Epidemiological investigations into the 2007 outbreak of equine influenza in Australia

By

Simon Matthew Firestone

BVSc, BSc, PGDipEpi, MAppEpi

A thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

Farm Animal and Veterinary Public Health,

Faculty of Veterinary Science,

The University of Sydney

2012

This thesis is submitted to the University of Sydney in fulfilment of the requirement is for the Degree of Doctor of Philosophy.

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Signature:  Date: 11 November 2012

馬
年

'Year of the Horse'

Firestone, S.M., 2012. 'Epidemiological investigations into the 2007 outbreak of equine influenza Australia.' Ph.D. thesis. Faculty of Veterinary Science, The University of Sydney, Sydney, Australia.

Abstract

Equine influenza is a highly contagious and widespread viral respiratory disease of horses and other equid species, characterised by fever and a harsh dry cough. Prior to August 2007, Australia was one of only three countries to have remained free of equine influenza. An incursion of equine influenza virus H3N8 in that month resulted in a four-month outbreak during which approximately 69,000 horses were infected on an estimated 9599 premises across two States. Most of the geographic spread occurred within the first 10 days and was associated with the movement of infected horses prior to the implementation of movement controls. The outbreak was contained through a series of interventions that ultimately led to the eradication of equine influenza from Australia. During and immediately after the outbreak, intensive epidemiological investigations, laboratory and retrospective analytical studies were conducted culminating in a series of detailed reports and publications, and the collation of a highly detailed outbreak dataset. Further research into the factors that contributed to the spread of the outbreak and the effectiveness of measures implemented to control and contain it was considered important. The aim of this thesis was therefore to investigate the factors that contributed to the spread of the 2007 equine influenza outbreak in Australia and to develop statistical methods and tools useful for informing the surveillance and control of future emergency animal disease events.

A case-control study was conducted to investigate premises-level risk factors, specifically whether compliance with advised biosecurity measures prevented the spread of equine influenza onto horse premises. Horse owners and managers on 200 properties across highly affected areas of New South Wales were interviewed. The proximity of premises to the nearest infected premises was the factor most strongly associated with case status. Case premises were more likely than control premises to be within 5 km and beyond 10km of an infected premises. Having a footbath in place on the premises before any horses were infected was associated with a nearly four-fold reduction in odds of infection (odds ratio = 0.27; 95% confidence interval:

0.09, 0.83). This protective association may have reflected overall premises biosecurity standards related to the fomite transmission of equine influenza: there was high correlation amongst several, generally protective, variables representing personal 'barrier hygiene' biosecurity measures (hand-washing, changing clothes and shoes, and having a footbath in place).

The movement of infected horses and local disease diffusion were known to be important mechanisms of spread early in this outbreak. A network analysis was conducted to investigate the relative contribution of each mechanism. The relationship between infected and susceptible horse premises (contact through animal movements and spatial proximity) was described by constructing a mixed transmission network. During the first 10 days of the 2007 equine influenza outbreak in Australia, horses on 197 premises were infected. A new likelihood-based approach was developed and it was estimated that 28.3% of early disease spread (prior to the implementation of horse movement restrictions) was through the movement of infected horses (95% CI: 25.6, 31.0%). Most local spread was estimated to have occurred within 5 km of infected premises. Based on a direct estimate of the shape of the spatial transmission kernel, the incidence beyond 15 km was very low. The median distance that infected horses were moved was 123 km (range 4–579 km).

In an extension of the network analysis, novel methods were developed to delineate spatial clusters of infected premises and describe the sequence of cluster formation and the widespread dispersal experienced during the first 30 days of the outbreak. Premises identified as infected by the movement of infected horses were found to be critical to the seeding of infection in spatial clusters. Combined analysis of spatial and contact network data demonstrated that early in this outbreak local spread emanated outwards from the small number of infected premises in the contact network, up to a distance of around 15 km. A purely spatial method of modelling epidemic spread (kriging) was imprecise in describing the pattern of spread during this early phase of the outbreak (explaining only 13% of the variation in estimated date of onset of

infected premises), because early dissemination was dominated by network-based spread.

Prior to this thesis, there was an abundance of anecdotal information regarding the role of meteorological factors and other environmental determinants in the spread of the 2007 equine influenza outbreak in Australia. A survival analysis was therefore conducted to empirically estimate the association between meteorological variables (wind, air temperature, relative humidity and rainfall) and time-to-infection in the largest cluster of the outbreak, in northwest Sydney. The equine influenza outbreak dataset was structured to enable generalised Cox regression modelling of the association between time-varying covariates representing premises-level meteorological conditions. The cumulative incidence in the northwest Sydney cluster was estimated to be 53.0% (95% CI: 51.4, 54.7%). Local spatial spread of equine influenza was found to be associated with relative humidity, air temperature and wind velocity. Meteorological conditions 3–5 days prior were strongly associated with hazard of influenza infection. Strong winds (>30 km $hour^{-1}$) from the direction of nearby infected premises were associated with influenza infection, as was low relative humidity (<60%). A nonlinear relationship was observed with air temperature: the lowest hazard was on days when maximum daily air temperature was between 20–25 °C.

Drawing on the findings of the above studies, a spatially-explicit stochastic epidemic model of equine influenza transmission was developed to investigate the underlying disease process, estimate the effectiveness of several control measures applied during the 2007 outbreak and to provide a dynamic modelling framework for rapid assessment of future equine influenza outbreaks in Australia. A reversible jump Markov chain Monte Carlo algorithm was used to estimate Bayesian posterior distributions of key epidemiological parameters based on data from two highly affected regions. A large amount of regional heterogeneity was observed in the underlying epidemic process, the estimated rate of decay of transmission by distance from infected premises, the intra-premises transmission rate and the effect of premises area. Model outputs were highly cross-correlated both temporally and spatially with data observed during

the 2007 outbreak, and with outputs of a previous model. Pseudo-validation of the model against data, not used in its development, demonstrated of how it may be applied to develop rapid assessments of future outbreaks affecting horse populations in comparable regions to those studied.

The study results documented in this thesis have elucidated the key factors underlying the spread of the 2007 equine influenza outbreak in Australia, and presented new methods of describing such rapidly spreading epidemics. The movement of infected horses, meteorological variables (air temperature, humidity and wind speed), on-farm biosecurity measures and intrinsic features of horse premises (proximity to other infected premises, numbers of horses held and premises area) were all important variables that influenced the spread of infection onto horse premises. These insights allow development of better policy and control programs in the event of a future equine influenza virus incursion.

Acknowledgements

A thesis is not possible without a team, and I have had a great team.

First and foremost, I acknowledge and thank my supervisory team: Navneet Dhand, my principal supervisor, mentor and biostatistician, Jenny-Ann Toribio, my inspirational colleague in veterinary public health education, and Michael Ward, my editor-in-chief; and my research partner Kathrin Schemann.

The equine influenza research project was Navneet's first as a principal investigator. My thesis was just one part of this project, and Navneet provided guidance and advice throughout my candidature, contributing to the design and conduct of all of my studies and providing me every opportunity to develop my skills in veterinary epidemiology. At many stages he trusted me to disappear, sometimes for months at a time, independently researching, collaborating internationally and frequently eloping to Melbourne (once to marry, once for employment, and once for reproduction). His tireless inputs into my research and well-considered guidance gave me room to grow, and ensured that we produced quality research together.

Throughout my PhD experience Jenny-Ann has been the most human character, someone that never saw or treated me as a student, but an equal. She entrusted me with more teaching opportunities that I could handle, and this provided me the chance to develop into a lecturer as well as a researcher. Her thoroughness whilst reviewing my research proposals and outputs was also truly appreciated. Few research students have access to such an insightful review editor as Michael Ward before they open themselves up to the challenges of peer-review. His experience and advice counted on a number of occasions. Michael always knew which angles to pursue, asked the hard questions during study design, and ensured that the writing ultimately portrayed the quality of the research, and on top of all this he always managed to keep the whole thing entertaining.

Thoroughly deserving of a paragraph of her own, is Kathrin Schemann, my research partner and partner in crime. It has been a pleasure travelling this road towards a PhD together with you. Somehow you remained in good humour even when my random number generator sent you off to the most remote of our on-farm case-control study interviews, and after hearing so many of my conference practice presentations that you could recite them all yourself. Thank you for your contributions to my research.

I also specifically thank my co-authors and collaborators: Rob Christley, Chris Jewell, Naomi Cogger, Barbara Moloney and Graeme Garner; Peter Thomson for statistical advice when it was most needed; Nigel Perkins and Evan Sergeant for comments on study design; the NSW and Queensland Departments of Primary Industries for making their equine influenza datasets available, and the following individuals for putting in so much effort to collate, clean and analyse the outbreak surveillance data before I came onto the scene: Brendan Cowled, Barbara Moloney, Nina Kung and Graeme Garner. I also am very grateful to the case-control study interviewees, and the DPI laboratory and field staff, veterinarians and horse owners whose efforts both during and immediately following the outbreak were captured in the outbreak dataset.

The Australian Biosecurity Cooperative Research Centre for Emerging Infectious Diseases (ABCRC) provided my PhD scholarship. The Rural Industries Research and Development Corporation (RIRDC) provided the project funding and a top-up scholarship. The William and Catherine McIlrath and the Barrenger overseas veterinary research travelling scholarships, and an ABCRC student placement enabled me to pursue international research collaborations.

I would also like to thank my fellow postgraduate students out at Camden, the Farm Animal and Veterinary Public Health group, and the staff of the Faculty of Veterinary Science, particularly Marion Saddington, Chris Stimson and Andrew Oulianoff. Last, but not least, I would like to thank Snitzel and the connections of Encosta De Lago.

Dedication

This thesis is dedicated to my beautiful wife Jen, without your love and support none of this would have been possible.

And to our little Charlie, who arrived on time, then waited, and waited, and waited,
and eventually I came home.

Publications

Parts of this thesis have been published and presented elsewhere:

PEER-REVIEWED PUBLICATIONS

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2011. A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. Prev. Vet. Med. 100, 53-63. (Chapter 3)

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2012. Adding the spatial dimension to the social network analysis of an epidemic: investigation of the 2007 outbreak of equine influenza in Australia. Prev. Vet. Med. 106, 123-135.
(Chapter 4)

Firestone, S.M., Ward, M.P., Christley, R.M., Dhand, N.K., 2011. The importance of location in contact networks: describing disease spread using spatial social network analysis. Prev. Vet. Med. 102, 185-195. (Chapter 5)

Firestone, S.M., Cogger, N., Ward, M.P., Toribio, J.-A.L.M.L., Moloney, B.J., Dhand, N.K., 2012. The Influence of Meteorology on the Spread of Influenza: Survival Analysis of an Equine Influenza (A/H3N8) Outbreak. PLoS ONE 7, e35284. (Chapter 6)

MANUSCRIPT IN PREPARATION FOR SUBMISSION FOR PEER-REVIEW

Firestone, S.M., Jewell, C.P., Garner, M.G., Moloney, B.J., Ward, M.P., Dhand, N.K., unpublished manuscript. Development of a dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia. (Chapter 7)

PUBLISHED CONFERENCE PROCEEDINGS AND ABSTRACTS

Firestone, S.M., Cogger, N., Moloney, B.J., Toribio, J.A., Ward, M.P., Dhand, N.K., 2011. The effect of humidity, air temperature and wind on the local spread of equine influenza. In, Faculty of Veterinary Science Postgraduate Conference, The University of Sydney, Camden, Australia.

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2011. Adding the spatial dimension to the social network analysis of an epidemic. In, Society for Veterinary Epidemiology and Preventive Medicine 2011 Annual Conference, Leipzig, Germany.

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2010. Applying social network analysis to contact-tracing and spatial data from the 2007 outbreak of equine influenza in Australia. In, GEOVET 2010, Sydney.

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2010. Combining social network analysis with spatial analysis of equine influenza epidemic contact-tracing data to describe early spread of the 2007 outbreak in Australia. In, 4th International Symposium on Geospatial Health, Melbourne.

Firestone, S.M., 2010. Social network analysis combining contact-tracing and spatial data to investigate a rapidly spreading epidemic. In, Australasian Epidemiological Association, Sydney.

Firestone, S.M., 2010. Epidemiological studies into the 2007 equine influenza outbreak in Australia. In, Australian College of Veterinary Scientists, Epidemiology Chapter, Science Week Conference, Surfers Paradise, Gold Coast.

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2010. Risk factors for the infection of horse premises by equine influenza in New South Wales. In, Global Biosecurity 2010: safeguarding agriculture and the environment, Brisbane.

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2010. Social network analysis of the early spread of the 2007 equine influenza outbreak in Australia. In, Faculty of Veterinary Science Postgraduate Conference, The University of Sydney.

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2010. On-farm biosecurity and other factors influencing equine influenza spread during the 2007 outbreak. In, Sydney Institute for Emerging Infectious Diseases and Biosecurity Conference, Sydney.

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2010. On-farm biosecurity and other factors influencing equine influenza spread during the 2007 outbreak. In, Australian Equine Veterinarians Association Conference, Hunter Valley.

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2010. On-farm biosecurity and other factors influencing equine influenza spread during the 2007 outbreak in Australia. In, Australian Biosecurity CRC for Emerging Infectious Diseases, Annual National Workshop, Fraser Island.

Firestone, S.M., Schemann, K.A., Dhand, N.K., Toribio, J.A., Ward, M.P., 2009. Epidemiological investigations into the 2007 equine influenza outbreak in Australia. In, Faculty of Veterinary Science Postgraduate Conference, The University of Sydney.

AUTHOR CONTRIBUTIONS TO PUBLISHED PAPERS

Chapter 3

***Firestone, S.M.**, Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2011. A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. Prev. Vet. Med. 100, 53-63.*

This study was a team effort. The case-control study and social network analysis of contact-tracing data from the 2007 equine influenza outbreak were originally planned to be conducted under the RIRDC-funded equine influenza project. As the primary author I was responsible for formulating the research questions, refining the study design, conducting the study and all statistical analyses, drafting the manuscript, responding to reviewers' reports and coordinating submission and publication of the original research paper. Kathrin Schemann and I worked together to pilot the questionnaire, and each conducted half of the interviews and entered half of the data. All co-authors were involved in planning the study, and commented on the questionnaire design, manuscript for submission, response to reviewers' reports and final version of the original research paper. Anonymous reviewers provided constructive comments on the submitted manuscript, which were incorporated into the paper. Further individual contributions to compilation of the dataset used to construct the sampling frame and provision of comments on study design are specifically acknowledged in the published paper.

Chapter 4

***Firestone, S.M.**, Christley, R.M., Ward, M.P., Dhand, N.K., 2012. Adding the spatial dimension to the social network analysis of an epidemic: investigation of the 2007 outbreak of equine influenza in Australia. Prev. Vet. Med. DOI: 10.1016/j.prevetmed.2012.01.020.*

The impetus for this paper arose from a workshop that Robert Christley presented whilst on an international visiting fellowship to The University of Sydney. Prior to the workshop, on the

application of social network analyses to veterinary epidemiological data, I formulated the research questions and prepared the contact-tracing and spatial datasets for analysis. Robert Christley and I worked together after the workshop to refine the research methods and implement the social network analyses incorporating spatial relationships between infected premises. I was responsible for the overall design and conduct of the study and statistical analyses, drafting the manuscript, responding to reviewers' reports and coordinating submission and publication of the original research paper. This paper was presented at the Society for Veterinary Epidemiology and Preventive Medicine 2011 Annual Conference, in Leipzig, Germany, and is part of a special conference issue of the journal Preventive Veterinary Medicine. Anonymous reviewers provided constructive comments on the submitted manuscript. In response to a specific detailed request I sought the guidance of Peter Thomson who advised on the development of the likelihood-based approaches presented in the published paper. All co-authors reviewed and commented on the design of the study, the manuscript for submission, response to reviewers' reports and the final version of the original research paper. Further individual contributions to compilation of the contact-tracing and spatial datasets and provision of comments on study design are specifically acknowledged in the published paper.

Chapter 5

***Firestone, S.M.**, Ward, M.P., Christley, R.M., Dhand, N.K., 2011. The importance of location in contact networks: describing disease spread using spatial social network analysis. Prev. Vet. Med. 102, 185-195.*

I was responsible for conceiving and conducting this follow-up spatial social network analysis. Michael Ward provided specific guidance on the geostatistical approach and spatial cluster detection methods, which I then implemented. I also refined the spatial social network analyses methods, developed with Robert Christley, to delineate clusters of infected premises and describe the sequence of cluster formation. All co-authors reviewed and commented on the manuscript that I drafted for submission, response to reviewers' reports and the final version of

the original research paper. This paper was presented at the GEOVET 2010 conference, in Sydney, and is part of a special issue of the journal Preventive Veterinary Medicine. Anonymous reviewers provided constructive comments on the submitted manuscript, which were incorporated into the paper. Further individual contributions to compilation of the data used in this analysis are specifically acknowledged in the published paper.

Chapter 6

***Firestone, S.M.**, Cogger, N., Ward, M.P., Toribio, J.-A.L.M.L., Moloney, B.J., Dhand, N.K., 2012. The Influence of Meteorology on the Spread of Influenza: Survival Analysis of an Equine Influenza (A/H3N8) Outbreak. PLoS ONE 7, e35284.*

A survival analysis based on the equine influenza outbreak dataset was originally proposed by Naomi Cogger at a research planning meeting held in Canberra on 29 July 2008. After receiving training in advanced survival analysis methods from Naomi Cogger and her colleagues at the Epicentre, Massey University (New Zealand), I refined the research question and planned the methods to include meteorological variables considered as potential risk factors for equine influenza spread. Naomi Cogger provided ongoing guidance in survival analysis dataset formulation whilst I implemented all statistical analyses. I was responsible for drafting the manuscript, responding to reviewers' reports and coordinating submission and publication of the original research paper. All co-authors were involved in planning the study, manuscript for submission, response to reviewers' reports and final version of the original research paper. Anonymous reviewers provided constructive comments on the submitted manuscript, which were incorporated into the paper. Further individual contributions to compilation of the data used in this analysis are specifically acknowledged in the published paper.

Chapter 7

Firestone, S.M., Jewell, C.P., Garner, M.G., Moloney, B.J., Ward, M.P., Dhand, N.K., unpublished manuscript. Development of a dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia.

This study was a truly collaborative effort, conceived and implemented by Chris Jewell and myself during a 1 month visit to his research facility at the University of Warwick (United Kingdom) and over the subsequent 9 months. I prepared the data for statistical analysis including multiple imputation of missing data. Chris Jewell and I formulated the model together, then Chris implemented the Bayesian inference of epidemiological parameters by adapting his previously published reversible jump Markov chain Monte Carlo algorithms. Meanwhile, I coded a corresponding stochastic simulation model, then conducted all further analyses including description and interpretation of the inferred parameter distributions, stochastic simulations, model pseudo-validation, inter-model comparison and sensitivity analyses. Graeme Garner made a specific contribution to this analysis by running the simulation model developed by his team at DAFF to provide outputs for inter-model comparison based on data that I sent to him. I then drafted the manuscript, which is in the latter stages of preparation for submission to a peer-reviewed journal.

Table of Contents

Table of Tables

Table of Figures

Acronyms & Abbreviations

ABCRC	Australian Biosecurity Cooperative Research Centre
Amber zone	Control area
CCEAD	Consultative Committee on Emergency Animal Disease
CI	confidence interval
CPEC	Centennial Parklands Equestrian Centre
CrI	credible interval
CV	coefficient of variation
DIVA	differentiation of infected from vaccinated animals
DPI	Department of Primary Industries
EAD	emergency animal disease
ECQS	Eastern Creek Quarantine Station
EI	equine influenza
EID_{50}/ml	50% egg infective dose per ml of swab extract
ELISA	enzyme linked immunosorbent assay
FMD	foot-and-mouth disease
Green zone	Protected area
H	haemagglutinin
HI	haemagglutination inhibition
HPAI	highly pathogenic avian influenza
HPD	highest posterior density region
ICC	Intra-class correlation
IP	Infected Premises
IQR	inter-quartile range
LRT	likelihood-ratio test
MNAR	missing not at random
N	neuraminidase
NSW	New South Wales
OIE	World organisation for animal health
OR	odds ratio
Purple zone	Special restricted area
QLD	Queensland
qRT-PCR	real-time reverse transcription-polymerase chain reaction assay
R	effective reproductive ratio
R_0	basic reproductive ratio
R^2	coefficient of determination
Red zone	Restricted area
RIRDC	Rural Industries Research and Development Corporation
RJ-MCMC	reversible jump Markov chain Monte Carlo
SIR	susceptible-infected-recovered
SLIR	susceptible-latent-infected-recovered
SNA	social network analysis
UIP	uninfected premises

Introduction

Equine influenza is a highly contagious acute respiratory disease of horses, capable of infecting all members of the family *Equidae* (horses, donkeys, mules and zebras). Most horses recover uneventfully, following a mild disease characterised by a harsh dry cough, fever, lethargy and nasal discharge. Rapidly-spreading outbreaks have caused major disruptions to horse racing and other equestrian events when equine influenza virus has gained access to immunologically naïve horse populations. Previous outbreaks have also occurred in vaccinated horse populations when antigenic drift has led to the emergence of new strains.

The 2007 outbreak of equine influenza was the largest emergency animal disease outbreak and response effort in Australia's history. It provided a rare example of a highly contagious infectious agent spreading through an almost completely immunological naïve population, impeded only by a series of interventions that ultimately enabled containment and eradication. During and immediately after the outbreak, a series of intensive epidemiological investigations, laboratory and retrospective analytical studies were conducted culminating in a series of detailed reports and publications, and the collation of a highly detailed outbreak dataset. However, further research was considered important into the factors that contributed to the spread of the outbreak and the effectiveness of measures implemented to control and contain it.

Aim of the thesis

The aim of this thesis was therefore to investigate the factors that contributed to the spread of the 2007 equine influenza outbreak in Australia and develop statistical methods and tools useful for informing the surveillance and control of future emergency animal disease events in horses.

Synopsis and structure of the thesis

The literature on the epidemiology of equine influenza is reviewed in Chapter 1, including an overview of the 2007 outbreak in Australia and a summary of epidemiological approaches to emergency animal disease events.

Chapter 2 describes the collation of the equine influenza outbreak dataset and an assessment of the spatial accuracy of the dataset. This exploratory analysis identified considerable amounts of duplication of premises records in highly affected areas. To correct for this discrepancy and prepare the dataset for further epidemiological analyses, premises records were aggregated by cadastral (property) boundaries and the cumulative incidence re-estimated.

Chapter 3 presents a case-control study conducted to investigate risk factors for the spread of equine influenza onto horse premises in highly affected areas during the first 7 weeks of the 2007 outbreak, a period prior to the implementation of vaccination and the relaxation of movement restrictions. A specific objective of this risk factor study was to investigate whether compliance with advised biosecurity measures prevented the local spread of infection. Multivariable analyses enabled estimation of the effectiveness of certain biosecurity measures after adjusting for risk associated with proximity to the nearest infected premises and other important confounding variables. Estimates of the risk of infection at various distances from an infected premises are also presented.

The movement of infected horses and local spatial spread were known to be important mechanisms of spread in this outbreak. To investigate the relative contribution of each of these mechanisms two network analyses were conducted. The first of these analyses, presented in Chapter 4, was restricted to the first 10 days of the outbreak to focus on disease spread prior to the implementation of horse movement restrictions. A mixed transmission network was constructed to describe two important relationships between infected and susceptible horse

premises: contact through animal movements and spatial proximity. Novel methods are presented for delineating spatial clusters of infected premises based on contact-tracing data ('spatial social network analysis'), characterising the risk of local spatial spread by distance, and estimating the probability that a premises was infected through local spatial spread or the movement of an infected horse. These methods enabled the description of the sequence of cluster formation and the widespread dispersal experienced in the early phase of this outbreak. The characterisation of local spread is compared to direct estimates of the shape of the spatial transmission kernel of equine influenza under Australian conditions achieved by applying methods developed for modelling the 2001 foot and mouth disease outbreak in the United Kingdom.

The network analysis is extended in Chapter 5 to describe disease spread in the first month of the outbreak in terms of the network and spatial location of individual premises. Purely spatial approaches to describing clusters are compared to the spatial social network analysis approach. The importance of incorporating contact network relationships into a spatial analysis framework for outbreak spread is ascertained by cross-validation of a geostatistical approach to cluster description.

There was an abundance of anecdotal information into the role of meteorological factors and other environmental determinants in the spread of the 2007 equine influenza outbreak in Australia. Chapter 6 presents a survival analysis that aimed to empirically estimate the association between meteorological variables (wind, air temperature, relative humidity and rainfall) and time to infection in the largest cluster of the outbreak. Recent experimental research has identified how relative humidity and temperature influence influenza A transmission under controlled conditions. Analyses of the contribution of wind to the spread of outbreaks of influenza are more limited. To enable generalised Cox regression modelling of the association between hazard of infection and time-varying meteorological covariates representing premises-level conditions, the equine influenza outbreak dataset was re-structured

and combined with concurrent daily meteorological data.

Finally, in Chapter 7, a spatially-explicit stochastic simulation model of equine influenza in Australia is developed, validated and tested in a formal sensitivity analysis to further investigate the underlying disease process, estimate the effectiveness of several control measures applied during the 2007 outbreak and to provide a dynamic modelling framework for rapid assessment of future equine influenza outbreaks in Australia. The simulation model was informed by Bayesian inference of epidemiological parameters, with the formulation based on the research findings of all other Chapters in this thesis. Model predictions were tested against outputs of an earlier validated model and actual outbreak data not used in the development of the model. Sensitivity analysis and interpretation of modelling outputs from several regions allowed development of an understanding of what is now known about the underlying disease process, and where further research should be directed.

A discussion of the significance of the key findings of this thesis is presented in Chapter 8, along with areas for future research and recommendations for outbreak surveillance and epidemiological analysis in the event of future incursions of emergency animal diseases into Australia.

Chapter 1: Literature review

"*Declare the past, diagnose the present, foretell the future; practice these acts.*"

Hippocrates (c.410 BC), 'Epidemics', Book I, Section XI

1. Literature review

1.1. Introduction

During and immediately after the 2007 outbreak of equine influenza in Australia, a series of intensive epidemiological investigations, laboratory and retrospective analytical studies were conducted culminating in a series of detailed reports (Britton, 2007, Gilkerson, 2007, Equine Influenza Epidemiology Support Group, 2008), publications (Bryant et al., 2010, Cowled et al., 2009b, Garner et al., 2011b, Davis et al., 2009, East, 2009, Taylor et al., 2008, Sergeant et al., 2009, Foord et al., 2009, Heine et al., 2007, Jeggo et al., 2008) and a special issue of the Australian Veterinary Journal (Jackson, 2011). The earliest of these studies and reports informed decision-making during the outbreak, several informed a subsequent government inquiry into the outbreak (Callinan, 2008), and most have informed a major update of Australia's emergency animal disease strategy for responding to future outbreaks of equine influenza (Animal Health Australia, 2011). Two reports in particular present comprehensive accounts of the 2007 equine influenza outbreak in Australia: the Inquiry's report (Callinan, 2008) and an epidemiology report prepared for Australia's key animal disease emergency policy and action committee (the Consultative Committee on Emergency Animal Disease, CCEAD) (Equine Influenza Epidemiology Support Group, 2008).

This literature review aims to present an overview of the epidemiology of equine influenza, the 2007 outbreak in Australia, epidemiological approaches to emergency animal disease events and additional research relevant to this thesis.

