potting mix /compost).

5.1.10 Clarify the use of buffers and the cluster/outbreak definition.

If the current cluster/outbreak definition remains unchanged, technically the buffer distance should be increased to 1km to capture all potential exposure sites within 500m of the same geographical locations. Such an adjustment may increase the sensitivity for detecting clusters and outbreaks, although is likely to adversely affect specificity as some exposures may lie within 1km of each other through chance alone. Increasing the buffer distance is also likely to significantly increase the resources involved in environmental case investigations as many more outbreaks and clusters may be detected this way. Alternatively, the case definition itself may be reworded to better reflect current practice.

5.2 Recommendations for Disease Surveillance System

Major Recommendations for the Disease Surveillance System

5.2.1 Develop a structured approach to PHO handovers

To ensure the complexities of the system are communicated in full at each disease rotation, handovers should be conducted in a structured fashion. Ideally, a structured

handover checklist would be developed. This could guide the handover and be signed by incumbent PHOs to ensure handovers are complete and thorough.

At a minimum, handovers should include the *Legionella* Epidemiologist and cover the following information:

- Basic outline of combined Legionella Surveillance System including the roles of CDPC and LT and introduction to important contacts
- Case definition including complexities for serology cases
- Diagnositic methods including importance of collecting a sputum sample
- Speciation – unreliable using serology alone; how to determine species using *Legionella* urinary antigen tests
- Collecting exposure information
- Geocoding and how to check geocode completeness and accuracy
- Identifying clusters/outbreaks
- Record keeping for clusters/outbreaks
- Communicating information to the LT
- Recording results of environmental investigations in the laboratory package

In addition, the following should be included in the handover process:

- A Geocortex training session with the *Legionella* Epidemiologist and/or mapping administrator to learn how to use the mapping software
- Ideally, an orientation session with the LT including a cooling tower site inspection.

5.2.2 Modify the PHESS interface to provide immediate feedback on geocoding accuracy and allow manual override of inaccurate geo-coded data.

A major alteration to the PHESS contract has been requested through the software provider to address this recommendation. Until this is implemented PHOs should do a manual check of each exposure site to ensure it is plausibly positioned.

5.2.3 Introduce a routine check for geocoding completeness

The completeness of geocoding for a case can be rapidly assessed using the IRR exposure site table, and this should be routinely performed while awaiting the improvements in PHESS discussed above. The geocoding 'confidence' field appears blank where sites are not geocoded. PHOs' should be encouraged to check this output to ensure geocoding completeness for every case prior to circulating the IRR (if IRRs are discontinued after transitioning to a shared database, the exposure table should still be generated for the purpose of checking geocode completeness).

5.2.4 Enter case related environmental detections in PHESS

The poor completeness of environmental detections in PHESS would be addressed by adopting a shared database that includes all case related environmental results, both detections and non-detections, as previously discussed. However in the short term the problem should be addressed by raising awareness amongst PHOs of the importance and approach to entering results in the appropriate field after receiving the "Notification of Legionella detection" form from LT.

5.2.5 Increase collection of sputum samples

Raise awareness amongst PHOs of the importance of sputum sample collection. Routinely report status of sputum samples for each case (ie collected or not) at surveillance meetings. Measure completeness of the field "sputum sample collected yes/no" during routine data quality checks (eg data quality manager monthly reports).

5.2.6 Improve outbreak records

Develop an outbreak package specific to legionellosis to improve data quality for cluster and outbreak records.

Develop an Outbreak Quick Entry Guide to provide guidance for PHOs completing legionellosis outbreak records. Include clear guidelines on recordkeeping where events overlap. Conduct quality checks to ensure outbreak records are complete.

Develop a formal reporting mechanism for outbreaks; for example an automated Incident Response Report. This can be used to communicate information including a unique identifier for outbreaks and clusters to the Environmental Investigation System and the Office of the Chief Health Officer.

5.2.7 Use Geocortex during weekly surveillance meetings

This will improve familiarity with the system, provide regular quality checks, and promote trouble shooting.

Minor Recommendations for the Disease Surveillance System

5.2.8 Adapt case definition for serology cases to capture the complexities in diagnosis

Currently, the case definition for serology cases does not reflect the complexities in diagnosis. Ideally the case definition would be adapted to succinctly summarise these details. Box 3 provides an example of how the laboratory component of the case definition could be adapted. Clarifying these rules would also improve transparency when comparing surveillance data between jurisdictions.

Box 3. Suggested adaptations to laboratory component of case definitions for serology cases to simplify case classification

Confirmed case:

- Seroconversion or a significant increase in antibody level or a fourfold or greater rise in titre to *Legionella, where the second bleed is collected 3-8 weeks following illness onset and tested in parallel with the first*

Probable case:

- Single high titre to *Legionella (≥512), where the sample is collected at least 2 weeks after illness onset*

5.2.9 Include suspect cases in the data transmitted to Geocortex

These cases could then be included in the workflow that identifies outbreaks and clusters. There are a few considerations to take into account:

- The PHO would need to assess case classification when cases with shared exposure sites are identified. Appropriate training and attention to detail would address this.

- The process would also require a small additional investment of time from the PHO, as they would need to add the case residential address to the list of exposure sites for the case. This process takes between one and two minutes to complete.

- The outbreak detection workflow in Geocortex requires onset dates to identify outbreaks and clusters. Onset dates are often not known for suspect cases as these cases are usually notified by laboratories. The workflow could overcome this by using a calculated onset date (based on the notification date) as a proxy for onset date for suspect cases.

- The legionellosis follow-up protocol should clarify how outbreaks involving suspect cases should be addressed. A clear position should be adopted regarding whether environmental investigation is conducted as soon as the suspected outbreak is identified, or after the diagnosis is first confirmed by urinary antigen or sputum analysis. The protocol should also specify whether a

response should be initiated if a case can't be confirmed or rejected within a pre-determined time frame.

5.2.10 Confirm cases based on primary laboratory urinary antigen results

To improve timeliness cases should be confirmed and public health action implemented as soon as possible based on primary laboratory urinary antigen results, rather than waiting for confirmation at VIDRL. A phone notification from a primary laboratory is a form of notification and should be sufficient to confirm a case. This should be clarified in the protocol to minimise confusion. At the time of writing a new urinary antigen test (SOFIA) had recently been adopted by two primary laboratories in Victoria. The test has been associated with at least three false positive results in the few months it has been used. The test is being further evaluated and may be exempted from this recommendation in future.

5.2.11 Improve timeliness of serology case follow-up

Develop guidelines for timely case classification for serology cases eg:

i. Contact the testing doctor to request clinical details within 3 days of receiving blood results. Strongly recommend a urinary antigen test and collection of specimen for PCR

ii. If no response after seven days, contact by phone and letter

iii. If no response after 3 weeks, contact the doctor again

iv. Aim to classify or reject all suspect cases within 6 weeks of notification

v. Classify probable cases as soon as possible after receipt of notification, do not await a second serology result if the first is sufficient to consider the case "probable"

5.2.12 Provide an initial IRR as soon as possible after receipt of notification

If the routine investigation of cooling towers near the home and workplace for sporadic case continues, the Disease Surveillance System should provide the Environmental Investigation System with key information as soon as possible rather than waiting for complete exposure histories. Eg the initial IRR could include residential address, and work address if available. Provide updates when they become available.

5.2.13 Include exposure sites for cases classified as "Not Notifiable" in Geocortex exposure tables

This will enable inclusion of exposure sites for non-Victorian residents in the cluster and outbreak detection workflow.

5.3 Recommendations for Environmental Investigation System

Major Recommendations for the Environmental Investigation System

5.3.1 Store case related data in a searchable database

The potential to use PHESS for this purpose has already been discussed. PHESS can provide a solution to assist with the management of non-cooling tower data as well as cooling tower data. Some effort will be required to ensure cooling tower data in PHESS is appropriately linked to cooling tower records in Emerald, where they are required for registration and licensing purposes.

5.3.2 Develop data-driven objectives

The environmental investigation system would benefit from the development of data-driven objective/s. An example could be *:*

> *"Generate useful data. Document sample results for environmental Legionellae exposures for cases of legionellosis, and integrate this data with disease surveillance information to inform programming and policy*".

Such an objective would ensure generation of quality data to drive research, programming and policy is prioritised and resourced; and continues to be prioritised during future decision-making. A data-driven objective should prompt training/appointment of a team member with data management skills required to oversee the collection and utilisation of high quality data.

5.3.3 Capture reason for sampling

Capture nature of case exposure for case-related cooling tower inspections to enable analysis of risk factors for *Legionella* detection. This may help inform program planning in future. Eg whether investigation is sporadic case or outbreak related; proximity to exposure site (metres), nature of exposure site eg. near home, at workplace, other.

5.3.4 Review the protocol for sporadic case investigations

The finding that there is no increase in risk of *Legionella* detection in systems currently targeted for inspection and disinfection suggests these sites pose no greater risk to the public than other sites. Targeting sites within 500m of the home or at the workplace seems difficult to justify for all cases, although may be reasonable for cases with limited other exposures. Investigative effort may be better directed toward clusters rather than sporadic cases. Consideration should be given to assessing other risk factors for *Legionella* and target inspections based on these. For example, sites with an outbreak or cluster in the near vicinity and a history of *Legionella* detection or non-compliance should be prioritised. Spatio-temporal clustering of cases that falls outside the rigid protocol definitions could also be identified through the mapping system and help direct inspection efforts to a particular area. These factors are discussed further under recommendations for Geocortex.

5.3.5 Specifically document outbreak investigations

Each cluster/ outbreak should be specifically recorded along with all risk assessment and sampling information for the incident. Such records would provide useful data for later analysis, as well as good documentation for legal review in relation to outbreaks. The single record would provide a mechanism for keeping track of which sites have been sampled, which have been disinfected and when these interventions were performed. As previously discussed, using the PHESS database to store case investigation data would address this recommendation.

5.3.6 Add a field for PHESS case ID to the form to notify CDPC of a Legionella detection

This will assist the Disease Surveillance System to keep complete records of Legionella detections in the PHESS database. Once the teams share a single database, this form may no longer be required.

Minor Recommendations for the Environmental Investigation System

5.3.7 Develop a mechanism to keep track of cooling tower inspections during large or complex investigations, when routine data entry is often delayed

The Geocortex mapping system could be leveraged for this purpose. When an outbreak is identified the CTS sites targeted for inspection could be retrieved from the Geocortex interface using the “buffer identify” function to retrieve all systems within a selected radius of the outbreak. Data can easily be exported into an excel spread-sheet or the cooling towers could be linked to the event in PHESS. As sites are inspected and samples collected the investigation coordinator could record inspection date directly into the spread-sheet (or into PHESS), along with the anticipated date for culture results. If a spread-sheet is used it could later be attached to the outbreak record in PHESS or filed in TRIM alongside an outbreak summary sheet.

5.3.8 Introduce regular monitoring of data quality to improve data completeness.

Consider a quarterly or bi-annual audit of 1-2 randomly selected case investigations and associated Case Summary Sheets/Emerald/TRIM records (or PHESS records, in future) to assess record completeness. Results should be made available to all LT members to raise awareness of the importance of complete record keeping.

5.4 Recommendations for Geocortex

Major recommendations for Geocortex

5.4.1 Increase the sophistication of the approach to selecting cooling towers for inspection

Geocortex is currently utilised to identify cooling towers for case-related investigations based on a simple spatial relationship - that is, the distance between cooling towers and selected case exposure sites. However, Geocortex has potential to introduce more complex and useful analysis to aid in cooling tower risk assessment and sample site selection. These include consideration of *Legionella* detection history for cooling tower systems, and the development of spatial clustering algorithms able to detect statistically significant clustering of disease to identify areas for targeted inspections.

History of *Legionella* detection: Currently, the Geocortex mapping system includes a field showing the most recent *Legionella* test result for each cooling tower. However

this information is currently of limited utility because a) it only shows results for sites inspected by the Legionella team; and b) it only displays the most recent test result. When *Legionella* is detected, disinfection is performed and a repeat sample collected every seven-ten days until the sample result is negative. Therefore if *Legionella* is detected the field only reflects this for a short period.

These problems could be overcome by a small adjustment to the current systems. Victorian cooling tower operators are required to test for *Legionella* at least every three months under the Victorian Public Health and Wellbeing Regulations 2009 (30) but are not required to notify the Department unless *Legionella* is detected in three consecutive tests. The annual auditors could be leveraged to capture whether *Legionella* had been detected during routine CTS monitoring each year. This would involve the addition of a simple field to the current audit report (*Legionella* detected this year: yes/no). The cooling tower database would require a small modification to capture this information, which could then be transmitted to Geocortex. Geocortex could include such fields as "*Legionella* ever detected: yes/no" and "Date of most recent *Legionella* detection". This would provide a means to prioritise cooling towers with a history of *Legionella* detection during case or cluster investigations.

Spatial clustering algorithms: Selection of cooling towers for disinfection and sampling may be further aided by the use of a spatial clustering analysis. A recent New Zealand study used a sophisticated spatial analysis to identify the likely geographic source of a large urban legionellosis cluster of 19 cases (53). A similar system could be developed as an automated workflow in Geocortex. For example a workflow could utilise a local Moran's *I* spatial autocorrelation test statistic (53, 54) or other spatial analysis to examine clustering of cases routinely. The local Moran's generates a test statistic (P value) measuring the likelihood that clustering of features (eg. cases) has occurred through chance alone. Cooling towers in areas associated with spatial clustering could then be targeted as part of routine regulatory activities. Note that any spatial analysis must take into account the fact that a single case can have multiple exposure sites by only identifying clustering of 'unique identifiers'. The system must also include an adjustment for "foot traffic" density.

5.4.2 Adapt the automated workflow to search within the 2 year data table

To ensure all relevant cases are considered during the search for shared exposure sites, the workflow should be reworked to search within the 2 year data table rather

than the 90 day data table, and adjusted to provide the capacity to search within a chosen *onset date range*, ie. cases with an onset date +/- 90 days from the *onset date* (not event date) from the case of interest.

5.4.3 Provide applied training in the use of the mapping system to staff in CDPC, CDES and LT

This training should include senior Public Health Officers and the Legionella Epidemiologist. Consider developing a structured training workshop in collaboration with key stakeholders, specific to the needs of both PHOs and Environmental Health Officers.

For the Disease Surveillance System stakeholders, training should include information on how data is transported to the maps, and requirements for accurate mapping.

Minor Recommendations for Geocortex

5.4.4 Highlight each cluster and outbreak in the workflow output

To minimise the risk of error, tailor the workflow output to provide some form of alert where outbreaks or clusters are identified. For cases involved in more than one outbreak/cluster, aim to highlight each event separately.

References

1. Centres for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations form the guidelines working group. MMWR: Morbidity & Mortality Weekly Report. 2001;50(RR-13).

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

3. Vaisrub S. My Name Is *Legionella*. JAMA. 1980;243(17):1747.

4. Weisse AB. A plague in Philadelphia: the story of Legionnaires' disease. Hosp Prac (Off Ed). 1992(6):151.

5. 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 Eng J Med. 1977;297(22):1197-203.

6. Glick TH, Gregg MB, Berman B, Mallison G, Rhodes WW, Jr., Kassanoff I. Pontiac fever. An epidemic of unknown etiology in a health department: I. Clinical and epidemiologic aspects. Am J Epidemiol. 1978;107(2):149-60.

7. Bartram J, Chartier Y, Lee J, Pond K, Surman-Lee S. *Legionella* and the Prevention of Legionellosis. Geneva: World Health Organisation, 2007.

8. Broadbent. National Enivoronmental Health Forum Monographs Water Series No. 1: Guidance for the Control of *Legionella*. Glenelg: National Environmental Health Forum; 1996.

9. Morris GK, Patton CM, Feeley JC, Johnson SE, Gorman G, Martin WT, et al. Isolation of the Legionnaires' disease bacterium from environmental samples. Ann Int Med. 1979;90(4):664-6.

10. Diederen BMW. *Legionella* spp. and Legionnaires' disease. J Infect. 2008;56(1):1-12.

11. Carson P, Mumford C. Legionnaires' disease: causation, prevention and control. Loss Prevention Bulletin. 2010(216):20-9.

12. O'Neill E, Humphreys H. Surveillance of hospital water and primary prevention of nosocomial legionellosis: what is the evidence? J Hosp Infect. 2005;59(4):273-9.

13. Den Boer JW, Euser SM, Brandsema P, Reijnen L, Bruin JP. Results from the National *Legionella* Outbreak Detection Program, the Netherlands, 2002-2012. Emerging infectious diseases. 2015;21(7):1167-73.

14. Potts A, Donaghy M, Marley M, Othieno R, Stevenson J, Hyland J, et al. Cluster of Legionnaires disease cases caused by *Legionella longbeachae* serogroup 1, Scotland, August to September 2013. Euro Surveillance: Bulletin Européen Sur Les Maladies Transmissibles. 2013;18(50):20656-.

15. Mykietiuk A, Carratalà J, Fernández-Sabé N, Dorca J, Verdaguer R, Manresa F, et al. Clinical outcomes for hospitalized patients with *Legionella* pneumonia in the antigenuria era: the influence of levofloxacin therapy. Clin Infect Dis. 2005;40(6):794-9.

16. Plouffe JF, Breiman RF, Fields BS, Herbert M, Inverso J, Knirsch C, et al. Azithromycin in the treatment of *Legionella* pneumonia requiring hospitalization. Clin Infect Dis. 2003;37(11):1475-80.

17. McNally C, Hackman B, Fields BS, Plouffe JF. Potential importance of *Legionella* species as etiologies in community acquired pneumonia (CAP). Diagnostic Microbiology and Infectious Disease. 2000;39:79-82.

18. Lieberman D, Porath A, Schlaeffer F, Lieberman D, Boldur I. *Legionella* species community-acquired pneumonia: a review of 56 hospitalized adult patients. CHEST. 1996;109(5):1243-9.

19. Arancibia F, Cortes CP, Valdés M, Cerda J, Hernández A, Soto L, et al. Importance of *Legionella pneumophila* in the etiology of severe community-acquired pneumonia in Santiago, Chile. Chest. 2014;145(2):290-6.

20. Decker BK, Palmore TN. The role of water in healthcare-associated infections. Curr Opin Infect Dis. 2013;26(4):345-51.

21. Roig J, Aguilar X, Ruiz J, Domingo C, Mesalles E, Manterola J, et al. Comparative study of *Legionella pneumophila* and other nosocomial-acquired pneumonias. Chest. 1991;99(2):344-50.

22. Bryner B, Miskulin J, Smith C, Cooley E, Grams R, Bartlett R, et al. Extracorporeal life support for acute respiratory distress syndrome due to severe *Legionella* pneumonia. Perfusion. 2014;29(1):39-43.

23. Engel MF, van Manen L, Hoepelman AIM, Thijsen S, Oosterheert JJ. Diagnostic, therapeutic and economic consequences of a positive urinary antigen test for *Legionella* spp. in patients admitted with community-acquired pneumonia: a 7-year retrospective evaluation. J Clin Path. 2013;66(9):797-802.

24. 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.

25. National Notifiable Disease Surveillance System. Notification Rate of Legionellosis [On-line data]. http://www.health.gov.au: Australian Government Department of Health and Aging; 2014 [cited 2014 April 13].

26. Beaute J, Zucs P, de Jong P. Legionnaires' disease in Europe, 2009-2010. Euro Surveill. 2013;18(10).

27. Australian Bureau of Statistics. 3101.0 Australian Demographic Statistics TABLE 4. Estimated Resident Population, States and Territories (Number) 2015 [cited 2015 11 November]. Available from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3101.0Mar%202015?OpenDocument.

28. Greig JE, et al. An outbreak of Legionnaires disease at the Melbourne Aquarium, April 2000: investigation and case-control studies: Sydney : Australasian Medical Pub. Co.; 2004.

29. Loff B, Cordner S. Legionnaires' outbreak in Australia highlights government's testing gaps. Lancet. 2000;355(9216).

30. Public Health and Wellbeing Regulations 2009, Victoria [Statute on the internet], (2009).

31. Communicable Disease Prevention and Control. Legionellosis (Legionnaires' Disease) Group A condition. Case Investigation Protocol (Internal Document). 2013.

32. Communicable Diseases Network Australia. Legionellosis. National Guidelines for public health units. In: Department of Health and Ageing, editor. online: Australian Government; 2009.

33. Public Health and Wellbeing Act 2008, Victoria, (2008).

34. Building (*Legionella* Act), Victoria, (2000).

35. Legionella Team. Legionnaires' disease environmental investigations. Standard Operating Procedure (internal document). 2011.

36. Public Health Laboratory Network. *Legionella* Laboratory Case Definition http://www.health.gov.au/internet/main/publishing.nsf/Content/cda-phlncd-*Legionella*.htm: Commonwealth of Australia; 2007 [cited 2015 31 March]. Available from: http://www.health.gov.au/internet/main/publishing.nsf/Content/cda-phlncd-*Legionella*.htm.

37. Murdoch DR. Diagnosis of *Legionella* infection. Clin Infect Dis. 2003;36(1):64-9.

38. 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. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2010;16(9):1405-7.

39. Communicable Disease Epidemiology and Surveillance. PHESS Quick Entry Guide: *Legionella* (internal document). In: Services VGDoHaH, editor. 2015.

40. Porta. Dictionary of Epidemiology. 5th ed. Oxford: Oxford University Press; 2008.

41. Arnow PM, Weil D, Para MF. Prevalence and significance of *Legionella pneumophila* contamination of residential hot-tap water systems. The Journal of infectious diseases. 1985;152(1):145-51.

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

43. Mentasti M, Fry NK, Afshar B, Palepou-Foxley C, Naik FC, Harrison TG. Application of *Legionella pneumophila*-specific quantitative real-time PCR combined

with direct amplification and sequence-based typing in the diagnosis and epidemiological investigation of Legionnaires' disease. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2012;31(8):2017-28.

44. Graham RM, Doyle CJ, Jennison AV. Real-time investigation of a *Legionella pneumophila* outbreak using whole genome sequencing. Epidemiology and infection. 2014;142(11):2347-51.

45. StataCorp. Stata Statistical Software: Release 13. 2013.

46. Formica N, Yates M, Beers M, Carnie J, Hogg G, Ryan N, et al. The impact of diagnosis by *Legionella* urinary antigen test on the epidemiology and outcomes of Legionnaires' disease. Epidemiology and infection. 2001;127(2):275-80.