1.2. The epidemiology of equine influenza

1.2.1. Overview - clinical aspects and risk factors

Equine influenza (EI) is a major cause of respiratory disease of horses worldwide (Daly et al., 2004), capable of affecting all members of the family *Equidae* (horses, donkeys, mules and zebras) (Spickler, 2009). Clinically, equine influenza is often a mild disease characterised in

immunologically naïve horses by a harsh dry cough, fever (up to 41 °C), lethargy and nasal discharge (Myers and Wilson, 2006, Paillot et al., 2006). Severe viral and secondary bacterial infections may develop, especially in foals and young horses (Miller, 1965), donkeys (Uppal and Yadav, 1987), and horses in poor health (Paillot et al., 2006). Old and stressed horses occasionally suffer complications including myositis, myocarditis, oedema and encephalitis (Wilson, 1993), and pregnant mares may abort or resorb the foetus (Myers and Wilson, 2006). Although most horses recover uneventfully, large rapidly-spreading outbreaks of equine influenza have occurred when highly susceptible horse populations have been exposed to novel virus strains (Uppal and Yadav, 1987, Dalglish, 1992, Guthrie et al., 1999) causing major disruptions to horse racing and other equestrian events (Daly et al., 2006). Previous studies of equine influenza outbreaks have identified increased risk of infection amongst young horses (Morley et al., 2000b, Gildea et al., 2011), horses kept in confined/high density housing (Powell et al., 1995, Morley et al., 2000b, Gildea et al., 2011), those with low levels of homologous antibodies (Morley et al., 2000b, Barquero et al., 2007), horses involved in equestrian events (Gildea et al., 2011), and those with higher frequency of contact with other horses both on the premises or through movements on and off multi-horse premises (Morley et al., 2000b, Gildea et al., 2011). In a study of an outbreak amongst horses previously vaccinated with a non-homologous strain, male horses were found to be at increased risk of infection (Barquero et al., 2007).

1.2.2. Virological aspects: emerging and re-emerging strains

The causative agent of EI is an influenza A virus of the *Orthomyxoviridae* family of enveloped RNA viruses (Hannant and Mumford, 1996). Influenza A viruses affect many animal species; most are avian influenza viruses, but several (human, swine, equine and canine influenza) have become adapted to non-avian host species and circulate predominantly in these species (Taubenberger and Kash, 2010). The accumulation of point mutations during the replication of influenza viruses (‘antigenic drift’) leads to changes in the structure of surface glycoproteins,

until they are no longer recognisable by antibodies generated through earlier infection or vaccination (Daly et al., 2011). In contrast, 'antigenic shift' may occur through genetic reassortment when two different influenza viruses invade the same cell (Taubenberger and Kash, 2010).

There are two known pathogenic subtypes of equine influenza viruses, denoted H7N7 and H3N8 according to their haemagglutinin (H) and neuraminidase (N) surface glycoproteins (Hannant and Mumford, 1996). Equine influenza virus A/eq/Prague/56 (H7N7) was first isolated in Czechoslovakia in 1956 (Sovinová et al., 1958), during a widespread outbreak of respiratory disease affecting horses across Eastern Europe. The last reported outbreak of H7N7 equine influenza was in Italy in 1979 (Daly et al., 2006) and the last known isolation in Egypt in 1989 (Ismail et al., 1990). In 1963, equine influenza virus A/eq/Miami/63 (H3N8) was isolated for the first time, in the United States of America (USA), following the importation of horses from Argentina (Scholtens et al., 1964). Large-scale outbreaks have continued to occur in Europe and North America in every decade, caused by H3N8 subtype viruses (Daly et al., 2006).

In the mid-1980s, H3N8 equine influenza viruses diverged into distinct 'American' and 'Eurasian' lineages (Daly et al., 1996). American lineage H3N8 equine influenza viruses now predominate, and continue to diverge (Daly et al., 2011). Phylogenetic analysis has shown that repeated sporadic incursions of virus have occurred from North America into Europe and other regions, followed by periods of localised antigenic drift (Daly et al., 2011) driven in part by immune selection pressures due to vaccination coverage (Murcia et al., 2011). The American lineage has subsequently diverged into 'Kentucky', 'South American' and 'Florida' sublineages (Lai et al., 2001). A novel strain from the latter was responsible for an outbreak amongst vaccinated and unvaccinated horses at Newmarket in the United Kingdom (UK) in 2003 (Newton et al., 2006), later spreading across Europe, before being re-introduced to South Africa causing that country's second major outbreak (Guthrie, 2006, Daly et al., 2006). Since then, the

Florida sublineage has further diverged into two clades which co-circulate in Europe, whilst the majority of recent isolates from North American are now from 'Florida Clade 1' (Bryant et al., 2011).

Adding to the complex pattern of H3N8 equine influenza virus emergence and re-emergence, a novel strain with avian-derived HA sequences emerged in China in 1989 (Webster and Guo, 1991), causing an unusual outbreak with 20% mortality rates reported in some herds. Horse deaths were associated with enteritis and pneumonia (Webster and Guo, 1991). The presumably avian-derived Chinese H3N8 strain (Guo et al., 1992) circulated locally for 5 years but has not been detected since 1990 (Guo et al., 1995). Equine influenza H3N8 has also recently crossed the species barrier from horses into dogs, causing limited numbers of cases, many subclinical, in the UK in 2002–03 (Daly et al., 2008, Smith et al., 2005, Newton et al., 2007). H3N8 equine influenza virus adapted to circulate independently in dogs following an outbreak of severe respiratory disease in the USA in 2004 (Crawford et al., 2005), then again in Korea in 2007 (Song et al., 2008, Song et al., 2009). Canine influenza virus spread across the USA in 2004–05 (Beeler, 2009), diverging sufficiently from equine influenza A/H3N8 to be capable of efficient dog-to-dog transmission (Payungporn et al., 2008).

1.2.3. Mechanisms of transmission, incubation, viral shedding and immunity

Equine influenza is contracted by inhalation of infectious virus aerosols (Bryant et al., 2010, Mumford et al., 1990). The typical incubation period is 1–3 days (Crouch et al., 2004, McQueen et al., 1966, Mumford et al., 1994, Paillot et al., 2006, Bryant et al., 2010). However, delayed onset of clinical signs of up to 5 days has been observed after low dose aerosol exposure (Mumford et al., 1990). Infection is more reliably achieved by inhalation of aerosolised virus than intranasal inoculation (Mumford et al., 1990), with a minimum infective dose of 10^2 EID_{50}/ml of inhaled aerosolised virus. Horses may shed the virus in nasal secretions 24 hours after infection (Crouch et al., 2004, Myers and Wilson, 2006, Mumford et al., 1990, Bryant et

al., 2010) and then for 5–7 days depending on dose and method of infection (Daly et al., 2011, Mumford et al., 1990, Bryant et al., 2010). Infected horses shed large amounts of virus (>10^3 EID_{50}/ml of swab extract) throughout the period that they are infectious (Wood et al., 2007). Horses infected through nasal inoculation excrete virus for longer than those exposed to low dose aerosols (Mumford et al., 1990), with upper estimates of the infectious period being 7–10 days (Myers and Wilson, 2006, Hannant and Mumford, 1996). There is no long-term asymptomatic carrier state and infection-induced immunity affords protection against re-infection by a homologous strain for over 12 months, even in the absence of high levels of circulating antibody (Daly et al., 2004, Hannant et al., 1988, Bryant et al., 2010). Vaccination-induced immunity is more short-lived. Whilst clinical signs are diminished following vaccination, sterile immunity is not achieved when challenged with a non-homologous strain, which is the most likely scenario in the field (Paillot et al., 2006, Bryant et al., 2010).

Equine influenza may be transmitted through both direct and indirect means. An infected coughing horse can spread the virus for up to 32 metres (Miller, 1965), and transmission on fomites (such as people, vehicles or equipment) was reported in two South African outbreaks (Guthrie, 2006, Guthrie et al., 1999, Hannant and Mumford, 1996). Although rapidly inactivated by sunlight, heat, cold, drying conditions and common disinfectants, equine influenza can survive in soil for up to 2 days and in tap water for 2 weeks (Yadav et al., 1993). Influenza A viruses have also been shown to remain viable for 2 days on hard non-porous surfaces (Bean et al., 1982).

Windborne spread of aerosolised equine influenza virus was suggested to have occurred over distances of 3.2 km in Jamaica (Dalglish, 1992) and 8 km in South Africa (Huntington, 1990). There is some consensus in the literature that airborne transmission of influenza viruses is at least possible; however, there is strong disagreement about its importance (Weber and Stilianakis, 2008). Despite an abundance of anecdote, very few studies have investigated the contribution of wind to the spread of equine influenza, or indeed any influenza viruses. Davis et

al. (2009) presented circumstantial evidence that the mean direction of epidemic equine influenza spread coincided with prevailing wind conditions at the time. A time-series analysis of human influenza A surveillance data identified an association between increasing wind velocity and increasing hospital admissions (du Prel et al., 2009), and a recent atmospheric dispersal modelling study estimated that 24% of transmission of highly pathogenic avian influenza within 25 km of an infected premises was via windborne aerosols (Ssematimba et al., 2012). Most of the research into the windborne spread of infectious diseases of animals has involved atmospheric dispersal modelling of foot-and-mouth disease spread from a limited number of sources (Gloster et al., 2011, Garner and Cannon, 1995, Gloster et al., 2010, Sanson et al., 2011a, Schley et al., 2009, Mikkelsen et al., 2003, Alexandersen et al., 2002).

1.2.4. Recent outbreaks of equine influenza

Between 1986–2006, the importation of horses by air from countries where equine influenza was endemic caused outbreaks in South Africa, India, Hong Kong, Dubai, the Philippines and Puerto Rico (Paillot et al., 2006, Daly et al., 2006, Daly et al., 2004). Prior to 2007, only three countries with substantial horse populations had remained free of equine influenza: Australia, New Zealand and Iceland (OIE, 2009). Australia's equine influenza free status was confirmed by paired serological testing conducted in 1997–98 (Christley et al., 2001, Animal Health Australia, 2007). Prior to the 2007 outbreak in Australia, Myers and Wilson (2006) commented that the equine influenza-free status of these geographically isolated countries was threatened by the global transportation of horses, with a risk of large outbreaks if quarantine measures failed.

1.3. The 2007 outbreak of equine influenza in Australia

1.3.1. Pre-movement ban: release from quarantine, initial spread and detection

Horses were first imported into Australia by sea at the time of European settlement in 1788; importation by air commenced in 1973 (Callinan, 2008). In 2007, the Australian equine population was estimated to comprise nearly 1 million domesticated horses, donkeys and mules

(Centre for International Economics, 2007), and between 300,000–400,000 feral horses (Moloney, 2011). Into this population, only 500–900 horses are imported annually from countries other than New Zealand (Callinan, 2008). Prior to 29 September 2007, vaccination of horses for equine influenza was not routinely practiced in Australia (Animal Health Australia, 2007). All horses imported into Australia from countries other than New Zealand had to be vaccinated against equine influenza within 4 months of entering pre-export quarantine (Animal Health Australia, 2007, Perkins et al., 2011). Therefore, other than those few horses that had recently undertaken or were preparing for international travel, the Australian horse population could be considered immunologically naïve to equine influenza (Moloney, 2011).

In early August 2007, a total of 79 horses were imported into Australia from Ireland, the UK, the USA and Japan (Callinan, 2008, Watson et al., 2011b). A consignment of 13 horses was imported Japan on 8 August 2007; nine were offloaded in Melbourne and entered Spotswood Quarantine Station, and 4 horses travelled on to Sydney where they entered Eastern Creek Quarantine Station (ECQS). Unbeknown to Australian quarantine authorities at the time (Britton, 2007), the first outbreak of EI in Japan since 1972 was already underway (Yamanaka et al., 2008). Affected areas included those around pre-export quarantine stations used for horses in the consignment that arrived in Australia from Japan on 8 August 2007 (Yamanaka et al., 2008, Callinan, 2008). Subsequent serological testing determined that at least one of the 4 horses that entered post-arrival quarantine in ECQS from Japan was subclinically infected at the time with a Florida sublineage clade 1 equine influenza A/H3N8 virus (Callinan, 2008). This triggered undetected transmission of EI amongst recently vaccinated horses within the Eastern Creek Quarantine Station (Watson et al., 2011b).

Clinical signs consistent with equine influenza were first observed in an Irish stallion in ECQS on 17 August 2007 (Callinan, 2008, Watson et al., 2011b), and confirmed by haemagglutination inhibition (HI) testing on 20 August 2007, and real-time reverse transcription-polymerase chain reaction assay (qRT-PCR) on 23 August 2007 (Jeggo et al., 2008). Subsequent serological

testing identified that this horse was seronegative on entry to the quarantine station, despite a documented vaccination history (Watson et al., 2011b). It is now clear, that by 17 August 2007 the virus had already escaped quarantine in Sydney and infected a horse/horses in the Australian equine population (Kirkland et al., 2011). The exact means of release of equine influenza virus from quarantine and the early infection of horses in the Australian population has been the subject of thorough epidemiological investigation and a government inquiry, with the findings contained in two separate detailed reports (Callinan, 2008, Equine Influenza Epidemiology Support Group, 2008). The Callinan Inquiry report states that the most likely explanation of events was that equine influenza *"escaped from Eastern Creek on the person, clothing or equipment of a groom, veterinarian, farrier or other person who had contact with an infected horse and who then left the Quarantine Station without cleaning or disinfecting adequately or at all."* Despite strenuous efforts, the Inquiry was unable to identify a specific person or piece of equipment that left ECQS and was therefore responsible for infecting the first horse/s in the Australian horse population (Callinan, 2008). Several alternative scenarios were considered possible but unlikely causes: airborne spread from ECQS or Sydney airport, transmission on fomites used to transfer the horses between Sydney airport and ECQS, and cross-infection by dogs, birds or some other vector or vehicle (Callinan, 2008).

What is known from retrospective epidemiological investigations is that transmission of equine influenza definitely occurred at an equestrian event held at Maitland (160 km north of Sydney, Figure 1.1) between 17–19 August 2007 (Britton et al., 2011, Kirkland et al., 2011). Two horses were observed with clinical signs consistent with equine influenza on 22 August 2007 in a large horse boarding and training facility in central Sydney (Centennial Parklands Equestrian Centre, CPEC). These two horses, the index cases in the Australian horse population, had attended the equestrian event at Maitland the previous weekend and their infection with equine influenza was suspected by a veterinarian and subsequently confirmed by laboratory testing late on 24 August 2007 (Britton, 2007, Kirkland et al., 2011). A national horse movement standstill was implemented from 25 August 2007 (Equine Influenza Epidemiology Support Group, 2008) in

accordance with the 2007 version of the Australian Veterinary Emergency Plan (Animal Health Australia, 2007). Of the 208 horses that competed at the Maitland equestrian event, a further 44 horses were observed with clinical signs consistent with equine influenza in the week following the event (Britton, 2007, Britton et al., 2011).

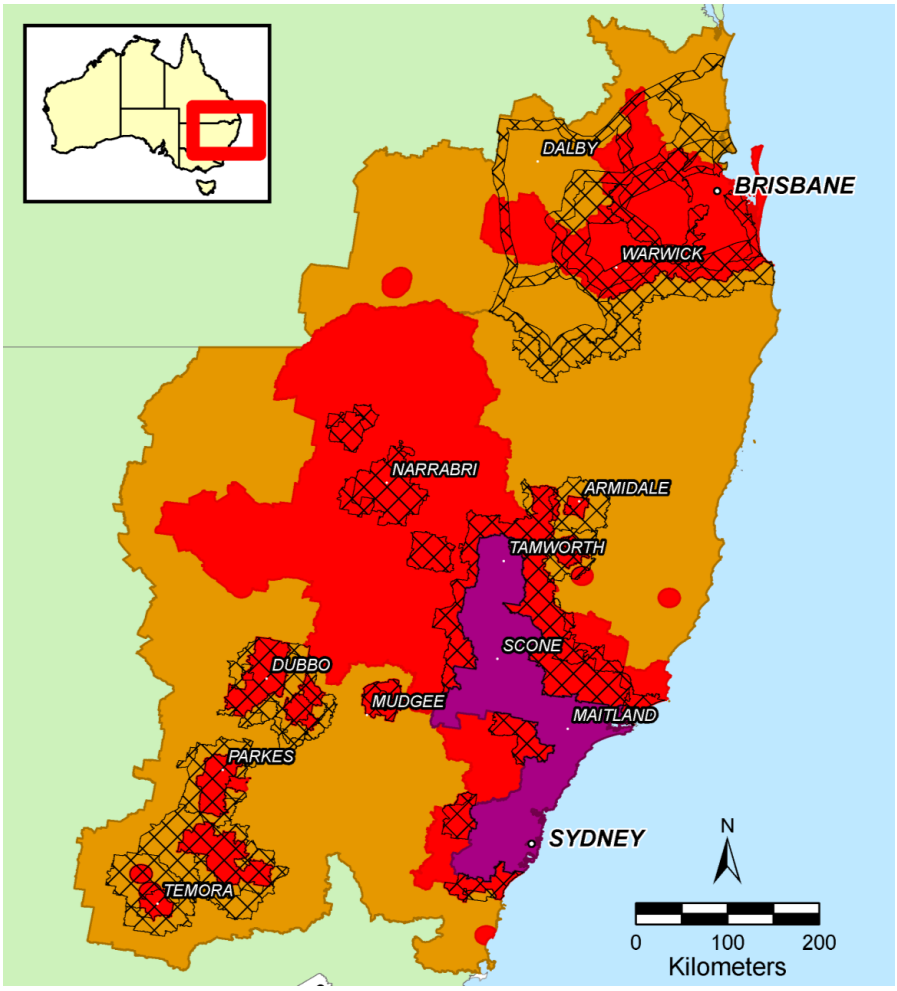


Figure 1.1 Peak movement restriction zones and vaccination buffers (hatched areas) from the 2007 equine influenza outbreak in Australia.
All infected premises were within the Restricted Areas (red and purple zones). After 21 September 2007 movement restrictions were relaxed in the Special Restricted Area (purple zone) according to the EI protection plan (NSW DPI, 2007b). The amber zone was designated a Control Area and green zone a Protected Area according to existing emergency animal disease management policy.

Nucleotide sequencing of the first Australian isolate of equine influenza virus (A/equine/Sydney/07) confirmed that it was similar to a 2003 Florida sublineage H3N8 isolate

from the USA (A/equine/Wisconsin/1/03) (Watson et al., 2011a, Jeggo et al., 2008), and follow-up testing showed that it was nearly identical to viral isolates obtained in August 2007 in Japan (A/equine/Ibaraki/07) and the USA (A/equine/Pennsylvania/07) (Ito et al., 2008, Watson et al., 2011a, Cook, 2008). This suggested that in August 2007 equine influenza arrived in Australia from Japan, having recently arrived there from the USA (Watson et al., 2011a).

1.3.2. Post-movement ban: Local spread, containment, vaccination and eradication

1.3.2.1. Descriptive analyses and field observations

The key components of the multi-faceted equine influenza control program were: (a) initial rapid response, outbreak investigation, contact-tracing and surveillance, (b) zone-based movement restrictions to contain spread, (c) rapid development, field testing and application of laboratory tests to guide risk-based zoning decisions and provide data for proof of freedom from disease, (d) a public awareness program that included outbreak communications targeting horse owners and the horse industry, providing updates and suggesting on-farm biosecurity measures to reduce local spread of disease, and (e) an emergency vaccination campaign.

The horse movement standstill did not manage to contain the outbreak to those premises where horses had returned from the Maitland equestrian event because infected horses had already been transported from Maitland to three further equestrian events, and from CPEC to an agricultural show, all held on the weekend of 25–26 August 2007 (Moloney et al., 2011, Equine Influenza Epidemiology Support Group, 2008). Two of these events and the agricultural show were held in New South Wales (NSW) and one event was held across the border in Queensland (QLD). Indeed, by the time the horse movement standstill had been completely implemented on 25 August 2007, infected horses had already left one of the events in northwest NSW and the event in southeast QLD (Equine Influenza Epidemiology Support Group, 2008, Moloney et al.,

2011, Kung et al., 2011). At this early stage in the response, when important strategic decisions were being made, authorities were only aware of three infected premises (IPs), but it has since been estimated that over 80 had already been infected (Garner et al., 2011b). These uncertainties undermined confidence during critical stages of the containment and eradication program (Glanville et al., 2009).

In the ensuing 19 week outbreak, equine influenza spread through the horse populations of the two affected Australian states (NSW & QLD). According to 2008 estimates, approximately 76,000 horses were infected on 10,651 IPs, whilst 281,000 horses were vaccinated and >100,000 laboratory tests conducted (Equine Influenza Epidemiology Support Group, 2008). A spatial and temporal analysis of the outbreak (Cowled et al., 2009b) describes the pattern of spread, delineates 37 regional clusters using spatial interpolation techniques (kriging based primarily on date of onset of clinical signs) and estimates epidemiological parameters (incidence, effective reproduction ratio, nearest neighbour spread distances and features of the epidemic curve) for each cluster. Cowled et al. (2009b) describe how spread across most of the eventual geographic extent of the outbreak occurred undetected within the first 10 days (17–26 August 2007) associated with the movement of infected horses prior to the complete implementation of horse movement restrictions. Their characterisation of the epidemic described regional heterogeneities in the course of the epidemic, especially when comparing rural and peri-urban regions (Cowled et al., 2009b). In different clusters, premises-level cumulative incidence ranged from 6–47%, whilst spread distances of 0.6–15.6 km were described (Cowled et al., 2009b). In rural areas, where the density of horse premises was lower than in peri-urban areas, equine influenza spread over greater distances and epidemic duration was shorter; however, region-specific estimates of cumulative incidence were similar. When Cowled et al. (2009b) estimated the effective reproduction ratio (R) at the premises-level, as the mean number of secondary IPs each individual IP transmitted infection to, using a nearest neighbour method (Ward et al., 2009b), they found similar values in rural and peri-urban clusters ($R \approx 2.0$).

Once the movement of infected animals ceased, 'infill' of already affected areas by local spread (transmission by direct and close contact, on fomites, and presumably via windborne aerosol) over short distances was noted as the main pattern of spread (Equine Influenza Epidemiology Support Group, 2008, Moloney, 2011). Cowled et al. (2009b) described how during this phase of local spread there was a large increase in the number of IPs but little expansion in the affected area. On 25 August 2007, Restricted Areas (or red zones) were delineated by merging 10 km buffers around IPs in both affected States (Equine Influenza Epidemiology Support Group, 2008). According to the existing policy (Animal Health Australia, 2007), the whole of the State of NSW was simultaneously declared a Control Area (amber zones). Restricted Areas are relatively small declared areas (compared to Control Areas) that are subject to intense surveillance and movement controls (Animal Health Australia, 2008). Control Areas are lower risk declared areas, surrounding all Restricted Areas, subject to less intense conditions (i.e. initially all horse movements were banned in Control Areas, however soon horse owners could apply for a permit to move their horses within this zone) (NSW DPI, 2007b).

From the 21 September 2007, gradual risk-based adaptation of the movement restriction policy was implemented (Garner et al., 2011c, Equine Influenza Epidemiology Support Group, 2008, NSW DPI, 2007b). Initially, two highly affected areas were designated Special Restricted Areas (purple zones). These were soon merged into a central corridor of the Restricted Area in NSW (see Figure 1.1). Unrestricted horse movements were allowed within this high horse density area because it was considered that further control of spread between properties was nearly impossible and the impact of continuing movement restrictions on the Thoroughbred horse breeding season would far outweigh the benefits of such restrictions (Equine Influenza Epidemiology Support Group, 2008, Animal Health Australia, 2011). Most of the remainder of NSW was redesignated as a Protected Area (green zone), free from disease, with a Control Area maintained as a wide buffer around all highly affected areas (Equine Influenza Epidemiology Support Group, 2008, Garner et al., 2011c).

Affected horses were typically febrile, with a distinctive cough and serous nasal discharge. There were variable degrees of lethargy and inappetance. Clinical signs typically resolved within 2–4 weeks (Gilkerson, 2011, Wong, 2011, Major and Jones, 2011, Faehrmann et al., 2011, Dups et al., 2011). Pyrexia was often the first clinical sign observed (Dups et al., 2011). Few horses exhibited coughing, pyrexia and nasal discharge simultaneously (Dups et al., 2011). More severe clinical signs were noted in horses in more intensive settings, and under stress, including competition and training (Anthony, 2011). Only a small number of horse deaths were attributed to equine influenza infection (Axon et al., 2008, Gilkerson, 2011, Smyth and Dagley, 2011). A case series of 35 deaths is reported (Begg et al., 2011), including a cluster of foal mortalities associated with bronchointerstitial pneumonia and cases of older horses suffering from secondary bacterial pneumonia.

Incubation periods of <48 hours were observed in horses that were systematically monitored for clinical signs at the quarantined equestrian event in QLD (Equine Influenza Epidemiology Support Group, 2008, Dups et al., 2011). Similar incubation periods have been observed in ponies experimentally infected with the outbreak strain (Bryant et al., 2010). Longer incubation periods (4–7 days) were observed in horses that attended the equestrian event at Maitland (Britton, 2007), suggesting exposure to lower doses of virus or measurement bias due to delays in recognising and reporting clinical signs. The index case at CPEC had an incubation period of 4–5 days (Britton, 2007, Wong, 2011). However, once infection started propagating in this high density, confined and naïve population all 161 horses in the complex were clinically ill within 4 days (Wong, 2011, Britton, 2007). Similarly, all 255 horses quarantined at the Warwick event showed signs of infection within 9 days of the presumed introduction of infection (Equine Influenza Epidemiology Support Group, 2008, Kung et al., 2011, Morton et al., 2011). Within 14 days of the onset of clinical signs detected in the first horse in a large racing complex in central Sydney, 623 of 641 Thoroughbred racehorses (97.2%) held in 19 facilities on the racecourse and in surrounding streets were showing ≥1 clinical sign of infection (Ryan, 2011). In this large population, all horses had positive serology and nasal swabs were negative for

qRT-PCR 53 days after onset in the first horse at the racecourse (Ryan, 2011).

A seroprevalence study that tested all horses on 5 horse studs in the Hunter Valley region, the major Thoroughbred breeding region of NSW, found a similarly high seroprevalence (93.3%) in extensive environs (Happold and Rubira, 2011). Research conducted during clearance testing found that all horses were seropositive on around 90% of IPs (Equine Influenza Epidemiology Support Group, 2008, Dhand and Sergeant, 2011b).

Five naïve ponies experimentally infected with a representative outbreak isolate (A/eq/Sydney/2888-8/07) shed live virus for 3–7 days; viral RNA was detected by qRT-PCR from 1–10 days post challenge (Bryant et al., 2010). Repeated testing of a population of police horses infected during the 2007 outbreak demonstrated that viral genome (of unknown viability) was continuously shed for 10 days, and could be detected with qRT-PCR in some horses as long as 35 days after the onset of clinical signs (Equine Influenza Epidemiology Support Group, 2008, Faehrmann et al., 2011).

1.3.2.2. Mechanisms of local spread

There were very few cases in which local spread to a particular IP could be definitively attributed to a particular mechanism (Equine Influenza Epidemiology Support Group, 2008, Kung et al., 2011). Only on certain isolated IPs and several closely monitored closed horse populations could spread via fomites be attributed (Kirkland et al., 2011, Ryan, 2011, Equine Influenza Epidemiology Support Group, 2008, Britton, 2007, Arthur and Suann, 2011, Faehrmann et al., 2011). Whether this was on peoples' clothes, vehicles or horse equipment remains unknown. Spread along major roads linking rural communities was considered evidence of poor biosecurity and human-associated spread (Equine Influenza Epidemiology Support Group, 2008). Some cases of spread were suspected to have resulted from the unauthorised movement of horses, although no specific instances were identified (Kung et al., 2011).

Airborne spread was suspected in high horse density areas over short distances (<2 km, and rarely up to 5 km) when there were no contiguous IPs and no known human contact between horses on different premises (Equine Influenza Epidemiology Support Group, 2008, Britton, 2007, Moloney et al., 2011). Horses with clinical signs of EI were observed on two IPs within 2 km of Eastern Creek Quarantine Station in late August 2007 (Britton, 2007, Moloney et al., 2011). Based on their onset dates and that no other epidemiological links could be identified, windborne spread from ECQS was suspected in both instances and it was considered that these were separate events from the original release from quarantine that led to the infection of horses at the Maitland equestrian event (Moloney et al., 2011). Further circumstantial evidence of windborne spread of equine influenza was presented in a detailed analysis of spatial spread in a single cluster of 437 mostly non-contiguous IPs in southeast Queensland (Davis et al., 2009). In their analysis, Davis et al. (2009) described the direction of spread based on the pattern of emergence of statistically significant space-time clusters of IPs and found this to be consistent with the predominant wind direction. Most local spread was observed within 2 km of existing IPs, with one instance of local spread over 9 km (Davis et al., 2009). The authors concluded that 10 km vaccination buffers around IPs during an outbreak appeared sufficient, and that whilst they did not conclusively identify the mechanism of spread, they stated that characterisation of the extent of local spread would be invaluable in guiding future response policies (Davis et al., 2009).