47. Bhopal RS, Fallon RJ, Buist EC, Black RJ, Urquhart JD. Proximity of the home to a cooling tower and risk of non-outbreak Legionnaires' disease. BMJ (Clinical research ed). 1991;302(6773):378-83.

48. Nguyen TM, Ilef D, Jarraud S, Rouil L, Campese C, Che D, et al. A community-wide outbreak of legionnaires disease linked to industrial cooling towers--how far can contaminated aerosols spread? The Journal of infectious diseases. 2006;193(1):102-11.

49. Graham FF, White PS, Harte DJ, Kingham SP. Changing epidemiological trends of legionellosis in New Zealand, 1979-2009. Epidemiology and infection. 2012;140(8):1481-96.

50. Helbig JH, Uldum SA, Luck PC, Harrison TG. Detection of *Legionella pneumophila* antigen in urine samples by the BinaxNOW immunochromatographic assay and comparison with both Binax *Legionella* Urinary Enzyme Immunoassay (EIA) and Biotest *Legionella* Urin Antigen EIA. J Med Microbiol. 2001;50(6):509-16.

51. Olsen CW, Elverdal P, Jorgensen CS, Uldum SA. Comparison of the sensitivity of the *Legionella* urinary antigen EIA kits from Binax and Biotest with urine from patients with infections caused by less common serogroups and subgroups of *Legionella*. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2009;28(7):817-20.

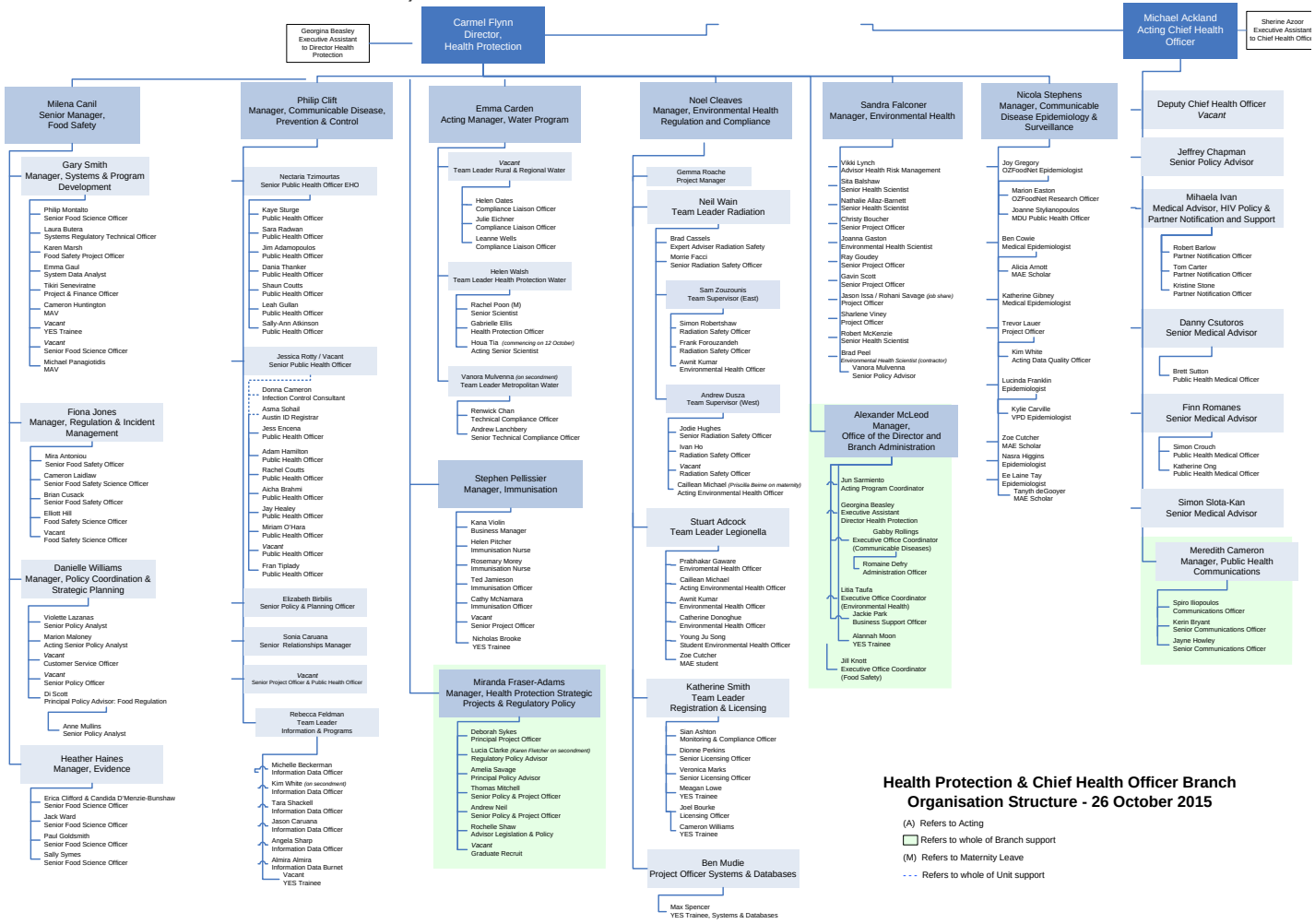
52. Benson RF, Tang PW, Fields BS. Evaluation of the Binax and Biotest urinary antigen kits for detection of Legionnaires' disease due to multiple serogroups and species of *Legionella*. Journal of clinical microbiology. 2000;38(7):2763-5.

53. White PS, Graham FF, Harte DJ, Baker MG, Ambrose CD, Humphrey AR. Epidemiological investigation of a Legionnaires' disease outbreak in Christchurch, New Zealand: the value of spatial methods for practical public health. Epidemiology and infection. 2013;141(4):789-99.

54. Fotheringham AS, et al. (2000). Quantitative Geography: Perspectives on Spatial Data Analysis. London, Sage.

Chapter 2 Appendices

Appendix 1. Structure of Health Protection Branch, 2015



Appendix 2. Example of an Incident Response Report (IRR)

Risk classification[1]	Public health officer	Incident update	PHESS reference
Confirmed	***Jay Healy Tel. 1300 651160***	***Update**	***320142219050***
Incident details			
Incident date/time	*17-Jun-2014 (reported to DH)*		
Reported by	*VIDRL (on: 17 June 2014)*		
Initial classification			
Event classification	*Confirmed*		
Organism identified	*Legionella pneumophila 1*		
Case details			
Demographics	*Male Birth date.* [redacted] *ATSI. Not Aboriginal or Torres Strait Islander*		
Residence	[redacted]		
Work	*Employed* [redacted] [redacted] — [redacted] *Port Melbourne VIC 3207* *NORTHERN AND WESTERN METROPOLITAN — Melbourne (C)* *Melway.* [redacted] *Private*		
Onset of illness	[redacted]*-2014*		
Incubation period	[redacted]*-2014*		
Diagnosis method	*Urinary Antigen*		
Clinical presentation	Fevers and lethargy from [redacted]. Admitted to Austin Hospital on [redacted] with ongoing fever, dry cough, shortness of breath, headache and extreme lethargy		
Predisposing conditions	*Chronic disease=No* *Immunocompromised=No* *Smoker=Current smoker* *Drinks alcohol=2-3 times per week*		
Linked cases	3003 – None within 500 metres 3012 – 320142214645 – [redacted]. Site 600 metres from [redacted] where current case attended on [redacted] 3018 – 320142214645 – attended [redacted] from [redacted]. Site is 1 km away from [redacted] where current case attended on [redacted] 3020 – None within 500 metres 3043 – Cluster identified – PHESS 320142210192. Attended [redacted] on [redacted] – within 500 metres of [redacted] which case attended sometime during incubation period [redacted] 3061 – None within 500 metres		

[1] Further information on incident management is available in "Health Protection Branch Incident reporting instruction — 2012" and "Communicable Disease Incidents—Emergency/Incident Report Distribution Checklist."

	3171 - None 3207 – Possible Outbreak – PHESS **320142218242** who worked at [redacted] on [redacted], [redacted]. Current case is based at [redacted] 3218 – None

Actions and comments

Case is employed as a [redacted] based at [redacted], [redacted]. Works Monday – Saturdays and [redacted] multiple sites. PHO has requested work run sheet from employer.

Cluster identified with PHESS 320142210192 – LP1 – Pascoe Vale who attended [redacted] on [redacted]

Possible Outbreak identified with PHESS **320142218242** who worked at [redacted] on [redacted], [redacted] and [redacted]. Current case is based at [redacted]. Further information regarding the first case and their exact exposures sites is being sought.

Exposure sites

Site	Street1	Street2	City	Postcode	Confidence	Exposure type	From	To	Notes
* exposure sites (n=20) have been removed to protect privacy *									

Appendix 3. Example of a RIEMS notification

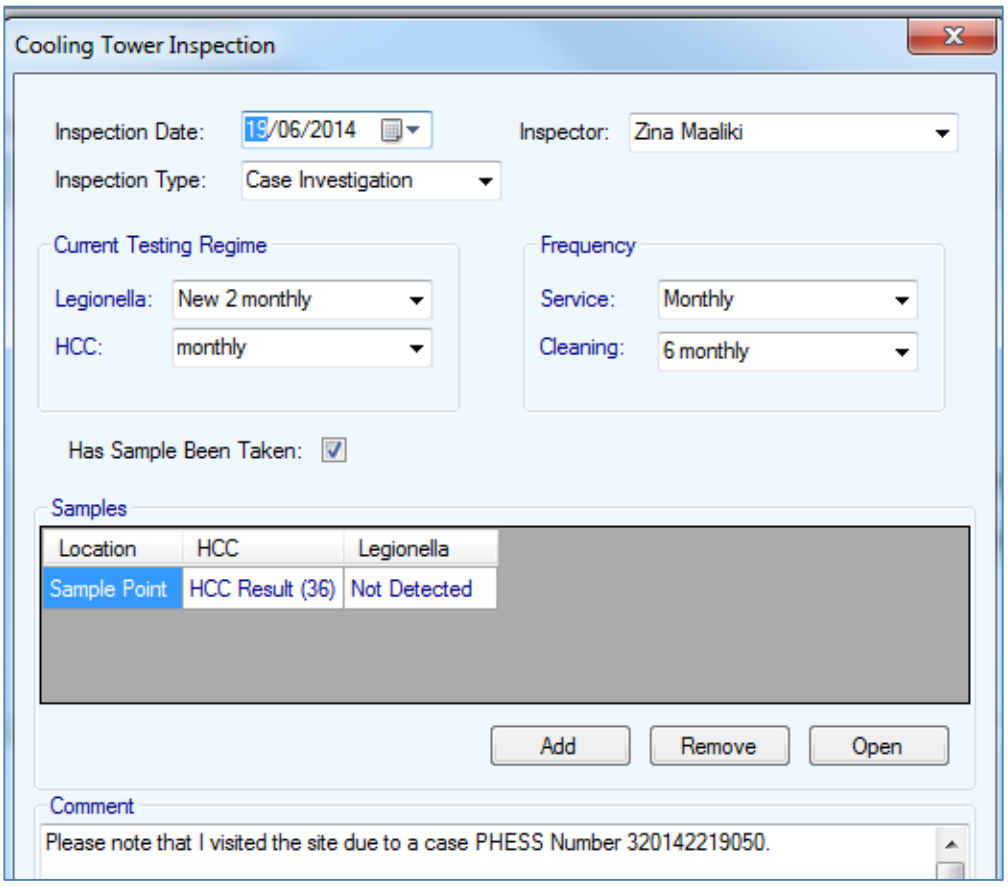
RIEMS.NET incident 54924
You are listed as having a special interest in the following incident.
You are being notified because this incident has been logged.

Ref No: 54924
Category: Disease\Legionella/Legionnaires disease - Other
Description: Legionella/Legionnaires disease(Other)- PHESS 320152302891. Onset: 2 June 2015.
Residence: 65 Legionella Street, Smithsville VIC 3333.
Reported By: Frances Tiplady
Reported At: 03/07/15 12:00 AM
Location: 65 Legionella Street,
SMITHSVILLE, VIC 3333 (SOUTHERN)
Action Officer: Stuart Adcock
Action Taken: There are no cooling towers within 500 m of residential address. Will wait for more information.

Appendix 4. Example Case Summary Sheet

Legionella Team Action				
Incident Name :	**RIEMS No.**	**Leading Officer**	**Update No.**	
Are there any cooling towers within 500m of the home?		**Y☐ N☐**		
Cooling Towers that have been sampled and disinfected				
SID-CTS/Address	**Date Sampled**	**Date Disinfected**	**Results**	**Date Resolved/Comments**
Are there any cooling towers at the workplace		**Y☐ N☐**		
Cooling Towers that have been sampled and disinfected				
SID-CTS/Address	**Date Sampled**	**Date Disinfected**	**Results**	**Date Resolved/Comments**
Have any other sites been inspected in response to case?		**Y☐ N☐**		
SID-CTS/Address	**Date Sampled**	**Date Disinfected**	**Results**	**Date Resolved/Comments**
Any further samples taken?		**Y☐ N☐**		
Details:				
Follow up action required		**Y☐ N☐**		
Details:				
Has this case been checked against possible linked cases		**Y☐ N☐**		
DHHS Contacts		**Stuart Adcock X65028**		

Appendix 5. Example of an Emerald cooling tower inspection record



Appendix 6. Stakeholders interviewed for the evaluation

Position	Stakeholder
Legionella Team	
Team Leader	Stuart Adcock
Environmental Health Officer	Awnit Kumar
Environmental Health Officer	Caillean Michael
Environmental Health Officer	Zena Maalaki
Environmental Health Officer	Catherine Donoghue
Environmental Health Officer	Prabhaker Gaware
Environmental Health Officer	Lachland Chapman
Manager, Environmental Health Regulation and Compliance	Noel Cleaves
Database Manager	Ben Mudie
CDCP	
Public Health Officer	Jay Healy
Public Health Officer	Jess Encena
Public Health Officer	Shaun Coutts
Public Health Officer	Frances Tiplady
Senior Public Health Officer	Nectaria Tzimourtis
Manager, CDPC	Philip Clift
OCHO	
Medical Advisor	Brett Sutton
CDES	
Legionella Epidemiologist	Lucinda Franklin
Database Manager	Trevor Lauler
GIS and Planning	
Database Administer	Clare Brazenor

Appendix 7. Cases reviewed for the case audit

PHESS ID	Event date	Organism	Classification	Cluster/ outbreak	Environ mental investiga tion	Number of cooling towers inspected	Other sites inspected	Legionella detected?
320142219050	17/06/2014	L. *pneumophila* 1	confirmed	outbreak and cluster	yes	11	spa	yes
320142220384	23/06/2014	L. *pneumophila* 1	confirmed	no	yes	1	no	no
320142220725	23/06/2014	L. *pneumophila* 1	confirmed	cluster	yes	0	car wash	no
320142222195	2/07/2014	L. *pneumophila* 1	confirmed	cluster	yes	2	no	no
320142255195	10/12/2014	L. *pneumophila* 1	confirmed	no	yes	11	no	no
320152266412	6/02/2015	Legionella	confirmed	no	yes	0	humidifier	yes
320142221037	24/06/2014	L. *pneumophila* 1	confirmed	no	no	0	no	n/a
320142221785	26/06/2014	L. *pneumophila* 1	confirmed	no	no	0	no	n/a
320142224372	7/07/2014	L. *pneumophila* 1	confirmed	no	no	0	no	n/a
320142247570	6/11/2014	L. *pneumophila*	probable	no	no	0	no	n/a
320142248538	11/11/2014	L. *longbeachae*	probable	no	no	0	no	n/a
320142257251	19/12/2014	Legionella	confirmed	no	no	0	no	n/a
320152272777	19/02/2015	L. *pneumophila* 1	confirmed	no	no	0	no	n/a
320152272992	20/02/2015	L. *longbeachae*	confirmed	no	no	0	no	n/a
320152274977	3/03/2015	L. *pneumophila* 1	confirmed	no	no	0	no	n/a

This page is intentionally blank

Chapter 3. Genomics links sporadic *Salmonella* Typhimurium Phage type 9 infections to a recurrent source of outbreak salmonellosis

Contents

Preface

Background to project

Large scale whole genome sequencing (WGS) is becoming increasingly feasible and may soon be adopted for routine surveillance purposes in Australia. The unprecedented discriminatory power provided by WGS will revolutionise investigation of disease sources and transmission routes. However there is much to be learnt regarding the best way to record, use and interpret the massive amounts of data generated by WGS.

In Victoria a single egg producer has been associated with recurrent Salmonella outbreaks dating back to 2004. Because sporadic cases are not followed up, the number of sporadic cases associated with the farm has remained unknown. This situation provided an ideal forum to explore the potential of WGS data, and gain an understanding of its strengths and limitations.

This epidemiological project is the first application of large-scale WGS data in Victoria to investigate relationships amongst Salmonella isolates for surveillance purposes.

My role

I was the lead investigator in this project, although I was privileged to receive regular consultation and advice from a project team with Nicola Stephens and Marion Easton (DHHS). I also received invaluable support from Dieter Bulach (MDU) who provided WGS data in a usable format, tutored me in WGS theory, and guided me through the use of various software programs for tree visualisation.

My role included:

- Scoped the project and developed the research question and analysis plan, in consultation with Nicola Stephens and Marion Easton
- Collated and analysed MLVA data provided by Karolina Dimovski and Mary Volcanis at the Microbial Diagnostic Unit (MDU).
- Selected isolates for WGS with assistance from Marion Easton
- Liaised with the Doherty Centre for Applied Microbial Genomics to arrange for sequencing of selected isolates, and monitored progress
- Analysed WGS data provided by Glenn Carter, Dieter Bulach, Timothy Stinear, and Benjamin Howden at the Doherty Centre for Applied Microbial Genomics

- Developed tables and figures including the phylogenetic tree, in consultation with Nicola Stephens and Marion Easton
- Lead author on an article for publication, with review of content by co-authors Karolina Dimovski, Mary Volcanis, Dieter Bulach, Glen Carter, Timothy P Stinear, Nicola Stephens, Martyn D. Kirk, Benjamin P Howden and Marion Easton

Lessons learnt

When I commenced this project I knew little about WGS or MLVA. Through collaborating with MDU I learnt some basic technical aspects of public health genomics. To assist readers, a brief outline of technical aspects of *Salmonella* typing and WGS is included in the appendix.

I learnt there is a strong need for epidemiologists to engage with this new technology to shape its adoption and to ensure it is used and presented in a relevant way.

Once again, I learnt the value of collaboration between professionals with different skills in public health, including (in this case), microbiologists, bioinformaticians, veterinarians, and epidemiologists.

I learnt that new technologies can be overwhelming and human resource skills take time to catch up.

And, not surprisingly, I learnt that the burden of disease due to recurrent outbreak sources is likely to be substantially greater than cases identified in outbreaks.

Public Health Implications

The project led to valuable collaboration and exchange of skills and knowledge between DHHS and MDU. The skills and knowledge gained will help shape the adoption of WGS for surveillance purposes in Victoria, including a deeper understanding of the likely genetic diversity seen in point source outbreaks compared to recurrent outbreak sources.

In addition, the findings provide support to the hypothesis that, when traced to a recurrent source, outbreaks likely form ‘the tip of the iceberg’. This highlights the need for comprehensive trace-back and public health action during outbreaks.

Project outputs

Project outputs include the following:

- Inclusion as a chapter in the Bound Volume submission for the Master of Philosophy in Applied Epidemiology
- Publication in a peer-reviewed journal. The late draft for peer reviewed publication makes up the body of this chapter.

Acknowledgements

I gratefully acknowledge the following parties for their role in this project:

- Nicola Stephens and Marion Easton from DHHS for their infectious enthusiasm and regular feedback along the way
- Ben Howden, Director of MDU and Tim Stinear, Scientific Director of the Doherty Centre for Applied Microbial Genomics, for their invaluable technical advice and support during the project
- Dieter Bulach, bioinformatician from MDU for patiently sharing his knowledge and gently introducing me to microbial genomics
- Anders deSilva for his generosity in developing the R script and tailoring it to my needs
- Glen Carter, Doherty Centre for Applied Microbial Genomics for co-ordinating the sequencing of historical isolates
- Mary Valcanis and Karolina Mercoulia for provision of data and advice regarding MLVA patterns
- Martyn Kirk for his excellent strategic advice
- Joy Gregory for reviewing the paper draft

Genomics links sporadic *Salmonella* Typhimurium Phage type 9 infections to a recurrent source of outbreak salmonellosis

Late draft of an article for peer reviewed publication

Article Summary: Genomics links a substantial number of “sporadic” salmonellosis cases to a recurrent outbreak source

Running title: Genomics links sporadic salmonellosis to outbreaks

Keywords: Salmonella, Salmonella Infections, genome, Bacterial Typing Techniques, Public Health, epidemiology, Public Health Surveillance***,*** Foodborne Diseases, Communicable Disease Control

Title: Genomics links sporadic *Salmonella* Typhimurium Phage type 9 infections to a recurrent source of outbreak salmonellosis

Authors: Zoe Cutcher, Dieter Bulach, Glen Carter, Torsten Seemann, Mary Valcanis, Karolina Mercoulia, Timothy P Stinear, Nicola Stephens, Martyn D. Kirk, Benjamin P Howden, and Marion Easton

Author Affiliations: Victorian Department of Health and Human Services, Melbourne, Victoria, Australia (Z.Cutcher, N.Stephens, M.Easton), The University of Melbourne, Parkville, Victoria, Australia (D Bulach, G Carter, T Seemann, M Valcanis, K Mercoulia, T Stinear, B Howden); Australian National University, Canberra, Australian Capital Territory, Australia (Z. Cutcher, MD. Kirk), and Victorian Life Sciences Computation Initiative (D. Bulach, T Seemann)

Abstract

The number of sporadic salmonellosis cases attributable to outbreak sources is often unknown. We used multilocus variable-number tandem-repeat analysis (MLVA) and whole genome sequencing to estimate the number of sporadic isolates related to an outbreak source. We categorised S. Typhimurium Phage type 9 isolates from 2009-2015 according to MLVA then sequenced 99 isolates from a single outbreak source, 161 sporadic isolates, and 41 from other outbreaks. We tabulated SNPs between sporadic isolates and those from the outbreak source according to MLVA. We then estimated the total number of sporadic cases potentially attributable to the outbreak source for each MLVA level. We estimate 72 of 1585 sporadic isolates between July 2009-June 2014 were within 5 SNPs, and almost 400 were within 20 SNPs of the predominant outbreak clade. Sporadic salmonellosis potentially attributable to the outbreak source eclipsed cases identified in outbreak investigations, highlighting a pressing need for trace back and intervention.