There are many anecdotal accounts associating local spread of EI with various environmental variables, including ambient temperature changes, air humidity, fog, wind conditions, season change and with mists along valley floors (Major and Jones, 2011, Wilson et al., 2011). Davis et al. (2009) tested for an association between relative humidity measurements and the time-series of reported IPs in the cluster of IPs that they investigated, however they found no clear association. Many clusters of infection were bounded by topographic features (Wilson et al., 2011), national parks or other land classes where horses were prohibited (East, 2009, Moloney

et al., 2008); therefore, caution is warranted when interpreting such circumstantial associations with the presumed directions of spread, which may reflect nothing more than the spatial distribution of the population at risk. Clarifying the relative importance of particular transmission mechanisms, particularly airborne spread, would inform risk-assessment about which biosecurity measures should be applied at quarantine stations (Equine Influenza Epidemiology Support Group, 2008). The epidemiology report to the CCEAD called for further, more complex analyses in this area (Equine Influenza Epidemiology Support Group, 2008).

Prior to this thesis, only preliminary risk factor studies (Spokes et al., 2009, Dhand and Sergeant, 2011b) had been conducted into local spread during this outbreak. Dhand and Sergeant (2011) found that on IPs, the proportion of horses at risk that became infected (the 'attack risk') differed by region and the number of horses held at the time of the outbreak. The authors suggested that further research would provide insights into the epidemiology of equine influenza, and how infection spread in such a naïve population (Dhand and Sergeant, 2011b).

The role of other animals (dogs, birds and insects) as vectors or mechanical vehicles of equine influenza spread after the movement ban was implemented remains unknown, but was considered of little importance to spread (Equine Influenza Epidemiology Support Group, 2008). Ten dogs from four IPs around Sydney (all large horse stables) were noted to have clinical signs consistent with cross-infection with equine influenza virus (Kirkland et al., 2010, Crispe et al., 2011). All recovered uneventfully. A serological survey of 40 dogs from these 4 IPs identified 28 dogs that were positive based on HI testing or blocking enzyme linked immunosorbent assay (ELISA) (Kirkland et al., 2010). Nasal swabs from one dog were positive based on qRT-PCR testing on 2 consecutive days, however, attempts to isolate virus from these swabs failed (Kirkland et al., 2010). Unlike the recent adaptation of equine influenza in the USA that lead to sustained dog-to-dog transmission (Payungporn et al., 2008), no changes in nucleotide gene sequences were identified when comparing isolates from an infected dog and the infected horse with which it had been in contact, and there was no evidence of lateral

transmission to dogs that did not have direct contact with infected horses (Kirkland et al., 2010). A small hypothesis-generating study presented very weak anecdotal evidence that transmission via birds may have been possible to one IP in southwest Sydney (Spokes et al., 2009).

Australia has one of the world's largest populations of feral horses. There was concern at the time of the 2007 outbreak that feral horses might become infected when horses were infected on IPs close to national parks and State forests where feral horses reside in both affected States (Gilchrist and Sergeant, 2011, Kung et al., 2011). Due to anticipated difficulties vaccinating this population, this may have led to the establishment of small foci of equine influenza infection that could not be easily eradicated (Kung et al., 2011). Risk assessments concluded that whilst the risk of domestic horses infecting feral horses was extremely low, and that infection was unlikely to be sustained if introduced into feral horse populations, there was a high risk that any infected feral horses would spread infection back to domestic horses at some later stage (Department of Agriculture Forestry and Fisheries, 2008, Gilchrist and Sergeant, 2011). Serosurveillance and monitoring of five populations of feral horses in close proximity to IPs was conducted, and there was no suspicion at any stage that the virus had infected feral horses (Department of Agriculture Forestry and Fisheries, 2008, Gilchrist and Sergeant, 2011, Kung et al., 2011). No detail is provided in the literature of the number of feral horses tested, Kung et al. (2011) note the locations of the populations tested in Queensland, and that sampling was conducted opportunistically during the outbreak period.

1.3.2.3. Vaccination – implementation and effectiveness

Temporal analysis of the weekly rate of growth of the epidemic (based on traces of the estimated dissemination 'rate', which compares the count of IPs reported in consecutive time periods) described how in both affected States the epidemic peaked in its seventh week (Equine Influenza Epidemiology Support Group, 2008). This was prior to the administration of any substantial numbers of vaccine doses (Equine Influenza Epidemiology Support Group, 2008), which commenced on 29 September 2007. The strict implementation of movement restrictions,

and the compliance of the horse industry and horse owners with on-farm biosecurity, were considered critically important in containing the spread of the outbreak, whereas vaccination was considered to have played a relatively minor role later in the outbreak ensuring the timely achievement of eradication (Kung et al., 2011, Cowled et al., 2009b, Perkins et al., 2011, Equine Influenza Epidemiology Support Group, 2008).

At the time of the outbreak, no commercially available vaccines contained the identified strain of virus (Equine Influenza Epidemiology Support Group, 2008). Limited supplies of killed equine influenza vaccine were available in the country, and authorisations were issued for vaccination of only 122 high value breeding stallions (Perkins et al., 2011). Based on experiences of apparent vaccine breakdown observed in the outbreak of the same strain in Japan, mismatch of killed vaccine and outbreak strains (Bryant et al., 2010), the 122 stallions were likely to have only had partial protection. Serial monitoring of 27 racehorses previously vaccinated with killed vaccines at a large racing complex in Sydney, found that when infected, 22 showed clinical signs, and viral RNA was detected by qRT-PCR for up to 20 days after disease onset (Ryan, 2011).

After regulatory and supply issues were addressed, approximately 140,000 horses were administered the recombinant canarypox-vectored equine influenza vaccine (ProteqFlu®/ProteqFlu TE®, Merial Ltd., UK) delivered in a combination of ring, predictive (targeted) and blanket (mass) vaccination (Perkins et al., 2011, Equine Influenza Epidemiology Support Group, 2008). Predictive vaccination was targeted at groups of horses considered as having a high potential to contribute to spatial transmission of infection, including: large aggregations of horses, certain racehorse, breeding and competition horse populations, and horses owned by the police and other essential industry workers (Perkins et al., 2011).

The specific vaccine administered was selected because:

- it was anticipated that vaccine-induced immunity would be sufficient to reduce virus shedding and clinical signs following exposure in the field within 14 days of the initial dose (Edlund Toulemonde et al., 2005),
- the recent example of its application as a component of the successful 2003 equine influenza outbreak control and eradication program in South Africa (Guthrie, 2006), and
- that its use would enable 'differentiation of infected from vaccinated animals' (DIVA techniques) with the blocking ELISA when testing to demonstrate freedom from disease (Perkins et al., 2011, Kirkland and Delbridge, 2011).

Recombinant canarypox vectored vaccines have been genetically-engineered to enable the vaccine-induced immune response to better mimic the mixed humoral and cell-mediated immune response that follows natural infection (Minke et al., 2007, Edlund Toulemonde et al., 2005). In an experimental study, where naïve ponies were vaccinated with two doses of ProteqFlu® administered 5 weeks apart, then challenged with the Australian outbreak strain, clinical signs and the amount and duration of viral shedding were markedly reduced compared to controls (Bryant et al., 2010). The authors considered that the partial protective immunity achieved to be sufficient to reduce the spread of disease as used in the face of the 2007 outbreak in Australia (Bryant et al., 2010). However, sterile immunity was not achieved, and a single dose of ProteqFlu® was considered to induce insufficient antibody levels to protect naïve horses against virus infection in an outbreak setting (Bryant et al., 2010). An accelerated ('off-label') vaccination schedule, based on a 2 week interval between doses, was used in some areas during the 2007 outbreak and has been shown in a vaccination trial to produce levels of immunity comparable to the standard schedule (El-Hage et al., 2009). In a number of the foals and mature horses studied, protective immunity was achieved after only a single dose of vaccine (Minke et al., 2011).

Several other factors are important when considering the effectiveness of vaccination in the face of such a large-scale outbreak. The immune responses achieved in experimental vaccination trials conducted under 'optimal conditions' are unlikely to be achieved 'in the field', where horses are exposed at different points in their vaccination schedules (Bryant et al., 2010). In the 2007 outbreak, many horses received only a single primary dose before being exposed to virus (Kannegieter et al., 2011). As a result of predictive and blanket vaccination, some horses were inadvertently vaccinated whilst incubating disease or when they were at very high risk of infection from nearby IPs (Perkins et al., 2011). This provided opportunity to observe the effect of ProteqFlu® vaccine amongst horses exposed to circulating virus at varying times post-vaccination (Perkins et al., 2011).

Anecdotally, little observable difference was noted in the severity of clinical signs between unvaccinated horses and horses with onset of disease within 48 hours of vaccination (Perkins et al., 2011, Garner et al., 2011b). Less prominent clinical signs and reduced viral shedding was observed in a population of horses exposed approximately 6 days after their primary vaccination (Kannegieter et al., 2011), and field reports indicated that little to no clinical signs of disease were evident in horses considered to have been exposed to circulating virus >7 days after primary vaccination (Perkins et al., 2011, Garner et al., 2011b). In a study of a cluster of infection in southeast Queensland (Davis et al., 2009), horses on only 3 of 52 vaccinated premises were found to have been infected following vaccination, and all within 14 days of their primary vaccination. Further research was suggested into the relative impact of vaccination and movement restrictions in containing and eradicating equine influenza from Australia (Cowled et al., 2009b), and to ascertain the effectiveness of vaccination with ProteqFlu® in the field (Perkins et al., 2011).

Stochastic simulation has been applied to study the effectiveness of several vaccination regimes earlier in the outbreak (Garner et al., 2011b). Garner et al. (2011b) developed and validated a model to replicate the spread of equine influenza in the 2007 outbreak, forwards from the point

of detection. They concluded that movement restrictions and on-farm biosecurity measures, as applied in 2007 were highly effective, and that if vaccination had been implemented early in the outbreak (according any of the modelled strategies) the clinical and economic impact of the outbreak would have been reduced (Garner et al., 2011b). Depending on resource constraints, a policy of ring vaccination of all horses within 1–3 km of IPs was found to be the more effective than vaccinating horses in a band 7–10km from IPs, or targeting vaccination at all premises holding >10 horses within 20 km of IPs (Garner et al., 2011b).

Early in the vaccination program there were a number of outbreaks of infection on horse properties within several days of vaccination teams having visited (Kung et al., 2011). The possible role of the vaccination teams in spreading infection was investigated, but insufficient information was available to determine whether the visiting vaccination teams contributed to the spread (Kung et al., 2011). In their in depth analysis of a single cluster, Davis et al. (2009) found that dates of vaccination did not particularly correlate with onset dates on nearby IPs, which they presented as evidence that vaccination teams were not directly responsible for a large proportion of spread.

A small number of deaths and other severe adverse consequences were reported in horses within days of vaccination, however it was not possible to definitively determine whether vaccination caused these incidents (Perkins et al., 2011). Minor adverse events were considered under-reported. Anecdotal evidence from vaccination teams indicated that 25% of vaccinated horses experienced transient localised swelling at the infection site, <5% were lethargic for approximately 24 hours and small numbers of horses were inappetant, had elevated temperatures, muscle stiffness and distal limb oedema in the days immediately following vaccination (Perkins et al., 2011).

1.3.2.4. Laboratory testing

Three diagnostic assays were the cornerstone of the massive outbreak surveillance effort that enabled the control and eradication of equine influenza from Australia in 2007: qRT-PCR assay conducted to test for viral RNA on nasal swabs, and two serological tests to detect antibodies (a blocking ELISA and HI testing)(Kirkland, 2011). Virus isolation was also conducted to inform vaccine selection (Watson et al., 2011a), and to aid the rapid adaptation of a qRT-PCR for a wide range of influenza A strains (Heine et al., 2007) into a novel H3-specific assay using partial HA sequences of A/equine/Sydney/2007 that aligned with 105 other equine influenza virus H3 sequences available in GenBank at the start of the outbreak (Foord et al., 2009). Virus isolation is not suited to rapid detection of infected horses in an outbreak setting as it requires multiple passages through chicken eggs or tissue culture (Watson et al., 2011a).

The qRT-PCR was utilised for rapid detection of infectious horses during early contact-tracing and epidemiological investigations (Kirkland, 2011). The diagnostic sensitivity of the qRT-PCR was considered almost 100%, and the diagnostic specificity was also very high (>98%) (Sergeant and Wilson, 2011). Where clinical onset of disease had not been observed, combined interpretation of ELISA and RT-PCR results allowed estimation of time-frames of when infection had occurred on an IP, considering that most viral shedding would occur within the first 7–10 days of infection, and that detectable antibodies would not be expected to be present in all horses until >7 post infection (Kirkland, 2011, Read et al., 2011).

Serological testing was conducted in ‘clearance’ and ‘proof of freedom’ testing to establish the immune status of horses (Kirkland, 2011, Sergeant and Wilson, 2011). Differentiation of whether horses had acquired immunity through natural exposure or from vaccination was possible (see Table 1.1) because the recombinant vaccine only induced antibodies to the H antigens, and these antibodies were not detected by the blocking ELISA (Kirkland and Delbridge, 2011). Field evaluation of the blocking ELISA, originally developed from an ELISA

for avian influenza virus, showed that it had very high diagnostic sensitivity (>99%) and specificity (97%) (Sergeant et al., 2011a, Sergeant et al., 2009). Where all horses were sero-negative on the first visit to a premises for clearance testing purposes, sampling would be repeated within 2–3 weeks, to confirm that active infection was not present (Kirkland, 2011).

Table 1.1 Differentiating vaccinated from naturally infected animals as conducted in the 2007 outbreak of equine influenza in Australia.

Individual horse status	HI test result	ELISA result
Naturally infected	+	+
Vaccinated with ProteqFlu® (Merial)[a]	+	–
Neither infected nor recently vaccinated	–	–
bELISA false positive	–	+

(Table adapted from Cowled et al., 2009a)
HI = Haemagglutination inhibition test, ELISA = blocking enzyme-linked immunosorbent assay
[a] Recombinant canarypox vectored equine influenza vaccine, that induces antibodies only to haemagglutinin surface glycoproteins (not detected by the influenza A blocking ELISA).

1.3.2.5. On-farm biosecurity

Engagement of the horse industry and public ensured community support for biosecurity measures, and was considered an important factor that enabled the eradication of equine influenza from Australia (Kung et al., 2011). There was a risk that disease control measures might cause severe social disruption to private horse owners, and both social and economic losses for those involved in horse breeding and racing because the peak of the outbreak coincided with the Thoroughbred horse breeding season and the Spring Racing Carnival (Arthur and Suann, 2011). Very stringent biosecurity measures were adopted at four horse training and racing precincts around the Sydney area, which allowed racing and training to continue for approximately 1400 horses (Arthur and Suann, 2011). Preventive measures included: finding alternative places of work or compensation for people in contact with infected horses elsewhere, road blocks for strict access control and disinfection of vehicles, reducing non-essential visitors, prescribing a period of non-contact with horses elsewhere and ensuring that people entering the

precinct showered, changed into clean clothes, cleaned and disinfected their boots, hands, forearms and all equipment on arrival (Arthur and Suann, 2011). Arthur and Suann (2011) estimate that following the implementation of these measures, infection was prevented for 3–5 weeks, despite highly affected horse populations surrounding the four biosecurity racing precincts, and even though trainers and other staff were residing in areas within the purple zone where disease was rapidly spreading.

A small retrospective cohort study to assess the effectiveness of biosecurity measures was conducted on people tending horses at the Warwick equestrian event in QLD, identified from entry and exit logs, who were also tending horses at their home properties (Frazer et al., 2011). These mostly recreational horse owners practiced stringent personal decontamination when leaving the quarantined premises, including showering and changing clothes, and minimising contact with horses elsewhere for 48 hours (Frazer et al., 2011). Of the 11 eligible respondents, six reported equine influenza infection in the horses on their home premises (55%), all practiced hand-washing and disinfection of footwear, however not all showered and changed clothes when leaving the quarantined horse facility. Nonetheless, the biosecurity measures appeared effective, as there was an interval of 21 days between the last date of onset of clinical signs in the quarantined facility and the first reported onset on any of the home premises. The authors considered that the six IPs were infected by local spread from surrounding IPs rather than by fomite transmission from the quarantined horse facility. Despite a very high prevalence of high psychological distress in horse owners and horse industry participants being observed in an online survey conducted during the outbreak (34%, compared to 10–12% in the general Australian population) (Taylor et al., 2008), respondents felt that biosecurity measures such as personal disinfection were likely to be effective and that they were committed to controlling and eradicating the disease (Frazer et al., 2011).

The effectiveness of recommended on-farm biosecurity was considered poor during the peak of the epidemic in the high horse density, highly affected purple zone (Equine Influenza

Epidemiology Support Group, 2008). In contrast to the few documented examples of fomite transmission over longer ranges, it was not considered possible to determine the contribution of poor biosecurity to local spread over short ranges (<4 km) (Equine Influenza Epidemiology Support Group, 2008), further research was clearly warranted.

1.3.3. Summary

Equine influenza was eradicated from Australia in less than 5 months (Department of Agriculture Forestry and Fisheries, 2008). The onset of clinical signs in the last confirmed case was on 25 December 2007 in Queensland (Equine Influenza Epidemiology Support Group, 2008). Movement restrictions were completely lifted on 28 February 2008, and serological surveying for proof of freedom purposes continued until declaration of freedom on 1 July 2008 (Department of Agriculture Forestry and Fisheries, 2008). The combination of early detection and diagnosis, rapid implementation of movement restrictions, on-farm biosecurity, vaccination and clearance testing, has been credited with enabling eradication to occur in such a timely manner whilst minimising the impacts on horse owners and the horse industry (Equine Influenza Epidemiology Support Group, 2008).

1.4. Epidemiological approaches to emergency animal disease events

An emergency animal disease (EAD) event may be caused by: a known *exotic* disease (present elsewhere) gaining entry to a highly susceptible population, an entirely new or previously unrecognised *emerging* disease, the *re-emergence* of a new variant of a known endemic disease, or an endemic disease gaining access to a previously unexposed population (Australian Animal Health Council Limited, 2001). Following detection of an EAD event, an emergency response is often required to prevent the serious social, economic, animal welfare or public health impacts that may be associated with a large-scale outbreak (Australian Animal Health Council Limited, 2001). In the context of EAD events, epidemiological analyses are used to inform decision-making, the consideration of the strategic objective(s) of future responses (eradication,

elimination or control) and how these objectives will be achieved (control options and implementation). Together with clinical, pathological, microbiological and experimental investigations, field epidemiological studies may aim to: estimate disease distribution, rates of morbidity and mortality; advance our understanding of disease causation and the underlying disease process; monitor the effectiveness and efficiency of interventions, predict the course of the outbreak or estimate the possible impacts of potential control options (Thrusfield, 2005). A range of approaches are used to achieve these objectives including risk factor studies, the temporal, spatial and network analysis of contact-tracing and outbreak surveillance data, and stochastic simulation modelling of outbreaks.

1.4.1. Risk factor studies

Epidemiological studies of disease causation may be classified as either descriptive or analytical depending on whether or not comparisons are made between the characteristics of sub-groups (Dohoo et al., 2009). Initial descriptive analyses of the 2007 equine influenza outbreak in Australia were conducted to generate working hypotheses on disease causation as a basis for rapid intervention (Britton, 2007), and to present clinical aspects of the disease (Wong, 2011, Begg et al., 2011, Faehrmann et al., 2011, Ryan, 2011, Major and Jones, 2011).

Analytical epidemiological studies are further categorised as either experimental or observational depending on whether treatments or interventions are allocated or the relationships are naturally observed (Dohoo et al., 2009). The key observational epidemiological study designs are cohort, case-control and cross-sectional studies. Cohort studies, in which exposed and unexposed groups are observed over time to measure and compare the occurrence of disease, are considered the most ideal observational epidemiological study design because exposure is known to precede outcome (Woodward, 2005). A cohort may be identified retrospectively (after disease detection), and either the whole cohort or a representative sample interviewed to identify subjects to exposure groups. After retrospectively

identifying the cohort of Maitland equestrian event attendees from registration lists, Britton et al. conducted interviews in an attempt to elucidate risk factors and the source of infection at this event (Britton et al., 2011). A prospective cohort study design is even more ideal, but harder to achieve in an outbreak setting. In 1990, Morley et al. conducted a cross-sectional study of risk factors for equine influenza infection on horses following an epidemic at a racetrack in Saskatoon, Canada, and then prospectively recruited cohorts at the same racetrack during influenza epidemics in the following 2 years (Morley et al., 2000b).

A case-control study design is a more efficient and cost-effective option for rare diseases and in large outbreaks (Dohoo et al., 2009). Groups of cases and non-cases ('controls') are sampled from the population, and exposure to putative risk factors are compared between these groups. The unit of interest in a case-control study may be the individual animal (Newton et al., 2003, Christley et al., 2001, Barquero et al., 2007) or organisational units such as farms or herds (Henning et al., 2009, Kung et al., 2007, Ferns et al., 1991, White et al., 2010). Recall bias and the method of control selection are major considerations in case-control studies. A case-control study may be conducted using incidence density sampling, whereby non-cases are selected and data collected as cases arise (Alford et al., 2001, Archer et al., 2008), thereby reducing the opportunity for differential recall bias. However it is more common in case-control studies of outbreaks for the disease event to have already occurred before the study commences. One method of preventing selection bias is by using a nested case-control study design (Newton et al., 2003), whereby case and controls are selected from a known source population, and the sampling fractions of each are known in advance. Differential selection bias may occur when cases are selected from a secondary study base (such as registry, hospital admission data, laboratory testing or passive surveillance data) and controls are sourced by random sampling of the source population. A suggested approach is therefore to select controls from the same secondary study base (Dohoo et al., 2009), ensuring that the exposure(s) of interest are independent of the diagnostic basis of the outcome (having the exposure is not associated with whether or not non-cases enter the secondary study base), as implemented in a number of

veterinary epidemiological case-control studies (Alford et al., 2001, Archer et al., 2008, Nodtvedt et al., 2007, Cohen et al., 2007).

In observational epidemiological studies, where subjects are non-randomly assigned to comparison groups, confounding bias is another major concern. Control of confounding can be achieved in case-control studies in design, by matching or restriction of the study to a population that is homologous with respect to a possible confounder; or in analysis, either through stratification or multivariable logistic regression. Matching involves selecting non-cases that are identical to the cases with respect to the distribution of one or more potential confounders. In case-control studies of risk factors for respiratory disease in racehorses, controls have been matched to cases by stable (Christley et al., 2001), and by trainer and proximity of horse yards (Newton et al., 2003). In case-control studies with aggregated units of interest, matching on location is often practiced to control for confounding and also for ease of sourcing population-based controls (Veling et al., 2002, Henning et al., 2009, Ferns et al., 1991). A limitation of this approach is that it may not be obvious exactly which factors (both known and unknown confounders) have been matched for. Once a factor is matched its effects cannot be studied separately, and where doubt exists it is suggested to control for confounding in analysis rather than to match (Miettine, 1970, Thrusfield, 2005).

The simplest observational epidemiological study design is a cross-sectional study, in which a population is sampled to seek data on disease status and risk factors at the same time. These studies cannot measure disease incidence, instead estimating differences in prevalence between groups, and often rely on subject recall in attempts to establish whether exposure preceded disease onset (Woodward, 2005). Recent examples of cross-sectional studies into outbreaks of equine influenza include the first in a series of three risk factor studies into influenza epidemics at a racetrack in Canada (Morley et al., 2000a, Morley et al., 2000b) and research into premises-level management and other environmental factors in outbreaks in Ireland between 2007–10 (Gildea et al., 2011).

1.4.2. Temporal analysis of outbreak surveillance data

The basic temporal analysis of an outbreak involves counting cases and constructing an epidemic curve (or time-series), then analysing for differences by region, as demonstrated by Cowled et al. (2009b). Related methods involve estimating the ratio of counts of cases in consecutive time intervals (the estimated dissemination 'rate' of an outbreak) (Equine Influenza Epidemiology Support Group, 2008, Kung et al., 2011, Thrusfield et al., 2005a), or plotting rolling averages of epidemic curves which are then compared to the timing of interventions (Taylor et al., 2004) or compared between sub-groups of premises with different characteristics (Ahmed et al., 2010, Mardones et al., 2009). These methods are relatively crude in that they depend only on counts of new cases and do not involve estimates of the population at risk. Surveillance data may also be analysed using time-series analysis (Murray and Morse, 2011), Poisson (Jensen et al., 2004) and logistic regression (Woodruff et al., 2002); however all of these methods are better suited to longer time-series than are typically available in a single EAD outbreak.

Survival analysis was used to describe the crude shape of the hazard over time in a closed population during the 2007 equine influenza outbreak (Morton et al., 2011), to compare weekly hazard for premises holding different types of livestock in the 2001 foot-and-mouth disease (FMD) outbreak in the UK (Wilesmith et al., 2003) and to test for an association between wind direction and outbreaks of epizootic haemorrhagic disease in cattle in Israel (Kedmi et al., 2010). Semi-parametric Cox regression modelling, as applied in an analysis of temporal aspects of the bovine spongiform encephalopathy epidemic in the United Kindgom (Wilesmith et al., 2000, Stevenson et al., 2000), allows multivariable estimation of the association between potential risk factors and the times to infection of individual subjects (animals or premises, depending on the unit of interest).

Time-dependent covariates (covariates that change in time, or whose effect changes in time) can

be included in the analysis by using a generalisation of the Cox proportional hazards model that requires restructuring the data into a counting process formulation (Hosmer et al., 2008, Therneau and Grambsch, 2000). This approach has been applied in a study of infectious salmon anaemia outbreaks in Canada (Hammell and Dohoo, 2005, Dohoo et al., 2009) and to investigate the association between time in training and the onset of musculoskeletal injuries in Thoroughbred racehorses (Cogger, 2006). In the counting process formulation, each subject contributes one observation for every day that it is at risk and each observation contains covariates for the subject at each time point of observation (Therneau and Grambsch, 2000) and the proportional hazards assumption does not apply (Allison, 2010). The partial likelihood specification for the counting process Cox regression model, including a term for each unique event time, is estimated by summing over those observations that are still at risk at each actual event time (Therneau and Grambsch, 2000).

The basic reproductive ratio (R_0) of an outbreak can be directly estimated based on the slope of the ascending branch of an epidemic curve (Ward et al., 2009b). R_0 describes the potential transmissibility of a disease in a totally susceptible population, and is defined as the number of secondary infections produced, on average, by a typical infective case assuming homogenous random mixing (i.e. equal likelihood of contact between premises) (Anderson and May, 1991). A related concept, the effective reproductive ratio (R) is the average number of secondary infections produced by each infected individual that enters a population that contains non-susceptible individuals or is subject to disease control measures (Matthews and Woolhouse, 2005). In veterinary epidemiological studies, R_0 and R may be estimated for individual animals (Satou and Nishiura, 2006) or functional groups of animals (Marquetoux et al., 2012, Mardones et al., 2011, Arroyo et al., 2011, Ward et al., 2009b), and can be adjusted to account for non-susceptible hosts, control measures such as vaccination, and non-homogenous mixing (Matthews and Woolhouse, 2005).

1.4.3. Spatial and spatio-temporal analysis of outbreak surveillance data

Numerous spatial and spatio-temporal analytical methods are available for describing the dynamics of epidemic spread in space and time, and identifying clusters of cases. Ideally, premises should be represented in spatial analysis by polygons of their boundaries (Pfeiffer et al., 2008). To reduce the complexity of analytical methods, centroids (the mean coordinates of a polygon's mass) (Cowled et al., 2009b) or the coordinates of a central feature such as the main farm building or farmhouse (Wilesmith et al., 2003) are commonly used to represent animal holdings as point data. This introduces a differential measurement bias for covariates that vary over the area of a premises or that relate to premises boundaries (such as adjacency and distance to other premises) which is most pronounced on larger or multi-site premises (Pfeiffer et al., 2008).