Background

Salmonellosis is a global public health problem often associated with consumption of contaminated food (1). Identifying sources of infection provides opportunity for public health intervention. *Salmonella* Typhimurium is the most common serotype in Australia, accounting for almost half of salmonellosis cases notified during 2011 (2). The majority of Australian jurisdictions use phage typing or multilocus variable-number tandem-repeat analysis (MLVA) (3, 4) or both to further type *S.* typhimurium isolates. MLVA uses the observed variation in the number of short tandem repeats at five loci; identical or highly similar patterns of repeats between isolates usually correlate with a close genetic relationship (5-12).

In Victoria, one of eight Australian states and territories, public health investigation occurs when salmonellosis notification identifies a potential source, an epidemiological link between cases, or spatiotemporal clustering. Other cases are considered 'sporadic' and trigger little or no investigation due to resource limitations and challenges in identifying a shared source using current typing (13). Because the source of infection remains unknown for the majority of cases ongoing sources of contamination may remain undetected, representing a missed opportunity for public health intervention.

Improved isolate discrimination can be achieved using whole genome sequencing (WGS) comparison of isolates (14-16). WGS has better discriminatory ability than Victoria's current typing standard, MLVA (16) and is sufficiently discriminatory to trigger investigation of genetically similar cases without an identified epidemiological link (17).

This discriminatory power means WGS can detect clusters not identified by traditional subtyping methods and is useful for attributing 'sporadic' isolates to recognised outbreaks (18, 19).

Large scale *Salmonella* WGS as a public health surveillance tool is increasingly cost effective and technically feasible (20-22); all *Salmonella* clinical isolates in the UK (23, 24) and a majority of food isolates across the USA (25) are now routinely sequenced. Australian laboratories have been exploring WGS to better define the nature of Salmonella outbreaks.

In February 2014 eight point source outbreaks of *S.* Typhimurium Phage type 9 at different food establishments, accounting for 303 confirmed and probable cases, were epidemiologically linked to a Victorian egg producer, Farm A. Farm A had been epidemiologically linked to five other *S.* Typhimurium Phage type 9 outbreaks since 2004 (one in 2004; two in 2005, and two in 2010) and an additional outbreak occurred in February

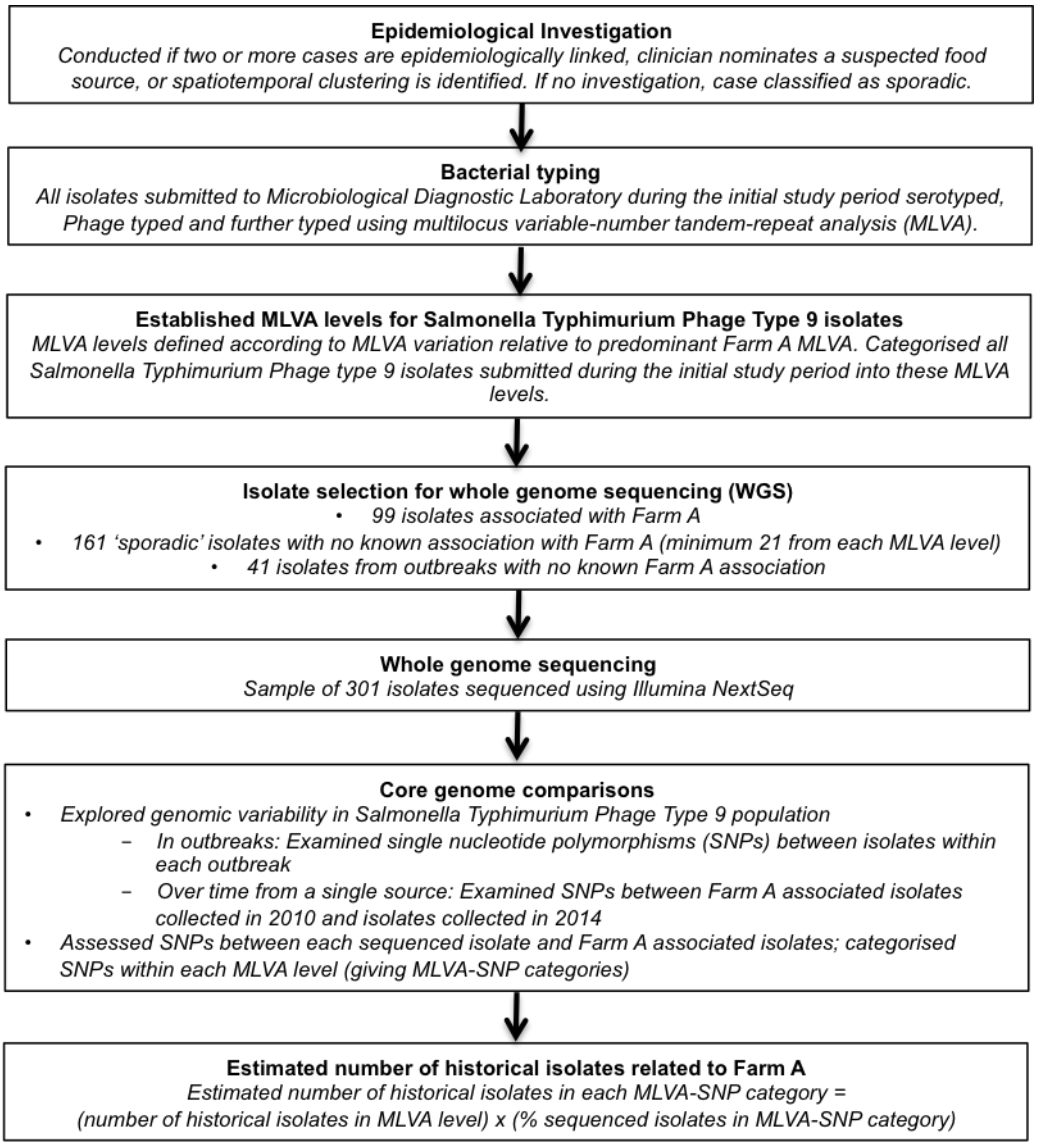
2015. Ongoing contamination of the farm, and therefore ongoing or intermittent contamination of eggs, is suggested by the persistence of a predominant MLVA pattern in outbreak isolates during 2004-2014. *S.* Typhimurium Phage type 9 is one of the top three notified Phage types in Victoria, but sources are usually only identified for outbreak associated cases. The number of sporadic cases associated with Farm A remains unknown.

We examined MLVA and WGS data for a sample of *S.* Typhimurium Phage type 9 isolates and estimated the number of clinical isolates not identified to be associated with an outbreak that were genetically related to Farm A over a five year period.

Methods

Briefly, we examined MLVA patterns for all *S.* Typhimurium Phage type 9 isolates submitted to Victoria's *Salmonella* reference laboratory during July 2009 – June 2015 (the initial study period). We identified the predominant MLVA pattern in Farm A associated isolates and categorised other isolates into MLVA levels according to variation from this predominant pattern. Next, we sequenced clinical isolates not known to be associated with outbreaks (sporadic isolates) clinical isolates from outbreaks associated with Farm A (Farm A isolates); and clinical isolates from cases linked to outbreaks with no known association with Farm A. We assessed WGS relatedness between Farm A associated and sporadic isolates, then examined the proportion of sequenced isolates in each MLVA level that were genetically similar to the dominant Farm A clade to estimate the number of sporadic isolates likely to be similarly related. These steps are represented in Figure 1, and further described below.

Figure 1. Study methods



Epidemiological Investigations

All salmonellosis cases in Victoria are reported to the Department of Health and Human Services (DHHS) by doctors and laboratories under public health legislation. Epidemiological investigation is undertaken using standard disease investigation questionnaires if two or more cases are epidemiologically linked, a clinician nominates a suspected food source, or where spatiotemporal clustering is identified. The majority of other cases are considered sporadic, and not routinely investigated.

Bacterial Isolates

Primary laboratories in Victoria voluntarily forward *Salmonella* clinical isolates to the Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL) for further typing. Virtually all isolates for the state are forwarded to MDU PHL. *Salmonella* isolates received at MDU PHL are serotyped (26) and Phage typed (27) according to published methods. Isolates received after 1 July 2014 have routinely been further typed using MLVA according to standard methods (28). Stored isolates received between 1 July 2009 and 30 June 2014 underwent retrospective MLVA analysis using the same method.

Establishing MLVA levels

We examined MLVA patterns for all *S.* Typhimurium Phage type 9 isolates submitted to the MDU PHL culture collection during the initial study period, and categorised isolates into MLVA levels. MLVA levels were determined by the number of repeat difference variations at each loci relative to the predominant Farm A pattern, as follows:

Level 1: Patterns indistinguishable from the predominant Farm A pattern (2-22-12-10-0212);

Level 2: Differ at one of the three central loci (ie STTR5, STTR6, or STTR10pl) by one or two repeat differences;

Level 3: Differ at two of the three central loci by one or two repeats;

Level 4: Differ at one of the three central loci by three or more repeats;

Level 5: Differ at two of the three central loci including three or more repeats at one of these loci;

Level 6: Unrelated to the predominant pattern by MLVA, ie any number of repeat differences at all three central loci or at one of the outer loci (ie STTR 9 or STTR3).

Isolate selection for WGS

Ninety-nine clinical isolates from 11 Farm A-associated outbreaks were selected for WGS. All available isolates were sequenced for 10 outbreaks (n=69); for the eleventh outbreak (associated with 116 isolates) 30 isolates were randomly selected.

One hundred and sixty one sporadic isolates with no known Farm A association were selected using a stratified sampling approach based on MLVA level. This included 45 isolates from MLVA level 1; 22 each from levels 2 and 3; 21 from level 4 and 25 from level 6. In addition, 41 isolates from outbreaks with no known Farm A association were randomly selected (regardless of MLVA pattern) to ensure the complete *S.* Typhimurium population structure was represented.

Whole genome sequencing

WGS of purified isolate DNA was performed on the Illumina NextSeq platform using Nextera XT libraries and protocols (Illumina, San Diego, California, USA). Raw sequence data has been uploaded to the European Nucleotide Archive (ENA) under the study accession PRJEB12216.

Read Sequence Quality

An initial quality check and filtering step was performed on each of the Illumina WGS read sets. Skewer (29) was used to remove Illumina Nextera adapters and low-quality sequence (Phred score <10). Kraken (30) was used to infer a taxonomic classification for each read and check for contamination.

Core Genome Comparisons

Filtered read sets were mapped to the reference genome *Salmonella enterica* subsp. enterica serovar Typhimurium strain LT2 (accessions: NC_003197 [chromosome] and NC_003277 plasmid pSLT]) and single nucleotide polymorphisms (SNPs) called using *Snippy* v2.5 (https://github.com/tseemann/snippy), requiring a minimum base quality of 20, minimum read coverage of 10x and 90% read concordance at a given nucleotide position for a variant to be reported. Those positions in the reference genome that were covered by at least one read from each and every genome defined a core genome (31). *Snippy* v2.5 was also used to construct a core genome multiple alignment file and this was the input for inferring phylogenetic relationships among isolates using FastTree (32). Bootstrapping was

performed by feeding 1000 resampled core genome alignments generated in *Seqboot* v3.69 (http://evolution.genetics.washington.edu/phylip/doc/seqboot.html) into *FastTree* using the –n option. The resulting core genome phylogeny was viewed alongside relevant data for each taxon using the Interactive Tree of Life (33). Core genome pairwise SNP distances were additionally analysed using a custom R script (https://github.com/MDU-PHL/pairwise_snp_differences) (34) and STATA (35).

To understand the extent of genetic variability within and between *S.* Typhimurium Phage type 9 outbreaks in our local context, pairwise SNPs comparisons were made for all groups of outbreak isolates in the dataset. To explore variability arising between Farm A associated isolates over time, we examined pairwise SNP differences between Farm A associated isolates collected in 2010 against those collected in 2014. To examine the relatedness of sequenced isolates to isolates associated with Farm A, we identified a dominant clade of Farm A isolates (separated by 0 SNPs) and examined SNP differences between sequenced sporadic isolates and this dominant Farm A clade. We tabulated the number and percent of sequenced isolates within each SNP category according to MLVA level producing MLVA-SNP categories.

Estimating the number of historical isolates related to Farm A

Next, we estimated the total number of sporadic isolates expected within each MLVA-SNP category. For each MLVA level we assumed the percentage of sequenced isolates in a given SNP category approximated the percentage of historic isolates from that MLVA level that would also fall in this SNP category. Therefore, for each MLVA level we multiplied the percentage of sequenced sporadic isolates in each SNP category by the total number of sporadic isolates in that MLVA level, to estimate the total number of isolates from the MLVA level expected in each SNP category. We calculated ninety-five percent confidence intervals in STATA (35) using confidence intervals for a proportion.

Results

We found that over 50 sporadic isolates are like to have indistinguishable core genomes, and almost 400 are likely within 20 SNPs of the dominant Farm A clade.

Epidemiological Investigations

Eleven outbreaks associated with consumption of raw or runny eggs were epidemiologically linked to Farm A during the initial study period (two during 2010, eight during 2014, and one

during 2015). These outbreaks accounted for 340 cases (187 laboratory confirmed) and were all associated with foods prepared in restaurants.

Table 1. Summary of epidemiological investigations for outbreaks associated with Farm A during the initial study period

Outbreak code	Year	Month	No. isolates (No. cases)	Predominant MLVA*	% Isolates with predominant MLVA	Suspected source (eggs from Farm A)	Environmental microbiology
Farm A Outbreak 1	2010	Jan	3 (13)	2-22-12-10-0212	100%	Raw egg dish	*Salmonella* not detected
Farm A Outbreak 2	2010	Feb	4 (8)	2-22-12-10-0212	100%	Scrambled eggs	*Salmonella* not detected
Farm A Outbreak 3	2014	Jan - Feb	6 (6)	2-22-12-10-0212	83%	Poached and scrambled eggs	*Salmonella* not detected
Farm A Outbreak 4	2014	Jan	3 (3)	2-22-12-9-0212	100%	Runny eggs	Samples not obtained
Farm A Outbreak 5	2014	Feb	116 (242)	2-22-12-10-0212	97%	Mayonnaise	*Salmonella* Typhimurium 9 isolated from mayonnaise spoon.
Farm A Outbreak 6	2014	Feb	3 (3)	2-22-12-10-0212	100%	Aioli	*Salmonella* not detected
Farm A Outbreak 7	2014	Feb	2 (2)	2-22-12-10-0212	100%	French Toast	Samples not obtained
Farm A Outbreak 8	2014	Feb	13 (13)	2-22-12-10-0212	100%	Scrambled eggs	Samples not obtained
Farm A Outbreak 9	2014	Feb	12 (14)	2-22-12-10-0212	100%	Scrambled eggs, French toast, and hollandaise sauce	Salmonella Typhimurium 9 isolated from Hollandaise sauce
Farm A Outbreak 10	2014	Feb	11 (20)	2-22-12-10-0212	100%	Menu includes aioli containing raw eggs (eggs from Farm A)	*Salmonella* not detected
Farm A Outbreak 11	2015	Feb	13 (16)	2-22-15-10-0212	85%	Chicken and mayonnaise mix (eggs from Farm A)	*Salmonella* Typhimurium 9 isolated from chicken and mayonnaise mix

* MLVA, multilocus variable-number tandem-repeat analysis. MLVA patterns provided using European nomenclature. See Supplementary table S1 for Australian nomenclature

MLVA Analysis

MLVA results were available for 1655 sporadic *S.* Typhimurium Phage type 9 isolates during the initial study period. Prior to 2015, the majority of isolates associated with Farm A shared the MLVA pattern 2-22-12-10-0212. Outbreak 4 was the only exception; the predominant MLVA pattern amongst clinical isolates was 2-22-12-9-0212 (differing by one repeat). During the 2015 Farm A associated outbreak (Outbreak 11) most isolates had an MLVA type of 2-22-15-10-0212, differing from previous outbreaks by three repeats at the STTR6 locus. The predominant MLVA pattern for each Farm A associated outbreak was detailed in Table 1.

MLVA results for sporadic clinical isolates received during the initial study period are summarised according to MLVA level in Table 2. There were 225 unique patterns, with 20 accounting for more than 1% of isolates. The predominant Farm A pattern accounted for 12% of all sporadic isolates. Overall, 14% of MLVA patterns differed from the predominant Farm A pattern by one or two repeat difference at one locus, and 8% differed by one or two repeats at two central loci.

Table 2. MLVA* level classification of S. Typhimurium 9 isolates, July 2009-June 2015

MLVA level†	MLVA	All sporadic isolates (n=1655)		Sequenced sporadic isolates (n=161)	
		No.	%	No.	%
Level 1	2-22-12-10-0212	198	12	45	28
Level 2	2-22-12-09-0212	94	6	14	9
	2-20-12-10-0212	42	3	0	0
	2-22-13-10-0212	31	2	3	2
	2-23-12-10-0212	27	2	0	0
	2-21-12-10-0212	25	2	0	0
	Other (<1%; n=5)	19	1	7	4
Level 2 subtotal		238	14	24	15
Level 3	2-20-14-10-0212	24	1	2	1
	Other (<1%, n=27)	115	7	20	12
Level 3 subtotal		139	8	22	14
Level 4	Other (<1%; n=11)	31	2	22	14
Level 5	2-22-23-09-0212	53	3	10	6
	2-25-15-10-0212	44	3	6	4
	Other (<1%, n=50)	88	11	7	4
Level 5 subtotal		185	17	23	14
Level 6	2-24-15-11-0212	105	6	1	1
	2-10-15-11-0212	103	6	6	4
	2-10-14-11-0212	86	5	4	2
	2-08-13-10-0212	60	4	0	0
	2-08-14-10-0212	40	2	0	0
	2-24-14-11-0212	35	2	0	0
	2-23-24-09-0212	33	2	0	0
	2-23-23-09-0212	31	2	1	1
	2-20-11-12-0212	29	2	3	2
	2-08-12-10-0212	18	1	0	0
	2-09-12-10-0212	18	1	1	1
	Other (<1%; n=117)	306	18	9	6
Level 6 subtotal		864	52	25	16

* MLVA: multilocus variable-number tandem-repeat analysis

† Level 1: indistinguishable from the predominant Farm A pattern; Level 2: one central loci differs by 1-2 repeat differences; Level 3: two central loci differ by 1-2 repeat differences; Level 4: one central loci differs by 3 or more repeat differences; Level 5: two central loci differ and at least one differs by 3 or more repeat differences; Level 6 unrelated.

Analysis of WGS data

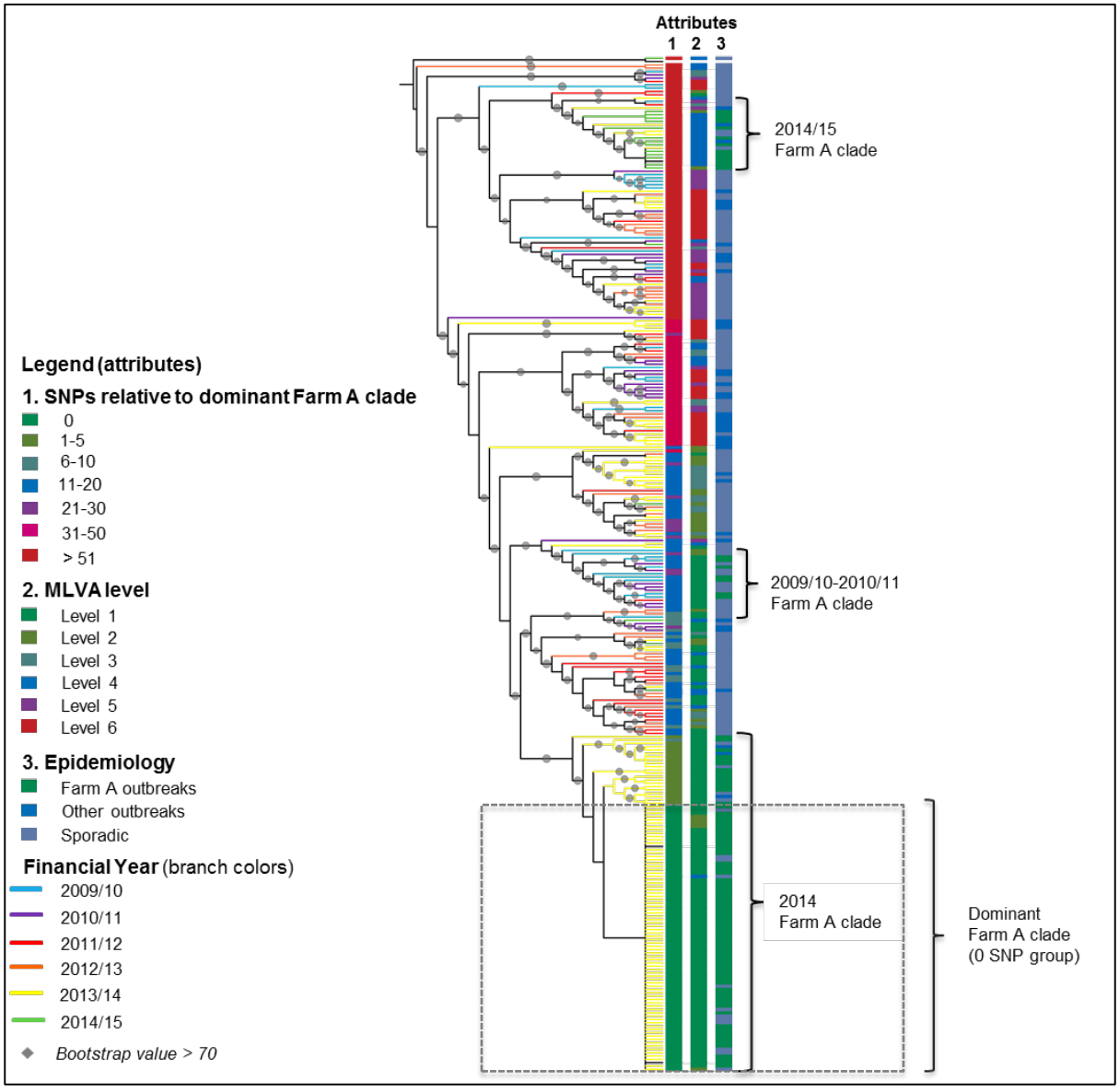
Mapping whole genome sequence reads from each of the 301 clinical isolates to the *S. Typhimurium* strain LT2 reference genome defined a 4473834 base pair (bp) core genome, spanning 90% of the 4951371 bp reference chromosome. Pairwise comparisons of SNPs among the 301 isolates revealed only 2193 variable nucleotide positions. All isolates differed by less than 200 SNPs and were between 662-7732 SNPs different to the *S. Typhimurium* strain LT2 reference genome sequence, apart from one outlier (MLVA 2-22-15-10-0212) that differed by up to 672 SNPs from the other isolates. Repeat wet lab analysis confirmed the Phage type and MLVA profile for this isolate.

Phylogenetic analysis reveals temporal clustering of Farm A isolates

The high-resolution phylogenetic relationship inferred among the isolates from the core genome alignment is presented in Figure 2. Farm A associated isolates clustered in three temporal groups aligned with epidemiologically confirmed outbreaks, including a clade featuring isolates from 2009/10 and 2010/11, one from 2013/14, and one from 2014/15 (Figure 2).

All Farm A associated clades include sporadic isolates, and one includes outbreak isolates with no identified Farm A association. Two such outbreak isolates from January 2014 were within five SNPS of the dominant Farm A clade. These two cases shared a variety of raw and runny egg exposures including foods purchased outside the home. The *Salmonella* source was not identified, but this finding suggests a likely association with Farm A.

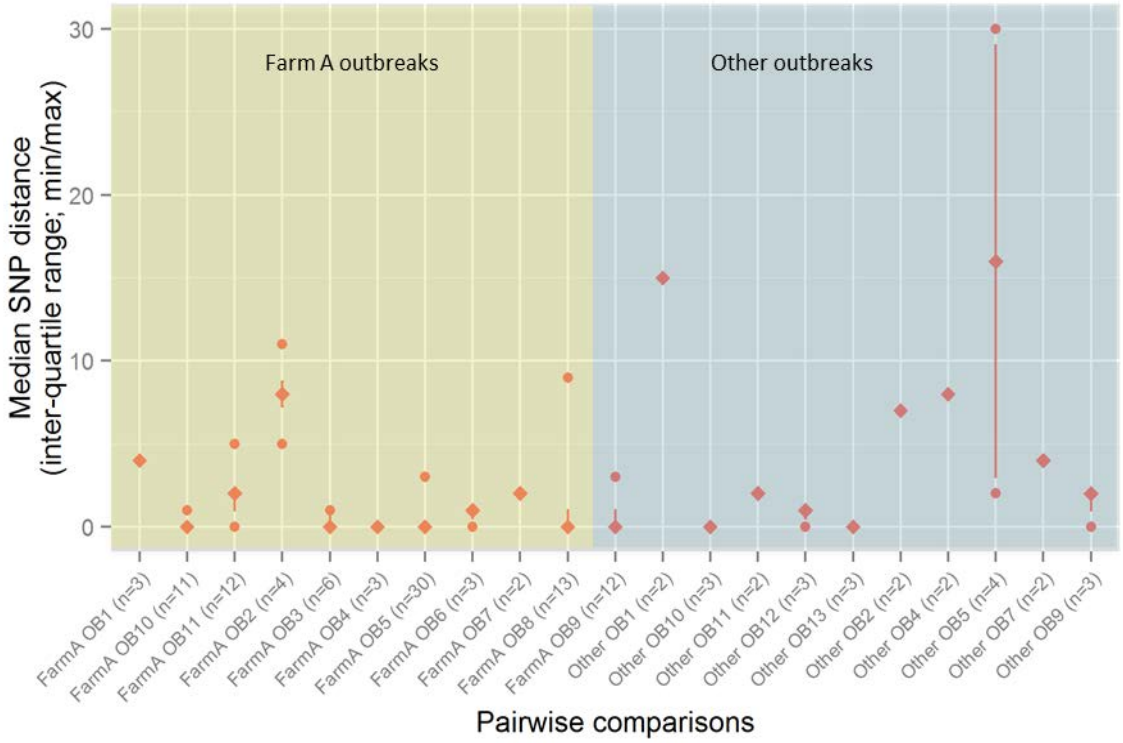
Figure 2. Phylogenomic relationships and MLVA patterns for Farm A and other isolates, July 2009-June 2015 Phylogenomic relationships between sequenced isolates are shown alongside key attributes. Attributes include: SNPs (single nucleotide polymorphisms) relative to the dominant Farm A clade (attribute column 1); MLVA level (multilocus variable-number tandem-repeat analysis variation from the Farm A pattern, defined below) (attribute column 2), and epidemiological details (Farm A outbreak, other outbreak, or sporadic) (attribute column 3). Branch colors show financial year at the time of sample collection. Grey circles denote branches with bootstrap values above 70%. The dominant Farm A clade is boxed, and features isolates separated by zero SNPs difference. Branch lengths are transformed and not to scale. Three 'Farm A associated clades' include both Farm A associated isolates and 'sporadic' isolates from similar points in time.



SNP differences

SNP differences within outbreaks are illustrated in Figure 3. Two Farm A outbreaks had a maximum of nine and 11 SNPs respectively, while nine were characterised by five or fewer SNPs. One non-Farm A outbreak exhibited 30 SNPs, three featured seven to 15 SNPs and seven outbreaks exhibited five or fewer SNPs. The outbreak with 30 SNPs was a point source outbreak of seven cases who all reported consuming Eggs Benedict at a single restaurant.

Figure 3. Single nucleotide polymorphism differences observed in outbreaks for Farm A outbreaks and those from outbreaks with no identified Farm A association



As shown in Figure 4, SNP differences between sporadic sequenced isolates and the dominant Farm A clade followed a bimodal distribution, with most isolates presenting less than 40 or 80-100 SNPs from the dominant clade.

Figure 4. Frequency of single nucleotide polymorphism differences amongst sporadic sequenced isolates, relative to the dominant Farm A clade

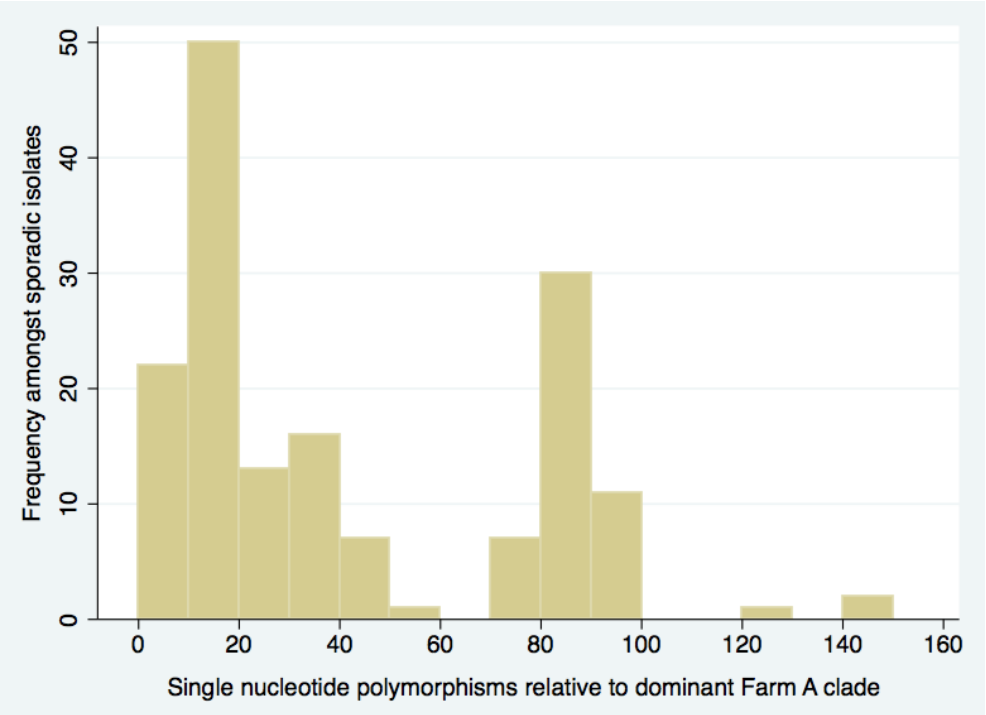
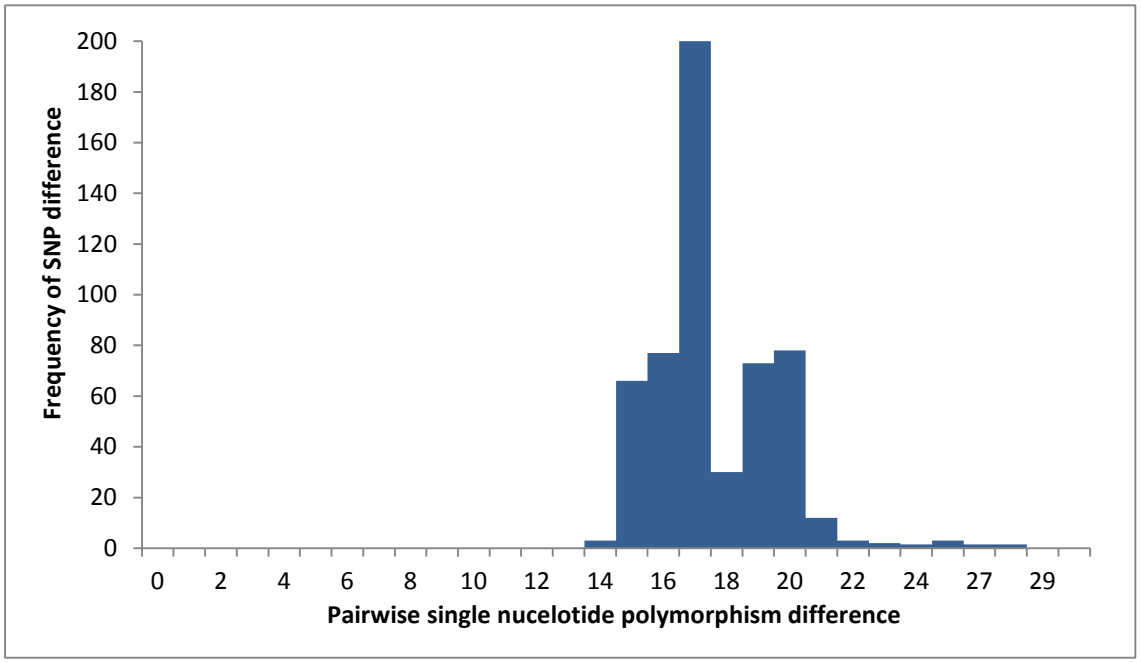


Figure 5 presents SNP differences between Farm A associated isolates from 2010 relative to 2014 isolates. The median difference between isolates from the two years (17 SNPs, range 14 to 28) is substantially larger than that within 2010 isolates (1 SNP), or 2014 isolates (0 SNPs), reflecting evolutionary change in the source population.

Figure 5. Frequency of single nucleotide polymorphism differences in pairwise comparisons between 2010 vs 2014 Farm A isolates



Estimated number of sporadic isolates in each SNP category

SNP differences between sequenced sporadic isolates and the main Farm A clade for the period July 2009-June 2014 (the estimation study period) are categorised in Table 3 according to MLVA level. The estimated number of historic sporadic isolates in each MLVA-SNP category is also presented. Calculations are provided in the column headers. Isolates from 2015 were excluded from calculations because a change in predominant Farm A MLVA pattern meant isolate selection utilising defined MLVA levels would not select Farm A isolates in a representative way.

During the estimation study period, over one fifth of all sequenced isolates from MLVA level 1 and one in twenty from level 2 had core genomes indistinguishable from the dominant Farm A clade. Therefore we estimated over one fifth of all historical isolates from MLVA level 1 and one in twenty from MLVA level 2 may also be genetically indistinguishable. Overall we estimate 55 of 1585 historical isolates may have indistinguishable core genomes, 72 may fall within 5 SNPs, and almost 400 may be within 20 SNPs of the predominant Farm A clade, substantially more than the 174 lab confirmed cases known to be associated with Farm A during the period.

Table 3. Estimates of the pairwise core genome SNP* differences between the typical Farm A clade isolates and isolates at varying MLVA levels† amongst sporadic *Salmonella* Typhimurium Phage type 9 isolates, July 2009-June 2014 (n=1585)

SNP* difference from dominant FarmA clade	MLVA Level†	Sequenced isolates in MLVA level	Sequenced isolates in MLVA-SNP category			Total sporadic isolates in MLVA level	Estimated sporadic isolates in MLVA-SNP category	
		No	No	%	95% CI (%)‡	No	No	95% CI
		s	n	p [=(n/s)x100]	(d, e)	t	(p/100)*t	((d/100)*t, (e/100)*t)
0 SNPs	Level 1	45	10	22	(12, 37)	195	43	(23, 72)
	Level 2	22	1	5	(1, 30)	230	10	(1, 69)
	Level 3	22	0	0		128	0	(0, 0)
	Level 4	21	1	5	(1, 31)	28	1	(0, 9)
	Level 5	23	0	0		181	0	(0, 0)
	Level 6	25	0	0		823	0	(0, 0)
0 SNPs subtotal							55	(25, 150)
Cumulative total							**55**	**(25, 150)**
1-5 SNPs	Level 1	45	4	9	(3, 22)	195	17	(6, 43)
	Level 2	22	0	0		230	0	(0, 0)
	Level 3	22	0	0		128	0	(0, 0)
	Level 4	21	0	0		28	0	(0, 0)
	Level 5	23	0	0		181	0	(0, 0)
	Level 6	25	0	0		823	0	(0, 0)
1-5 SNPs subtotal							17	(6, 43)
Cumulative total							**72**	**(31, 193)**
6-10 SNPs	Level 1	45	9	20	(10, 35)	195	39	(20, 68)
	Level 2	22	1	5	(1, 30)	230	12	(2, 69)
	Level 3	22	0	0		128	0	(0, 0)
	Level 4	21	2	10	(2, 34)	28	3	(1, 10)
	Level 5	23	0	0		181	0	(0, 0)
	Level 6	25	0	0		823	0	(0, 0)
6-10 SNPs subtotal							53	(22, 147)
Cumulative total							**126**	**(53, 340)**

Table 3. continued

11-20 SNPs	Level 1	45	19	42	(28, 57)	195	82	(55, 111)
	Level 2	22	11	50	(29, 71)	230	115	(67, 163)
	Level 3	22	12	55	(33, 75)	128	70	(42, 96)
	Level 4	21	4	19	(7, 44)	28	5	(2, 12)
	Level 5	23	0	0	(0, 0)	181	0	(0, 0)
	Level 6	25	0	0		823	0	(0, 0)
11-20 SNPs subtotal							272	(166, 383)
Cumulative total							**398**	**(219, 722)**
21-30 SNPs	Level 1	45	2	4	(1, 2)	195	9	(2, 4)
	Level 2	22	7	32	(15, 55)	230	73	(35, 127)
	Level 3	22	1	5	(1, 30)	128	6	(1, 38)
	Level 4	21	0	0		28	0	(0, 0)
	Level 5	23	0	0		181	0	(0, 0)
	Level 6	25	1	4	(0, 26)	823	33	(0, 214)
21-30 SNPs subtotal							120	(38, 383)
Cumulative total							**518**	**(256, 1105)**
31-50 SNPs	Level 1	45	0	0	(0, 0)	195	0	(0, 0)
	Level 2	22	1	5	(1, 30)	230	10	(2, 69)
	Level 3	22	5	23	(9, 46)	128	29	(12, 59)
	Level 4	21	5	24		28	7	(3, 13)
	Level 5	23	3	13		181	24	(7, 65)
	Level 6	25	7	28	(13, 50)	823	230	(107, 412)
31-50 SNPs subtotal							300	(131, 618)
Cumulative total							**818**	**(387, 1723)**
>51 SNPs	Level 1	45	1	2	(0, 15)	195	4	(0, 29)
	Level 2	22	1	5	(1, 30)	230	10	(2, 69)
	Level 3	22	4	18	(6, 42)	128	23	(8, 54)
	Level 4	21	9	43		28	12	(6, 18)
	Level 5	23	20	87		181	157	(116, 174)
	Level 6	25	17	68	(46, 84)	823	560	(379, 691)
>51 SNPs subtotal							767	(511, 1036)
Cumulative total							**1585**	**(898, 2759)**

* SNP single nucleotide polymorphism

† MLVA: multilocus variable-number tandem repeat analysis level - Level 1: indistinguishable from the predominant Farm A pattern; Level 2: one central loci differs by 1-2 repeat differences; Level 3: two central loci differ by 1-2 repeat differences; Level 4: one central loci differs by 3 or more repeat differences; Level 5: two central loci differ and at least one differs by 3 or more repeat differences

‡95% CI; 95% confidence interval, calculated using confidence interval for a proportion

Discussion

We examined WGS and MLVA data for 301 *S.* Typhimurium Phage type 9 clinical isolates, calculated pairwise SNP differences between sporadic clinical isolates and a reference clade associated with a recurrent outbreak source, and estimated the number of sporadic isolates likely to be related to this dominant clade over a five year period. We found that over 50 sporadic isolates may have had indistinguishable core genomes and almost 400 are likely within 20 SNPs of the clade indicating numerous 'sporadic' cases likely share a common source or reservoir. The estimated number of isolates within 20 SNPs eclipses the number of isolates with known epidemiological links to the outbreak source.

Our finding that primary sources of *Salmonella* outbreaks may be associated with a substantial number of additional cases for whom epidemiological links are not currently recognised is consistent with previous studies. Octavia *et al* (36) observed a 107% increase in *S.* Typhimurium outbreak size when WGS was performed on sporadic isolates with indistinguishable MLVA notified within two weeks of the outbreaks. Similarly, den Bakker *et al* (19) sequenced isolates with similar PFGE patterns received within three weeks of an *S.* Enteritidis outbreak. A distinct outbreak clade included seven identified outbreak cases, plus nine 'sporadic' isolates differing by no more than one SNP.

Because there is uncertainty regarding SNP variation arising from a persistently contaminated source, we estimated numbers of isolates for a range of SNP differences from the dominant Farm A clade. Point source outbreaks of *S.* Typhimurium are frequently characterised by 5 or fewer SNPs (6, 24, 36). Thus, the 72 isolates estimated to be within five SNPs of the dominant clade are likely related to the outbreak source.

At least some isolates with greater than five SNPs are also likely related to the outbreak source. Around one third of outbreaks we reviewed were characterised by over five SNP differences, including one with a maximum of 30 SNPs. The potential for genetic divergence in a source population (eg on a primary production property or feed supply) resulting in large SNPs between epidemiologically linked isolates has been recognised (19) and is consistent with a previous report documenting up to 30 SNPs in an outbreak of *S.* Typhimurium Definitive Type 12 (37). Single Nucleotide Polymorphisms within a single food borne outbreak may not reflect the diversity in the source population, because the majority of outbreak isolates arise from a small number of contaminating organisms (17). Where multiple outbreaks arise from a single source population this 'bottleneck' effect may be overcome because separate 'bottlenecks' contribute to each outbreak.

We documented up to 28 SNPs between isolates associated with Farm A collected four years apart. Other researchers have reported similar findings, for example 22 SNPs were observed between *S.* Typhimurium isolates linked to a single farm over a five year period (6), and 15 SNPs was identified between related *S*. Dublin isolates collected two years apart (38). The genetic diversity associated with the source population in this study is unknown. However, given Farm A has been linked to outbreaks with indistinguishable MLVA patterns for over ten years, a previously estimated rate of change for *S.* Typhimurium of 3-5 SNPs per year (39) suggests a difference of 30-50 SNPs may be reasonable.

Distinguishing between linked isolates and background population is a vital step in outbreak investigations (24). Knowledge of the population structure through careful sampling of the background population is necessary for WGS to discriminate between outbreak-related and sporadic cases (14). We included sporadic isolates from a range of MLVA patterns as well as outbreaks with no identified Farm A association to document the population structure of *S.* Typhimurium Phage type 9. A shift in WGS observed in Farm A isolates during 2015 coincided with enhanced control measures and perhaps increased evolutionary pressure at the farm, demonstrating the importance of supplementing WGS with epidemiological data.

The study is limited by knowledge gaps regarding the genetic diversity of egg production sites in Victoria, and thus care must be taken in interpreting these findings. It is possible genetically similar isolates arose from other sites sharing a source of contamination with Farm A, such as parent stock or food supply. For example, two separate egg producers who shared a contaminated feed supply were both implicated in an outbreak of *S.* Enteritidis in the USA, leading to record egg recalls (15).

We acknowledge that WGS data cannot be the sole basis for determining isolates originating from a common source (18), and epidemiological investigation to confirm such links is imperative to establish causation. Real-time WGS results that identify genetically similar isolates may soon be routinely available in Victoria. Increased discriminatory power will prompt increased recognition of potential outbreaks and better detection of outbreak associated cases, leading to increased public health follow up to confirm epidemiological links between isolates previously regarded as ‘sporadic’. Our findings suggest the burden of disease attributable to persistent sources will be demonstrably larger than cases observed in outbreaks, highlighting the need for comprehensive trace back and intervention at the primary production level.

Acknowledgments

We are grateful for bioinformatic support provided by Anders Goncalves de Silva, and laboratory support provided by Tomita Takehiro and Susan Ballard. Thanks also to Joy Gregory for providing feedback on the final manuscript. Zoe Cutcher is supported by a Master of Applied Epidemiology (MAE) scholarship provided by Victorian DHHS. Doherty Applied Microbial Genomics is funded by the Department of Microbiology and Immunology at The University of Melbourne. The National Health and Medical Research Council, Australia funded a Senior Research Fellowship GNT1008549 to T.P.S. and Career Development Fellowship GNT1023526 to B.P.H.).

Biographical Sketch

Zoe Cutcher is a Master of Philosophy (Applied Epidemiology) scholar with Australian National University, Canberra (Australia's field epidemiology training program). She is completing a field placement at the Victorian government Department of Health and Human Services, Melbourne where her research includes exploring implications of genomics for communicable disease surveillance.

References

1. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal Salmonella gastroenteritis. Clin Infect Dis. 2010;50(6):882-9.

2. OzFoodNet Working Group. Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: annual report of the OzFoodNet network, 2011. Communicable diseases intelligence. 2015;39(2):E236-64.

3. Lindstedt BA, Heir E, Gjernes E, Kapperud G. DNA fingerprinting of Salmonella enterica subsp. enterica serovar typhimurium with emphasis on phage type DT104 based on variable number of tandem repeat loci. Journal of clinical microbiology. 2003;41(4):1469-79.

4. Lindstedt BA, Vardund T, Aas L, Kapperud G. Multiple-locus variable-number tandem-repeats analysis of Salmonella enterica subsp. enterica serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. J Microbiol Methods. 2004;59(2):163-72.