When point data on individual premises are available, smoothed estimates of spatial relative risk can be produced using kernel density estimation (Kelsall and Diggle, 1995). Areas of elevated risk are identified by estimating the ratio of Gaussian-smoothed kernel density surfaces of infected premises and dividing this by the surface of uninfected premises, as demonstrated in a study of the spread of FMD in two counties of the United Kingdom during the 2001 outbreak (Wilesmith et al., 2003). These methods are known to be sensitive to the amount of smoothing applied (bandwidth selection issues) (Bowman and Azzalini, 1997). There is a trade-off between over- and under-smoothing: large bandwidths may produce estimates that hide local features of the estimated density surface, whereas small bandwidths may introduce spurious peaks. A recent advance has seen the implementation of spatially-adaptive smoothing within widely available statistical software, whereby the amount of smoothing applied is varied across the study extent in inverse proportion to the density of the population at risk (Davies and Hazelton, 2009).

Spatial risk maps may also be plotted using a variety of methods that rely on estimating spatial

transmission kernels, which describe the decline in probability of infection with increasing distance from IPs, based on direct estimates from contact-tracing data, as demonstrated in analyses of the 2001 FMD outbreak in the United Kingom (Ferguson et al., 2001b, Keeling, 1999, Taylor et al., 2004, Savill et al., 2006, Ferguson et al., 2001a, Haydon et al., 2003, Kao, 2002) and the 2003 outbreak of highly pathogenic avian influenza (HPAI) in the Netherlands (Boender et al., 2007), or by fitting models to epidemic data (Diggle, 2006, Rorres et al., 2010). Transmission kernel estimates may then be used to parameterise stochastic simulation models (Keeling et al., 2001, Tildesley et al., 2011, Savill et al., 2006).

A spatial method of estimating the basic reproductive ratio of an outbreak that involves identifying the nearest infectious neighbour to each newly infected premises has been developed and applied to describe the spread of HPAI in Romania (Ward et al., 2009b). This 'nearest neighbour' method was implemented by Cowled et al. (2009b) to estimate the effective reproductive ratio of the 2007 outbreak of equine influenza in Australia, and has since been used to estimate the reproductive ratio of the HPAI outbreak in Thailand (Marquetoux et al., 2012).

Variography, a geostatistical technique, can be used to model the trend and underlying spatial dependency (autocorrelation) in spatially continuous data based on a covariance function (Matheron, 1963). Interpolation based on the semivariogram model ('kriging') has been applied to epidemic data to produce risk maps based on similarity in date of onset of IPs and provide insight into disease spread (Wilesmith et al., 2003, Ward et al., 2008), and also to delineate clusters of infected premises (Cowled et al., 2009b). The spatial dependency (or clustering) of IPs in space may also be estimated globally across the extent of an outbreak , or characterised in more detail through methods that involve estimating K-functions (Diggle and Chetwynd, 1991). K-functions express the expected number of events occurring within a given distance from a random event (Dohoo et al., 2009) and will be relatively high for clustered patterns because IPs are likely to be surrounded by other IPs (Stevenson, 2009). These methods have been applied in analyses of clustering of infected premises in a study of two small area affected during the 2001

FMD outbreak in the United Kingdom (Wilesmith et al., 2003), and epidemics of low pathogenic avian influenza in Italy (Mulatti et al., 2010). Others have constructed mixed-effects geostatistical models to analyse spatially-dependent epidemic data (Diggle and Ribeiro, 2007) and quantify disease risk by distance from IPs (Stevenson et al., 2005a).

Clusters, defined as 'aggregations of disease events in space and time unlikely to have occurred by chance' (Knox, 1989) may be described, or detected (tested for their statistical significance), depending on the method used. There are numerous methods of delineating clusters. Pfeiffer et al. (2008) discusses the relative merit of a range of approaches. The space-time scan statistic (Kulldorff and Nagarwalla, 1995) is the most-widely used method for detecting clusters across a study region. The scan statistic uses a circular scanning window to identify clusters as areas of elevated risk, the statistical significance of detected clusters is then tested with Monte Carlo simulations, and has found application in a number of veterinary epidemiological studies (Arroyo et al., 2011, Davis et al., 2009, Mulatti et al., 2010, Mardones et al., 2009, Ward and Farnsworth, 2009). To improve the power of detecting irregularly-shaped clusters a recent advance has seen implementation of a scan statistic for both circular and non-circular clusters (Takahashi et al., 2008).

Describing the dependence among outbreak data in space *and* time is important because disease transmission is more likely to occur between infectious and susceptible individuals that are close together in these two dimensions (Ward and Carpenter, 2000a, Pfeiffer et al., 2008). Ward and Carpenter (2000) present four approaches that may be used to describe space-time clustering in veterinary epidemiological data: the Mantel test (Mantel, 1967), Barton's method (Barton et al., 1965), a nearest-neighbour test (Clark and Evans, 1954) and Knox's test (Knox, 1964). The spatio-temporal interaction of infection risk may also be estimated using the space-time *K*-function (Diggle et al., 1995), as demonstrated in the investigation of outbreaks of FMD (Wilesmith et al., 2003) and HPAI (Minh et al., 2011). Spatial dependence has been incorporated into the Bayesian survival analysis of the 2001 FMD outbreak in Devon (UK) to

develop space-time predictions (McKinley, 2007). Individual farm-level spatio-temporal epidemic processes have also been statistically modelled using a probabilistic Markov chain Monte-Carlo model (Fenton et al., 2009) and by a partial likelihood estimation approach (Diggle, 2006).

Spatial and spatio-temporal analyses may be conducted on outbreak surveillance data that have been spatially-aggregated by areal units such as regions, districts or sub-districts (Stevenson et al., 2005b, Ahmed et al., 2010, Lawson and Zhou, 2005), or into data aggregated for reporting purposes into localised 'outbreaks' (Ward et al., 2008). However when premises-level polygon or point data are available, it has been shown that fine-scale spatial structure may be lost in the process of aggregation (Diggle, 1983, Henderson et al., 2002). Given the arbitrary nature of boundaries typically used for aggregation, it is highly likely that selecting other boundaries or methods of aggregation would produce different results (Dohoo et al., 2009), and the relationships observed in aggregated data may not hold at the individual level (ecological fallacy) (Woodward, 2005).

1.4.4. Social network analysis of outbreak surveillance data

Early in an EAD outbreak the network component of proximity may be important when describing the dynamics of disease spread through space. Such outbreaks are often first observed as several clusters of disease in widespread locations, with initial spread seemingly occurring through a contact or other network (such as market chains or major transportation routes) (Thrusfield et al., 2005b). When a complex contact network structure underlies an outbreak, traditional approaches may be insufficient to appropriately describe the spatio-temporal pattern of spread and estimate key parameters (Small et al., 2007, Kiss et al., 2009, Keeling and Eames, 2005). Social network analysis (SNA) is an approach suited to the investigation and modelling of interactions between animal and farms (Martinez-Lopez et al., 2009b, Dubé et al., 2011). The 'social' refers to the context of this method's early application in

infectious disease research (Klovdahl, 1985, Klovdahl et al., 1977, Klovdahl et al., 1994). When applied to contact-tracing data to build a putative transmission network, how the contact links ('potential transmissions') are defined must be clearly specified (Keeling and Eames, 2005). A non-representative sample of the full mixing network is constructed, with many dead-ends consisting of uninfected individuals. The putative transmission network also depends on animal owners providing complete and accurate data on contact frequency (Keeling and Eames, 2005).

Networks can be used to represent the patterns of connectivity of populations and therefore describe aspects of disease transmission that depart from the mean field model (Diekmann et al., 1998). Nodes may be used to represent individual animals (Porphyre et al., 2008) or farms (Ortiz-Pelaez et al., 2006), whilst connections can represent a wide range of context-dependant relationships and may be either directed or undirected depending on whether or not the object of interest can travel in both directions between a pair of nodes (Wasserman and Faust, 1994). The strength of social network analysis is its generalisability. Dynamic networks can be constructed to demonstrate temporal aspects of contact patterns (Kao et al., 2007, Saramäki and Kaski, 2005, Heath et al., 2008), and spatial proximity (Webb, 2005), and these can be statistically analysed in the same framework as other relationships between premises or individuals.

Several recent reviews describe in detail the application of network analysis techniques in epidemiological modelling (Keeling and Eames, 2005, Danon et al., 2011) and preventive veterinary medicine (Martinez-Lopez et al., 2009b, Dubé et al., 2009, Dubé et al., 2011). SNA has been used to study contact-tracing data from the 2001 outbreak of FMD in the UK (Green et al., 2006, Shirley and Rushton, 2005, Ortiz-Pelaez et al., 2006), the global spread of H5N1 avian influenza (Small et al., 2007), estimate the contribution of local spread and animal movements in the transmission of bovine tuberculosis (Green et al., 2008), model the efficacy of contact-tracing in a range of scenarios (Kiss et al., 2008, Kiss et al., 2006a, Kiss et al., 2005) and simulate EAD incursions (Christley et al., 2005, Sharkey et al., 2008). A growing number of network analysis studies investigate animal movements and contacts, either linked to

laboratory testing data or irrespective of disease status, to develop models of how a future outbreak might spread through animal populations (García Álvarez et al., 2011, Turner et al., 2008, Christley and French, 2003, Bigras-Poulin et al., 2007, Bigras-Poulin et al., 2006, Webb, 2005, Webb, 2006, Martinez-Lopez et al., 2009a, Lockhart et al., 2010, Woolhouse et al., 2005, Kiss et al., 2006b, Robinson and Christley, 2007, Robinson et al., 2007, Brennan et al., 2008, Dent et al., 2008, Natale et al., 2009, Heath et al., 2008, Nöremark et al., 2011, Dommergues et al., 2011). Very few of these network analyses have included analysis of the spatial proximity between infected premises (García Álvarez et al., 2011, Webb, 2005, Martinez-Lopez et al., 2009a, Lockhart et al., 2010), with only one such study investigating spatial proximity and contact relationships between infected premises during an actual epidemic in animals (Green et al., 2008).

1.4.5. Stochastic simulation modelling of outbreaks

Models of outbreaks incorporating random variables ('stochastic simulation models') have been under development since the 1950s (Carpenter, 2011). These models have been used to advance our understanding of underlying disease processes (Glass et al., 2002, Satou and Nishiura, 2006), to support planning of outbreak responses by comparing and evaluating control scenarios on simulated data (de la Rua-Domenech et al., 2000, Bates et al., 2003, Garner et al., 2011b), and for prediction to support decision-making during actual outbreaks (Keeling et al., 2001, Morris et al., 2001). Whilst models should only be as complex as required to fulfil the purpose for which they were developed, Carpenter (2011) noted that it is often necessary to explicitly incorporate non-random aspects of mixing (heterogeneity) into premise-level stochastic epidemic models to better reflect reality. Approaches have included the explicit representation of relationships between premises in terms of time *and* space (Sanson, 1994, Keeling et al., 2001, Morris et al., 2001, Bates et al., 2003, Garner and Beckett, 2005, Tildesley et al., 2011, Harvey et al., 2007, Schoenbaum and Disney, 2003), and explicitly representing the movement of animals and potential fomites between premises (Jewell et al., 2009b, Sharkey et al., 2008).

Following the application of stochastic simulation models to inform decision-making during the 2001 FMD outbreak in the UK, debate has surrounded the question of whether models should be used strategically in outbreak response planning, or as predictive tools to inform the response during an actual outbreak (Kao, 2002, Kitching et al., 2006, Taylor, 2003). A review of the use of models to inform the 2001 FMD outbreak response (Taylor, 2003) recommended that when predictive models were used to inform decision-making in the face of an outbreak, they should only be used to augment, rather than replace, detailed epidemiological investigations by providing comparisons of observed and modelled outcomes for the purpose of alerting epidemiologists to unexpected events, and aiding the targeting of interventions. The importance of empirical data quality has also been emphasised (Ward et al., 2009a).

Recent research into the Bayesian statistical inference of epidemics (Neal and Roberts, 2004) has presented novel methods of predicting undetected infections (Jewell et al., 2009a) and risk of infection (Jewell et al., 2009b). The benefit over previous predictive models is that models fitted by Bayesian inference can be populated with field data from an ongoing outbreak, and incorporate this with prior estimates, expert opinion, contact-tracing and other animal premises data into a single analysis framework (Jewell et al., 2009c). These models can then be used to develop meaningful forward predictions on the future course of the epidemic, including estimates of the level of uncertainty in these predictions (Jewell et al., 2009c).

Previous modelling of the 2007 equine influenza outbreak in Australia has been conducted, and validated, with the specific strategic purpose of recreating the spatial and temporal patterns of spread to enable retrospective evaluation of different control strategies (Garner et al., 2011b). Outbreaks of equine influenza at horse racing facilities in the United States, United Kingdom and Japan have been modelled at the individual horse level, to specifically describe intra-premises epidemiological processes and the effects of vaccination (Satou and Nishiura, 2006, Glass et al., 2002, de la Rua-Domenech et al., 2000), with extensions to encompass the

influences of seasonal changes in the population at risk (Park et al., 2003), and antigenic drift in vaccinated horse populations (Park et al., 2004). A larger more general metapopulation model of between yard spread of equine influenza has also been developed based on the 2003 outbreak at the Newmarket facility in the United Kingdom (Baguelin et al., 2010).

1.5. Conclusions

Although a large amount of epidemiological analysis was completed prior to commencement of this thesis, further research was suggested in a number of areas, including:

- Further cleaning of the equine influenza outbreak dataset to improve the accuracy of estimates of the population at risk and other key epidemiological parameters.
- A premise-level risk factor study to investigate property type and management factors, and estimate whether compliance with advised on-farm biosecurity measures prevented spread.
- Spatio-temporal analysis or stochastic simulation modelling to estimate the relative contribution of the various mechanisms of spread, including: determining the relative importance of local spread and animal movements prior to the movement ban, and a more detailed analysis into airborne spread and the contribution of wind.
- Research into the effect of vaccination with ProteqFlu® in the field, given that horses were exposed to varying amounts of a non-homologous virus at varying times post-vaccination.
- Development of methods of cluster mapping to support rapid epidemiological responses and outbreak communications. Such methods would need to be able to make use of contact-tracing data on infected animal movements.
- Estimating how far equine influenza spread spatially to enable assessment of whether control zones and vaccination buffers were set at an appropriate distance from IPs.
- Further characterisation of the epidemiological parameters of equine influenza spread in a naïve population, including the infectious period on multi-horse premises and the

intra-premises reproductive ratio (R_0) to ascertain whether appropriate IP quarantine periods were applied.

- Modelling the equine influenza outbreak dataset to determine what predictions could be made concerning future equine influenza outbreaks in Australia.

Chapter 2: The equine influenza outbreak dataset

– spatial assessment and preparation for epidemiological analyses

"*Not everything that counts can be counted, and not everything that can be counted counts.*"

Albert Einstein (attributed)

2. The equine influenza outbreak dataset – spatial assessment and preparation for epidemiological analyses

2.1. Introduction

During the 2007 outbreak of equine influenza in Australia, approximately 130,000 polymerase chain reaction assay, blocking enzyme linked immunosorbent assay and haemagglutination inhibition test results were collated with population at risk, contact-tracing, vaccination and outbreak surveillance data to form the equine influenza outbreak dataset. Although not collected for the purpose of epidemiological research, this data resource is the most complete animal outbreak dataset ever collated in Australia. Prior to this thesis, a series of retrospective studies of the spread of the 2007 outbreak were based on the equine influenza dataset (Cowled et al., 2009b, Kung et al., 2011, Moloney et al., 2011, Garner et al., 2011b, Wilson et al., 2011, Dhand and Sergeant, 2011b, Dhand and Sergeant, 2011a, East, 2009, Davis et al., 2009, Sergeant et al., 2009, Sergeant et al., 2011b, Sergeant and Wilson, 2011).

Detailed data on the distribution of horses in Australia were not available at the time of the 2007 outbreak (Kung et al., 2011, Moloney et al., 2011). Data on all horse premises in the States of New South Wales (NSW) and Queensland (QLD), the population at risk dataset, were compiled during the course of the outbreak by the government Departments of Primary Industries (DPIs) of the two affected States (Moloney, 2011). Both jurisdictions implemented web-based self-registration systems for horse owners and collated these records with lists of investigated horse premises, routine farm-animal population survey data, vaccination and horse movement permit data, horse industry organisation databases and address lists (Kung et al., 2011, Moloney et al., 2011). The population at risk dataset was aggregated to the premises-level, and contained data on the following covariates: IP status (infected or uninfected during the 2007 outbreak), suburb, geocoded coordinates (based on property centroid), number of horses, premises area, vaccination status, date of vaccination (*not supplied for QLD premises*), and for infected

premises, the date of onset of first clinical signs of the first horse affected ('onset date') (Cowled et al., 2009a). Early in the outbreak IP status was based on a case definition that included both classical signs of equine influenza (cough, rectal temperature >38.5 °C, nasal discharge and lethargy) and laboratory confirmation using real-time reverse transcription-polymerase chain reaction assay (Moloney, 2011). As the outbreak neared its peak, a case definition change was implemented to reduce the laboratory resource requirement involved in confirming all highly suspect premises. From the fifth week of the outbreak (10 September 2007), a premises could be classified as infected based on a report of classical clinical signs without laboratory confirmation (Moloney et al., 2011, Cowled et al., 2009b).

Some early analyses were based on only a subset of the population at risk data collated, by combining only the known IPs and vaccinated premises datasets from early 2008 (Wilson et al., 2011). Using this approach, Wilson et al. (2011) estimated that 58.4% of premises in the purple zone were naturally infected over the course of the outbreak, and that after vaccination, herd immunity was achieved on 84–87% of premises.

In preparation for a series of epidemiological analyses, Cowled et al. (2009a, 2009b) conducted rigorous cleaning to: remove duplicates, include further IPs identified from laboratory records and correct wrongly geocoded coordinates and typographic errors. Following this process, the population at risk dataset for the States of NSW and QLD contained around 120,000 premises holding an estimated 950,000 horses (Cowled et al., 2009a, Cowled et al., 2009b), and was considered a quasi-census of the population at risk in outbreak affected areas in 2007 (Garner et al., 2011b, Moloney et al., 2011). Of these premises 9599 were designated as IPs, holding around 69,000 horses infected during the outbreak (Garner et al., 2011b, Cowled et al., 2009a). Based on all available data, Cowled et al. (2009b) re-estimated the cumulative incidence in clusters within the purple zone to be between 12–42%, and estimated cumulative incidence, presented epidemic curves and the premises-level effective reproductive ratio for all other clusters of IPs in the 2007 outbreak.

The contact-tracing dataset included 1034 reported horse movements onto and off premises investigated by NSW and Queensland animal disease authorities prior to complete implementation of the horse movement standstill on day 10 of the outbreak (26 August 2007). Each movement record included the date of the movement, and the unique identifiers of the origin and destination premises. Most of these records were traced backwards during interviews of horse owners and managers whilst attempting to establish the source of infection and implement quarantine and horse movement restrictions.

Despite all of the considerable effort and resources that were devoted to collecting, collating and cleaning the equine influenza outbreak dataset, biases are still known to exist. These include underestimation of the population at risk because combined data sources could not be expected to provide 100% coverage (Cowled et al., 2009a), differential ascertainment of infected premises due to under-reporting (Dhand and Sergeant, 2011a, Cowled et al., 2009a) and incentives to clear infection in certain areas, missing data, misclassification and case definition changes (Cowled et al., 2009b, Moloney et al., 2011).

Whilst preparing the sampling frame for a premises-level case-control study (Chapter 3), spatial exploratory investigations identified that a large amount of duplication of uninfected premises (UIP) remained in the equine influenza outbreak dataset (see Figure 2.1). A number of premises were noted to have multiple entries. An example includes an uninfected premises where horses were kept for other owners (agistment), that had 10 duplicate entries, each with near identical coordinates and similar numbers of horses. Presumably each of these records had a different owner's name or address details (*these data were de-identified*). Further issues identified included premises that were geocoded on the road outside a property boundary, and a large number of premises that had duplicate records with discordant IP status (i.e. ≥2 records listing a property as both an IP and an UIP), often with different reporting dates, presumably due to change in status over time. It was considered that this duplication would bias the estimation of

key epidemiological parameters and planned spatio-temporal analyses in a highly unpredictable manner. The aims of this analysis were therefore to assess the accuracy of the equine influenza outbreak dataset using data on property boundaries, and to prepare the dataset for spatio-temporal analyses.

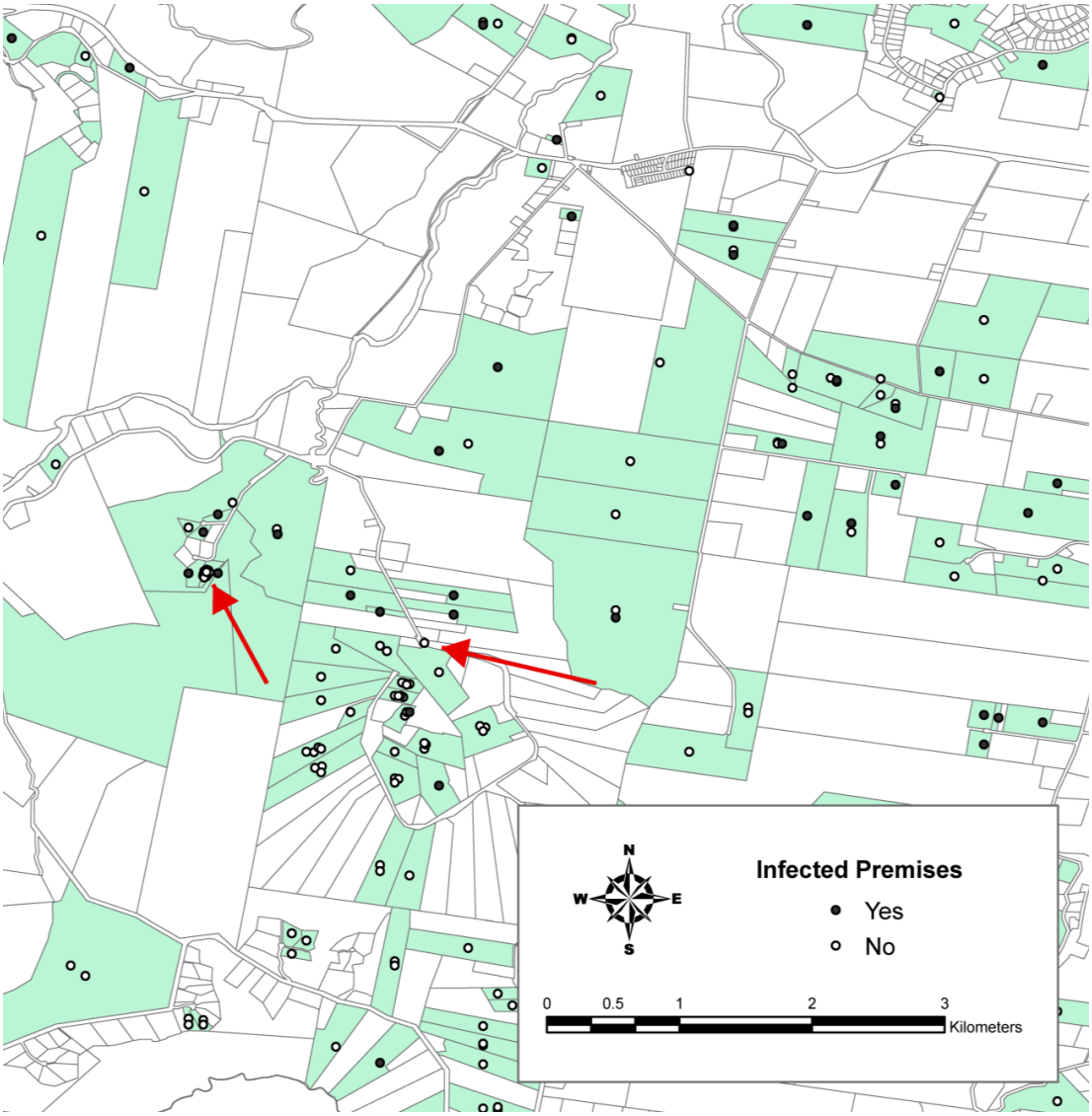


Figure 2.1 Duplication of premises records in the equine influenza outbreak dataset before aggregation by cadastral polygon.
This small area was from within the highly affected 'purple zone' (Special Restricted Area) of the 2007 outbreak in Australia. Most duplication involved uninfected premises records (white circles). The red arrows indicate two important errors in the data, a premises with coordinates that are on a road (not contained by a cadastral property boundary supplied by the NSW Government) and a single property boundary that contains 10 duplicated uninfected premises records.

2.2. Materials and methods

2.2.1. State property boundaries (cadastral) datasets

On request, the NSW Land and Property Management Authority and the QLD Department of Environment provided highly detailed spatial representations of all property boundaries (also known as cadastral or 'lot' polygons) within their jurisdictions, updated in 2009. Both datasets were originally developed between 1983–1990 by digitising the best available cadastral maps and are being continuously updated from registered plans of subdivision and from any other appropriate updates from government gazettes and other administrative notifications (NSW Land and Property Management Authority, 2009, QLD Department of Environment and Resource Management, 2009). Together, the NSW and QLD cadastral datasets contained around 5.5 million non-overlapping polygons, supplied in ESRI ArcGIS® shapefile format in the Geocentric Datum of Australia 1994 (Geoscience Australia, 2007), the same coordinate reference system used to collect the point data. For accuracy of area-based calculations, all coordinate data were projected into the Australian Albers conic equal-area projection. The equine influenza dataset was then linked to these spatial layers in ArcMap 9.3 (ESRI, Redlands, CA, USA) to create a relational geodatabase. All further spatial analyses were conducted in ArcMap 9.3.

2.2.2. Spatial data linkage and aggregation of premises records by cadastral polygon

A spatial join was made to link each premises record in the equine influenza outbreak dataset using its centroid coordinates (point) to the nearest cadastral polygon (without any spatial tolerance). This resulted in two fields being added to each premises record in the equine influenza outbreak dataset: the unique identifier of the nearest cadastral polygon ('CAD_ID') and the distance in metres to the nearest cadastral polygon ('CAD_DIST'). For premises records whose existing centroid was contained within a cadastral polygon 'CAD_DIST' was set to zero. For reasons of computational tractability, spatial linkage was conducted separately for each

State.

Data were then imported into the R statistical package version 2.12.0 (R Development Core Team, 2011), and a script was run to identify 'duplicated records' in the equine influenza outbreak dataset: premises records with centroid coordinates either contained within or nearest to the same cadastral polygon as another premises record. Data from duplicated records were consolidated into a new single premises record assigning the values of fields according to the following rules:

- A count field ('count') was added to each record and used to identify how many duplicated records had been consolidated to create this record (count =1 if no duplication, otherwise count >1).
- A text field containing a list of the original unique premises ID of each premises record consolidated into the new record was retained in a new field. This allowed linkage to original records and laboratory testing data.
- 'IP status': If any of the original records listed the premises as an IP, the revised record was listed as an IP.
- 'Date of onset': For IPs, the earliest recorded date of onset was retained from all records.
- 'Number of horses': For each duplicated set of premises records, the median of all non-zero entries was entered into a new field and used for all further analyses.
- 'Vaccination status': If any of the original records listed the premises as vaccinated, then the revised record was listed as vaccinated. The earliest date of vaccination was entered as the revised 'Date of vaccination'.

The number of premises cleaned from the dataset was tabulated, by State and IP status, and cumulative equine influenza incidence was recalculated at the individual horse and premises level, for comparison with earlier estimates, and the revised dataset assessed for missing data.

The revised dataset was then linked to the cadastral polygon data, now with only one premises record within each property polygon, and no premises outside an existing property polygon in either of the States' datasets. Centroid coordinates were recalculated based on the cadastral polygon, and fields were added to the dataset listing the number of adjacent premises that held horses at the time of the outbreak and the length of shared fence with such premises. Crude exploratory visualisation of the spatial distribution of the duplicated entries and the revised population at risk was conducted by kernel density estimation (Kelsall and Diggle, 1995) of the revised point data based on a grid of cell size 5 km × 5 km using an arbitrarily selected bandwidth of 20 km, in ArcMap 9.3.