5. Barua H, Lindblom IL, Bisgaard M, Christensen JP, Olsen RH, Christensen H. In vitro and in vivo investigation on genomic stability of Salmonella enterica Typhimurium DT41 obtained from broiler breeders in Denmark. Vet Microbiol. 2013;166(3-4):607-16.

6. Dimovski K, Cao H, Wijburg OL, Strugnell RA, Mantena RK, Whipp M, et al. Analysis of Salmonella enterica serovar Typhimurium variable-number tandem-repeat data for public health investigation based on measured mutation rates and whole-genome sequence comparisons. Journal of bacteriology. 2014;196(16):3036-44.

7. Paranthaman K, Haroon S, Latif S, Vinnyey N, de Souza V, Welfare W, et al. Emergence of a multidrug-resistant (ASSuTTm) strain of Salmonella enterica serovar Typhimurium DT120 in England in 2011 and the use of multiple-locus variable-number tandem-repeat analysis in supporting outbreak investigations. Foodborne Pathog Dis. 2013;10(10):850-5.

8. Kuhn KG, Sørensen G, Torpdahl M, Kjeldsen MK, Jensen T, Gubbels S, et al. A long-lasting outbreak of Salmonella Typhimurium U323 associated with several pork products, Denmark, 2010. Epidemiology and infection. 2013;141(2):260-8.

9. Inns T, Lane C, Peters T, Dallman T, Chatt C, McFarland N, et al. A multi-country Salmonella Enteritidis phage type 14b outbreak associated with eggs from a German

producer: near real-time application of whole genome sequencing and food chain investigations, United Kingdom, May to September 2014. Euro Surveill. 2015;20(16).

10. Friesema IH, Schimmer B, Ros JA, Ober HJ, Heck ME, Swaan CM, et al. A regional Salmonella enterica serovar Typhimurium outbreak associated with raw beef products, The Netherlands, 2010. Foodborne Pathog Dis. 2012;9(2):102-7.

11. Garvey P, McKeown P, Kelly P, Cormican M, Anderson W, Flack A, et al. Investigation and management of an outbreak of Salmonella Typhimurium DT8 associated with duck eggs, Ireland 2009 to 2011. Euro Surveill. 2013;18(16):20454.

12. Petersen RF, Litrup E, Larsson JT, Torpdahl M, Sørensen G, Müller L, et al. Molecular characterization of Salmonella Typhimurium highly successful outbreak strains. Foodborne Pathog Dis. 2011;8(6):655-61.

13. Moffatt CR, Musto J. Salmonella and egg-related outbreaks. Microbiol Aust. 2013;34(2):94-7.

14. Bakker HC, Switt AI, Cummings CA, Hoelzer K, Degoricija L, Rodriguez-Rivera LD, et al. A whole-genome single nucleotide polymorphism-based approach to trace and identify outbreaks linked to a common Salmonella enterica subsp. enterica serovar Montevideo pulsed-field gel electrophoresis type. Appl Environ Microbiol. 2011;77(24):8648-55.

15. Allard M, Luo Y, Strain E, Pettengil J, Timme R, Wang C, et al. On the Evolutionary History, Population Genetics and Diversity among Isolates of Salmonella Enteritidis PFGE Pattern JEGX01.0004. PloS one. 2013;8(1).

16. Deng X, Shariat N, Driebe EM, Roe CC, Tolar B, Trees E, et al. Comparative analysis of subtyping methods against a whole-genome-sequencing standard for Salmonella enterica serotype Enteritidis. Journal of clinical microbiology. 2015;53(1):212-8.

17. Sintchenko V, Holmes EC. The role of pathogen genomics in assessing disease transmission. BMJ (Clinical research ed). 2015;350:h1314.

18. Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ, Medus C, et al. Characterization of Foodborne Outbreaks of Salmonella enterica Serovar Enteritidis with Whole-Genome Sequencing Single Nucleotide Polymorphism-Based Analysis for Surveillance and Outbreak Detection. Journal of clinical microbiology. 2015;53(10):3334-40.

19. den Bakker HC, Allard MW, Bopp D, Brown EW, Fontana J, Iqbal Z, et al. Rapid whole-genome sequencing for surveillance of Salmonella enterica serovar enteritidis. Emerging infectious diseases. 2014;20(8):1306-14.

20. Koser CU EM, Cartwright EJP, Gillespie SH, Brown NM, et al. Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. PLOS Pathogens. 2012;8(8):e1002824.

21. PHG Foundation. Pathogen Genomics Into Practice. www.phgfoundation.org: 2015.

22. Dunne WM, Jr., Westblade LF, Ford B. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2012;31(8):1719-26.

23. Ashton PM, Satheesh Nair, Tansy Peters, Janet Bale, David G. Powell, Anaïs Painset, Rediat Tewolde et al. . Identification and typing of Salmonella for public health surveillance using whole genome sequencing. PeerJPre-Prints. 2015;3(e1778):https://dx.doi.org/10.7287/peerj.preprints.1425v1.

24. Ashton PM, Peters T, Ameh L, McAleer R, Petrie S, Nair S, et al. Whole Genome Sequencing for the Retrospective Investigation of an Outbreak of Salmonella Typhimurium DT 8. PLoS currents. 2015;7:http://www.ncbi.nlm.nih.gov/pubmed/25713745.

25. US Food and Drug Administration. Genome Trakr Fast Facts http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm403550.htm: US Department of Health and Human Services; 2015 [cited 2015 27 October]. Available from: http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm403550.htm.

26. Grimont PAD, F. W. Antigenic formulae of the Salmonella serovars. WHO Collaborating Centre for Reference and Research on Salmonella: Insitut Pasteur, France, 1986.

27. Anderson ES, Ward LR, Saxe MJ, de Sa JD. Bacteriophage-typing designations of Salmonella typhimurium. The Journal of hygiene. 1977;78(2):297-300.

28. European Centre for Disease Prevention and Control. Laboratory standard operating procedure for MLVA of Salmonella enterica serotype Typhimurium. Stockholm: ECDC; 2011.

29. Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC bioinformatics. 2014;15:182.

30. Davis MP, van Dongen S, Abreu-Goodger C, Bartonicek N, Enright AJ. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. Methods. 2013;63(1):41-9.

31. Chen C, Zhang W, Zheng H, Lan R, Wang H, Du P, et al. Minimum core genome sequence typing of bacterial pathogens: a unified approach for clinical and public health microbiology. Journal of clinical microbiology. 2013;51(8):2582-91.

32. Price MN, Dehal PS, Arkin AP. FastTree 2-approximately maximum-likelihood trees for large alignments. PloS one. 2010;5(3):e9490.

33. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007;23(1):127-8.

34. R Core Team. R: A language and environment for statistical computing. 2015.

35. StataCorp. Stata Statistical Software: Release 13. 2013.

36. Octavia S, Wang Q, Tanaka MM, Kaur S, Sintchenko V, Lan R. Delineating community outbreaks of Salmonella enterica serovar Typhimurium by use of whole-genome sequencing: insights into genomic variability within an outbreak. Journal of clinical microbiology. 2015;53(4):1063-71.

37. Leekitcharoenphon P, Nielsen EM, Kaas RS, Lund O, Aarestrup FM. Evaluation of whole genome sequencing for outbreak detection of Salmonella enterica. PloS one. 2014;9(2):e87991.

38. Mohammed M, Delappe N, O'Connor J, McKeown P, Garvey P, Cormican M. Whole genome sequencing provides an unambiguous link between Salmonella Dublin outbreak strain and a historical isolate. Epidemiology and infection. 2015;FirstView (Epub 13 July 2015):http://dx.doi.org.ezp.lib.unimelb.edu.au/10.1017/S0950268815001636.

39. Hawkey J, Edwards DJ, Dimovski K, Hiley L, Billman-Jacobe H, Hogg G, et al. Evidence of microevolution of Salmonella Typhimurium during a series of egg-associated outbreaks linked to a single chicken farm. BMC genomics. 2013;14:800.

Chapter 3 Appendices

Appendix 1. Supplementary Table

S1. Australian and European nomenclature for MLVA patterns

MLVA (European)	MLVA (Australian)
2-22-12-10-0212	03-24-13-11-523
2-22-12-9-0212	03-24-13-10-523
2-20-12-10-0212	03-22-13-11-523
2-22-13-10-0212	03-24-14-11-523
2-23-12-10-0212	03-25-13-11-523
2-21-12-10-0212	03-23-13-11-523
2-20-14-10-0212	03-22-15-11-523
2-22-23-9-0212	03-24-24-10-523
2-25-15-10-0212	03-27-16-11-523
2-24-15-11-0212	03-26-16-12-523
2-10-15-11-0212	03-12-16-12-523
2-10-14-11-0212	03-12-15-12-523
2-08-13-10-0212	03-10-14-11-496
2-08-14-10-0212	03-10-15-11-496
2-24-14-11-0212	03-26-15-12-523
2-23-24-09-0212	03-25-25-10-523
2-23-23-9-0212	03-25-24-10-523
2-20-11-12-0212	03-22-12-13-523
2-08-12-10-0212	03-10-13-11-496
2-09-12-10-0212	03-11-13-11-496

Appendix 2. Technical Descriptions

This appendix is provided as supplementary background information regarding Salmonella typing and whole genome sequencing.

Traditional Typing Methods

Traditional *Salmonella* typing includes (in order of discriminatory power): serotyping, phagetyping, and more recently multilocus variable-number tandem-repeat analysis (MLVA). The first two examine phenotypic (ie physical) characteristics of *Salmonella* isolates, while the latter examines genetic characteristics.

Briefly, serotyping involves serological tests that use antibodies to detect antigens on the body (somatic or "O" antigens) and tail (flagella, or "H" antigens) of the bacteria. It categorises *Salmonella* into over 2000 serovars or serotypes, including *S.* Typhimurium.

Phagetyping is used to further differentiate between *S.* Typhimurium (and some other serovars). The method introduces a range of different viruses with differing capacity to kill bacteria ("bacteriophages") to a bacterial lawn. The phagetype is allocated according to the pattern of lysis produced.

More recently, MLVA has been adopted for routine subtyping of *S.* Typhimurium in Victoria. This method utilises PCR to examine five pre-defined regions on the genome where there are repeating units of DNA (base pairs). The number of base pair repeats at each location is generally identical or very similar for related isolates. However, as the regions are highly variable it is possible they may be identical through chance alone.

Whole Genome Sequencing

Whole genome sequencing involves a pipeline of pathogen preparation, sequencing and data analysis. The following brief description of the pipeline employed at MDU is adapted from Pathogen Genomics into Practice[1].

First *Salmonella* (or any other pathogen) is isolated in culture, and DNA is extracted and purified. DNA is broken down into shorter fragments to enable sequencing, and the ends of DNA fragments are linked to synthesized DNA molecules called "adapters".

Next, sequencing is performed using an Illumina NextSeq, which uses “sequencing by synthesis” technology. In this process, DNA fragments are fixed to a glass plate using the “adapters”. DNA is then amplified, producing dense clusters of identical DNA fragments across the plate. This increases the quality of sequence data. Sequencing begins with the introduction of a universal primer, DNA polymerase and four nucleotides (A, T, C and G), each fused to a colour label and terminator. Incorporation of a complementary nucleotide onto the fragment terminates the reaction. Unincorporated nucleotides are washed off, and colour imaging is used to identify the incorporated base (the outermost complementary base on each strand). The raw signal is converted into a nucleotide (A, T, C, or G) based on its colour. The dye label and terminating group are washed off, and the cycle repeated until all nucleotides along the template fragment have been paired and identified.

The fragments are then aligned into a whole genome sequence by orienting against a “reference” genome. This step involves significant computational effort.

Quality measures

Read depth. Fragments are sequenced multiple times, producing overlapping “reads”. Read depth refers to the number of times a particular sequence is “read”. Greater read depth increases the confidence that a given nucleotide has been inferred correctly. Generally, at MDU a read depth of 80 or greater is considered sufficient.

Read length. Read length refers to the length of the sequenced fragments. Increasing length improves the accuracy when mapping against a reference genome, as it is easier to misplace smaller fragments. The maximum read length available using the NextSeq is 150 base pairs.

Phred score. The Phred-like quality score measures the probability that a nucleotide in a particular position has been identified correctly. A score of 10 suggests a 1 in 10 chance of error; 20 suggests there is a 1 in 100 chance of error, 30 suggests a 1 in 1 000; 40 suggests 1 in 10 000 and 50 suggests 1 in 100 000.

1. PHG Foundation (2015). Pathogen Genomics Into Practice. www.phgfoundation.org.

This page is intentionally blank

Chapter 4. Outbreak investigation: Salmonellosis at a school function

Contents

Preface

Background to project

Investigation of an outbreak or an acute public health event is a core requirement of the Master of Philosophy in Applied Epidemiology (MAE) program. This outbreak came to the attention of Communicable Disease Prevention and Control through a doctor's notification of a case of *Salmonella* Typhimurium 44 infection. The doctor reported the patient had been a guest at a function and that other guests who attended the function were also unwell. OzFoodNet epidemiologists considered the reported outbreak to be an ideal incident for an MAE student investigation and the investigation was assigned to me. I interviewed the first notified case to confirm it was an outbreak and commenced an outbreak investigation.

My role

I was the sole investigator in this outbreak, though I sought advice as required from OzFoodNet epidemiologists, Public Health Officers, and my field supervisor.

My role included:

Interviewing the first notified case and confirming the situation was an outbreak.

Liaising with event organisers and school administrators to obtain names and contact details for all attendees at the function

Liaising with event organisers to obtain a copy of the menu

Designing the menu-based questionnaire

Conducting all telephone interviews

Entering data into the Public Health Event Surveillance System (PHESS) database

Extracting and analysing the data using STATA

Writing the investigation report.

Lessons learnt

The investigation highlighted the value of epidemiological approaches for investigating outbreaks. During the investigation I identified a number of previously unknown cases

amongst the guest list. Although many guests could name a few cases, no single guest (nor any of the organisers) was aware of every case. This meant that until the investigation occurred no one was aware of the size of the outbreak, even amongst this group who all worked together. It taught me of the importance of active case finding to quantify the size of an outbreak, and in future outbreaks I will consider this a priority. It also illustrated how routine surveillance only captures a fraction of disease in the population, as only two out of the ten cases were reported to the routine notification system.

The investigation demonstrated the importance of timely follow up for outbreaks. Unfortunately, by the time the guest list and menus were received and the investigation was underway it was too late to obtain faecal samples from most of the cases, and there were no food samples left for microbiological analysis. This meant the investigation was entirely reliant on the epidemiological investigation and the suspected food vehicles could not be verified by laboratory evidence. At the same time, this reinforced the value of having multiple streams of evidence when trying to determine disease causation. In future outbreak investigations I will ensure all possible streams of evidence are pursued as far as practicable.

I learnt it can be difficult to identify a food source in an outbreak setting with a buffet style menu, because many people eat multiple items. This can be particularly problematic with a small cohort where the sample size limits the utility of stratified analysis. Microbiological evidence may have assisted in confirming the source if it were available.

From a practical perspective I learnt how to conduct an outbreak investigation from start to finish. I also learnt how to set up an outbreak in Victoria's Public Health Surveillance System Database, how to link cases to it, design and complete an outbreak questionnaire, and extract this data for analysis.

Another valuable lesson was the human face of the outbreak. The cases reported here are more than simply 'cases'; they are living, breathing people with busy lives, families, senses of humour and often an incredibly generous spirit. Their recollections of salmonellosis were of a debilitating and painful illness, yet many expressed concern for the organisers of the event and a desire to protect them against any negative ramifications following the incident. The other side of the outbreak has a human face,

too: the school staff who catered for the event in their home kitchens felt a strong sense of responsibility and were deeply troubled by news of the outbreak.

Public Health Implications

The public health implications were limited because the likely food vehicle was prepared in a private kitchen. The staff who catered for the event were provided information regarding *Salmonella*, safe food handling, and egg safety with the aim of preventing recurrence.

The findings from the investigation will contribute to the evidence base regarding *Salmonella* Typhimurium epidemiology. Eggs are frequently suspected as the source of *Salmonella* Typhimurium infections in Australia, however, there are few robust studies documenting complete epidemiological investigation of outbreaks. For example there are currently only three published articles reporting investigations into *Salmonella* Typhimurium phage type 44 outbreaks in Australia despite it being a common cause of egg-associated outbreaks.

Project outputs

Project outputs include the following:

- Inclusion as a chapter in the Bound Volume submission for the Master of Philosophy in Applied Epidemiology
- Presented at the Communicable Disease Control Conference, 2015
- Publication in a peer-reviewed journal (after thesis submission).

Acknowledgements

I gratefully acknowledge the following parties for their role in this project:

- Joy Gregory, Karin Lalor and Marion Easton from OzFoodNet for their support and advice throughout the investigation
- Supervisors Nicola Stephens and Martyn Kirk for provision of excellent guidance and for their patient and dedicated supervision
- Mary Valcanis and Karolina Mercoulia for provision of data and advice regarding MLVA patterns

- Nectaria Tzimourtas for allowing me the opportunity to conduct the investigation
- Public Health Officers Kay Sturge, Jay Healy, Jim Adamopoulos, and Sara Radwan for welcoming my questions, sharing their knowledge and for their good humour
- The principal and administrative staff at the school for providing contact lists for attendees at the function
- The organisers of the event for provision of menus and detailed accounts of food preparation
- Study participants for their assistance with the investigation.

Abstract

Background: In November 2014, a Salmonella Typhimurium phage type 44 outbreak occurred following a function in Victoria. An epidemiological investigation was conducted to characterise the outbreak by person, place and time, and identify the source of illness.

Methods: A retrospective cohort study was conducted of all guests at the function using a menu-based questionnaire and descriptive statistics and relative risks were calculated for all food items. Private laboratories in Victoria performed microbiological analysis on faecal samples. Salmonella isolates were forwarded to the Microbiological Diagnostic Unit for typing including multilocus variable-number tandem-repeat analysis (MLVA).

Results: Twenty-nine out of thirty guests were interviewed. Ten cases were identified giving an attack proportion of 30%. Risk of illness increased with consumption of an appetiser and frittata but after adjusting for confounding the roast beef appetiser was the most likely food vehicle. Both items were prepared in the same domestic kitchen. Cross contamination from eggs was suspected.

Conclusion: A roast beef appetiser was the most likely food vehicle for this outbreak, but cross contamination from eggs appears likely. Appropriate food handling and hygiene is essential to minimise the risks of food borne illness.

Background

Salmonella enteritis is a severe bacterial gastroenteritis. Transmission occurs through the faecal oral route, and contaminated food is the most frequent source of illness (1) although transmission from contaminated water, infected animals or person-to-person transmission is also possible (2). Symptoms commonly include diarrhoea, abdominal pain, nausea and fever lasting between 4 – 7 days (3). Around 20% of culture confirmed cases require hospitalisation, and death is rare (4). It is estimated around 94 million cases of salmonellosis occur globally each year, including 155,000 deaths (2). In Australia it is the second most commonly notified gastrointestinal illnesss (5). Australia's National Notifiable Diseases Surveillance System received 16358 notifications of salmonellosis (a notification rate of 70.7 per 100,000) during 2014 (5). However because only a fraction of cases present to a doctor, a fraction of those presenting are tested, a fraction of tests detect Salmonella due to intermittent shedding, and not all cases of Salmonella isolated at the laboratory are reported to the disease surveillance system, this figure underestimates the true incidence of disease at the population level. It is likely the true rates of disease are around seven fold higher (6).

To assist with identifying outbreaks, *Salmonella* bacteria can be subtyped using a variety of laboratory techniques (7). Subtyping is invaluable for monitoring trends, detecting outbreaks and identifying sources of illness. Current subtyping methods include serotyping, phage-typing, and more recently multilocus variable-number tandem-repeat analysis (MLVA). Serotyping is the first level of subtyping and categorises *Salmonella* into approximately 2,000 servoars or serotypes based on serological tests which detect antigens on the body (somatic or "O" antigens) and tail (flagella, or "H" antigens) of the bacteria (8). Phage typing is used to further differentiate bacteria within some serovars, including *S*. Typhimurium. Phage type is determined by introducing a range of bacteriophages (viruses which kill bacteria) to the isolate and examining patterns of lysis, which vary between phage types (8). In recent years, the use of phage typing has declined due to the difficulty in testing and the increasing popularity of molecular tools. MLVA is a more recent genetic fingerprinting technique with higher discriminatory power (9). It examines pre-defined regions of DNA to assess the number of repeats of base pairs and is useful for determining the degree of relatedness of different isolates (10).

Salmonella Typhimurium is the most commonly reported serotype in Australia, accounting for over one-third of all gastro-intestinal outbreaks (11) and is frequently associated with raw or runny eggs (11-14). For example, *S*.Typhimurium associated with raw and runny eggs was responsible for close to half of all outbreaks in Australia with a food source identified during 2010 (11).

On the 3rd December 2014, Communicable Disease Prevention and Control at the Victorian Government Department of Health was notified by a doctor that a patient had been diagnosed with a *Salmonella* Typhimurium phage type 44 infection (*S.* Typhimurium 44) following attendance at a school function and that several other guests were unwell. An outbreak investigation was initiated to characterise the outbreak by person, time and place; identify the most likely source of the infection; and guide public health intervention to prevent further cases.

Methods

A retrospective cohort study was conducted to characterise the outbreak and identify the source of illness. The cohort included all guests who attended the function on Friday 22nd November 2014. The school administration provided a list of all attendees and their contact details. School staff members who had catered for the function provided details of all foods served. A structured outbreak questionnaire was developed and included demographics, details of illness both before and after the event, attendance at other school functions in the week prior to the function, and menu items consumed at the function. Interviews were conducted over the telephone by a single interviewer. As this was a public health investigation under the Victorian Public Health and Wellbeing Act (2008), ethics approval was not required.

The case definition for the outbreak included any person who consumed food at the function and experienced diarrhoea within 48 hours of the function. Cases were invited to provide faecal samples within 14 days of the last day of symptoms if they had not already provided a specimen to a general practitioner.

Questionnaire data were entered into Victoria's secure Public Health Event Surveillance System (PHESS) database. Statistical analysis was conducted using STATA 13 (Statacorp™). Descriptive statistics characterised the illness and demographics of cases. Univariate analysis calculated crude relative risk and 95% confidence intervals for illness associated with each menu item. Mantel-Haenszel

stratification was used to examine for evidence of confounding. We considered results statistically significant if they were $p <= 0.05$.

Environmental investigation was not undertaken because there were no food samples remaining for microbiological analysis and the suspected food vehicles were prepared in a home kitchen rather than a licensed food premises.

Faecal samples requested by general practitioners were submitted to private laboratories in Victoria. These laboratories forwarded isolates of *Salmonella* to the Microbiological Diagnostic Unit (MDU) for typing, including multilocus variable-number tandem-repeat analysis (MLVA). MLVA patterns for the outbreak isolates were compared against other notified cases within the preceding 12 months to identify other cases with indistinguishable isolates.

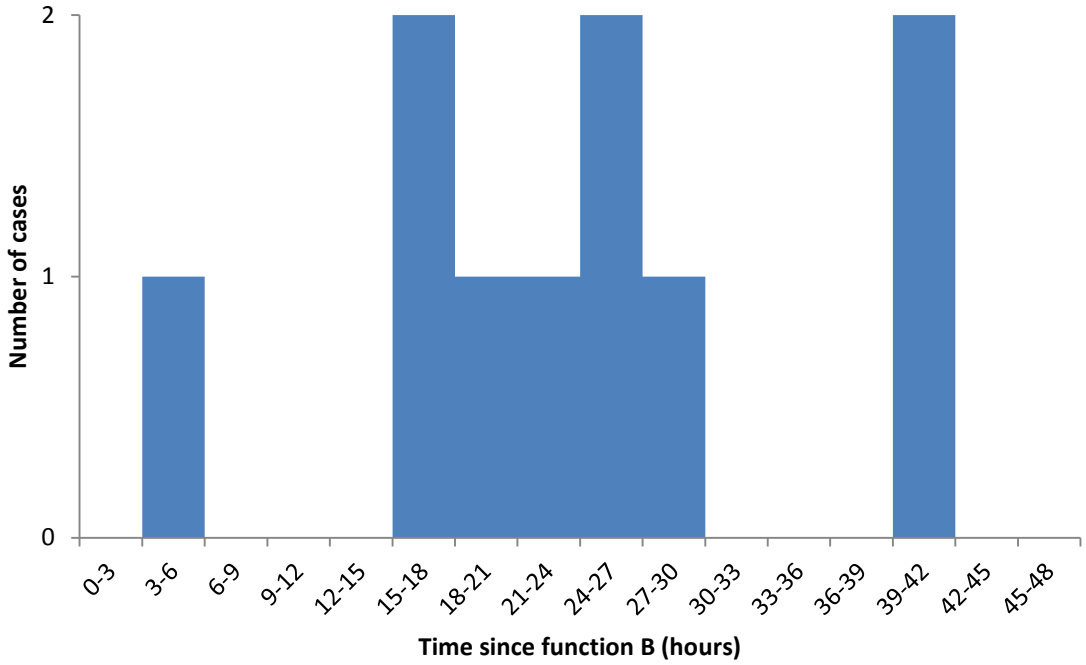
Results

All food served at the function was prepared at home by staff either the day before or on the morning of the event and transported to the school on the day of the lunch, apart from a pre-prepared quiche which was purchased frozen from a local supplier.

Twenty-nine of thirty attendees were interviewed, a response rate of 96%. All attendees at the function were staff at the school. Ten cases were identified including 7 males and 3 females, an overall attack rate of 34%. Median age of cases was 42 (range 28 to 58), slightly younger than the median age of all guests which was 47 (range 25 to 64). Cases were interviewed a median of 24.5 days after the event (range 10 to 27), compared to a median of 25 for non-cases (range 19 to 30).

The median incubation period was 25 hours (range 3 to 41), while median duration of illness was 7 days (range 1-8 days). An epidemic curve is presented in Figure 1. All cases experienced watery diarrhoea, and most (90%) also reported fever, nausea, abdominal pain and headaches. Only three cases (30%) reported vomiting. No blood in stools was reported. Half of all cases saw a doctor; one presented to an emergency department but was not hospitalised.

Figure 1. Epidemic curve showing time to symptom onset following attendance at Function B



The results of the cohort investigation are summarised in Table 1. Amongst those interviewed, 30% attended an alternative event (Function A) at the school on the evening prior to Function B. There was no association between illness and attendance at Function A.

Regarding food consumed at Function B, both roast beef appetiser and frittata had statistically significant crude associations with illness. The attack rates were 64% amongst those who ate roast beef appetiser and 53% of all those who ate frittata. The relative risk of illness was elevated for those who consumed roast beef appetiser (relative risk 5.09, 95% confidence interval 1.29 to 20.02) and for those who ate the frittata (relative risk 3.20, 95% confidence interval 0.83 to 12.35). Both were statistically significant associations ($p<0.05$). The two dishes were associated with each other; 73% of guests who ate roast beef appetiser also ate the frittata, and 61% of those who ate the frittata also ate roast beef appetiser. After adjusting for confounding between these two items the combined risk ratio for each item remained elevated, but only the appetiser (stratified by frittata) approached statistical significance (95% confidence interval 0.81 to 15.36 and $p=0.057$).

Table 1. Attack rates and risk ratios for menu items and salmonellosis in attendees of Library Lunch (n=29)

	Exposed	Not Exposed	Crude Analysis			Mantel-Haenszel Adjusted Analysis		
	Ill/Total (%)	Ill/Total (%)	Risk Ratio	95% CI	P Value*	Risk Ratio	95% CI	P Value**
Function A	3/7 (43%)	7/22 (32%)	1.35	(0.47, 3.86)	0.593			
Function B								
Roast beef Appetiser (n=27)	7/11 (64%)	2/16 (13%)	5.09	(1.29, 20.02)	0.006	3.53[a]	(0.81, 15.36)	0.057
Salmon pie	5/12 (42%)	5/17 (29%)	1.42	(0.52, 3.83)	0.494			
Sour Cream Sauce	1/5 (20%)	9/24 (38%)	0.53	(0.08, 3.32)	0.454			
Quiche (n=27)	8/23 (35%)	2/4 (50%)	0.70	(0.23, 2.15)	0.561			
Frittata (n=27)	8/15 (53%)	2/12 (17%)	3.20	(0.83, 12.35)	0.049	2.1[b]	(0.53, 8.38)	0.234
Salad	8/35 (23%)	2/6 (33%)	1.04	(0.30, 3.69)	0.947			
Bread (n=27)	6/13 (46%)	2/14 (14%)	3.23	(0.79, 13.24)	0.070			
Cake	5/16 (31%)	3/13 (38%)	0.81	(0.30, 2.21)	0.684			
Juice (n=27)	7/15 (47%)	3/13 (25%)	1.87	(0.61, 5.72)	0.247			
Water (n=28)	3/6 (50%)	7/22 (32%)	1.57	(0.57, 4.30)	0.410			
Tea	0/2 (0%)	10/27 (37%)	-	-	-			
Coffee	1/4 (25%)	9/25 (36%)	0.69	(0.12, 4.10)	0.667			
Milk (n=28)	1/4 (25%)	9/24 (38%)	0.67	(0.11, 3.93)	0.629			

[a.] adjusted for frittata; [b.] adjusted for roast beef appetiser

* P from chi square analysis; ** P value from Fisher's exact

The frittata contained eggs, cream, potatoes, olive oil, red onion, baby spinach, cherry tomatoes, red capsicum, and grated cheese. It was baked in a domestic kitchen the night preceding the event, refrigerated overnight and transported to school in an insulated bag with frozen bricks an hour before the event. Guests reported the frittata was firm, with no apparent undercooking. The roast beef appetiser consisted of a French stick sliced and baked on the same evening in the same kitchen as the frittata was prepared. It was topped with pureed tinned cannellini beans, lemon juice and rind, olive oil, roast beef and semi dried tomato pesto on the morning of the function. Both dishes were served cold.

Three cases submitted a faecal sample to a general practitioner. Two were positive for *Salmonella* Typhimurium phage type 44 with indistinguishabble MLVA patterns. No pathogens were isolated from the remaining faecal specimen. A review of MLVA patterns found three additional notified cases in Victoria with indistinguishable MLVA patterns and onset dates within 30 days of the current outbreak. These cases were from an unrelated family outbreak and declined interview. There were no other cases with identical MLVA pattern during the preceding 12 months.

Discussion

Outbreak investigations provide essential insight into the causation of *Salmonella* infections, because the large number of sporadic cases mean it is not feasible to investigate the source of individual infections (15). Previously published outbreaks of *Salmonella* Typhimurium phage type 44 in Australia have been associated with roast pork and apple sauce (16), eggs (14) and aioli (17).

The investigation found a statistically significant crude association between *S.* Typhimurium 44 infection and two foods, a frittata and a roast beef appetiser which were frequently eaten together. Neither association retained statistical significance following stratification, probably a consequence of insufficient numbers in this small cohort. However, the adjusted risk ratio for roast beef appetiser approached significance (p=0.057) suggesting it was the most likely food vehicle for the outbreak.

Salmonella Typhimurium is frequently associated with raw or undercooked eggs (10-15, 17, 18) and kitchen utensils contaminated with raw egg can spread *Salmonella* (3). As both the frittata and roast beef appetiser were prepared in the same domestic kitchen at the same time it is plausible raw egg contamination of preparation surfaces

and implements (eg cutting boards, knives) occurred during the preparation of the frittata, with subsequent cross-contamination to the roast beef appetiser.

Salmonella have been found on 3.5% of commercially available eggs in Australia (18). Incidence and shedding of *Salmonella* in laying hens is associated with a complex range of factors including husbandry practices, flock size and age, bird stress, weather, transport, initiation of egg lay and moulting (19). *S.*Typhimurium gain entry into eggs horizontally, that is by penetrating the shell after ovipositioning, or less commonly directly from infected ovaries during production of the egg (20). Alternatively, bacteria present on the external surface of the shell can lead to contamination of the egg proper when the egg is opened or cross contamination of other foodstuffs. Commercial egg washing is conducted in Australia on a large scale as a means to reduce bacterial contamination on the surface of eggs, but must be undertaken with care to avoid damaging the egg shell, as this increases the risk of bacteria entering the egg (21). A recent Australian study has shown the rate of horizontal *Salmonella* Typhimurium penetration was higher in washed eggs compared with non-washed eggs (20). Ultimately, appropriate food storage, hygienic food handling and thorough cooking are required to ensure any potential pathogens associated with eggs are destroyed prior to consumption.

Limitations

This study is subject to a number of potential limitations due to chance, bias and confounding. The investigation could have been strengthened by the inclusion of environmental and microbiological evidence to corroborate the epidemiological findings.

Chance refers to the likelihood that the results generated by the study have occurred coincidentally rather than due to a true association at the population level. It is a function of sample size in epidemiological studies, with larger sample sizes reducing the likelihood that results are due to chance. In this study, many participants ate both foods crudely associated with illness, raising the possibility that one of the associations may be a result of confounding. A stratified analysis was performed to address this. However, due to the smaller sample size within each strata, the precision of the adjusted estimates was reduced. The adjusted risk ratio for roast beef appetiser approached but did not reach statistical significance. A larger sample size would have

produced more precise estimates and increased the confidence that this adjusted association was more than simply a chance finding.

Bias is a systematic error in study design or execution. Like all retrospective cohort studies the results of this study may be affected by recall bias. This phenomenon occurs when there is a preconceived idea about the likely source of illness with the result that cases are more likely to recall certain exposures than controls. Attempts were made to minimise recall bias through the use of a structured menu-based questionnaire. Specific and targeted questions were designed to improve recall amongst all guests. Nonetheless it is possible recall bias has some influence on the findings, particularly if guests had formed their own theories regarding the likely food source prior to interview. Measurement bias can also be introduced by the interviewer if questions are asked differently of cases and controls. The use of a structured questionnaire and trained interviewer helped minimise this possibility. The study had an excellent response rate of 96%, which ensures selection bias will have very minimal, if any, effect on the measured association.

Confounding occurs when the association between two variables is altered due to the effect of a third variable that is associated with both the exposure and the outcome but not on the causal pathway. It is likely that only roast beef appetiser was truly associated with illness while the crude association with frittata arose due to confounding, as both items were frequently eaten together. Stratification addressed this limitation in the statistical analysis. Stratification considers the association between exposure and outcome separately in two strata: those exposed to the third variable, and those not exposed to the third variable. This approach removes the effect of that third variable from the measured association, and thus controls for its confounding effect. Mantel-Haenszel stratification was used to calculate the association between illness and frittata for those who consumed roast beef appetiser and for those who did not, and likewise to calculate the association between illness and roast beef appetiser separately for those who consumed the frittata and for those who did not. The stratified risk ratios were then weighted to reflect the uncertainty associated with the different samples size in each strata, and pooled to produce an overall estimate of the risk ratio. After adjusting for roast beef appetiser the association between illness and frittata did not persist, indicating the crude association was due to confounding.

This investigation faced limitations due to a lack of environmental and microbiological evidence. An environmental investigation was not possible as no food samples were

left over for sampling and the foods were prepared in a private kitchen. Insights gained through environmental investigation such as unsafe food handling practices or evidence of inadequate temperature controls can supplement and verify epidemiological evidence, and assist with identifying the source of an outbreak. Microbiological investigation has the potential to strengthen evidence of causation by isolating pathogens from food samples or preparation surfaces. However, in practice *Salmonella* is isolated from food in less than half of all *Salmonella* outbreak investigations (15). If food samples had been available, and if *Salmonella* had been isolated, this may have helped confirm epidemiologic findings. More importantly, isolating *Salmonella* from an ingredient in one of these foods could have provided opportunity for further trace-back and public health intervention, with the potential to prevent further cases of disease.

Conclusion

This outbreak was one of approximately fifty *Salmonella* outbreaks occurring in Australia each year (11, 22). Outbreaks provide an excellent opportunity for public health investigation to generate evidence regarding epidemiology and causation of food borne illness, which can have important implications for public health policy. In this outbreak a roast beef appetiser was epidemiologically associated with *S.* Typhimurium infection. It was not possible to identify the specific source for this outbreak, but cross contamination from raw eggs is a likely explanation for the findings. More timely public health follow-up including analysis of environmental and microbiological evidence may have assisted in identifying a food source, but left over food samples were not available.

References

1. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. Emerging infectious diseases. 1999;5(5):607-25.

2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal Salmonella gastroenteritis. Clin Infect Dis. 2010;50(6):882-9.

3. Centres for Disease Control and Prevention. Salmonella serotype Enteritidis Atlanta2010. Available from: http://www.cdc.gov/nczved/divisions/dfbmd/diseases/salmonella_enteritidis/#keys.

4. Voetsch AC, Van Gilder TJ, Angulo FJ, Farley MM, Shallow S, Marcus R, et al. FoodNet estimate of the burden of illness caused by nontyphoidal Salmonella infections in the United States. Clin Infect Dis. 2004;38 Suppl 3:S127-34.

5. Number of notifications for all diseases by year, Australia, 1991 to 2014 and year-to-date notifications for 2015 [Internet]. Australian Government Department of Health and Aging. 2015 [cited 20 February 2015].

6. Hall G, Yohannes K, Raupach J, Becker N, Kirk M. Estimating community incidence of Salmonella, Campylobacter, and Shiga toxin-producing Escherichia coli infections, Australia. Emerging infectious diseases. 2008;14(10):1601-9.

7. Kirk M. Foodborne disease surveillance needs in Australia: Harmonisation of molecular laboratory testing and sharing data from human, animal, and food sources. New South Wales public health bulletin. 2004;15(2):13-7.

8. Lightfoot D, Veitch M. Naming Salmonellae. Victorian Infectious Disease Bulletin. 1999;2(2):27-8.

9. Wang Q, Chiew R, Howard P, Gilbert GL. Salmonella typing in New South Wales: current methods and application of improved epidemiological tools. New South Wales public health bulletin. 2008;19(1-2):24-8.

10. Stephens N, Sault C, Firestone SM, Lightfoot D, Bell C. Large outbreaks of Salmonella Typhimurium phage type 135 infections associated with the consumption of products containing raw egg in Tasmania. Communicable diseases intelligence quarterly report. 2007;31(1):118-24.

11. OzFoodNet Working Group. Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: annual report of the OzFoodNet network, 2010. Communicable diseases intelligence quarterly report. 2012;36(3):E213-41.

12. Gregory J. A Case Control Study of *Salmonella* Typhimurium 44 in Victoria 2007: The University of Melbourne; 2007.

13. Tay E. Cluster investigation of Salmonella Typhimurium 44 in Victoria, 1 December 2012 to 30 June 2013 (Thesis chapter 2). Canberra: Australian National University; 2013.

14. Dyda A, Hundy R, Moffatt CR, Cameron S. Outbreak of Salmonella Typhimurium 44 related to egg consumption. Communicable diseases intelligence quarterly report. 2009;33(4):414-8.

15. Moffatt CR, Musto J. Salmonella and egg-related outbreaks. Microbiol Aust. 2013;34(2):94-7.

16. Tribe IG, Walker J. An outbreak of Salmonella typhimurium phage type 44 linked to a restaurant in South Australia. Communicable diseases intelligence. 2000;24(11):347.

17. Denehy EJ, Raupach JC, Cameron SA, Lokuge KM, Koehler AP. Outbreak of Salmonella typhimurium phage type 44 infection among attendees of a wedding reception, April 2009. Communicable diseases intelligence quarterly report. 2011;35(2):192-6.

18. Fearnley E, Raupach J, Lagala F, Cameron S. Salmonella in chicken meat, eggs and humans; Adelaide, South Australia, 2008. International journal of food microbiology. 2011;146(3):219-27.

19. Whiley H, Ross K. Salmonella and eggs: from production to plate. International journal of environmental research and public health. 2015;12(3):2543-56.

20. Gole VC, Chousalkar KK, Roberts JR, Sexton M, May D, Tan J, et al. Effect of Egg Washing and Correlation between Eggshell characteristics and Egg Penetration by Various Salmonella Typhimurium Strains. PloS one. 2014;9(3-12).

21. European Food Safety Authority. Opinion of the scientific panel on biological hazards on the request from the commission related to the microbiological risks on washing of table eggs. EFSA Journal. 2005;269(39).

22. OzFoodNet Working Group. Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: annual report of the OzFoodNet network, 2011. Communicable diseases intelligence. 2015;39(2):E236-64.

Chapter 4 Appendices

Appendix 1. Conference abstract

Abstract Title: *Salmonella* Typhimurium phage type 44: A Victorian outbreak and review of MLVA patterns

Presenter: Zoe Cutcher. Australian National University, ACT, and Department of Health and Human Services, Victoria.

Authors: Cutcher, Z. Gregory, J. Valcanis, M. Mercoulia, K. Kirk, M. Stephens, N. and Easton, M.

Background: In December 2014, a *Salmonella* Typhimurium phage type 44 (STm44) outbreak occurred following a function in Victoria. We investigated the outbreak to determine a cause and compared Multi Locus Variable-number Tandem Repeat Analysis (MLVA) patterns to previous cases.

Methods: We conducted a cohort study using a menu based questionnaire and calculated relative risks for all food items. We compared MLVA patterns for the outbreak strain against other Victorian STm44 cases and reviewed outbreak investigations from 2009-2014 to examine potential sources.

Results: There were 10 cases among 29 guests interviewed. Risk of illness increased with consumption of the appetiser and frittata. Cross contamination from eggs was suspected. The outbreak strain was indistinguishable from 1.7% (7/ 392) of MLVA patterns since 2009. A predominant historical pattern accounted for 45% of all patterns; another 51% were closely related including the outbreak strain. There were 5 historical STm44 outbreaks (78 cases) and 1 cluster (102 cases); all were related to the predominant MLVA pattern. Previous investigations all implicated or suspected eggs as the source.

Conclusion: We were unable to identify a specific source for this outbreak, but cross contamination from eggs appears likely. MLVA provided limited differentiation between STm44 isolates.

Appendix 2. Conference Presentation: Brisbane Communicable Disease Control Conference June 2015

Salmonella Typhimurium phage type 44: A Victorian outbreak and review of MLVA patterns

Cutcher Z[1,2], Gregory J[1], Valcanis M[3], Mercoulia K[3], Kirk M[2], Stephens N[1], and Easton M[1].

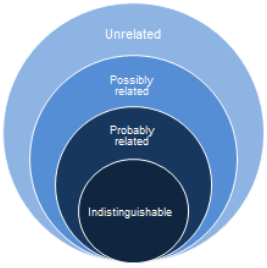
1. Victorian Government Department of Health and Human Services
2. Australian National University
3. The University of Melbourne

Department of Health & Human Services

Background & Methods

Salmonellosis

- Typing critical to identify outbreaks
- Cohort study on *Salmonella* Typhimurium 44 outbreak, 2014
- Reviewed MLVA patterns to identify related strains 2009-2014

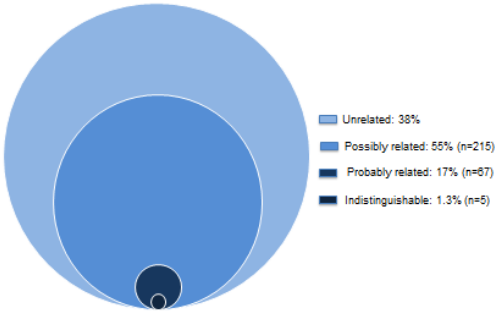


Results

- Interviewed 29/30 (96%) guests
- Attack rate 10/29 (34%)
- 2 faecal samples positive for STM44
- Roast beef appetiser: RR 3.53 (0.81, 15.36; P 0.057)
  - Self catered - domestic kitchen
  - Possible cross contamination from egg frittata

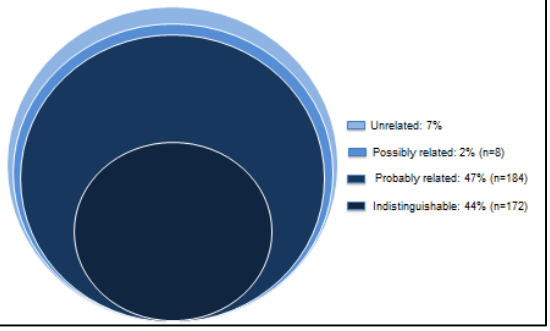
Outbreak versus historical isolates 2009-2014 (n=392)

Outbreak MLVA: 03-10-09-09-523



Predominant versus all isolates 2009-2014 (n=392)

Predominant MLVA 03-10-08-10-523



Conclusion

- Roast beef appetiser most likely vehicle
- Possible cross-contamination from eggs
- MLVA
  - some differentiation for outbreak isolate pattern
  - limited differentiation for predominant historical pattern
- Further work is required to investigate heterogeneity of MLVA for different Salmonella phagetypes

Chapter 5. Antimony testing in Victorian residents living near a mine

Contents

ꕥ

Due to it's sensitive nature, the content in this chapter has been restricted from publication at the request of the Victorian Government Department of Health and Human Services

ꕥ

This page is intentionally blank

Appendices to the Thesis

This page is intentionally blank

Appendix A. Research Protocol: Melbourne Hot Water System (HoWS) Survey

Research Protocol: Melbourne Hot Water System (HoWS) Survey

Study Team:

Zoe Cutcher, Stuart Adcock, Noel Cleaves, Lucinda Franklin, Nicola Stephens with the assistance of Loretta Vaughan, VPHS.