2.3. Results

The supplied centroid coordinates of 94.8% of the 120,891 premises records in the equine influenza dataset were contained by an existing cadastral polygon (Table 2.1). Of the 6281 'uncontained' premises records identified, 5167 (82.2%) were uninfected premises from NSW. The median distance between the nearest cadastral polygon boundary and the original centroid coordinates of the uncontained premises was 7 m (IQR: 4–12 m, maximum 8.3 km).

A total of 17,169 'duplicated' premises records were identified (Table 2.2). Again, nearly all of these duplicated premises records (87.9%) were uninfected premises from NSW. After aggregating records by cadastral polygon, the population at risk was re-estimated to be 103,722 horse premises, of which 9359 premises were estimated to have held horses infected in the 2007 outbreak of equine influenza.

A smoothed kernel density surface of the distribution of horse premises in the two affected States, after aggregation by cadastral property boundaries to remove duplication, is presented in Figure 2.2a. Most of the duplication appeared to be localised to regions that were highly affected during the 2007 outbreak (Figure 2.2b).

The total number of horses in NSW and Queensland at the time of the 2007 outbreak of equine influenza was re-estimated to be nearly 800,000. Of these, around 67,000 were estimated to have been infected with equine influenza (Table 2.3). Aggregating records by cadastral polygon increased the cumulative (premises-level) incidence of equine influenza in the most highly affected control zone (the Special Restricted Area, or purple zone) (Table 2.4). There was little change in the estimated cumulative incidence of equine influenza in the Restricted Area (red zone) after aggregation. The revised equine influenza outbreak dataset was mostly complete for premises-level variables except for self-reported property type ('SITE_TYPE') (Table 2.5).

Table 2.1 Premises records from the original equine influenza outbreak dataset, by whether or not their centroid coordinates were contained by a property polygon in the digital cadastral database of NSW or QLD.

State	**Premises infection status**	**Original premises centroid[a] contained within a cadastral polygon**				
		Yes		No		Totals
NSW						
	Infected	6 152	(*97.5%*)	160	(*2.5%*)	6 312
	Uninfected	46 448	(*90.0%*)	5 167	(*10.0%*)	51 615
QLD						
	Infected	3 281	(*99.8%*)	6	(*0.2%*)	3 287
	Uninfected	58 729	(*98.4%*)	948	(*1.6%*)	59 677
Total		114 610	(*94.8%*)	6 281	(*5.2%*)	120 891

NSW = New South Wales, QLD = Queensland
[a] Based on data supplied by Cowled et al., 2009a.

Table 2.2 Duplicated premises records in the original equine influenza outbreak dataset, by State and infected premises status.

	Total premises (original estimates [a])			Duplicated premises records			Total premises (revised)		
State	IPs	UIPs	Total	IPs	UIPs	Total	IPs	UIPs	Total
NSW	6 312	51 615	57 927	220	15 093	15 313	6 092	36 522	42 614
QLD	3 287	59 677	62 964	20	1 836	1 856	3 267	57 841	61 108
Total	9 599	111 292	120 891	240	16 929	17 169	9 359	94 363	103 722

IPs = Infected premises, UIPs = Uninfected premises
[a] Based on data supplied by Cowled et al., 2009a.

Table 2.3 Comparison of estimates of horse numbers, by State and infection status during the 2007 outbreak of equine influenza in Australia, before and after aggregation of records by cadastral polygon to reduce duplication.

State	**Original estimates (horses)**[a]			**Revised estimates (horses)**		
	IPs	UIPs	All premises	IPs	UIPs	All premises
NSW	45 799	233 876	279 675	42 763	153 455	196 218
QLD	23 367	623 272	646 639	24 284	577 588	601 872
Total	69 166	857 148	926 314	67 047	731 043	798 090

IPs = Infected premises, UIPs = Uninfected premises
[a] Based on data supplied by Cowled et al., 2009a.

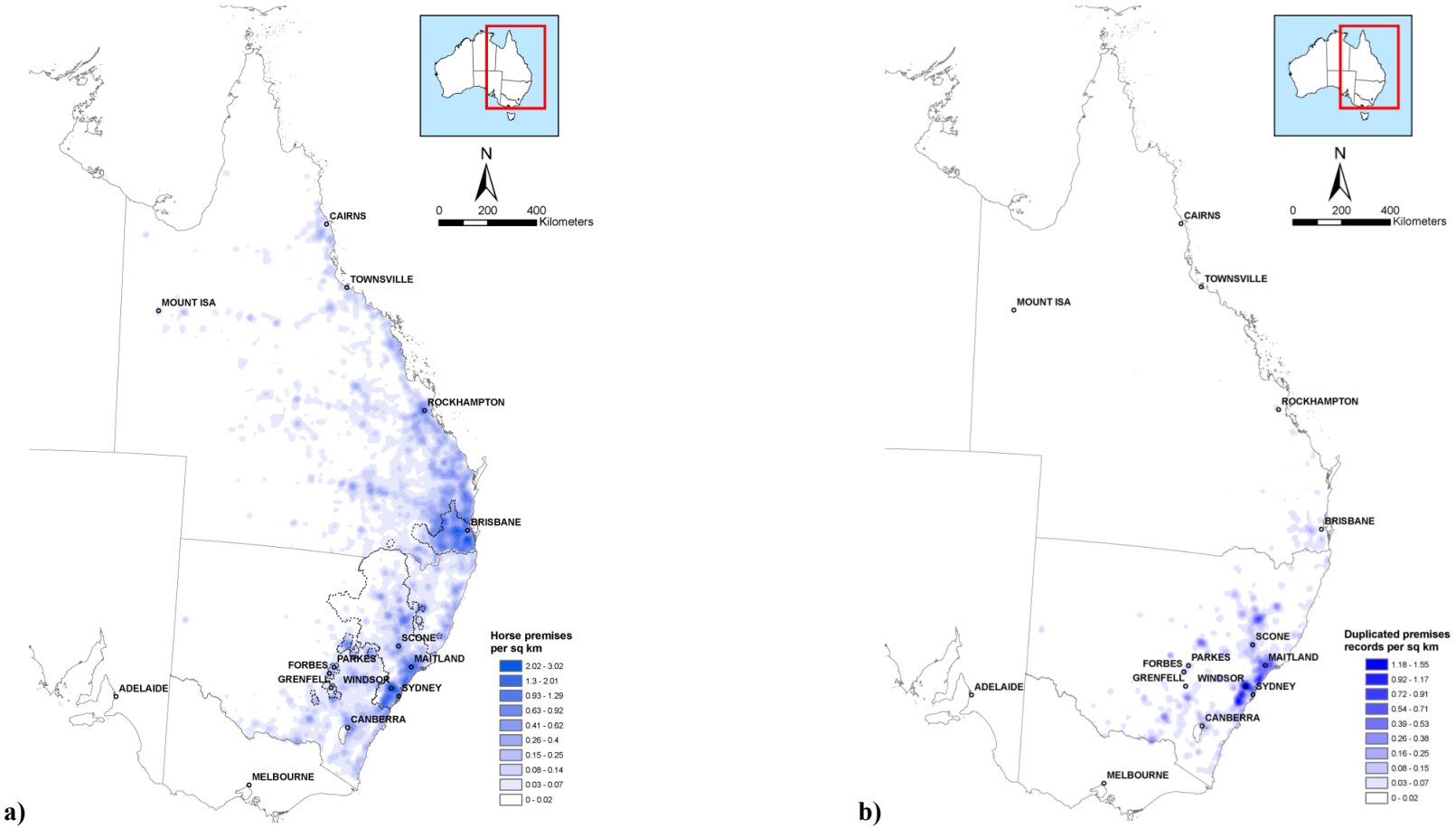


Figure 2.2 Kernel density smoothed surfaces of the distribution of horse premises and duplicated premises errors in the two equine influenza affected States. (a) Horse premises per square kilometre after aggregation by cadastral property boundaries to remove duplication, and (b) duplicated horse premises per square kilometre Kernel density estimation was based on a bandwidth of 20 km and presented on a grid of 5 km × 5 km cells. Dashed black lines denote the area affected by the 2007 outbreak of equine influenza.

Table 2.4 Comparison of estimates of cumulative incidence during the 2007 outbreak of equine influenza in Australia, before and after aggregation of records by cadastral polygon to reduce duplication.

	Cumulative incidence (%)		
	Original estimate[a]	Revised estimate	Difference (95% CI)
Special Restricted Area ('purple zone')	24.4	35.2	10.8 (9.8, 11.7)
Restricted Area ('red zone')	12.3	14.1	1.8 (1.2, 2.3)

[a] Based on data supplied by Cowled et al., 2009a.

Table 2.5 Summary of data completeness in the revised equine influenza outbreak dataset.

Variable name	Description (units / levels)	Data completeness	
		NSW (n=42,614)	**QLD** (n=61,108)
ID	Unique premises identifier (linkable to laboratory testing data)	100%	100%
CAD_ID	Unique identifier of containing or nearest cadastral polygon boundary	100%[a]	100%[a]
IP_STATUS	Infected premises status ('Yes' or 'No')	100%	100%
DATE_CX	Earliest date that clinical signs were observed in the first horse infected on an IP	94.6%[b]	94.8%[b]
N_HORSES	Number of horses (median of all non-zero consolidated duplicate records)	94.2%	97.4%
VACC_STATUS	Vaccination status from linked to vaccination database ('Yes' or 'No')	100%	100%
VACC_DATE	Earliest date of vaccination on the premises	100%[b]	NS
SITE_TYPE	Self-reported property type	2.2%	27.3%
SUBURB	Suburb, shire or postcode. All other address details withheld for privacy purposes	80.8%	97.4%
SOURCE	Data source (e.g. epidemiological investigations database, Rural Lands Protection Board survey, breeding or equestrian society, Travelling Horse Statement, etc.)	91.1%	NS
ORIG_CENT	Original centroid coordinates (decimal degrees)	100%	100%
CAD_CENT	Centroid coordinates of containing or nearest cadastral polygon boundary (decimal degrees)	100%	100%
CAD_DIST	Distance from original centroid coordinates to nearest cadastral polygon boundary (m)[c]	100%[a]	100%[a]
CAD_AREA	Area of containing or nearest cadastral polygon boundary (acres)	100%[a]	100%[a]
CAD_FENCE	Length of shared fence-line with contiguous horse-holding properties (m)	100%[a]	100%[a]

NSW = New South Wales, QLD = Queensland, IP = infected premises, NS = not supplied

[a] Variable calculated in geographical information system for all premises.

[b] Completeness only assessed amongst the subset of infected and vaccinated premises, where appropriate.

[c] If the original centroid coordinates were contained by an existing cadastral polygon, 'CAD_DIST'=0.

2.4. Discussion

The level of data completeness and spatial data quality of the equine influenza outbreak dataset was remarkable given its size. A low proportion (5.2%) of records were geocoded outside existing property boundaries, and most of these were within metres of a premises boundary. Of the horse premises included in the original dataset, 14.2% were geocoded to within the same property boundary as another premises record. In a dataset of >100,000 observations, such discrepancies sum to large numbers of errors. It was therefore considered necessary to aggregate premises records by cadastral polygon (to reduce duplication) before detailed spatio-temporal analyses were undertaken.

The total domesticated horse population of Australia has been estimated at >932,000 horses (Centre for International Economics, 2007, Animal Health Australia, 2011). After revising the dataset, it was estimated that nearly 800,000 horses were held on over 100,000 premises in NSW and Queensland alone. Given that highly detailed data has only been compiled in two States in 2007, it remains unclear how many horses are actually in Australia. However, the number is likely to be greater than previous estimates. Further research is required into the distribution of horses across Australia to inform risk assessments and modelling of emergency animal diseases involving horses, especially those that may affect humans (such as West Nile disease and Hendra virus infection).

Most of the issues identified during early exploratory spatial analysis of this dataset related to records of uninfected premises in highly affected areas of NSW. This analysis did not attempt to investigate the reasons for the identified data discrepancies, rather how to address them. That such discrepancies occurred is understandable considering the absence of reliable population at risk data prior to the outbreak, the rapid collation of data from a range of data sources early in the outbreak, several groups collating and geocoding data separately, and repeated visits or other contacts (phone calls) with horse owners/managers from a premises throughout the

outbreak resulting in non-identical details being recorded on separate occasions. When collating such data in the midst of a large outbreak the objective is to maximise data capture on the premises at-risk. Issues such as duplication can be resolved during field investigations and during other contacts with horse owners, however such errors may introduce considerable bias into estimates in retrospective epidemiological analyses.

Given the size of this dataset, manual cleaning of these errors (hand-cleaning) was considered unfeasible, so an automated (scripted) coded alternative was implemented. As most errors were identified to be associated with uninfected premises, these records were consolidated into the existing premises records contained with the same or the nearest cadastral polygon, conserving links to laboratory testing data. The cadastral datasets, maintained by the Governments of the two affected States, are high quality spatial resources on property boundaries. Most of the original premises records had centroid coordinates that were either contained by an existing cadastral polygon or very close to a premises boundary (typically on the road outside a premises). Discrepancies of several metres may be explained by the use of handheld Global Positioning System (GPS) devices to geocoded premises at their entrances, differences between actual property boundaries and fencelines ('on the ground') and irregularly shaped polygons (such as 'U' shaped premises) for which correctly geocoded centroid coordinates are surrounded by the polygon but not necessarily contained by it. The few original centroid coordinates that were large distances from premises boundaries were in remote and rural areas. Some of these premises may have been identified to the wrong cadastral polygon, which could affect high resolution spatial analyses in remote and rural areas of the affected States.

Cadastral polygon boundaries may not always reflect how several property 'lots' are grouped or the location of fences that bound herds of horses, and changes may have occurred in cadastral boundaries between the period of the outbreak (August–December 2007) and collation of the digital cadastral datasets released in 2009. Therefore, the cadastral boundaries cannot be considered a 'gold standard'. As the data were de-identified (not even including address details),

this analysis was limited in that it was not possible to test the accuracy of the automated cleaning process across the data extent. Some premises records were undoubtedly identified as duplicates when in fact they were separate premises. It was considered that the bias associated with not removing the >17,000 premises records identified as duplicated records would be relatively much greater than any bias associated with incorrectly consolidating premises records (even though this would result in a marginal under-estimation of the population at risk).

On large agistment premises (those that hold horses for other owners or rent paddocks to horse owners) the premises owner, the horse owner(s) and the horse manager(s) may all be different individuals, and each person could provide different details, especially when asked about the number of horses on the premises on separate occasions. When consolidating records identified as duplicates, the median of all of the respective entries for the number of horses recorded for a premises (in all consolidated records) was retained as the final value. On a small number of these multi-owner premises the sum rather than the median could have been taken. The measurement bias introduced would only be an issue in analyses involving the number of horses on a premises, and even then, this would only involve a small proportion of the 14% of original records that were duplicated.

Identifying duplication and errors when data are being generated, and updating these straight away, is clearly preferable to retrospective cleaning. The duplication identified in this analysis could have been avoided by 'spatially checking' whenever a 'new' premises was identified that it did not already exist in the dataset (potentially with different address, ownership, animal numbers, coordinates or other data). Furthermore, spatial data could be more accurately collected and collated in future emergency animal disease outbreaks by identifying premises by their cadastral polygon rather than geocoded coordinates. As shown in Figure 2.1, geocoded coordinates can be very close together spatially, but non-identical, whilst still referring to the same premises. Automated methods are available to rapidly geocode addresses and identify coordinates to a containing (or the nearest) polygon.

Differences were noted in the density of horse premises reported in the two affected States based on visual assessment of the kernel smoothed surface in Figure 2.2a. In NSW, there was a more marked difference between the density of horse premises in outbreak-affected compared to unaffected areas. This may in part relate to the different reporting practices and data sources collated on the population at risk in the two affected States (Moloney et al., 2011, Kung et al., 2011), leading to higher levels of under-reporting in unaffected areas of NSW compared to unaffected areas of Queensland. The population at risk dataset was considered reasonably reliable in outbreak affected areas only (Garner et al., 2011b), further research would be required to validate the population at risk data elsewhere. Bandwidth selection is a known issue that may affect the interpretation of kernel density smoothed point data. Adaptive kernel density estimation (Davies and Hazelton, 2009) may be implemented to vary the amount of smoothing across the study extent in inverse proportion to the density of the population at risk. This was considered unnecessary for the exploratory spatial visualisation undertaken in this analysis.

2.5. Conclusions

The spatial data quality and data completeness of the equine influenza dataset was found mostly to be very high. Duplication of uninfected premises in highly affected regions of NSW was identified as an important discrepancy that required amendment in preparation for further epidemiological analyses. After aggregating premises records by cadastral polygon, the number of horses infected in the 2007 outbreak was re-estimated to be around 67,000 horses, held on 9359 infected premises. Spatial data could be more accurately collected and collated in future emergency animal disease outbreaks by identifying premises by their cadastral polygon.

Chapter 3: Premises-level case-control study

"*An ounce of prevention is worth a pound of cure.*"

Benjamin Franklin (1736), advice to Philadelphia's Union Fire Company, the first in the city.

This chapter appears as the following published paper in Preventive Veterinary Medicine:

Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2011. A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. Prev. Vet. Med. 100, 53-63.

3. Premises-level case-control study – biosecurity

3.1. Introduction

Equine influenza is a highly contagious acute respiratory disease of horses and other equid species. Clinically, infection with the influenza A/equine/2 (H3N8) virus often results in a mild disease characterised by a harsh dry cough, fever, lethargy and nasal discharge (Myers and Wilson, 2006), but severe viral and secondary bacterial infections may develop, especially in foals and young horses (Miller, 1965, Myers and Wilson, 2006). Previous introductions into immunologically naïve horse populations have led to explosive outbreaks (Uppal and Yadav, 1987, Dalglish, 1992, Guthrie et al., 1999). Equine influenza has an incubation period of 1–3 days (Myers and Wilson, 2006). It is transmitted through both direct and indirect means. Horses infected for the first time may shed the virus in nasal secretions 24 hours after infection, and continue shedding for up to 10 days (Myers and Wilson, 2006). An infected coughing horse can spread the virus for up to 32 m (Miller, 1965), and transmission on fomites (such as people, vehicles or equipment) was reported in two South African outbreaks (Guthrie, 2006, Guthrie et al., 1999). Although rapidly inactivated by sunlight, heat, cold, drying conditions and common disinfectants, equine influenza can survive in soil for up to 2 days and in tap water for 2 weeks (Yadav et al., 1993). Human influenza A viruses have been shown to remain viable for 2 days on hard non-porous surfaces, and then be transferred by contact on human hands (Bean et al., 1982). Windborne spread of aerosolised equine influenza virus was also suggested to have occurred over 8 km in South Africa (Huntington, 1990) and 3.2 km in Jamaica (Dalglish, 1992).

Prior to August 2007, Australia was one of only three countries to have remained free of equine influenza (OIE, 2009). Introduction to Australia resulted in a 4-month epidemic during which approximately 67,000 horses were infected on 9359 premises across the State of New South Wales (NSW) and the south-eastern corner of the State of Queensland (Callinan, 2008). Equine influenza inadvertently entered a quarantine facility in Sydney on 8 August 2007 in infected

vaccinated horses imported for breeding from Japan (Callinan, 2008). Similar to the situation in the South African epidemics of 1986 and 2003 (Guthrie, 2006, Guthrie et al., 1999), the disease spread undetected, possibly on fomites, from the infected horses in quarantine to local horses. The exact means of escape from quarantine was not identified despite being the subject of a thorough legal inquiry (Callinan, 2008). The first local horses infected were competing in an equestrian event at Maitland, near Newcastle NSW, between 17 and 19 August. Some of these horses were then transported long distances whilst incubating the disease. At least one infected horse likely infected horses competing at another event at the Narrabri showground (nearly 400 km away) the following weekend (25–26 August 2007). Most of the resulting cases were either linked to these two events, or have been found to be the result of local spread over distances of around 2–8 km from premises infected in the first few weeks of the epidemic (Cowled et al., 2009b, Davis et al., 2009).

Control measures were effective in restricting this rapidly spreading epidemic to defined areas, and enabling eradication of equine influenza from the Australian horse herd. The key control measures included implementing a zone-based horse movement control system; contact-tracing and quarantine of suspect premises; targeted vaccination; and on-farm biosecurity measures implemented by horse owners, managers, veterinarians and others visiting horse premises during the epidemic. On-farm biosecurity measures were widely advised in communications from animal health authorities to horse owners (NSW DPI, 2007a). As there are relatively few studies into the effectiveness of on-farm biosecurity measures in preventing epidemic spread in animals, this advice was based on expert opinion.

The aim of this study was to investigate risk factors for the spread of equine influenza onto horse premises during the 2007 epidemic in Australia, specifically non-compliance with advised biosecurity measures. The findings are intended to improve our understanding of the epidemiology and transmission of equine influenza under Australian conditions. This information is of value for disease control authorities in implementing appropriate control

strategies in the event of future epidemics of equine influenza and other similarly transmitted diseases of horses.

3.2. Materials and methods

3.2.1. Study design

We conducted a case-control study of horse premises from 'at risk' regions of New South Wales. 'At risk' areas were defined as restricted areas and special restricted areas according to the risk-based zoning system implemented by the NSW Department of Primary Industries (NSW DPI) in its Equine Influenza Protection Plan (NSW DPI, 2007b). Restricted areas were designated 'infected areas' (within 10 km of an infected premises) where equid movements were prohibited (Figure 3.1). The special restricted area was a part of the restricted area with high horse density where horse movements were allowed in later phases of the epidemic, once spread of infection within that area was considered inevitable. Permitting movements within the

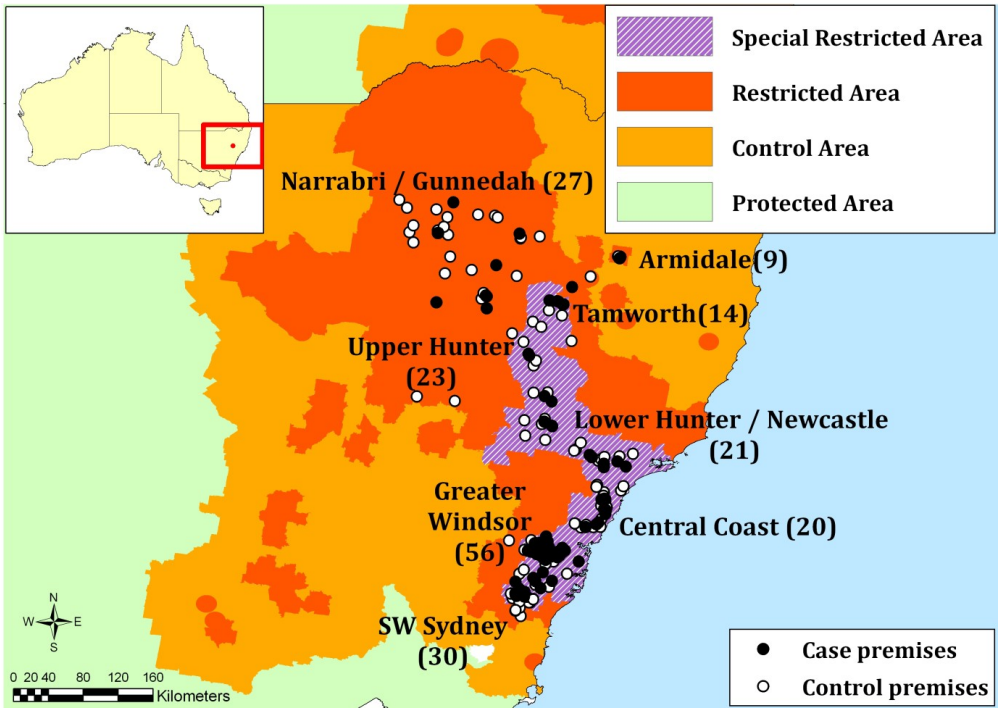


Figure 3.1 Premises interviewed in a case-control study of equine influenza in New South Wales, Australia, classified by regional cluster , as at 28th September 2007.

special restricted area enabled disease control activities to continue whilst minimising the impact on the start of the Australian Thoroughbred breeding season (officially beginning September each year).

We anticipated that relaxation of horse movement restrictions and vaccination of horses in response to the outbreak would confound the detection of factors that facilitated or prevented spread of infection onto horse premises and, therefore, restricted the study period to the first seven weeks of the epidemic (17 August–5 October 2007). This period starts at the 'Maitland equestrian event', the first known spreading event in the epidemic, and ends prior to the commencement of vaccination and the relaxation of movement controls.

3.2.2. Sample size

A sample size of 100 case and 100 control premises was calculated to provide 95% confidence of detecting an odds ratio of three with 80% statistical power, assuming a 1:1 ratio of case to control premises and a minimum of 10% of control premises exposed to the factor of interest (Thrusfield et al., 2001).

3.2.3. Selection of horse premises

A dataset of horse premises laboratory tested for equine influenza virus (H3N8) was supplied by the NSW DPI. A detailed description of this dataset, its compilation and cleaning is provided by Cowled et al. (2009b). The dataset contained results of more than 130,000 tests using real-time reverse transcription polymerase chain reaction (qRT-PCR) (Foord et al., 2009) and an antigen-detecting enzyme-linked immunosorbent assay (Heine et al., 2007). The list of all premises tested during the outbreak was separated into two populations according to disease status.

Case premises were included in the study if they were: recorded as infected premises (IP) in the NSW DPI dataset, were from within the restricted or special restricted areas, and had one or

more horses recorded as having tested positive for equine influenza with a date of onset of clinical signs for the first horse infected on the premises and a date of sampling within the study period. Control premises were included in the study if they were: recorded in the dataset as not being an IP, were from within the restricted or special restricted areas, had no positive test results within the study period, and at least one horse had been tested and found negative by qRT-PCR.

A sample of 270 potential case and control premises were randomly selected from the two eligible populations using computer generated pseudo-random numbers. Managers and owners of these horse premises were asked to participate in the study by a personally addressed letter and follow-up telephone calls. When first contacted, 38 control premises were deemed ineligible because the horse owner/manager: reported that their horses had clinical signs consistent with a diagnosis of equine influenza during the study period (n = 35) or could not recall their horses being tested for equine influenza during the study period (n = 3).

3.2.4. Questionnaire and interviewing

Two interviewers collected data through structured face-to-face interviews. These interviews were based on a questionnaire that was piloted on four horse owners/managers and modified according to their comments. The interviewers conducted the pilot interviews together to ensure similarity in interview method.

Interviews were conducted on the horse premises between July and October 2009, and took around 60 minutes to complete. The intended respondent was the person directly responsible for the management of horses on the premises at the time of the equine influenza epidemic. If this person could not be identified and interviewed, then the owner of the horses, or the owner of the premises, or another person familiar with the management of the horses during the time of interest was interviewed. The questionnaire (see Appendix) contained 44 closed questions that

captured data on:

- the location of the premises,
- premises layout and proximity to other horse premises,
- the horses and the types of equestrian activities undertaken,
- signs of equine influenza in horses on the premises,
- movements of horses and people on and off the premises,
- any contacts with horses elsewhere,
- the implementation of bio-security practices, and
- other factors likely to be associated with equine influenza spread.

For each premises, all data were collected with reference to one week within the study period as it was recalled by the interviewee. For cases, this was the week immediately prior to the first horse on the interviewee's premises being noticed to have clinical signs consistent with equine influenza. For controls, this was the week prior to a veterinarian or DPI officer visiting the premises to collect samples for laboratory testing from horses either considered at the time to be suspect cases or that had been in contact with infected horses. The time frame of one week was chosen to include the likely incubation period of equine influenza, and was intended to improve recall and the accuracy of responses by providing a memorable event and time landmark on which to base questioning. Prior to identifying the specific week, interviewees were engaged in a series of recall-improving activities. These included: discussion of events leading up to the epidemic reaching their local area; inspection of diaries, horse records, veterinary receipts and notifications from the DPI and other agencies; and discussion of the horses and people on the premises at the time, their movements and contact with other horses and other horse premises.

Interviewees were provided a supermarket gift voucher worth AUD 50 to compensate for their time involved in the interview process and in preparing records for answering questions about horse numbers and movements on their premises at the time of the epidemic. Ethical approval

for all procedures used in this study was obtained from the Human Research Ethics Committee of the University of Sydney (07–2009/11840).