Research Questions

“What is the prevalence of Legionella bacteria colonisation in residential hot water systems in Metropolitan Melbourne? How does the measured prevalence differ between culture and PCR detection techniques?”

Study Aims and Objectives

The aim of this study is to act as a pilot study for a planned case-control study assessing the association between *Legionella* contamination in residential potable water and the occurrence of Legionnaires’ disease in Victoria.

The study objectives are to:

1. Determine the prevalence of *Legionella* contamination in residential potable water sources in metropolitan Melbourne, to inform sample size calculations for the above study.

2. Determine the risk factors for *Legionella* contamination in residential potable water sources in metropolitan Melbourne.

3. Develop capacity for PCR detection of *Legionella* in environmental samples at Victoria’s State Reference Laboratory for Legionella.

4. Compare the sensitivity of Legionella detection in environmental samples using the Standard culture method, an ultrafiltered culture method and PCR.

5. Refine sampling strategies, laboratory methods and participant resources for the planned case-control study.

Background

Legionella bacteria are ubiquitous organisms that, under certain conditions, cause severe pneumonia known as Legionnaires’ disease (LD). *Legionella* thrives in wet environments and can colonize cooling towers, fountains, spas and plumbing

systems(1). Transmission occurs through aspiration or inhalation of the bacteria. LD mainly occurs when those exposed have predisposing factors such as immune-compromise, lung disease, male sex or smoking (2). It accounts for 3 – 15% of community-acquired pneumonia (3-5).

The public health response to LD is focused on preventing outbreaks by identifying contaminated cooling towers. However approximately 80% of notified LD in Victoria are sporadic cases, the source of which is largely unknown. Contamination of potable water is a likely source for at least some of these cases (6), and some argue potable water may cause more cases of disease than cooling towers (7).

Numerous case reports (8-15), and case series (6, 16) document cases of sporadic LD where *Legionella* bacteria from clinical isolates matched *Legionella* from potable water in patient's homes. Outbreaks in apartment blocks (17), hotels (18) and suburbs (19) have been caused by contaminated potable water. Despite this the magnitude of the association between *Legionella* contaminated potable water and the occurrence of LD is not known. An outbreak in an apartment building attributed to potable water contamination had an attack rate of 4.7% (17); however as this was an outbreak setting it cannot be generalized to estimate the overall risk associated with contamination of residential potable water. Previous case-control studies (20, 21) failed to document a significant association between *Legionella* contamination in residential potable water and LD but these studies were flawed due to inadequate sample size calculation (20, 21), poor timeliness for sample collection (20) and poor control participation (21). Well-designed epidemiological studies including sample size estimation are needed to fill this knowledge gap.

Sample size calculations require knowledge of the prevalence of *Legionella* contamination in the study setting. International estimates of *Legionella* prevalence in residential water range between 0-30% by culture (20, 22-27) and up to 50% by PCR (27, 28). Prevalence varies geographically and is influenced by numerous factors including sampling and laboratory methods, water temperature (22, 23, 26), hot water system type (20, 26), and water composition (20, 23, 25). To our knowledge the prevalence of *Legionella* in residential potable water in Australia has not been reported.

Legionella's ability to adopt a "viable but non-culturable" state limits the capacity to detect the bacteria with culture alone. PCR has been proposed as a tool to overcome this; however the correlation between culture and PCR is still unclear (29). Further

research is required to investigate the usefulness of PCR for routine environmental testing.

Rationale

The study is intended as a pilot study to refine sampling and laboratory methods for a future case-control study investigating the association between water-saving showerheads and LD in Melbourne, Victoria. Findings will inform sample size calculations for the case-control study to increase the likelihood of producing statistically significant results.

Lay Summary

This project aims to investigate how commonly a type of bacteria called *Legionella* are found in household hot water systems in Melbourne. It is a pilot study for a future case-control study to find out whether there is any association between Legionella contamination in household hot water and the occurrence of Legionnaires' disease. The pilot study will allow us to refine sampling methods. It will also provide information needed for a sample size calculation that will ensure the future case control study mentioned above is large enough to be meaningful. An additional aim is to explore new laboratory techniques for the detection of *Legionella* in environmental samples in Victoria.

The study will collect hot water samples from around 80 homes in Melbourne and analyse them for the presence of *Legionella* bacteria. Participants will be owner-occupiers recruited from the Victorian Population Health Survey Control Bank. This is a database of people who have previously participated in a health survey conducted by the Victorian Government Department of Health and agreed to be invited to participate in other departmental research studies.

Hot water samples will be collected from up to three showers in each household, and a swab collected from the opening of a tap in the bathroom. Additional factors influencing Legionella colonisation will be recorded including type of hot water system, temperature of the hot water system, and type of shower fittings.

Microbiologic analysis will be undertaken at Victoria's State Reference Laboratory for *Legionella* using the current Australian Standard unfiltered culture method (currently the default testing method for environmental Legionella testing in Victoria); an enhanced version of this method that uses filtration to concentrate the sample before

culture; and DNA detection technology (PCR). Results achieved by each method will be compared and contrasted to explore the correlation between the newer technology and the unfiltered Australian Standard method.

Methods

Study Design

A cross sectional study design will be employed. Water samples will be collected over a 10-month period from showers in private residences located within metropolitan Melbourne.

Study Parameters

Study parameters include study factors and outcome factors.

Outcome factors are laboratory measures of Legionella colonisation in domestic water systems:

1. Positive culture of *Legionella* from domestic shower water samples using Victoria's current standard method, which has a lower detection threshold of 10,000 cfu/L.

2. Positive culture of *Legionella* from shower water samples using an enhanced culture method, which has a lower detection threshold of 100 cfu/L.

3. Detection of *Legionella* from domestic water shower samples using PCR.

4. Detection of *Legionella* from direct swabs of bathroom taps using standard culture.

Study factors are factors which are known to predispose water systems to *Legionella* growth:

1. Type of hot water system (electric storage, gas storage, instantaneous gas, solar)

2. Storage capacity of hot water system (for storage systems)

3. Approximate age of hot water system (based on make and model)

4. Approximate age of dwelling (estimated by resident)

5. History of plumbing work within the preceding 4 weeks
6. Hot water temperature at shower
7. Storage temperature of hot water (storage systems only)
8. Chlorine and pH levels
9. Concentration of iron, copper and zinc
10. Apartment vs freestanding dwelling
11. Type of showerhead (water-saving or other)
12. Type of shower fitting (fixed or flexible pipe)

Sample and collection procedures

Sample Size

A sample size calculation has been performed under the assumption the true prevalence of *Legionella* in domestic potable water in Melbourne is 5%. Under this assumption samples from 73 residences will be sufficient to estimate the prevalence to within a 5% margin of error and a confidence level of 95%. If the true prevalence is 20% the required sample size increases to 246.

Due to practical and resource limitations, the maximum number of water samples the laboratory is able to process for this study is set at 200. Thus, the study will recruit approximately 80 participants to give a total of 200 water samples.

Study population

Participants will be recruited from the Victorian Population Health and Wellbeing control bank. This is a register of participants from the Victorian Health and Wellbeing telephone survey who agreed to be invited to participate in additional health-related research. Participants of the Victorian Population Health Survey (VPHS) are sourced from the Victorian resident population using an extensive random selection process based on random digit dialing. The sample size for the VPHS survey is 34,000 with the majority of participants agreeing to be contacted again for health related research. Until recently participants were recruited exclusively through landline telephones but the survey is currently trialing recruitment using mobiles to minimize any potential bias

associated with access to landlines. The control bank participants are all over 18 years of age.

Recruitment

A list of telephone numbers for recently recruited control bank participants residing in metropolitan Melbourne will be sourced from the Victorian Population Health and Wellbeing Survey control bank. Initial recruitment will be conducted via telephone; five attempts will be made to contact each participant using a recruitment script. A Participant Information Form will be mailed to interested potential participants along with an introductory letter. Approximately one week later potential participants will telephoned to confirm participation and book in sampling. Contact will be maintained between recruitment and sampling with the aid of an additional letter. Written informed consent for participation will be sought at the sampling visit, prior to proceeding with sample collection.

Eligibility Criteria: Participants will be eligible to participate if they are over 18 years of age, reside in a private residence in Melbourne, own their own homes and provide written informed consent prior to sampling.

Sample collection

Written informed consent will be sought prior to proceeding with sampling. At this time a short survey will also be conducted to determine the secondary study factors. Age of premises and history of plumbing work will be determined through interview. Type of hot water system will be determined by examining the system (where possible) or through discussion with the resident. The approximate age of the hot water system will be estimated by recording the make and model in order to later investigate manufacture dates.

A direct swab will be collected from inside the bathroom sink tap, by inserting a sterile cotton tip inside the opening and gently agitating to remove organic matter. This will be placed in a sterile specimen container with 10ml of water from the same tap. Shower water samples will be collected into sterile containers pre-treated with thiosulphate to de-activate chlorine, as it inhibits bacterial growth. Two and a quarter litres of hot water will be collected from each shower in the house immediately after turning on the hot tap. This “first flush” sample is designed to collect bacteria colonising the shower head

and proximal pipes. Mineral, chlorine and pH levels will be assessed using an onsite detection kit on a separate 50ml sample collected from the most commonly used shower immediately after collection of the primary sample. Once all showers have been sampled the hot water in the most frequently used shower will be allowed to run continuously until a maximum temperature is achieved (as measured by a laser thermometer), and an additional two and a quarter litre sample collected. This "second flush" sample is designed to collect bacteria colonizing the hot water system and distal pipes. The temperature of the second flush sample will be measured and recorded using an infrared thermometer. Many residences have a tempering valve that lowers the temperature of hot water prior to delivery at the showerhead. Therefore, for homes with storage hot water systems, the storage temperature of hot water system will also be measured by assessing water temperature at the outlet tap.

All staff will be provided with heat proof gloves and training to minimise risk of scalding during sample collection. All samples will be transported in eskys immersed in ice and delivered to the lab within 12 hours of collection.

Laboratory analysis

Laboratory analysis will be performed by the Microbial Diagnostic Unit (MDU, Victoria's State Reference Laboratory for environmental detection of *Legionella*) at the University of Melbourne.

Three approaches will be used to detect and quantify *Legionella*.

The first approach is the current standard in Victoria which follows the Australian Standard AS/NZS 3896:2008. The approach utilizes samples of 100ml and has a detection threshold of 10,000 cfu/L. The method includes direct inoculation of water samples onto *Legionella* selective medium (both MWY and BMPA plates) in addition to preparation of heat and acid treated samples. The heat treated sample is inoculated on a separate MWY plate following heat treatment at 50^0 for 30 minutes; the acid treated sample is inoculated onto a BMPA plate after mixing with an HCl/KCL buffer.

The second approach is an ultrafiltered culture technique. This requires samples of 1 litre and has a detection threshold of 100 cfu/L. Samples will be concentrated using membrane filtration and re-suspended in 10ml for optimum coverage, before proceeding with the culture method outlined above.

The third approach will be a real-time polymerase chain reaction (PCR) to detect DNA fragments using commercial primers. Organisms in a one litre sample will be concentrated prior to analysis using membrane filtration. The filter membrane will then be processed for PCR. The method will enable estimation of the number of organisms present and include both detection of *Legionella* species, and a specific detection kit for *Legionella pneumophila serogroup 1.*

Data storage and reporting

Sample results and questionnaire data will be stored in the PHESS database. Data will not be used for any purpose apart from that for which it was collected. Access to the data will be restricted to departmental staff directly involved in the study.

Results of all Legionella testing using Standard methods will be reported to all participants by mail. Where Legionella is detected by culture, participants will be notified of results over the telephone with the guidance of a results notification script and a factsheet will be provided. Professional advice will be provided to help participants interpret the results of testing and decide whether remedial action is required. In most cases remedial action will only be required if residents of the household are predisposed to Legionella. A study doctor will be available to consult with participants to help determine their risk of illness.

At the request of the Laboratory results obtained through PCR will not be reported to participants, as these methods are not currently validated in Victoria and lack the support of an Australian Standard or National Association of Testing Authorities (NATA) accreditation.

Data Analysis

Data will be analysed using Stata™ version 13 statistical software (Statacorp, Texas). Analysis will include:

Descriptive analysis of *Legionella* prevalence as identified by the various detection methods

- according to species
  - *Legionella pneumophila serogroup1* versus
  - *Legionella* species
- according to site
  - first flush versus

- second flush

Chi squared comparisons

- Detection by standard culture, enhanced culture and PCR

Sensitivity of each culture method, comparing against PCR as the gold standard

Correlation between concentration of *Legionella* as detected by enhanced culture and that estimated by PCR

Univariate analysis of association between risk factors and *Legionella* detection

Multivariate analysis of association between risk factors and *Legionella* detection for those factors found to be significantly associated on univariate analysis

Ethics

Benefits to the community

The research will benefit the wider community by beginning to investigate to what extent Legionella contamination of domestic potable water contributes to the burden of disease due to Legionella in Victoria. The pilot study will inform a future case-control study to investigate the association between contamination of domestic potable water and the occurrence of Legionnares' disease. Once this issue is understood targeted educational and other interventions could be designed to reduce the burden of disease due to residential exposure for high risk groups.

In addition, this study will provide an opportunity to develop and explore PCR detection of Legionella from environmental samples at Victoria's State Reference Laboratory for Legionella. This method will improve the sensitivity of detection and may be helpful for investigating future outbreaks of disease in situations where the current Standard method is not able to identify a source.

Benefits to participants

Legionella test results obtained using the Australian Standard culture method will be provided to all participants unless they elect not to receive them. Where Legionella is detected by culture participants will be informed of results by telephone. For participants who are at risk of disease, identifying Legionella in their residential potable water provides them the opportunity to take steps to mitigate their risk (eg. arrange

treatment by a plumber, or modify their behaviour e.g. switch to baths instead of showers). Participants with Legionella contaminated samples who elect to pursue remedial action will be offered free follow up testing to ensure the action has been effective.

Informed consent

The Informed consent process is described under recruitment.

Strategies to reduce risk to participants

There is a risk that participants may experience anxiety if Legionella is detected in their home. However in most cases the actual risk of illness will be extremely small, as Legionnaires' disease is exceptionally rare in people who are not predisposed to disease. Participants will be informed of this possibility in the participant information form (Appendix 3), which they will be required to read prior to providing written consent to participate (Appendix 5).

Where Legionella is detected in participant homes extensive support will be available to help participants make an informed decision regarding what action to take, if any. This will include telephone consultation with Environmental Health Officers (Appendix 7), provision of a detailed fact sheet (Appendix 6), a letter explaining what the results mean (Appendix 9) and the option to consult with a public health physician if required. If the public health physician is concerned about a participant’s mental health referral to a counselling service will be arranged.

For participants who are at risk of disease, identifying Legionella bacteria in their potable water system provides them with the opportunity to address the problem either by arranging for a plumber to treat the system or by modifying their activities (e.g., bathing instead of showering).

Budget and Resources

The project has been planned within existing Department of Health resources for staffing and laboratory analysis. In kind support includes 600 hours of staff time, comprising:

120 hours for study design, review and ethics submission

60 hours for participant resource development and piloting

160 hours for participant recruitment

60 hours for sampling

80 hours travel

80 hours for data entry

40 hours for reporting results

Dissemination Plan

Results will be disseminated internally in the Department through publication of a report and presentation of a seminar.

It is anticipated results of the survey will be disseminated through conference presentations and publication in a peer-reviewed journal. Target conferences include the Australian Communicable Disease Control Conference and the Australian Environmental Health Conference. Target journals are those with a local communicable disease focus, including the Medical Journal of Australia or Communicable Disease Intelligence.

Information about the findings will also be fed back to the public through media release. The report of the findings will be made available on the Legionella team website.

References

1. Broadbent. National Enivoronmental Health Forum Monographs Water Series No. 1: Guidance for the Control of Legionella. Glenelg: National Environmental Health Forum; 1996.

2. Carson P, Mumford C. Legionnaires' disease: causation, prevention and control. Loss Prevention Bulletin. 2010(216):20-9.

3. McNally C, Hackman B, Fields BS, Plouffe JF. Potential importance of Legionella species as etiologies in community acquired pneumonia (CAP). Diagnostic Microbiology and Infectious Disease. 2000;39:79-82.

4. Lieberman D, Porath A, Schlaeffer F, Lieberman D, Boldur I. Legionella species community-acquired pneumonia: a review of 56 hospitalized adult patients. CHEST. 1996;109(5):1243-9.

5. Vila-Corcoles A, Ochoa-Gondar O, Rodriguez-Blanco T, Raga-Luria X, Gomez-Bertomeu F, Group ES. Epidemiology of community-acquired pneumonia in older adults: a population-based study. Respiratory medicine. 2009;103(2):309-16.

6. Stout JEMS, Yu VLMD, Muraca PME, Joly JMD, Troup NBS, Tompkins LSMDP. Potable Water as a Cause of Sporadic Cases of Community-Acquired Legionnaires' Disease. The New England Journal of Medicine. 1992;326(3):151-5.

7. Yu VL, Stout JE. Editorial Commentary: Commumiy-Acquired Legionnaires Disease: Implications for Underdiagnosis and Laboratory Testing. 2008:1365.

8. Chen YS, Lin WR, Liu YC, Chang CL, Gan VL, Huang WK, et al. Residential water supply as a likely cause of community-acquired Legionnaires' disease in an immunocompromised host. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2002;21(10):706-9.

9. Luck PC, Schneider T, Wagner J, Walther I, Reif U, Weber S, et al. Community-acquired Legionnaires' disease caused by Legionella pneumophila serogroup 10 linked to the private home. J Med Microbiol. 2008;57(Pt 2):240-3.

10. Hood J, Hay J, Gillespie T, Seal DV. A fatal case of community-acquired Legionnaires' disease acquired from potable water. J Infect. 1994;29(2):231-3.

11. Leverstein-van Hall MA, Verbon A, Huisman MV, Kuijper EJ, Dankert J. Reinfection with Legionella pneumophila documented by pulsed-field gel electrophoresis. Clin Infect Dis. 1994;19(6):1147-9.

12. Turner D.P.J, Boswell T.C, Lee J.V, Slack R.C.B, Burden R.P. Community-acquired Legionnaires' disease in an immunocompromised patient masquerading as a hospital acquired-infection. J Hosp Infect. 2000;47:76-7.

13. Sax H, Dharan S, Pittet D. Legionnaires' disease in a renal transplant recipient: nosocomial or home-grown? Transplantation. 2002;74(6):890-2.

14. Skogberg K, Nuorti JP, Saxen H, Kusnetsov J, Mentula S, Fellman V, et al. A newborn with domestically acquired legionnaires disease confirmed by molecular typing. Clin Infect Dis. 2002;35(8):e82-5.

15. Castellani Pastoris M, Vigano EF, Passi C. A family cluster of Legionella pneumophila infections. Scandinavian journal of infectious diseases. 1988;20(5):489-93.

16. Stout JE, Yu VL, Muraca P. Legionnaires' disease acquired within the homes of two patients. Link to the home water supply. JAMA. 1987;257(9):1215-7.

17. Silk BJ, Foltz JL, Ngamsnga K, Brown E, Muñoz MG, Hampton LM, et al. Legionnaires' disease case-finding algorithm, attack rates, and risk factors during a residential outbreak among older adults: an environmental and cohort study. BMC Infectious Diseases. 2013;13(1):1-8.

18. Silk BJ, Moore MR, Bergtholdt M, Gorwitz RJ, Kozak NA, Tha MM, et al. Eight years of Legionnaires' disease transmission in travellers to a condominium complex in Las Vegas, Nevada. Epidemiology and infection. 2012;140(11):1993-2002.

19. Cohn PD, Gleason JA, Rudowski E, Tsai SM, Genese CA, Fagliano JA. Community outbreak of legionellosis and an environmental investigation into a community water system. Epidemiology and infection. 2014:1-10.

20. Straus WL, Plouffe JF, File TM, Jr., Lipman HB, Hackman BH, Salstrom SJ, et al. Risk factors for domestic acquisition of legionnaires disease. Ohio legionnaires Disease Group. Arch Intern Med. 1996;156(15):1685-92.

21. Codony F, Alvarez J, Oliva JM, Ciurana B, Company M, Camps N, et al. Factors promoting colonization by legionellae in residential water distribution systems: an environmental case-control survey. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2002;21(10):717-21.

22. Arnow PM, Weil D, Para MF. Prevalence and significance of Legionella pneumophila contamination of residential hot-tap water systems. The Journal of infectious diseases. 1985;152(1):145-51.

23. Stout JE, Yu VL, Yee YC, Vaccarello S, Diven W, Lee TC. Legionella pneumophila in residential water supplies: environmental surveillance with clinical assessment for Legionnaires' disease. Epidemiology and infection. 1992;109(1):49-57.

24. Dimitriadi D, Velonakis E. Detection of Legionella spp. from Domestic Water in the Prefecture of Arta, Greece. Journal of pathogens. 2014;2014:407385.

25. Leoni E, De Luca G, Legnani PP, Sacchetti R, Stampi S, Zanetti F. Legionella waterline colonization: detection of Legionella species in domestic, hotel and hospital hot water systems. Journal of applied microbiology. 2005;98(2):373-9.

26. Alary M, Joly JR. Risk factors for contamination of domestic hot water systems by legionellae. Appl Environ Microbiol. 1991;57(8):2360-7.

27. Bates MN, Maas E, Martin T, Harte D, Grubner M, Margolin T. Investigation of the prevalence of Legionella species in domestic hot water systems. The New Zealand medical journal. 2000;113(1111):218-20.

28. Donohue MJ, O'Connell K, Vesper SJ, Mistry JH, King D, Kostich M, et al. Widespread molecular detection of Legionella pneumophila Serogroup 1 in cold water taps across the United States. Environmental science & technology. 2014;48(6):3145-52.