3.2.5. Data entry and management

Data from interviews were entered into a database using a purpose-built form in Microsoft Access 2007 (Microsoft Corporation, Redmond, WA , USA). Basic data manipulation, range checks and data cleaning were conducted in this database.

3.2.6. Data analysis

Unless otherwise specified, all statistical analyses were conducted using Stata Version 10.0 (Stata-Corp, College Station, TX, USA). The outcome variable of interest was the binary status of the premises according to the presence or absence of a case of equine influenza.

3.2.6.1. Explanatory variables

The complete list of 66 explanatory variables investigated in this study is presented in Table 3.1. Spatial variables were derived through linking the database to spatial layers in ArcMap 9.3 (ESRI, Redlands, CA, USA) using an Albers conic equal-area projection based on the Geocentric Datum of Australia 1994 (Geoscience Australia, 2007). Premises locations were designated as points at the centroid of each premises polygon. These polygons were obtained through a spatial join to a layer of premises polygons obtained from the NSW Land and Property Management Authority, confirmed by interviewees as correct at the time of the epidemic. The locations of premises entry points and shared fences with any neighbouring horse premises were collected by hand-held GPS devices during interviews for further validation. A measure of local horse density at the time of the outbreak was calculated by totalling the horses on all premises in the NSW DPI dataset that were within a 2 km buffer each premises' polygon. Premises were assigned to the 'regional clusters' delineated by Cowled et al. (2009b), who used an interpolated surface of date of onset of clinical signs, geographic data and location of the IPs

from the NSW DPI dataset. Premises that were not within the minimum convex hull of any regional cluster were assigned to the nearest cluster.

Table 3.1 Explanatory variables analysed for associations with equine influenza status of horse premises in a case-control study of the 2007 outbreak in Australia.

Variable group	**Variables** (Units)
Premises locality	Proximity to nearest infected premises (km); Local horse density (horses per acre within 2 km); Number of horses on nearest infected premises; Regional cluster (as per Cowled et al. 2009b).
Premises type	Enterprise type; Entrance on a private road[a]; Frequency of farrier usage; Frequency of worming; Horse breeders (non-Thoroughbred horses)[a]; Horse breeders (Thoroughbred horses)[a]; Horses kept for other owners[a]; Horses kept only for recreation[a]; Horses used for farm work[a]; Horses vaccinated for tetanus[a]; Horses vaccinated for strangles[a]; Income derived from horses[a]; Involved in equestrian events[a, c]; Involved in horse breaking-in[a]; Involved in pre-race training[a]; Involved in horse-racing[a]; Involved in horse showing[a]; Involved in polo/polocrosse[a]; Involved in pony club[a]; Involved in rodeo-style horse events[a, d]; Mean horse age (years); Number of horses; Premises area – horse accessible (acres); Premises area – total (acres); Premises horse density (horses per acre); Proportion of entire male horses; Shared fence with other horse premises [a]; Stabled indoors or outdoors; Stable type.
Biosecurity – *Contact and access control*	A veterinarian visited[a]; A farrier visited[a]; Another horse professional visited[a]; All vehicles disinfected prior to entry[a]; Horses monitored daily for signs of equine influenza; Records kept of horse monitoring[a]; Fences were checked[a]; Horses kept away from fences[a]; Horses attended an event[a]; Horses moved onto the premises[a]; Horses moved off the premises[a]; Interviewee contacted horses elsewhere[a]; Interviewee's vehicle entered another horse premises[a]; Number of people tending the horses; Number of visitors tending the horses; Visitors entered the premises[a]; Visitors contacted horses elsewhere[a]; Visitors' vehicles entered the premises[a].
– *Barrier hygiene*	How often did the interviewee wash their hands before horse contact[b]; How often did the interviewee change clothes before horse contact[b]; How often did the interviewee change shoes before horse contact[b]; How often did the interviewee shower before horse contact[b]; How often did visitors wash their hands before horse contact[b]; How often did visitors change clothes before horse contact[b]; How often did visitors change shoes before horse contact[b]; How often did visitors shower before horse contact[b]; Was a footbath in place[a].
– *Other measures*	Horse bedding was removed from the premises[a]; Horse manure was removed from the premises[a], Was a horse float driven off the premises[a]; Was feed shared with another horse premises[a]; Was horse equipment shared with another horse premises[a]; Was feed obtained from off the premises[a].

[a] Binary variables (1=yes, 0=no). [b] Ordinal variables (3=every time, 2=most times, 1=sometimes, 0=never). [c] Equestrian events include dressage, eventing, showjumping and cross-country. [d] Rodeo-style events include camp-drafting, cutting, barrel-racing.

The proximity of premises to the nearest infected premises was calculated using the spatstat library (Baddeley and Turner, 2005) within the R statistical package (R Development Core Team, 2011). This involved selecting the nearest IP from a subset of those considered infectious during the specific week of interest for each study premises. IPs were considered to pose a contagious disease risk from the day before the first horse on the premises reportedly showed clinical signs of equine influenza until the end of the study period. This allowed for several potential cycles of transmission on larger premises, assuming infected horses shed for up to 10 days in each cycle.

Several factors were considered *a priori* as likely to confound the association between the case status of a premises and risk factors likely to either facilitate or prevent spread of influenza virus onto horse premises. Data were collected at interview on potential confounders—such as type of enterprise, size of premises (in area and horse numbers) and involvement in certain types of horse activities—to enable adjustment for any effects during multivariable analysis. Data were also collected on the signalment (age, gender and breed) of horses on the premises, even though these factors were considered *a priori* to be unlikely to be confounders given the largely immunologically naïve status of the Australian horse population at the start of this epidemic.

Vaccination for equine influenza was not practiced in Australia prior to the 2007 outbreak, given the disease-free status of the horse population. Only around 500 vaccinated horses are imported annually into a domestic horse population estimated to be around one million (Biosecurity Australia, 2010), with immunity from inactivated vaccines lasting less than 1 year (Paillot et al., 2006). Nevertheless the interview included questions on the importation and vaccination history of horses on each premises, and the study period was restricted so as to end prior to the commencement of vaccination in response to the outbreak and the relaxation of movement restrictions.

3.2.6.2. Descriptive analysis

To assess the distribution of all explanatory variables, we generated frequency distributions and bar charts for all categorical variables, and histograms and summary measures for all continuous variables. To assess recall bias, dates provided by interviewees were compared with sampling dates and dates of first clinical signs recorded in the NSW DPI laboratory testing dataset.

3.2.6.3. Univariable analysis

Continuous variables were evaluated for differences between case and control premises using the Student's t-test or Wilcoxon rank sum test depending on fulfilment of the assumptions of normality and homoscedasticity. The linearity of the relationship between continuous variables and the log odds of being a case was assessed visually by fitting restricted cubic splines and lowess smoothed-line estimates of the relationship (Dohoo et al., 2009). If there was a considerable departure from linearity, then either a logarithmic transformation or a spline of the variable was used for further logistic regression analyses, as appropriate.

For binary and categorical variables we constructed contingency tables and calculated unadjusted odds ratios (OR), corresponding 95% confidence intervals and *P*-values using Fisher's exact test. On the basis of this analysis, the ordinal biosecurity variables were reclassified by grouping the levels 'Every time' with 'Most times' as 'Frequently', and 'Sometimes' with 'Never' as 'Infrequently', due to small numbers in certain strata. All variables unconditionally associated with the log odds of being a case at a *P*-value of <0.20 were tested for collinearity in pairs by calculating Spearman's rank correlation coefficient (ρ). Amongst highly correlated pairs of variables (ρ>|0.70|), only the variable most strongly associated with the outcome was retained for further analysis. All remaining explanatory variables were assessed for missing values; those with <10% missing values were eligible for multivariable analyses.

3.2.6.4. Multivariable analysis

We fit logistic regression models to the data using a manual forward stepwise approach (Hosmer and Lemeshow, 2000). At each step, all eligible candidate variables were individually tested for addition to the model using the likelihood-ratio test (LRT). The most strongly associated of these was selected for inclusion if its LRT-derived *P*-value was <0.05, and prior to the next step, any terms already included with $P \geq 0.15$ were removed and returned to the pool of candidate variables (Lee and Koval, 1997). First order interaction terms were added to the final main effects model and tested for significance at $P<0.05$. We examined the final logistic regression model for goodness-of-fit using the Hosmer-Lemeshow technique (Hosmer and Lemeshow, 2000). Outliers and points of influence on the final model were evaluated using standardised Pearson residuals, hat matrix leverage and delta-beta values. To control for effects of clustering of observations from the same region, we refit the final model using multi-level mixed effects logistic regression, with a regional cluster-level random effects term. An interviewer-level random effect term was also added to the final model to test for interviewer bias (Hox, 1994). To allow for inclusion of spline terms in this mixed-model we used the Davidon-Fletcher-Powell algorithm for maximising the likelihood function (Gould et al., 2006). Intra-class correlation (ICC) was then calculated for each random effect using the latent variable approach (Dohoo et al., 2009).

3.3. Results

3.3.1. Descriptive analysis

Of the 232 eligible horse premises, 32 premises declined to participate (13 case and 19 control premises) for the following reasons: too busy (22), not interested (5), in poor health (2), considered their premises never at risk (2), and no reason provided (1). Ultimately, 100 case and 100 control premises were willing to participate and were interviewed (response rate 86.2%), these are shown by location and case status in Figure 3.1. Most premises were small acreages with pleasure horses (46.5%), farms (18.5%), and equestrian centres or riding schools (18.0%)

(Table 3.2). The most frequently reported clinical signs of equine influenza amongst the first horse infected on case premises were: presence of nasal discharge, coughing, lethargy and a rectal temperature of over 39°C.

The respondent was the owner, manager or trainer directly responsible for the horses at the time of the equine influenza epidemic on 98% of premises interviewed. Of the respondents from case premises, 87% were able to recall the exact date of infection or testing, compared to 79% from control premises (P=0.13). Respondents from two premises (1 case and 1 control) were unable to recall this date to within a 2-week period. The mean magnitude of the difference between recalled date of infection or testing and that recorded in the NSW DPI laboratory data was 3.1 and 3.6 days (SD=4.9 days and 5.7 days) in all respondents from case and control premises respectively.

3.3.2. Univariable analyses

A total of 31 variables were unconditionally associated with the case status of a premises at P<0.20 (Tables 3.3–3.5). The proportion of case premises was very high amongst those interviewed in the Greater Windsor region (89.3%), compared with all other regions (range 29.6–41.5%). Horses movements (either on or off) were only reported from 19 premises, all such movements occurring prior to the implementation of movement controls. Of these, horses from 12 premises attended equestrian events where equine influenza transmission is now known to have occurred, 10 such premises were subsequently infected.

A non-linear association was observed between the log odds of being a case premises and five continuous variables: ‘Proximity to nearest infected premises’, ‘Number of horses’, ‘Premises area – total’, ‘Premises area – horse accessible’ and ‘Premises horse density’. Logarithmic transformations of the latter four variables returned these relationships to relative linearity, and they were entered in multivariable analysis as such. A cubic spline of the relationship between ‘Proximity to nearest infected premises’ and the log odds of being a case was constructed with

knots based on quintiles of the data (Figure 3.2). Two effects were apparent, a rapidly increasing likelihood of cases being found compared to controls within 5 km of an infected premises and a similar increase beyond 10 km.

Table 3.2 Case and control premises by enterprise type in a case-control study of the 2007 equine influenza outbreak in Australia.

Premises enterprise type	Cases	Controls
Small acreages/homes with horses	60	33
Equestrian centres or riding schools	19	17
Spelling or agistment paddocks	9	7
Farms – cattle, sheep or cropping	8	29
Commercial studs	4	14
Total	100	100

Table 3.3 Descriptive statistics for continuous exposure variables associated with case status (P<0.20) on 100 case and 100 control premises in a case-control study of the 2007 equine influenza outbreak in Australia.

Variable	Group	N	min	Q1	median	Q3	max	*P*-value[a]
Number of horses	Cases	100	1	2	4	8	70	0.017
	Controls	100	1	3	6	16.5	150	
Mean horse age (years)	Cases	100	3.2	7.8	11.0	14.5	25.0	0.024
	Controls	100	2.4	6.0	8.9	12.7	21.5	
Premises area – horse accessible (acres)	Cases	100	0.8	3.3	9.0	24	1500	0.001
	Controls	100	0.1	5.2	17.0	75	1400	
Premises area – total (acres)	Cases	100	1.0	5.0	10.3	30	4600	<0.001
	Controls	100	0.5	9.4	40.0	202	6000	
Premises horse density (horses acre^{-1})	Cases	100	0.02	0.29	0.54	1.13	11.51	0.045
	Controls	100	0.01	0.16	0.44	0.84	64.00	
Local horse density[b] (horses acre^{-1})	Cases	100	0.00	0.05	0.09	0.14	0.28	<0.001
	Controls	100	0.00	0.01	0.04	0.07	0.21	
Proximity to nearest infected premises (km)	Cases	100	0.04	0.3	0.8	2.0	291.7	<0.001
	Controls	100	0.03	1.9	5.8	10.8	51.7	

Q1= first quartile, Q3 = third quartile.
[a] *P*-values based on Wilcoxon rank-sum test of significance. [b] Estimated within a 2 km buffer.

Table 3.4 Contingency tables for categorical premises type variables associated with case status (P<0.20) and odds ratios based on univariable logistic regression analyses amongst 100 case and 100 control premises in a case-control study of the 2007 outbreak in Australia.

Variable	Category	Cases	Controls	Odds ratio	95% CI	*P*-value[a]
Regional cluster	Windsor	50	6	19.79	6.06, 64.60	<0.001
	Lower Hunter[b]	17	24	1.68	0.60, 4.73	
	SW Sydney	11	19	1.38	0.45, 4.18	
	Upper Hunter[c]	14	32	1.04	0.37, 2.93	
	Narrabri[d]	8	19	1		
Entrance on a private road	Yes	5	15	0.30	0.08, 0.91	0.032
	No	95	85	1		
Frequency of farrier usage	Every 2 months	78	64	1.99	1.02, 3.93	0.042
	Less frequently	22	36	1		
Horse breeders of Thoroughbreds	Yes	4	12	0.31	0.10, 0.98	0.065
	No	96	88	1		
Horses kept for other owners	Yes	12	20	0.55	0.23, 1.26	0.18
	No	88	80	1		
Horses kept only for recreation	Yes	50	23	3.35	1.75, 6.47	0.001
	No	50	77	1		
Horses used for farm work	Yes	5	15	0.30	0.08, 0.91	0.032
	No	95	85	1		
Income derived from horses	Yes	31	44	0.57	0.31, 1.06	0.079
	No	69	56	1		
Involved in rodeo-style events	Yes	11	22	0.44	0.18, 1.02	0.056
	No	89	78	1		
Involved in pony club	Yes	5	12	0.39	0.10, 1.24	0.13
	No	95	88	1		
Shared fence with other horse premises	Yes	66	55	1.59	0.86, 2.93	0.15
	No	34	45	1		

[a] *P*-values calculated using Fisher's exact test. [b] Includes Newcastle and Central Coast clusters.
[c] Includes Tamworth and Armidale clusters. [d] Includes Gunnedah cluster.

Table 3.5 Contingency tables for categorical biosecurity variables associated with case status (P<0.20) and odds ratios based on univariable logistic regression analyses amongst 100 case and 100 control premises in a case-control study of the 2007 outbreak in Australia.

Variable	Category	Cases	Controls	Odds ratio	95% CI	*P*-value[a]
A veterinarian visited	Yes	6	15	0.36	0.11, 1.05	0.063
	No	94	85	1		
Horses monitored daily for signs of equine influenza	Yes	87	78	1.89	0.84, 4.36	0.14
	No	13	22	1		
Records kept of horse monitoring	Yes	36	16	2.95	1.44, 6.20	0.002
	No	64	84	1		
Fences were checked	Yes	55	43	1.62	0.89, 2.94	0.12
	No	45	57	1		
Horses kept away from fences	Yes	18	24	0.42	0.18, 0.95	0.035
	No	54	30	1		
Horses attended an event	Yes	10	2	5.44	1.11, 52.04	0.033
	No	90	98	1		
Horses moved off the premises	Yes	12	88	2.59	0.81, 9.73	0.13
	No	5	95	1		
Horse manure was removed from the premises	Yes	3	0	Undef.	0.80, Undef.	0.12
	No	96	100	1		
Interviewee changed clothes[b]	Frequently	39	52	0.56	0.30, 1.03	0.061
	Infrequently	59	44	1		
Visitors washed their hands [b]	Frequently	24	36	0.48	0.23, 1.00	0.040
	Infrequently	46	33	1		
Visitors changed their shoes [b]	Frequently	30	42	0.48	0.23, 1.00	0.042
	Infrequently	40	27	1		
Visitors changed their clothes[b]	Frequently	25	40	0.40	0.19, 0.84	0.011
	Infrequently	45	29	1		
Was a footbath in place	Yes	17	34	0.40	0.19, 0.82	0.009
	No	82	66	1		

Undef.=Undefined.
[a] *P*-values calculated using Fisher's exact test. [b] Variable regrouped (Frequently='every time' or 'most times', Infrequently='sometimes' or 'never').

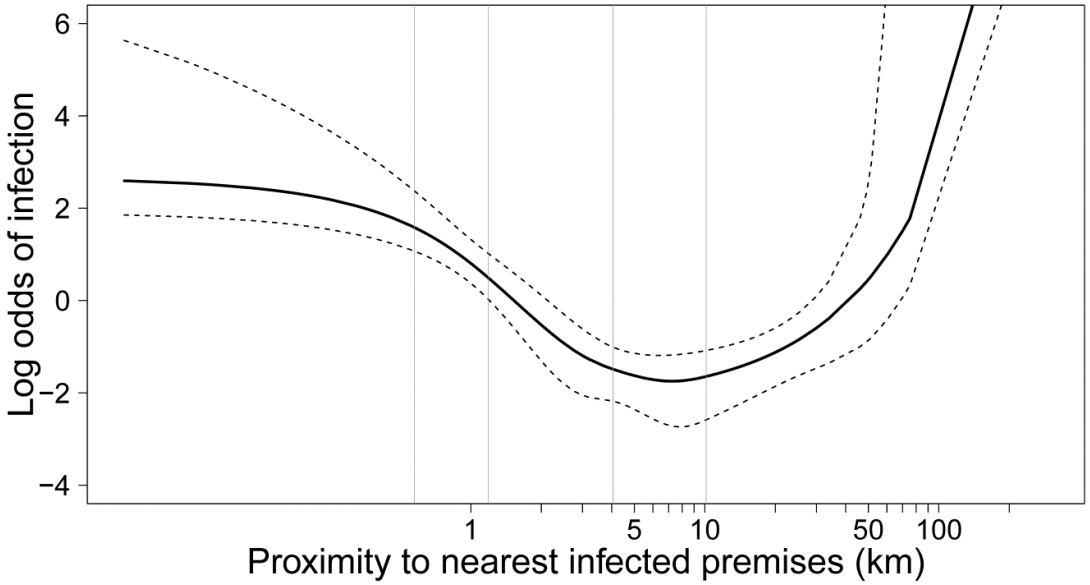


Figure 3.2 Estimated spline transformation and 95% confidence limits of crude association between proximity to nearest infected premises and equine influenza 'case' status of horse premises amongst 100 case and 100 control premises in a case-control study of the 2007 outbreak in Australia.
Reference lines depict knots in the spline, at quintiles in the data.

Barrier hygiene biosecurity practices of interviewees and visitors to the premises during the week of interest (i.e. hand-washing, changing clothes and shoes, and having a footbath in place) were generally protective (Table 3.5). The three variables representing visitor biosecurity practices ('Visitors washed their hands', 'Visitors changed their shoes' and 'Visitors changed their clothes') were highly correlated ($0.64<\rho<0.76$, all with $P\leq0.001$). These three variables contained numerous missing values because on 61 premises there was no visitor contact with the horses in the specific week; thus they were excluded from stepwise logistic regression but later individually added to the final model and tested for significance at $P<0.05$.

Two further pairs of variables were highly correlated, 'Premises area – horse accessible' with 'Premises area – total' ($\rho=0.88$, $P\leq0.001$), and 'Horses attended an event' with 'Horses moved off the premises' ($\rho=0.83$, $P\leq0.001$). After excluding 'Premises area – total' and 'Horses moved off the premises', 26 variables were included in multivariable analyses.

3.3.3. Multivariable analyses

The final model for risk factors associated with equine influenza infection is presented in Table 3.6. 'Proximity to nearest infected premises' was included as a cubic spline. Its form after adjusting for the other variables in the final model was similar to its crude form (Figure 3.2). The two effects noted during univariable analysis—increased likelihood of cases being found within 5 km and beyond 10 km—remained unaltered. A single outlier was found to have high leverage over the effect beyond 10 km. This premises was infected in the first week of the epidemic by horses returning from an equestrian event, and thereby was a long distance from the few IPs in the NSW DPI database at this early stage of the epidemic.

Case premises were nearly four times less likely than control premises to have had a footbath in place in the week before infection or testing for equine influenza. On case premises, horses were more likely to be kept for recreational purposes, and records were more likely to be kept of daily horse monitoring for signs of disease.

Intra-class correlation for clustering by region accounted for 13.6% of the total variability. When region was included in the multivariate model as a fixed effect, case premises were seven times more likely than controls to be from the Greater Windsor region (OR=7.3; 95% confidence interval: 2.39, 22.1). All other estimates were almost identical between the mixed effects model (Table 3.6) and the fixed effects model (not shown).

The final model displayed good fit to the observed data (Hosmer-Lemeshow goodness-of-fit chi-squared test statistic P-value=0.74, ROC area under curve=0.92). No first order interaction terms were significant at $P<0.05$. Evaluation of residual patterns identified only the single outlying observation noted earlier. Exclusion of this observation had minimal effect on the odds ratio estimates in the final model, but some effect on the shape of the cubic spline at its upper extreme. The interviewer-level random effect term was not significant, and when forced into the final model did not account for an appreciable proportion of the total variability (ICC<1%).

Table 3.6 Final mixed-effects logistic regression model for equine influenza infection amongst 100 case and 100 control premises in a case-control study of the 2007 outbreak in Australia.

Variable	Category	*b*	SE(*b*)	OR	95% CI	*P*-value[a]
Constant		2.02	1.32	–	–	–
Proximity to nearest infected premises		Included as cubic spline				<0.001
Was a footbath in place	Yes	-1.32	0.58	0.27	0.09, 0.83	0.011
	No	–	–	1		
Records kept of horse monitoring	Yes	1.96	0.60	7.08	2.17, 23.1	0.001
	No	–	–	1		
Horses attended an event	Yes	2.90	1.13	18.1	1.97, 166.7	0.006
	No	–	–	1		
Horses kept only for recreation	Yes	1.33	0.47	3.78	1.50, 9.50	0.004
	No	–	–	1		
Random Effect						
Regional cluster[b]		0.52	0.48	–	0.09, 3.14	0.007
Intra-class correlation (cluster)[c]		0.14	–	–	0.03, 0.49	

N=199, Log likelihood= -76.6; d.f.=6; $P<0.001$; Goodness-of-fit χ^2-test statistic P-value=0.737.
[a] P-values based on likelihood ratio test.
[b] The variance of the regional-level random effect may be interpreted qualitatively, where a value of zero means no variation between regions and a large positive value means a high degree of regional clustering (Dohoo et al., 2009).
[c] Conversion of the intra-class correlation (ICC) into a cluster-median OR (in this case, 1.14 with 95% CI: 1.03, 1.60) enables interpretation as the median OR between (identical) subjects from two randomly selected clusters (Dohoo et al., 2009).

3.4. Discussion

In this study, factors important in the spread of equine influenza onto a horse premises could be separated into two broad groups: intrinsic premises factors and biosecurity compliance factors. Intrinsic factors describe the locations and types of premises where horses were most at risk, and were not changeable once the epidemic had commenced. Biosecurity compliance factors concern management actions that may be taken to reduce the risk of introducing infection onto a premises.

3.4.1. Intrinsic premises factors

Proximity of a premises to the nearest infected premises was the most strongly associated factor, and the two effects observed likely represent local and long distance spread mechanisms. Local spread may occur through a combination of direct horse contact, cough droplets, windborne aerosols and transfer of virus particles on fomites such as horse handlers' clothing and vehicles (Guthrie, 2006).

Our finding of increased likelihood of cases being within 5 km of an infected premises is consistent with other estimates of the distance of local spread of equine influenza during this epidemic (Cowled et al., 2009b, Davis et al., 2009) and from previous epidemics overseas (Huntington, 1990, Dalglish, 1992). Our analysis also suggests that 10 km buffers around infected premises, such as those underpinning the zone-based movement control system, were appropriate for containing local spread during this outbreak.

Long distance spread, as observed beyond 10 km, was documented to have occurred before horse movement controls were implemented, through the transport of horses incubating equine influenza whilst returning from events where transmission is known to have occurred (Callinan, 2008). This was the case for the single outlier premises that had a large influence on the spline at its distal extreme and also for another nine of the 12 premises whose horses attended events in the week prior to the horse standstill.

The two other important intrinsic factors, 'Regional cluster' and 'Horses kept only for recreation', both relate to pleasure horse premises. The cluster of infected premises in the Greater Windsor region was previously identified as the earliest infected, containing the most IPs, and one of the highest cumulative incidence rates (42%) of any cluster during the Australian epidemic (Cowled et al., 2009b). Most horse premises in this area are small acreages holding low numbers of pleasure horses in close proximity to other horse premises. There are

relatively few Thoroughbred horse studs and race training premises, and those premises involved in racing are concentrated around the main horse racing venues.

In the NSW DPI dataset, of 114 premises from Greater Windsor that met the inclusion criteria only 25 were control premises. Eighteen of these were randomly selected, and twelve subsequently excluded as owners reported horses having had clinical signs of equine influenza during the study period. Of the six control premises interviewed in Greater Windsor, three were relatively isolated from other horse premises (>6.7 km from their nearest infected premises), and the respondents from the other three premises reported complying with most of the biosecurity measures investigated.

3.4.2. Biosecurity compliance factors

Two factors in the final model related to management actions taken on premises: having a footbath in place before horses were infected, and maintaining a record of daily monitoring of horses for clinical signs of equine influenza. Having a footbath in place for cleaning shoes before contacting horses, relates to the risk of introducing equine influenza virus to a premises on contaminated shoes, and thereby infecting horses or their immediate environment. The variable 'Was a footbath in place' may have been a proxy for general premises biosecurity standards, and may relate more broadly to fomite transmission of equine influenza. Spread on fomites was suggested as a pathway in the escape from quarantine of equine influenza virus in the two South African epidemics (Guthrie, 2006, Guthrie et al., 1999), and could not be ruled out as an intermediate pathway between quarantine and the first horse infected outside of quarantine in this epidemic (Callinan, 2008).

The association between case premises and the practice of keeping records of horse monitoring is not straightforward. Of the 165 premises that monitored their horses daily only 52 kept records. One explanation for the observed association is that horse owners and managers

monitored their horses and maintained records more vigilantly when they knew of infection in their local area. Alternatively, the maintenance of records might be a proxy for more attentive monitoring and thereby increased handling of the horses.

3.4.3. Control of confounding

To minimise confounding bias we restricted the study period and area, and tested all factors considered a priori as likely confounders by forcing them into the final model. A region-level random effect term was also included in the final analysis to account for clustering. The mean age of horses on the premises was not included in the final model as it was considered *a priori* not to be a confounder. Horse age did not significantly change other estimates, or improve model fit to the observed data when forced into the final model. It is however worth noting from our univariable analysis that a higher median age of horses was associated with case premises. This tendency for fewer infections among younger horses was previously described in a study of vaccinated racehorses during an epidemic in the United Kingdom (Newton et al., 2006). Subsequently, this age effect was explained by differences in the vaccination histories of the horse age cohorts studied (Barquero et al., 2007). In our analysis, of a mostly unvaccinated horse population, this age effect could not be similarly explained. As horse age was not important in the final model we consider this a spurious finding.