29. Whiley H, Taylor M. Legionella detection by culture and qPCR: Comparing apples and oranges. Critical reviews in microbiology. 2014.

Department of Health

Incorporating: Health, Mental Health and Ageing

50 Lonsdale St
Melbourne
Victoria 3000
GPO Box 4541
Melbourne
Victoria 3001
Telephone: 1800 248 898
www.health.vic.gov.au

Our Ref: PHESS ID

Enter Date HERE
Enter Name HERE
Enter Address HERE
ENTER SUBURB STATE PCODE HERE

Dear Name,

Your household recently participated in the Melbourne HoWS Survey. This important survey will increase our understanding of community-acquired Legionnaires' disease and depends on the support of volunteers like you – thank you.

Below, please find results of the Australian Standard testing methods performed on samples from your house.

Sample Location	Culture Result
Main Bathroom Tap	*Not detected*
Main Shower	*Not detected*
Second shower	*Not detected*
Third shower	* delete entire row if no third shower *

We can report that we **did not detect** *Legionella* bacteria in any water samples collected from your residence using the Australian Standard testing methods for detection of *Legionella*.

If you would like to discuss your results with our study team, or have any questions about the Melbourne HoWS survey please contact us on **1800 248 898**.

Thank you for contributing to this important work which will not only help improve the health of Victorians, but could also help save lives – and we couldn't do it without you.

Yours sincerely

Noel Cleaves, Manager Environmental Health Regulation and Compliance

Department of Health

Incorporating: Health, Mental Health and Ageing

50 Lonsdale St
Melbourne
Victoria 3000
GPO Box 4541
Melbourne
Victoria 3001
Telephone: 1800 248 898
www.health.vic.gov.au

Our Ref: PHESS ID

Enter Date HERE
Enter Name HERE
Enter Address HERE
ENTER SUBURB STATE PCODE HERE

Dear Name

Your household recently participated in the Melbourne HoWS Survey. This important survey will increase our understanding of community-acquired Legionnaires' disease and depends on the support of volunteers like you – thank you.

Below, please find laboratory results listed for each sample from your house. To understand what your results mean please see the explanation below.

Sample Location	Result
Bathroom Tap	*Not detected*
Main Shower	*Not detected*
Second shower	***Legionella* detected at low levels**
Third shower	* delete entire row if no third shower *
Hot water storage temperature - if measured*	*xxx ºC (NB if not measured = "not measured")*

** nb hot water storage temperature was only measured for storage systems. In some cases it could not be measured due to difficulty accessing the system or other practical barriers.*

Legionella bacteria have been **detected** in samples collected from a tap/ a shower/ both a tap and a shower in your house using Australian Standard detection methods.

What does this mean for my household? *Legionella* bacteria are widespread in wet environments and are found in plumbing systems around the world. The absolute risk of illness due to these bacteria in domestic homes is not well understood, but in general we do know the overall risk of disease for individuals exposed to *Legionella* bacteria is extremely low. Although the overall risk is extremely low, for particular people the risk may be slightly higher. Please see the attached fact sheet on *Legionella* to help decide if you or anyone in your household could be at risk of illness due to *Legionella* bacteria.

**** Insert appropriate hot water system paragraph if hot water system STORAGE temperature available. Otherwise, delete this line ****

If you would like to discuss your results with our study team or if you have any questions about the Melbourne HoWS survey please contact us on **1800 248 898**. If you wish to speak with a public health physician regarding your risk of illness our study team would be happy to arrange this for you. Alternatively, you may wish to consult your local doctor.

Thank you for contributing to this important work which will not only improve Victorian's health, but help save Victorian's lives – we couldn't do it without you.

Yours sincerely,

Noel Cleaves,
Manager
Environmental Health Regulation and Compliance

Additional comments for insertion into results letter, as appropriate based on hot water system temperature for culture positive households:

HWS Type	HWS Temperature	Copy and paste the following:
Storage (Electric or Gas)	**Inadequate** (<60^0)	**Hot water system temperature:** Your hot water system temperature was xxx <60^0 xxx. The Plumbing Regulations (2008) requires hot water be stored at or above 60^{0}C to minimise the growth of *Legionella* bacteria in the system. This is because hot water storage temperature is one of the main factors affecting *Legionella* growth. We recommend you arrange a plumber to adjust the storage temperature on your system to above 60^{0}C. In most cases the increase in storage temperature will be sufficient to clear the system of *Legionella*. Please note if you increase the temperature of your storage hot water system a tempering valve should also be installed to minimise the risk of scalding (if not already present) – we recommend you talk to your plumber about this. A repeat sample should be taken a week later to ensure the adjustment has been effective. Please call us on **1800 248 898** if you would like to arrange a repeat sample which can be done free of charge.
Storage (Electric or Gas)	**Adequate** (=>60^0)	**Hot water system temperature:** Your hot water system temperature was xxx =>60^0 xxx, which is consistent with the Plumbing Regulations (2008) requirements. The Regulations require hot water be stored at or above 60^{0}C to minimise the growth of *Legionella* bacteria in plumbing systems. This is

		because hot water system temperature is one of the main factors affecting *Legionella* growth. Because the hot water system temperature was adequate, treating the *Legionella* bacteria in your system would require treatment by a plumber. In most cases (but not all) this treatment will be sufficient to resolve the problem. As this type of work is not routine we suggest you ask your plumber to speak with our experts before commencing the work. Please ask them to call us on **1800 248 898**. If you elect to address the *Legionella* bacteria in your system we recommend a repeat sample should be taken a week after the treatment to ensure it has been effective. This repeat sample is free of charge. Please call us when your plumber treats the system if you wish to arrange a repeat sample.
Storage (Electric or Gas)	**Unmeasured**	If you decide to address the *Legionella* in your hot water system this would involve arranging a plumber to check the system, including the storage temperature, as storage temperature is one of the main things affecting *Legionella* growth**.** The plumber may also do a heat treatment on the system. We were unable to measure the temperature in your hot water system, but the Plumbing Regulations (2008) require hot water be stored at or above 60^0C to minimise the growth of *Legionella* bacteria in plumbing systems. We would be happy to collect another sample at no charge a week after the system is treated to make sure the treatment has worked. Please call us on **1800 248 898** when your plumber treats the system if you wish to arrange a repeat sample.

Gas Instant, Solar or Other	n/a	If you decide to address the *Legionella* in your hot water system this would involve arranging a plumber to check the system is functioning properly and possibly do a heat treatment. We would be happy to collect another sample at no charge a week after the system is treated to make sure the treatment has worked. Please call us on **1800 248 898** when your plumber treats the system if you wish to arrange a repeat sample.

This page is intentionally blank

Appendix B. Vanuatu mission report

Mission Report: SPC Assistance to Vanuatu Ministry of Health in Strengthening Surveillance System Post Cyclone Pam

Prepared by:	**Zoe Cutcher** *Epidemiology Intern and MAE Scholar* *Research Evidence and Information Programme* *Public Health Division, Secretariat of the Pacific Community*
Report Date:	8 June 2015
Date of deployment:	3 May to 22 May 2015
Area of Deployment:	Penama Province, Vanuatu
Objectives:	To support the Penama Provincial Surveillance Unit (PSU) to further develop the EWARN surveillance system by: • Establishing, via a consultative process, a template feedback report to be disseminated from the PSU to Penama Province's sentinel sites reporting on EWARN. • Harmonizing syndromic case definitions between the hospital Out-patient Division (OPD) registry and EWARN reporting. • Act as technical lead for the Provincial EpiNet Team, if need be, to respond to public health emergencies/threats, including verifying EWARN signals, conducting outbreak investigations and initiation of mitigation/control measures.

Background

On 13 March 2015 Category 5 Tropical Cyclone Pam impacted Vanuatu. The cyclone led to significant infrastructure damage in the developing island nation. Potential threats to public health included structural damage to health care facilities, impaired health service delivery, damage to housing, food security, sanitation and loss of clean water supply.

The Vanuatu Ministry of Health (MOH) (Health Cluster Lead), with assistance from the World Health Organization (WHO) (Co-Lead), expanded Vanuatu's existing Pacific Syndromic Surveillance System to establish an Early Warning Alert and Response

Network (EWARN) comprising 24 sentinel sites across Vanuatu. The system was designed to provide timely surveillance of epidemic prone diseases during the emergency phase and enable rapid implementation of control measures if required.

The introduction of EWARN was an opportunity to build capacity for disease surveillance and response in Vanuatu. Vanuatu's surveillance system prior to the cyclone comprised eight sentinel sites reporting on four syndromes (acute fever and rash, prolonged fever, influenza like illness, and diarrhoea). However reporting was unreliable and there was no established system for response to surveillance signals or formal feedback material and mechanism to sentinel sites (George Worwor, National Surveillance Unit, MOH, personal communication). The pre-existing sentinel sites were located on 6 out of over 60 inhabited islands and are listed in Table 1.

Table 1. Syndromic surveillance sites in operation in Vanuatu since 2012.

Site	Province	Island
Lenakel	Tefea	Tanna
Vila Central Hospital (Vila)	Shefa	Efate
Neil Thomas Ministry Hospital (Vila)	Shefa	Efate
Welu Hospital (Vila)	Shefa	Efate
Norsup Hospital	Malakula Island	Malampa
Lolowai	Ambae	Penama
Northern Provincial hospital	Santo	Sanma
Qatvaes hospital	Vanua Lava	Torba

The MOH and the WHO requested assistance from the Secretariat of the Pacific Community (SPC) in (i) the establishment and implementation of EWARN in Penama Province, and (ii) to assure rapid diagnosis of disease with epidemic potential or public health importance, specifically at the two reference hospitals. In addition to these initial Terms of Reference SPC was also requested, upon arrival in Vanuatu, to operate and maintain the EWARN system recently established in Tanna for 4 weeks. SPC committed to an initial term of 2 months in Penama and met this commitment by deploying four epidemiologists on successive missions, each of 3-week duration on average, and one laboratory specialist for 3 days of intensive training in Lolowai Hospital. My deployment was the third and final mission for this commitment.

EWARN was intended to be a short-term disaster-response surveillance system. At the time of writing, it is planned that reporting will be wound back to previous levels over

the next few weeks as response efforts enter the recovery and rehabilitation phase. The imminent withdrawal of EWARN from the 16 additional participating sites provides an opportunity to adapt and continue elements of the system in a selection of sites to strengthen disease surveillance and response in Vanuatu. SPC's activities in Penama Province were undertaken with a view to developing capacity at Provincial and local levels and establishing a sustainable surveillance and response system for the province.

The main activities of the previous two deployments to Penama province are listed below:

1. Onofre Edwin A. Merilles Jr. (Jojo) (28/03/2015 -17/04/2015).
    a. Enrolled and trained staff at three Penama EWARN sites: Melsisi (Pentecost); Mauna (Pentecost); and Kerepei (Maewo).
    b. Enrolled staff and initiated training at Lolowai provincial hospital (Ambae) where provincial data will be aggregated.
    c. Prepared for the establishment of a Provincial EpiNet Team
2. Dr Salanieta Taka Saketa (13/04/2015 -27/04/2015).
    a. Rapid epidemic risk assessment for Penama province.
    b. 2 day training course on EWARN for Lolowai and provincial healthcare workers.
    c. Recruited 4 additional sentinel sites from Ambae (these report only to the provincial level)
    d. Developed TORs and supported the appointment of a Provincial Surveillance Co-ordinator
    e. Supported the establishment of Vanuatu's first EpiNet team for investigation and control of outbreaks in Penama Province

The purpose of this final SPC deployment to Penama province was to continue strengthening the newly established surveillance and response system for outbreak-prone diseases in Penama, in order to facilitate the transition of EWARN including the EpiNet team into an ongoing, sustainable system for the province.

Activities

Key activities are described below. All were conducted with a focus on capacity building to strengthen the newly established system and improve sustainability.

1. Development of a surveillance feedback report template

Through close consultation with local staff, I developed a template for the provision of surveillance feedback to reporting sites. Currently eight sentinel sites in Penama submit weekly reports to the provincial surveillance coordinator [2]. Reports are aggregated at the provincial level and investigations and response undertaken by the EpiNet team if indicated. The provincial surveillance coordinator will be responsible for providing regular and timely feedback of surveillance data to reporting sites. Such feedback will strengthen the newly established system by:

- increasing sentinel site engagement
- increasing preparedness through awareness of outbreaks and other public health threats occuring within the province and nationally
- providing information to assist with preparedness and response to sentinel sites.

The feedback template was developed through consultation with the provincial surveillance coordinator, Epi-Net team members and representatives from sentinel sites. Its implementation is immediately feasible with the current capacity and resources available in Penama. However the document is expected to evolve as the system strengthens, capacity increases, and staff gain ownership of the system.

A no-cost fortnightly distribution route was identified for the feedback. Accordingly the document was designed for fortnightly reporting. The template and guidelines for its use and dissemination are included in Appendix 2b.

Challenges

- Limited access to power at the hospital restricts access to PCs, printing and copying to a few hours per day
- Provincial surveillance coordinator does not have access to a computer, internet or office space
- Limited capacity in data analysis and data management

[2] Note that only 4 sites are participants in the National EWARN. Data for these 4 sites are relayed up to national level

- Limited capacity in basic epidemiology
- Limited/no baseline information makes it difficult to interpret surveillance data and identify alerts
- Enrolment of four additional reporting sites in recent weeks increases resources required to receive, analyse and feedback surveillance data.

Documentation of reporting sources

I mapped syndromes currently reported under EWARN against syndromes/conditions reported through the monthly national Health Information System (HIS) reports to identify potential double reporting and duplication of effort. HIS reports are submitted monthly to MOH by all hospitals and health centres and are designed for planning and monitoring rather than surveillance and response purposes. However if elements of surveillance and response are adopted and sustained throughout Vanuatu in the future the two systems would ideally be integrated to reduce duplication. Reporting of EWARN syndromes under HIS at both hospital and health centres was documented to provide guidance for potential future integration of the two systems, and to provide guidance to the EpiNet team for case verification and case finding from clinic records. I produced a data matrix demonstrating the relationship between syndromes under each system, and provided it to MOH and SPC.

Challenges

- There are no case definitions for diseases/syndromes reported under HIS
- Health centres report more specific diagnoses than hospitals – eg Typhoid and Hepatitis B are reported by health centres, but not hospitals whereas diagnosis may not be available at either level
- Potential for different cases of the same disease to meet case definitions for different syndromes under EWARN[3]
- Some duplicate reporting was recognised within the EWARN system[4].

[3] Important for surveillance staff to be aware of when interpreting surveillance report

[4] At Lolowai hospital EWARN cases were being counted from both the nurse OPD log book and the doctors log book; however all cases seen by the doctor are first seen by the nurse so appear in both log books.

Capacity Building and Outbreak Investigation

During the deployment the newly established EpiNet team investigated two surveillance signals. These were a) increased cases of influenza like illness at Melsisi Health Centre in Central Pentecost and b) a rumor of conjunctivitis and rash in Northern Pentecost. These were the first investigations undertaken by the Penama EpiNet team which had no previous experience in outbreak investigation. Preparation and investigation of the two surveillance signals was used to build EpiNet team capacity through workshops and practice. Capacity building included:

- verifying the signal
- logisitcs and planning for a field investigation
- constructing and interpreting a line list and basic analysis of line list data to identify risk groups
- finding cases and gathering additional information through simple questionnaires
- collecting blood samples, separating serum, packaging and shipping samples for laboratory analysis from a field investigation
- providing guidance for implementing control measures
- producing an outbreak report

The increase in influenza like illness was considered likely to reflect a reporting issue rather than a true increase in cases following a phone call and visit to Melsisi Health Centre for verification. The rumour of conjunctivitis and rash was verified by a dispensary in Angoro as an outbreak of conjunctivitis and arthritis. The EpiNet team conducted an outbreak investigation on-site, during which I provided technical support. An outbreak report is provided in Appendix 2c.

Challenges

- EpiNet team has a fairly new and limited understanding of basic epidemiology
- EpiNet team inexperienced in outbreak investigation
- Clinical and laboratory EpiNet team members have limited availability for field investigations due to hospital duties; this includes the current EpiNet team leader

- Only draft questionnaires available for use
- No disease/syndrome specific protocols for outbreak investigation/response
- Limited health promotion materials available to assist with intervention, and none specific to the disease/syndromes
- Logistical challenges in maintaining cold chain and freighting samples to laboratory (infrequent flight schedule)
- Cost of outbreak investigations likely to be a significant barrier to sustainability
- Maintenance of field epidemiology and laboratory skills is also likely to be a challenge due to frequent staff turnover

Recommendations

The following general recommendations address the challenges encountered during the deployment:

- Provision of ongoing capacity building in basic epidemiology and data analysis for the provincial surveillance coordinator and key EpiNet team members, ideally using a structured approach such as the Data for Decision Making (DDM) training courses.
- Provision of a lap-top, a reliable and permanent internet access and office space for the surveillance coordinator
- Improved provision of power to the Provincial Hospital
- Review the structure of the EpiNet team to ensure members are available for field investigations. Consider seeking clinical expertise from the affected health centers to relieve the need for Lolowai clinicians to attend investigations
- Train or refresh health centre and dispensary nurses in clinical skills in accordance with EWARN case definitions, as well as in sample collection, packaging and shipping to reduce the expenses for the EpiNet team's operations.
- Develop province-specific alert thresholds after one year of data collection
- Develop protocols and questionnaires for each of the 8 syndromes, including targeted health promotion and risk communication resources for each syndrome/common disease under surveillance
- Develop innovative financial solutions to ensure sustainability. Consider sharing cost of investigations among public health stakeholders,eg. contributions from the national programs such as environmental health, malaria, and health promotion, plus assistance from the central level; encourage provincial

authorities to support accommodation costs for the EpiNet Team during on-site investigations

- Care with sustainability – ensure staged approach with consolidation of capacity and sustainability at each level before advancing to the next
- Once EpiNet team is stable and sustainable, consider assisting the MOH in developing similar systems in other provinces.

This page is intentionally blank

Appendix C. Teaching Activity

Lesson from the field 5 – MLVA as a surveillance tool for Salmonella

The LFF teleconference will be conducted on **Tuesday 30th of June 2015 between 14:00 – 15:00 AEST.**

To join the teleconference, dial **1800 153721** and then enter the conference PIN code **604499.** If you have any trouble please call me on 0434 021 396 or email me on Zoe.Cutcher@health.vic.gov.au

Please save your responses to the questions in a word file and send back to Zoe.Cutcher@health.vic.gov.au by **COB Wednesday 24th of June 2015.**

Learning Objectives

After completing this exercise, you should be able to:

- Explain in basic terms why typing is important for surveillance of Salmonella (and other common communicable diseases)
- Describe some advantages of using MLVA for Salmonella surveillance
- Describe some challenges of using MLVA for Salmonella surveillance
- Use a basic relatedness schema to determine relatedness of Salmonella isolates based on MLVA patterns

Scenario:

You have recently started working as an OzFoodNet epidemiologist in Queensland. During routine review of Salmonella notifications during January you notice an increase in notified cases, many of which have identical MLVA patterns. You decide to investigate further, but first need to brush up on MLVA patterns.

Questions:

Read Torpdahl et al (2007) and answer the following:

1. Why is typing of Salmonella useful for surveillance?

2. List three essential criteria for typing methods and explain why they are important.

3. Explain the major advantage of MLVA analysis compared to simple serotyping/phagetyping or PFGE analysis.

4. Examine Figure 3.
    a. Describe Figure 3a.
    b. Describe Figure 3b.
    c. How would an investigation based on the information in figure 3a differ from an investigation based on the information in 3b?
    d. Discuss how bias might impact the results of a case-control investigation into the data shown in Graph 3a. What type of bias would this be, and what would be the effect on the investigation?

Read Wang, 2007 before completing the exercise below. Note that there are other methods for assessing relatedness using MLVA, for example some authors consider any number of repeat differences at a given loci acceptable, whereas Wang allows only one or two repeat differences for related strains. For this example please use the method proposed by Wang.

Examine the dataset provided “mlva activity dataset.xls” and identify possible outbreaks during the period. (*Hint, create a stacked column chart displaying MLVA patterns over time)*.

5. Select a surveillance signal you feel may warrant further investigation. Describe the signal and list any relevant MLVA patterns. What additional information would you ideally consider before deciding whether or not to investigate?

6. Next, categorise all the MLVA patterns for the month of January according to their relatedness to the predominant pattern. This information is useful when

deciding which cases should be counted when quantifying the size of the outbreak. It also helps in deciding which cases should be investigated.
Hint: Use the guidelines provided in Wang 2007, page 3. First identify the predominant pattern in the signal you wish to investigate, then categorise each isolate in the dataset as "Indistinguishable, Probably related, Possibly related, or Unrelated" to this pattern. This can be done in STATA, in Excel (if/then statements), or by eye.

Relatedness to predominant outbreak pattern	**Number of isolates**
Indistinguishable	
Probably related	
Possibly related	
Unrelated	
Total	

7. Create an epi curve showing occurrence of MLVA patterns over time that are Indistinguishable, Probably or Possibly related to the outbreak, and paste it below.

8. In the absence of any exposure information, how likely do you think it is that the "probably related" isolates relate to the same source as the outbreak? Justify your answer.

9. You decide to conduct a case-control study to investigate this outbreak. Among notified cases, which cases would you interview? Would you handle the groups differently in your analysis? If so, how?

10. Now that you've revised the process of simple MLVA analysis, list some challenges that need to be overcome to successfully use MLVA for public health surveillance.

11. Use the internet to identify some other organisms that can be analysed using MLVA.

Reading

1. Torpdahl M, Sorensen G, Lindstedt BA, Nielsen EM. Tandem repeat analysis for surveillance of human Salmonella Typhimurium infections. Emerging infectious diseases. 2007;13(3):388-95.
2. Wang Q. A coding/naming system for molecular typing of *Salmonella* Typhimurium using multi-locus variable number tandem repeat analysis (MLVA) and the initial guideline for the interpretation of MLVA data in NSW. 2007

Further Reading

Sintchenko V, Wang Q, Howard P, Ha CW, Kardamanidis K, Musto J, et al. Improving resolution of public health surveillance for human Salmonella enterica serovar Typhimurium infection: 3 years of prospective multiple-locus variable-number tandem-repeat analysis (MLVA). BMC Infect Dis. 2012;12:78.

Lindstedt BA, Vardund T, Aas L, Kapperud G. Multiple-locus variable-number tandem-repeats analysis of Salmonella enterica subsp. enterica serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. J Microbiol Methods. 2004;59(2):163-72.