3.4.4. Study validity

A key consideration in the design of this study was how to minimise the potential for recall to bias the results. Recall bias can lead to unpredictable imprecision if the memory of one group is systematically different from that of the comparison group. Ideally, measurement of on-farm management factors would have been conducted through observation on premises during the epidemic period, with prospective follow-up until onset of infection. This was not possible at the time of the epidemic as all available resources were devoted to control and eradication of the disease.

Several aspects of this study were intentionally designed to minimise the potential for poor recall. A face-to-face on-farm interview was specifically chosen to increase cooperation, rapport, consistency and reliability of responses, and the completeness of data. Open discussion leading up to questioning was conducted to stimulate memory and provide memorable anchor points in time with reference to official documentation and veterinary records. Questions were focussed on horse management practices during a specific week prior to a memorable event, and at interview maps were developed of the premises and local area including dates when nearby premises were infected or quarantined. Validation rules were also included in the questionnaire and the face-to-face approach enabled immediate confirmation of any responses that did not agree with earlier statements.

Respondents from case and control premises were similarly precise in their ability to recall the date of testing or infection of their horses, and agreement of these responses with DPI data collected at the time of the epidemic was very good. Although recall bias cannot be completely eliminated from such a retrospective study it was considered to have a negligible effect on the observed results.

A recent study has shown that dogs were infected with equine influenza during the 2007 outbreak (Kirkland et al., 2010). There is no evidence as yet that dogs may be capable of transferring virus between horses on different premises, but it may be possible. Similarly, birds and flies have been suggested as possible mechanical vectors of influenza (Sawabe et al., 2009, Spokes et al., 2009). We considered including questions in the interviews on the movements of dogs on and off premises and the presence of wild birds, but ultimately decided that recall and accuracy of responses to such questions would be too unreliable to provide meaningful analysis.

In this study, outcome status was based on the diagnostic test results of at least one horse on each premises. On first contact with potential respondents we further confirmed this status by checking whether test results matched the observations of horse owners and managers during

the study period. Therefore, misclassification bias is unlikely to have had significant influence on the observed results.

3.5. Conclusions

The results of this study suggest that compliance with certain on-farm biosecurity practices prevented horses on premises in high risk areas being infected with equine influenza during the 2007 outbreak in Australia. Local spread mostly occurred within 5 km of an infected premises, whilst long distance spread, over distances greater than 10 km, was associated with horse events where transmission was known to occurred. In future outbreaks, in addition to broader disease control measures, on-farm biosecurity practices should be adopted by horse owners and managers to prevent local spread. This research will inform the manner in which control strategies are implemented in the event of future epidemics of equine influenza and other infectious horse diseases in Australia.

3.6. Appendix

Equine influenza case control study - Questionnaire

Interviewer: ☐ Date of interview: ☐

Object ID: ☐

Day_Cx: ☐ OR Sampling date (Lab) ☐

GPS X: ☐ GPS Y: ☐

Entrance on private road ☐ Yes, length ☐ Meters
☐ No

SECTION A: Responder and property details

1. Property Name: ☐

2. Property street address:

☐

3. Surname of person interviewed: ☐

4. First name of person interviewed: ☐

5. Telephone &/or mobile: ☐

6. Email address: ☐

7. During the 2007 equine influenza outbreak, what was your role on the premises:

☐

7.1. Did your work involve visiting other horse properties?

☐ Yes ☐ No ☐ Not applicable

8. At the time of the equine influenza outbreak, how would you have best described your **property type**? (tick all that apply)

- ☐ Home with horses on site
- ☐ Spelling / Agistment / Grazing paddocks
- ☐ Horse training facility
- ☐ Equestrian centre
- ☐ Racetrack
- ☐ Commercial stud
- ☐ Riding school
- ☐ Pony club
- ☐ Other

Please specify:

9. At the time of the equine influenza outbreak, what activities did your horses mostly participate in **on your property**?

- ☐ Hobby / recreational riding only
- ☐ Agistment / spelling only
- ☐ Racing / race training:
  - ☐ Thoroughbred
  - ☐ Harness
  - ☐ Other (please specify: _____)
- ☐ Equestrian / event training:

Please specify type of activity:

- ☐ Breeding:
  - ☐ Thoroughbred
  - ☐ Other (please specify: _____)
- ☐ Other

Please specify:

10. Do you have a **map** of your property that we could use for reference in the following questions. If not, can we draw one now for reference purposes including : Shared fence-lines with other horse properties, Access points, and
Distance and direction to the nearest horse property, and where horses were kept.

11. What is the area of your property (please include units)?

	Units

12. How much of the total area is utilised for all horses (including foals)?
Please provide units or estimate as a per cent of your total property area

	Units or estimated %

13. How much of the total area is utilised for foals only?

	Units or estimated %

14. Does your property share a fence with other horse properties?

☐ Yes, please answer parts 14.1 to 14.3.
☐ No, skip to Question 15.

14.1. If Yes, how many are horse properties: ☐

14.2. If Yes, what is the total shared fence length with other horse properties:

☐ Metres

14.3. For each neighbouring horse property, please provide the following details (or estimates) at the time of the outbreak:

	Type of horse premises	**Direction of shared fence**	**Approx no. of horses**	**Infected with EI during outbreak**
e.g.Property1	*TB stud*	*NW*	*6*	*Y / N*

15. In the 2007-08 financial year, did your property have any income from a horse-related activity?

- ☐ Yes, the main source of income - please answer 15.1
- ☐ Yes, as an additional / secondary source of income - please answer 15.1
- ☐ No.
- ☐ Decline to answer this question.

15.1. If Yes, approximately how much income did your property derive from horse-related activities?

under $50,000	$50,000-$99,999	$100,000-$249,999	$250,000-$499,999	Over $500,000	uncertain	decline to answer
○	○	○	○	○	○	○

16. During 2007, prior to the equine influenza outbreak, how often did you use the services of a **farrier** on horses on your property:

- ☐ More than once a month
- ☐ Every one to two months
- ☐ Less frequently than every two months
- ☐ Cannot recall
- ☐ Not at all: reason:
  - ☐ Do it yourself
  - ☐ Other: _____

17. During 2007, did you vaccinate horses on your property for:

	Yes	No	Cannot recall
17.1. **Tetanus** (2in1 or Equivac T, not Equivac TAT)			
17.2. **Strangles** (2in1 or Equivac S)			

18. During 2007, how often did you worm horses on your property:

- ☐ Every three months (or more often)
- ☐ Every four to six months
- ☐ Less frequently than every six months
- ☐ Not at all
- ☐ Cannot recall

19. Were any of the horses on your property infected with equine influenza?

☐ Yes, please answer parts 19.1 and 19.2 only.
☐ No, please answer parts 19.3 and 19.4 only.

19.1. If Yes, on what date did the first horse on your property show any sign of equine influenza? (see list in part 19.2)

19.2. What signs of disease did this horse show over the course of its illness?

		Occurred	
19.2.1	Coughing	Yes	No
19.2.2	Increased rectal temperature	Yes	No
19.2.3	Discharge from the eyes	Yes	No
19.2.4	Discharge from the nose	Yes	No
19.2.5	Reduced appetite	Yes	No
19.2.6	Reduced alertness / temperament	Yes	No

19.3. If No, on what date did the NSW DPI visit your property for testing / vaccination?

19.4. **All properties:** Did any horses on your property receive EI vaccine before this date from any source other than NSW DPI?

☐ Yes, who and when:

☐ No.

SECTION B: Specified week

All further questions relate **to the specific week up to and including** either:

For Infected Premises:

The date the first horse on your property showed signs of being diseased,

(refer to Question 19.1)

For premises that were not infected:

The date your property was tested &/or vaccinated by the NSW DPI

(refer to Question 19.3)

Action to improve recall (not analysed):

Tell me about what happened on your property during the specified week:

[ ] To [ ]

20. How many horses (including foals) were on your property in the specified week?

[ ] Horses (total)

20.1. How many horses or foals on your property were in each of the following age groups? (approximate as a per cent if necessary)

Foals <1 year olds	Yearlings (1year olds)	2-4 year olds	5-10 yos	11-20 yos	>20 yos

[ ] Totals, or [ ] %

20.2. How many of these horses and foals were?

Colts (<3 yo)	Geldings	Stallions
Fillies (< 3yo)	Mares	

☐ Totals, or ☐ %

20.3. What were the breeds of these horses?

21. How were horses on your property predominately stabled during the specified week:

21.1 ☐ Indoors, or ☐ Outdoors.

21.2 ☐ In Boxes, or ☐ In Yards, or ☐ In Paddocks.

22. During the specified week, how often did you **check the horses on your property for signs of equine influenza** and other infectious diseases.

(signs of equine influenza: coughing, increased rectal temperature, discharge from the eyes or nose, reduced appetite, alertness or temperament.)

Two or more times per day	Every Day	Every other day	Once or twice	Never	Cannot recall
○	○	○	○	○	○

23. Did you keep a **record** of any such monitoring of the horses on your property?

☐ Yes ☐ No ☐ Cannot recall

23.1. If No and Not infected property, would you have kept a record if horses were infected on your property?

☐ Yes ☐ No

Part 1: Horse contact

24. During the specified week, **did other horses visit or arrive at your property** for any reason?

☐	Yes,	please answer part 24.1 and Question 25.
☐	No	go to Question 26
☐	Cannot recall	go to Question 26.

24.1. If Yes, please list the following details for these horses, by group that arrived:

Date of arrival	No. of horses	Source(s)	Purpose of visit / activities whilst on your property

25. On arrival at your property, were these visiting or new horses isolated from all other horses?

☐	Yes,	please answer parts 25.1 to 25.3
☐	No,	go to Question 26.
☐	Cannot recall,	go to Question 26.

25.1 If Yes, for how many days were they isolated: ☐

25.2 Please mark on the map of your property where they were isolated.

25.3 **IPs only:** Was the first horse to get sick on your property amongst these horses?

☐ Yes ☐ No ☐ Cannot recall

26. Did horse(s) leave your property for any reason during the specified week?

☐	Yes,	answer part 26.1 and Question 27.
☐	No,	go to Question 28.
☐	Cannot recall,	go to Question 28.

26.1 If Yes, please list the following details for these horses, by group that left:

Date of departure	No. of horses	Destination(s)	Purpose of trip / activities whilst on destination property

27. Did any of these horses **return** to your property during the specified week?

☐	Yes,	please answer parts 27.1 to 27.4
☐	No,	go to Question 28.
☐	Cannot recall,	go to Question 28.

27.1 If Yes, **on arrival** at your property, were these returned horses isolated from all other horses?

☐	Yes,	answer parts 27.2 to 27.4
☐	No,	go to Question 28.
☐	Cannot recall,	go to Question 28.

27.2 For how many days were they isolated: ☐

27.3 Please mark on the map of your property where they were isolated.

27.4 **IPs only:** Was the first horse to get sick on your property amongst these horses?

☐ Yes ☐ No ☐ Cannot recall

28. During the specified week, were any horses from your property involved in any **meetings** involving other horses (e.g. **events / shows / races / pony club** etc)?

☐	Yes,	please answer parts and 28.1 and 28.2.
☐	No,	go to Question 29.
☐	Not sure,	go to Question 29.

28.1 If Yes, please complete the details below for each event attended:

Name of the event	Location	Type	No. of your horses that attended

28.2 Whilst at any of these events or meetings, did you share any of the following horse equipment with other horse owners:

Saddles, bridles, harnesses or rugs	☐ Yes	☐ No	☐ Cannot recall
Feed or water bins	☐ Yes	☐ No	☐ Cannot recall
Other horse-related tools or equipment	☐ Yes	☐ No	☐ Cannot recall

(please specify):

Part 2: People contact

29. How many people **tended or rode** the horses on your property during the specified week? ☐ If none, go to Question 30.

Of these people, how many:

29.1	lived on the property?	☐
29.2	came from off the property?	☐

30. During the specified week,		Yes	No	Cannot recall
30.1	Did you have contact with horses on another property?	○	○	○
30.1.1	If Yes, did any of these horses have EI during the specified week?	○	○	○
30.2	Did others who worked or lived on your property have contact with horses on another property?	○	○	○
30.3	Did other people visit your property?	○	○	○
30.3.1	If Yes, did these other people who came to your property have contact with the horses on your property?	○	○	○
30.3.2	Did you ask these people to stay away from the horses on your property?	○	○	○
30.3.3	Did you ask these people to stay away from your horse yards, stables or paddocks?	○	○	○

31. During the specified week,		Yes	No	Cannot recall
31.1	Did you have a footbath for people to disinfect their shoes on arrival at your property?	○	○	○
31.2.1	If Yes, did all visitors to your property use the footbath **on arrival?**	○	○	○
31.2.2	Did you use disinfectant in the footbath.	○	○	○

32. During the specified week, **before each contact** with the horses on your property, how often did **you**:

		Every time	Most times	Sometimes	Never	Cannot recall
32.1	Wash your hands.	○	○	○	○	○
32.2	Use soap (or another detergent) when washing your hands.	○	○	○	○	○
32.3	Use hot water when washing your hands.	○	○	○	○	○
32.4	Change into freshly cleaned clothes.	○	○	○	○	○
32.5	Change into a clean pair of shoes.	○	○	○	○	○
32.6	Shower.	○	○	○	○	○

33. During the specified week, before **other people** contacted the horses on your property, how often did **they**:

		Every time	Most times	Sometimes	Never	N/A	Cannot recall
33.1	Wash their hands.	○	○	○	○	○	○
33.2	Change into freshly cleaned clothes.	○	○	○	○	○	○
33.3	Change into a clean pair of shoes.	○	○	○	○	○	○
33.4	Shower.	○	○	○	○	○	○

34. Horse professionals: During the week before testing,

	Yes	No	Not applicable Not Needed	Cannot recall
Did you ask any of the following not to come to your property?				
34.1 Veterinarians?	○	○	○	○
34.2 Farriers?	○	○	○	○
34.3 Other horse professionals ? (dentists, chiropractors, etc)	○	○	○	○

During the week before testing:

	Yes	No	Cannot recall
35. Did any **veterinarians** visit your property?	○	○	○
35.1 If Yes, did they change into clean clothes before contact with the horses on your property?	○	○	○
35.2 Were their clothes actually dirty before contact with the horses on your property?	○	○	○
36. Did any **farriers** visit your property?	○	○	○
36.1 If Yes, did they change into clean clothes before contact with the horses on your property?	○	○	○
36.2 Were their clothes actually dirty before contact with the horses on your property?	○	○	○
37. Did any **other horse professionals** visit your property?	○	○	○
37.1 If Yes, did they change into clean clothes before contact with the horses on your property?	○	○	○
37.2 Were their clothes actually dirty before contact with the horses on your property?	○	○	○

37.3 Who were these **other horse professionals**?

Part 3: Other contacts **(vehicles, equipment, feed, bedding)**

38. During the specified week, did you check if your boundary fences were secure?

- ☐ Yes, checked and certain all were secure.
- ☐ Yes, checked, but uncertain whether they were secure or not.
- ☐ No, did not check, but certain they were all secure.
- ☐ No, did not check, uncertain about whether secure or not.
- ☐ Cannot recall.

39. During the specified week, did you try to keep the horses on your property away:

39.1. From fences that bordered other horse properties?

- ☐ Yes, and successfully achieved this.
- ☐ Yes, but not certain whether achieved or not.
- ☐ No, did not attempt to.
- ☐ Cannot recall
- ☐ Not applicable

39.2. From fences that bordered roads or other public spaces?

- ☐ Yes, and successfully achieved this.
- ☐ Yes, but not certain whether achieved or not.
- ☐ No, did not attempt to.
- ☐ Cannot recall
- ☐ Not applicable

40. Vehicles: During the specified week,

		Yes	No	Cannot recall
40.1	Did you drive any of your vehicles onto other premises that held horses?	○	○	○
40.1.1	If Yes, did you disinfect these vehicles on arrival at your property?	○	○	○
40.2	Did others drive vehicles onto your property?	○	○	○
40.2.1	If Yes, were all of these vehicles disinfected on arrival at your property?	○	○	○
40.3	Was an area designated for disinfection of vehicles?	○	○	○
40.4	Did you use a horse float for any particular reason?	○	○	○

41. Horse gear: During the specified week, did you:

		Yes	No	Cannot recall
41.1	Share any horse gear (saddles, bridles, harnesses, rugs, halters, leads or brushes) with another property?	○	○	○

If Yes, before using your horse gear, how often did you:

		Every time	Most times	Sometimes	Never	Cannot recall
41.1.1	**Clean** this gear	○	○	○	○	○
41.1.2	Use **disinfectant or soap** when cleaning this gear?	○	○	○	○	○

		Yes	No	Cannot recall
41.2	Share any horse-related tools or other equipment with another property?	○	○	○

42. Feed: During the specified week, did you:

		Yes	No	Cannot recall
42.1	Share feed with another horse property.	○	○	○
42.2	Share feed or water bins with another horse property.	○	○	○
42.3	Obtain any **feed** from off your property:	○	○	○
42.4	If Yes, was this feed transported in vehicles that were decontaminated **before** entering your property?	○	○	○

43. Bedding: During the specified week, did you:	Yes	No	Cannot recall
43.1 Obtain any **bedding** from off your property:	○	○	○
43.2 If Yes, was this bedding transported in vehicles that were decontaminated **before** entering your property?	○	○	○

44. Effluent: During the specified week:	Yes	No	Cannot recall
44.1 Was horse **effluent** (**manure** and /or **bedding**) removed from your property?	○	○	○

44.2 If Yes, how was this done? How often? And by who?

Agreement to follow-up

Would you be willing to be contacted again by our research team? ☐ Yes ☐ No

Preferred contact person & details:

Chapter 4: Descriptive spatial and social network analysis

"*My it's a small world…*
In some sense, we are all bound together in a tightly knit social fabric."

Stanley Milgram (1967), 'The Small-World Problem', Psychology Today, Vol.1 (May): 61-67.

Popularising the findings of the 'small world experiment' (Travers and Milgram, 1969), the first study to be associated with the term: '**six degrees of separation**'.

This chapter appears as the following published paper in Preventive Veterinary Medicine:

Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2012. Adding the spatial dimension to the social network analysis of an epidemic: Investigation of the 2007 outbreak of equine influenza in Australia. Prev. Vet. Med. 106, 123-135.

4. Descriptive social network analysis and spatial kernel estimation

4.1. Introduction

Emerging infectious disease outbreaks commonly spread undetected for a period of weeks, involving contact networks that seed clusters of cases in widespread locations (Cohen, 2000). The networks driving initial epidemic spread often follow major transportation or trade routes between large cities, or between gatherings of animals such as markets. Severe acute respiratory syndrome (SARS) was first identified in Vietnam in late February 2003 after already spreading undetected through international air travel from China and Hong Kong, and seeding clusters of infection in hospital patients and health-care workers in Singapore and Canada (World Health Organization, 2003). In mid-April 2009, pandemic influenza A/H1N1 was first detected in the United States having already spread from Mexico (World Health Organization, 2009). Within 3 weeks, cases were reported from 30 countries in three continents. The index cases of foot-and-mouth disease (FMD) in the 2001 United Kingdom outbreak were detected in southeast England nearly a week after infected sheep had entered the market chain in the Northeast of the country, leading to a widespread epidemic (Gibbens and Wilesmith, 2002).

Equine influenza is a highly contagious and widespread viral respiratory disease affecting all members of the horse family (*Equidae*). The disease is characterised in naïve horses by a harsh dry cough, pyrexia, lethargy, anorexia and occasionally a nasal discharge (Newton et al., 2006). The incubation period is 1–3 days, and infected horses may shed the virus for 7–10 days commencing as early as 24 hours before disease onset (Myers and Wilson, 2006). Transmission to other horses may occur through direct contact, or via transmission on fomites and in windborne aerosols (Myers and Wilson, 2006). Although most horses recover uneventfully, outbreaks in susceptible populations can cause substantial economic and social impacts.

Until 2007, Australia was one of only three countries to have never experienced an outbreak of

equine influenza. Only small numbers of horses imported for breeding had been recently vaccinated, leaving almost the entire horse population susceptible. During an outbreak lasting 4 months in 2007, equine influenza (A/H3N8) spread rapidly through the horse populations of two Australian states, infecting approximately 67,000 horses on 9359 premises. Most of the geographic spread of this epidemic occurred undetected within the first 10 days and was associated with the movement of infected horses prior to the complete implementation of movement restrictions (Cowled et al., 2009b). Clinical signs were first observed in a recently imported horse in quarantine 40 km west of Sydney (Figure 4.1a) and 5 days later in two horses kept in a large horse boarding and training facility in central Sydney (Centennial Parklands Equestrian Centre). These two horses had attended an equestrian event at Maitland, 160 km north of Sydney, on the previous weekend (17–19 August 2007). Contact-tracing revealed numerous infected premises linked to horses returning from either the event at Maitland, or another horse event the following weekend, at the Narrabri showground nearly 400 km away (Callinan, 2008). In accordance with the Australian Veterinary Emergency Plan horse movement restrictions were triggered immediately on 25 August 2007 (Callinan, 2008), subsequent spatial spread was believed to be driven by local spread from those premises infected in the first weeks of the epidemic (Cowled et al., 2009b). Previous analyses identified two spatial effects in the first month of the epidemic: local spread within 5 km of infected premises, and spread over much greater distances associated with the attendance of horses at equestrian events where transmission was known to have occurred (Firestone et al., 2011a).

When a complex contact network structure underlies an epidemic, traditional approaches may be insufficient to appropriately describe the spatio-temporal pattern of the epidemic and estimate key parameters (Small et al., 2007). For example, methods of estimating the reproductive ratio of an epidemic often rely on assuming homogenous probabilities of contact amongst infected and susceptible individuals in the population at risk (i.e. random mixing). However, the underlying contact network structure constrains contact leading to violation of this assumption (Anderson and May, 1991). To this end, social network analysis (SNA) has been

used to study contact-tracing data from epidemics in both human (McElroy et al., 2003, Sena et al., 2007) and animal populations (Ortiz-Pelaez et al., 2006, Shirley and Rushton, 2005). These approaches have largely ignored spatial proximity between infected premises. Recently, Green et al. developed a spatially explicit model to estimate the importance of cattle movements in the spread of bovine tuberculosis (Green et al., 2008), whilst García Álvarez et al. investigated the importance of animal movements, contacts and spatial relationships between dairy farms in the context of endemic disease spread (García Álvarez et al., 2011). Network analysis has also been applied to investigate animal movements, irrespective of disease status, to develop hypothetical models of how disease might spread through animal populations (Martinez-Lopez et al., 2009a, Webb, 2005).

A network, social or otherwise, is a set of connections amongst a group of nodes (Hanneman and Riddle, 2005). Networks can be used to represent the patterns of connectivity of populations, and therefore describe aspects of disease transmission that depart from the mean field model (Diekmann et al., 1998). Depending on the disease context, a network can be constructed from contact-tracing data representing different units of interest, with nodes representing either individual animals (Porphyre et al., 2008) or farms (Ortiz-Pelaez et al., 2006), schools, communities or households (Giebultowicz et al., 2011). The connections can represent a wide range of context-dependant relationships or links including actual physical contact, movement of a resource from one node to another, or the sharing of a space at a given point in time. The links thus formed may be either directed or undirected depending on whether or not a resource can travel in both directions between the pair of nodes (Wasserman and Faust, 1994). Undirected links can be represented in directed networks by two directionally opposed directed links. Depending on whether the magnitude of flow is contextually important, links may also be weighted by values or dichotomised.

The aim of this study was to apply social network analysis to describe spread of equine influenza between horse premises infected in the early phase of the 2007 Australian outbreak,

differentiating spread that occurred through a known contact network from secondary local spatial spread. The methods used may be applied to describe other epidemics in which disease transmission occurs initially through networks (such as transport networks or market systems) followed by local spread to adjacent farms.

4.2. Materials and methods

Three networks were constructed based on epidemiological data collected during the outbreak: a 'contact network' of horse movements between all premises infected during the first 10 days of the outbreak; a 'proximity network' based on a distance matrix between these infected premises; and a 'contact-and-proximity network' constructed by combining the contact network with the proximity network. Characteristics of these networks were analysed and compared to identify key characteristics of the spread during the early phase of this epidemic.

4.2.1. Data sources

The networks were constructed from three datasets provided by the New South Wales (NSW) and Queensland state governments: a contact-tracing dataset, an infected premises dataset and an uninfected premises dataset. Premises were defined as infected if they held horses that had been observed with the classical clinical signs of equine influenza (cough, elevated temperature, nasal discharge and lethargy) during the study period, and if the diagnosis had been confirmed by laboratory testing based on real-time reverse transcription polymerase chain reaction assay (Foord et al., 2009).

The contact-tracing dataset included 1034 reported horse movements onto and off premises investigated by NSW and Queensland animal disease authorities during the outbreak. Each movement record included the date of the movement, and the addresses and unique identifiers of the origin and destination premises. Most of these records were traced backwards during interviews of horse owners and managers whilst attempting to establish the source of infection

and implement quarantine and horse movement restrictions. Data on eighteen further horse movements were obtained and added to the dataset based on a review of reports from epidemiological investigations of this outbreak (Equine Influenza Epidemiology Support Group, 2008, Callinan, 2008).

The infected and uninfected premises datasets contained node attribute data, including: address, geocoded coordinates (based on property centroid), number of horses, premises area, and for infected premises, the date of onset of first clinical signs of the first horse affected ('onset date'). These datasets were linked to a database of laboratory testing records and thus a quasi-census of outbreak-affected regions in 2007 was created.

4.2.2. Data management and preparation

The infected premises dataset had previously been prepared for epidemiological analyses as described elsewhere (Cowled et al., 2009b). The uninfected premises and contact-tracing datasets were less complete and required extensive checking for typographic errors, removal of duplicates and geocoding of missing coordinates. Where possible, originating and destination premises were matched to the infected premises and uninfected premises datasets based on identical unique identifiers or addresses. Records that could not be matched were excluded from further analyses. All data were collated into a relational Microsoft Access 2007 database (Microsoft Corporation, Redmond, WA , USA) and then exported for further analyses into the R statistical package version 2.12.0 (R Development Core Team, 2011).

4.2.3. Network construction

In all networks, the nodes represented all premises holding horses infected within the first 10 days of the epidemic (corresponding to the period until complete implementation of the horse movement restrictions). These premises were selected from the infected premises dataset, allowing for the typical 1–3 day incubation period of equine influenza, and a 1 day margin for

error in observation (i.e. infected premises with onset date prior to 31 August 2007, the fourteenth day of the epidemic).

4.2.3.1. The contact network of infected horse movements

Directed links – representing horse movements between infected premises until the tenth day of the outbreak – were added to the contact network. As data were not available on actual numbers of horses moved, these links were binary (movement of a single horse was treated equivalently to several horses moving between two premises). To maintain the temporal dimension the date of movement was included as an attribute of all such contact links.

All of the back-traced horse movements occurred prior to disease detection, with infected horses being moved whilst incubating the disease. Infected horses can shed virus as early as 24 hours after infection, and 1 day before developing clinical signs, and they may continue to shed virus for 7–10 days (Myers and Wilson, 2006). Consequently, movements were only included if they occurred on or after the day before clinical signs were first observed on the originating premises, and if clinical signs were observed at the destination premises on or after the day of movement. A 1 day margin of error was included in these calculations to allow for inconsistent observation and reporting practices.

4.2.3.2. The proximity network of infected premises in space

The proximity network was constructed based on a matrix of the distances between each pair of infected premises in the contact network. The spatial coordinates of the centroid of each premises were converted to the Albers conic equal-area projection based on the Geocentric Datum of Australia 1994 (www.ga.gov.au/geodesy/datums/gda.jsp), and pair-wise distances were calculated between all premises. The resulting distance matrix was dichotomised at a distance cut-off used to represent an assumption of the maximum distance over which local spread might have occurred. A 5 km cut-off was selected for initial analyses based on previous

empirical research which suggested that local transmission mostly occurred within this distance (Firestone et al., 2011a). Although the proximity network was symmetrical it was considered as a directed network to facilitate combination with the contact network. Spatial coordinates were included as node attributes in all networks to maintain the spatial dimension. Rather than applying a set distance cut-off, we could have applied a nearest-neighbour method (creating a less dense and less clustered network) but this would have required implicitly assuming that the nearest of a group of proximate infectious premises was the actual source of infection.

4.2.3.3. The combined contact-and-proximity network

To produce the contact-and-proximity network, first the contact and proximity networks were transformed into valued networks by nominally weighting their links. All links in the transformed contact network were valued '1', similarly all links in the transformed proximity network were valued '2'. When combined, this allowed differentiation in the contact-and-proximity network between pairs of nodes that were connected by contact only ('1'), proximity only within the respective distance cut-off ('2') or contact and proximity ('3'). The contact-and-proximity network was therefore valued and directed, with proximity (within the distance cut-off) represented by two opposing links between a pair of nodes.

4.2.4. Network analyses

All network analyses were conducted in the R statistical package using the 'statnet' library (Handcock et al., 2003). Networks were described by their size, centrality and cohesion (Table 4.1). Graphs were constructed for each network incorporating temporal and spatial dimensions, and identifying equestrian events where disease transmission was known to have occurred from other horse-holding premises (such as studs, farms and homes on small acreages).

Two classes of network sub-structures were described to identify horse movements and infected premises that were likely to be of importance to disease spread: components and cutpoints. A

component is a subset of nodes connected to each other but disconnected from all other nodes, and cutpoints are nodes that if deleted would fragment the network into a larger number of smaller components (Hanneman and Riddle, 2005). In directed networks, two types of components may be defined: 'strong components' where every node within the subset can be reached from every other node obeying the direction of links, and 'weak components' where the directions of the links are disregarded (Hanneman and Riddle, 2005). From random network theory, it is known that if links are added randomly to an empty network a tipping-point is rapidly reached above which one very large

('giant') component is created that is much larger than the next largest component (Erdős and Rényi, 1961). The relative proportions of premises infected by horse movements (as described by the contact network), local spatial spread (described by the proximity network), and unexplained ('isolates', unconnected premises) were estimated by calculating the number of nodes in the giant weak component (GWC) of each network.

The centrality of each individual node (which reflects its potential importance to disease spread) was analysed by calculating the betweenness, degree and reach of each node in each network. The betweenness of a node is the frequency that it lies along the shortest path (the 'geodesic') between other nodes in the network (Freeman, 1979). The degree of each node, being the number of connections incident upon that node, was also differentiated into 'in-degree' and 'out-degree'. Two nodes are 'reachable' if there is a set of connections between them, and the reach of a node is the longest geodesic distance to another reachable node. 'Reach-to' and 'reach-from' each node was calculated as the three networks were directed.

Table 4.1 Description of network measures calculated for contact, proximity and contact-and-proximity networks of the first 10 days of the 2007 equine influenza outbreak in Australia.

Parameter	Description
Network size:	
Number of nodes	Each representing a premises holding horses infected in the first 10 days of the epidemic.
Number of directed links	Links are connections between nodes, in this study representing horse movements and/or premises proximity.
Number of isolates	The number of unconnected nodes.
Average path length	The shortest path (geodesic) between each pair of nodes, measured only amongst reachable pairs of nodes.
Network diameter	The largest geodesic distance between all reachable pairs of nodes in the network.
Network centrality: Betweenness Degree Reach	Heterogeneity in a node-level parameters, used to describes the extent that the network revolves around any single node. Calculated as a normalised proportion, summing the differences between the largest value in the network and all other observed values, divided by the theoretical maximum possible sum of differences for a hypothetical network of the same size (Wasserman and Faust, 1994).
Network cohesion:	
Density	General level of cohesion, calculated as the proportion of all possible links actually present using the formula: $\text{Density}_{\text{directed network}} = \frac{L}{k(k-1)}$ where: L = number of directed links present, and k = number of nodes.
Clustering coefficient	In a directed network, for each node (i) with out-degree >1, the mean proportion of possible links actually present amongst the neighbourhood (N_i) of nodes that i outwardly connects to (Turner et al., 2008): $\text{CC}_{\text{average}}^{\text{out-degree}} = \frac{1}{n}\sum_{i=1}^{n}\frac{L_{Ni}}{j(j-1)}$ where: L_{Ni} = number of arcs present in N_i, and j = out-degree of node i.

Heterogeneity in degree demonstrates that non-homogeneous mixing is occurring in the population, and this is important epidemiologically. The basic reproductive ratio of an epidemic (R_0) describes the potential transmissibility of a disease in a totally susceptible population, and is defined as the number of secondary infections produced, on average, by a typical infective case assuming homogenous random mixing (i.e. equal likelihood of contact between premises) (Anderson and May, 1991). R_0 may be estimated at the individual or premises-level, and can be adjusted to account for non-susceptible hosts, control measures such as vaccination, and non-homogenous mixing (Matthews and Woolhouse, 2005). A related concept, the effective reproductive ratio (R) is the average number of secondary infections produced by each infected individual that enters a population that contains non-susceptible individuals or is subject to disease control measures (Matthews and Woolhouse, 2005). R is often directly estimated based on actual epidemic data, and may approximate R_0 early in an epidemic. Cowled et al. (2009b) applied purely spatial methods to estimate R at the premises-level for this epidemic, assuming homogenous random mixing (equal likelihood of contact between premises). We corrected their naïve estimate (ρ) for the early outbreak period, using the following formula (May et al., 2001):

$$R = \rho\left(1 + \left[CV_{deg}\right]^2\right) \quad (4.1)$$

where the coefficient of variation of degree (CV_{deg}) represents heterogeneity in contact as the ratio of the standard deviation of degree to the mean degree, which was estimated based on the contact network.

4.2.5. Estimating the range of local spatial spread

The distance cut-off used to dichotomise the proximity networks can be considered an assumption of the distance over which local spread of equine influenza occurred. As the distance cut-off is increased, fragmentation in the proximity network is reduced leading to fewer isolates in the combined contact-and-proximity network. Sensitivity analysis was conducted to

test the influence of varying the distance cut-off on the number of nodes included in the largest weak component of the contact-and-proximity network (Webb, 2005). The point when further increases in the distance cut-off did not lead to inclusion of further isolates was considered to approximate the effective range at which local spread was likely to have occurred.

A simple transmission kernel was then directly estimated for the first 2 weeks of the epidemic, using an approach similar to that applied to contact-tracing data from the 2001 FMD outbreak in the United Kingdom (Keeling et al., 2001), and compared to the shape of the combined contact-and-proximity network's fragmentation curve. Briefly, for each infected premises, every other premises that it may have infected (a 'possible transmission') was identified based on corresponding dates of onset of clinical signs and an assumed premises-level infectious period starting the day before clinical signs were first observed on a premises, and continuing until the end of the study period, given that the 10 day study period was relatively short compared to the typical 7–10 day infectious period of individual horses (Myers and Wilson, 2006). The probability of infection was then estimated in narrow distance bands (250 m) as the proportion of possible transmissions out of the total sum of potential transmissions (susceptible population at risk at the commencement of the premises' infectious period within that band), excluding premises included in the contact network from all calculations.

For each premises infected in the first 10 days of the epidemic ($i = 1, \ldots, n$) we estimated the probability that it was infected by local spatial spread (π_{Si}) or through the contact network (π_{Ni}), considering that:

$$\pi_{Si} = 1 - \pi_{Ni} \qquad (4.2)$$

For each of these premises that was not included in the contact network, $\pi_{Ni} = 0$, as they must have been infected by local spatial spread. For each premises included in the contact network, all premises considered infectious in the period when premises i was presumably infected were identified ($j = 1, \ldots, m$), and the Euclidean distance to each such potential source premises was

calculated ($\boldsymbol{d_{ij}}$). The probability that premises *i* was infected by each of these potential source premises, $p(\boldsymbol{d_{ij}})$, was estimated using the transmission kernel. The complementary probabilities, $[1-p(\boldsymbol{d_{ij}})]$, that each potential source premises did not actually infect premises *i*, were multiplied together to estimate the probability that none of the potential source premises proximate to premises *i* actually infected premises *i* through local spatial spread. Then, π_{Si} was estimated as:

$$\pi_{Si} \approx 1 - \prod_{j=1}^{m}\left[1 - p(\boldsymbol{d_{ij}})\right] \quad (4.3)$$

Next, the expected number of premises that were likely to have been infected through local spatial spread, $E[\pi_S]$, and its variance were estimated:

$$E[\pi_S] = \sum_{i=1}^{n} 1 - \pi_{Si} \quad (4.4)$$

$$var(\pi_S) = \sum_{i=1}^{n} \pi_{Si}(1 - \pi_{Si}) \quad (4.5)$$

allowing calculation of the complementary proportion of premises infected directly through the contact network, $E[\pi_N]$, and 95% confidence intervals (CI) based on its expected variance, where:

$$var(\pi_N) = \sum_{i=1}^{n} \pi_{Ni}(1 - \pi_{Ni}) = \sum_{i=1}^{n} \pi_{Ni}(\pi_{Si}) = \sum_{i=1}^{n} (1 - \pi_{Si})\pi_{Si} = var(\pi_S) \quad (4.6)$$

4.2.6. Cluster description

To describe spatial clusters of infected premises whilst also considering network topology, we selected the proximity and contact-and-proximity networks with the distance cut-off equal to the estimated maximal distance of local spread (the 'maximal proximity network' and 'maximal contact-and-proximity network'). These networks were then analysed using block-modelling, a technique that involves grouping nodes into 'blocks' (also termed 'clusters') based on some

similarity (Wasserman and Faust, 1994). This may be either a structural similarity related to position in the networks, or based on some other concept. Networks were reconstructed representing the relationships between these groups. In this analysis, we grouped nodes into 'clusters' according to the weak components in the maximal proximity network, then block-modelled the maximal contact-and-proximity network based on these clusters. Such analyses emphasise movements that introduced infection into new spatial clusters.

4.3. Results

In the first 10 days of the equine influenza outbreak in Australia, horses on 197 premises were infected. Of the 1052 horse movements in the contact-tracing dataset, 978 occurred during the first 10 days of the epidemic, with 300 originating from infected premises. Only 70 of these horse movements were to subsequently infected premises with appropriately corresponding dates of movement and onset on the originating and destination premises. These infected horse movements (Figure 4.1) were all backwards traced, and covered the 10 days prior to the complete implementation of movement restrictions. The median distance that infected horses were moved was 123 km (range 4–579 km).

4.3.1. The contact network of infected horse movements

Over 85% of the infected horse movements originated from the equestrian events at Maitland and Narrabri. The contact network, comprising 197 nodes and 70 infected horse movements, was dominated by these two events where disease transmission is known to have occurred. Each event functioned as a 'hub' in the early spread of this epidemic (Figure 4.2a). The Maitland and Narrabri events were potential cutpoints in the network with out-degrees of 26 and 34 premises, respectively. Only five other nodes, representing the equestrian events at Warwick, the Centennial Parklands Equestrian Centre, and three smaller private horse premises (nodes 26, 113 and 388), served as the origins of infected horse movements (out-degree >0). Each of these nodes functioned as an additional cutpoint in the contact network. This heterogeneity in degree

is quantified in the contact network's relatively high out-degree centralisation. Conversely, the contact network's in-degree centralisation was very low, with only one node (Centennial Parklands Equestrian Centre) being the destination of infected horse movements from multiple premises (in-degree = 2). The latter of these movements originated from the Narrabri event and is represented in Figure 4.2b by a dotted line.

A single node (113) with high betweenness was intermediate to the Maitland and Narrabri events and to a third event across the Queensland state border in Warwick. Horses moving from the Maitland event to the Warwick event were held midweek at premises '113' (Figure 4.1b). In-contact horses from this premises travelled to the event at Narrabri on the following weekend, thereby introducing infection to that event. Overall, the contact network was sparsely connected with negligible network density and clustering (Table 4.2). It consisted of a single component, including 36% of the nodes, and 127 unconnected premises (these isolates are excluded from Figure 4.2). The heterogeneous structure of this contact network, dominated by several nodes with relatively high out-degree, produced a seven-fold increase in the estimated effective reproductive ratio of this epidemic from $\rho = 2.0$ (calculated assuming random mixing) to $R = 14.6$.

4.3.2. The proximity of infected premises in space

The proximity network was highly fragmented irrespective of the distance cut-off (Figure 4.3). With the distance cut-off dichotomised at 5 km, the proximity network comprised 28 relatively small components and 54 isolated premises (Figure 4.3a). Clustering was very high within these highly connected components (clustering coefficient = 90.2%). Each component represented several premises infected in the same short time period and tightly grouped in space ('a cluster'). The six cutpoints in this proximity network were completely different from those in the contact network, each representing a premises located spatially within 5 km of two non-adjacent premises (more than 5 km apart). Compared to the contact network, the proximity

network with cut-off at 5 km contained many more links (n = 852), and these links were more evenly distributed across the network (in-degree and out-degree centralisation indices = 0.07).

4.3.3. The combined contact-and-proximity network

The combined network displays how the premises infected within the first 10 days of the epidemic were connected through the movement of infected horses and spatial proximity (Figure 4.4). With the distance cut-off set at 5 km, this network contained 6 components and 19 isolated nodes (Figure 4.4a). The largest of these components, the GWC, contained 85% of the nodes (n = 167), representing all infected premises described by the known contact-tracing data (36%) and those infected premises described by local spread within 5 km (a further 49%). The five smaller components each comprised between two and three spatially adjacent infected premises (within 5 km of each other) that were neither connected to any premises in the GWC through known infected horse movements nor through spatial proximity within 5 km. The 19 isolates were similarly disconnected from the GWC. Overall, this contact-and-proximity network (with distance cut-off set at 5 km) inherited a relatively high out-degree centralisation from the contact network, reflecting the relative importance of several nodes with high out-degree in the contact network to the combined network structure. This network also displayed two features characteristic of a small world network: short average path length (L = 2.22) and a high degree of clustering (clustering coefficient = 86.1%). The 22 cutpoints in this network included all of the cutpoints from the contact network, and all except one of the cutpoints from the proximity network.

4.3.4. Estimates of the range of local spatial spread

The proportion of infected premises included in the GWC of the contact-and-proximity network (the inverted fragmentation index) varied according to the distance cut-off (Figure 4.5). With the distance cut-off set at zero, local spread was not described; the 36% of nodes included in the GWC of this network are those nodes explained by the structure of the contact network alone.

When the cut-off was increased to 5 km, 85% of nodes were included in the GWC, corresponding to the network in Figure 4a, and when the cut-off was increased to 15.3 km, all of the remaining isolates were included in the GWC of the contact-and-proximity network (Figure 4.4b). Suggesting that most of the early spread during this epidemic was effectively 'explained' with a cut-off distance of 15.3 km, providing an estimate of the effective range over which local spread occurred (through direct contact, transmission on fomites and windborne transmission). Figures 4.3b and 4.4b show the proximity and combined contact-and-proximity networks with distance cut-off set to 15.3 km. These networks were termed the 'maximal networks', because beyond this cut-off all nodes were incorporated into the single GWC of the contact-and-proximity network.

The directly estimated transmission kernel was highly comparable in shape to the fragmentation index (Figure 4.5). The risk of infection decreased rapidly within 5 km of an infected premises, then continued to gradually decline, with little risk beyond 15 km. Of the 70 premises in the contact network, 14 premises were likely to have been infected through local spatial spread from nearby infected premises rather than directly through the contact network (95% CI: 9, 20), suggesting that 28.3% of spread in the early epidemic period was 'network-associated' (95% CI: 25.6, 31.0%).

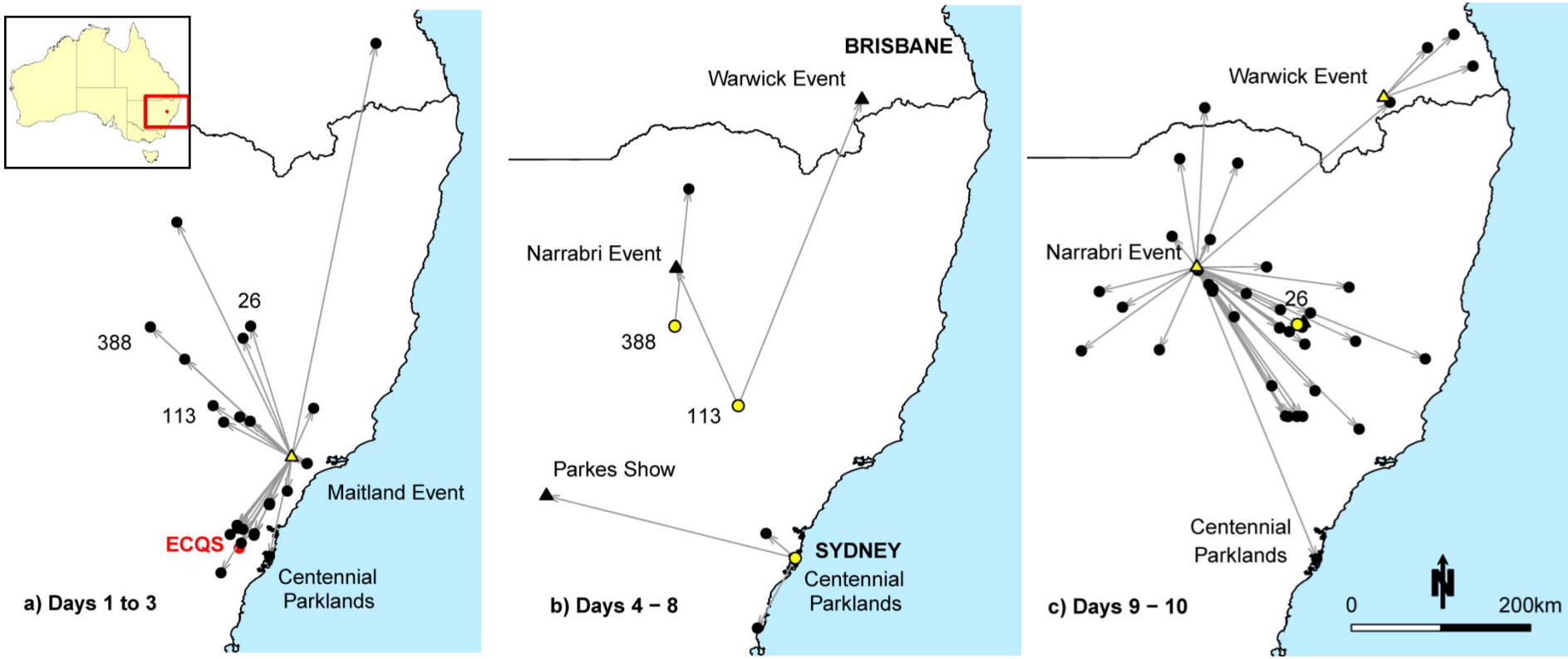


Figure 4.1 Infected horse movements during the first 10 days of the equine influenza outbreak of 2007 in Australia.
(a) 17–19 August 2007, (b) 20–24 August 2007, (c) 25–26 August 2007. Nodes are coloured yellow and labelled if their out-degree ≥1. Horse events where transmission is known to have occurred are denoted by triangles. Numerical labels are unique node identifiers. The red node (ECQS) denotes Eastern Creek Quarantine Station (Sydney) where clinical signs were first observed in a horse in quarantine on 17 August 2007.

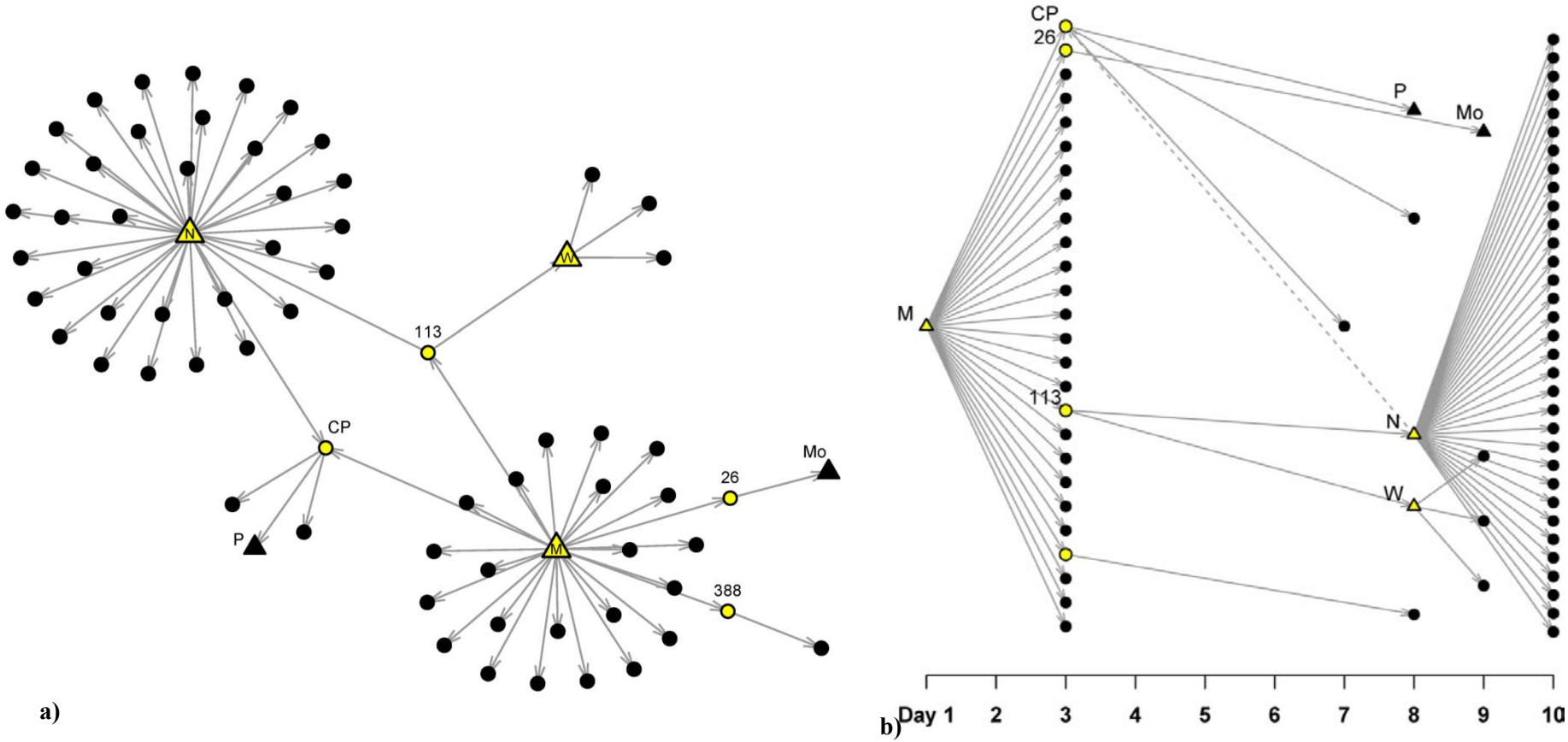


Figure 4.2 Contact network representing infected horse movements between premises holding horses infected in the first 10 days of the 2007 equine influenza outbreak in Australia.
(a) Graphed in arbitrary space, excluding 135 isolates. (b) Time-ordered dendrogram, excluding isolates. Nodes are coloured yellow if their out-degree ≥1. Horse events where transmission is known to have occurred denoted by triangles. Node labels: Centennial Parklands Equestrian Centre Sydney (CP), Maitland event (M), Moonbi event (Mo), Narrabri event (N), Parkes show (P), Warwick event (W), other labels are unique premises identifiers.

Table 4.2 Network-level parameters calculated for contact, proximity and contact-and-proximity networks of the first 10 days of the 2007 equine influenza outbreak in Australia.

Parameter	Contact network	Proximity network		Contact-and-proximity network	
		5 km[a]	15.3 km[b]	5 km[a]	15.3 km[b]
Network size:					
Number of nodes	197	197	197	197	197
Number of directed links	70	852	2518	921	2586
Number of isolates	127	54	24	19	0
Average path length	1.80	1.63	1.45	2.22	1.92
Network diameter	3	4	4	6	6
Network centralisation:					
Betweenness centralisation	0.002	0.005	0.004	0.008	0.015
In-degree centralisation	0.01	0.07	0.20	0.07	0.19
Out-degree centralisation	0.17	0.07	0.20	0.15	0.19
CV of total degree	2.51	–	–	–	–
Effective reproduction ratio[c]	14.6	–	–	–	–
Network cohesion:					
Density	0.002	0.022	0.065	0.024	0.067
Clustering coefficient	0.000	0.902	0.906	0.861	0.882
Network substructures:					
Number of components	1	28	20	6	1
Size of largest component (number of nodes)	70	28	58	167	197
Number of cutpoints	7	6	6	22	19

CV = coefficient of variation.
[a] Distance cut-off dichotomised at 5 km based on empirical research Firestone et al. (2011a). [b] Maximal network' with distance cut-off of 15.3 km, all nodes are incorporated into the largest weak component. [c] Based on estimate of herd-level $R \approx 2.0$ from Cowled et al. (2009b).

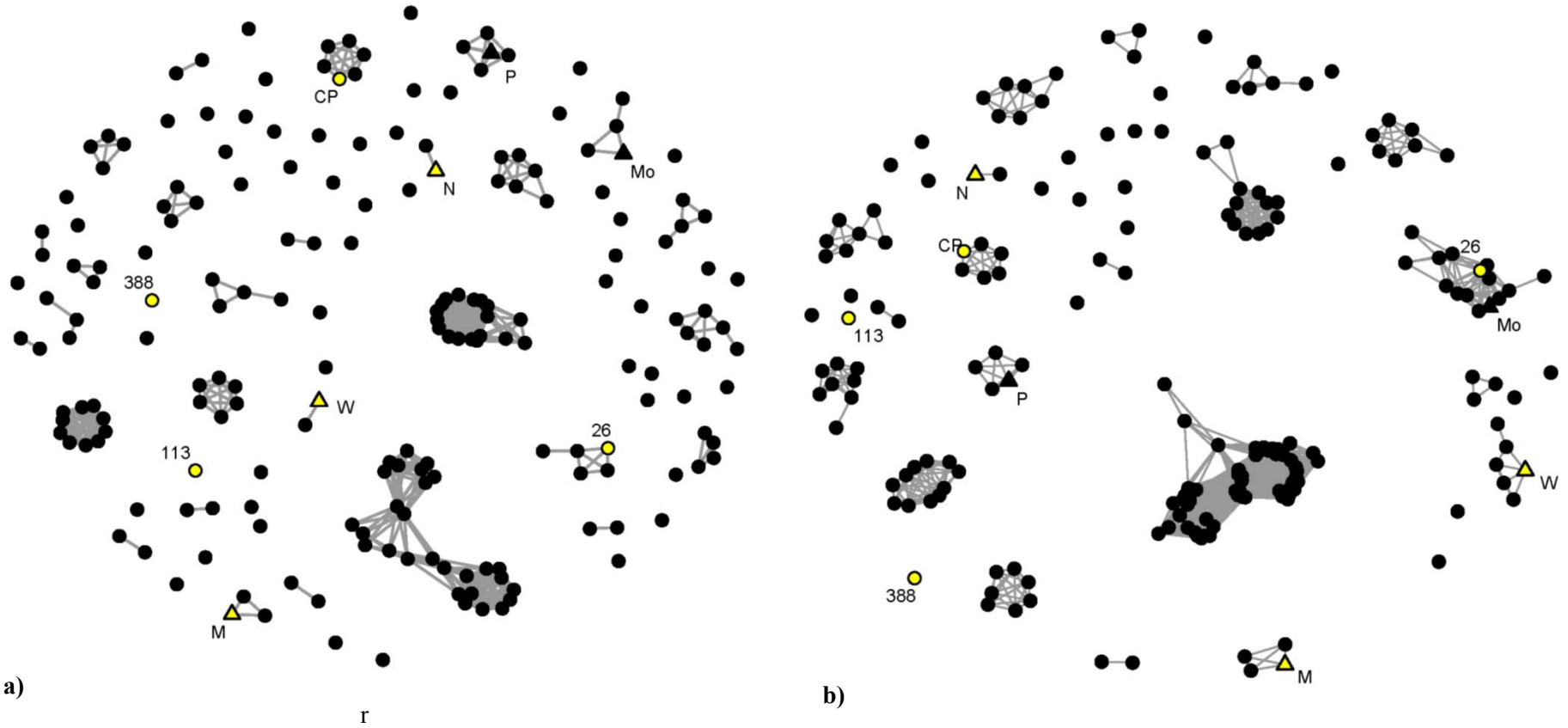
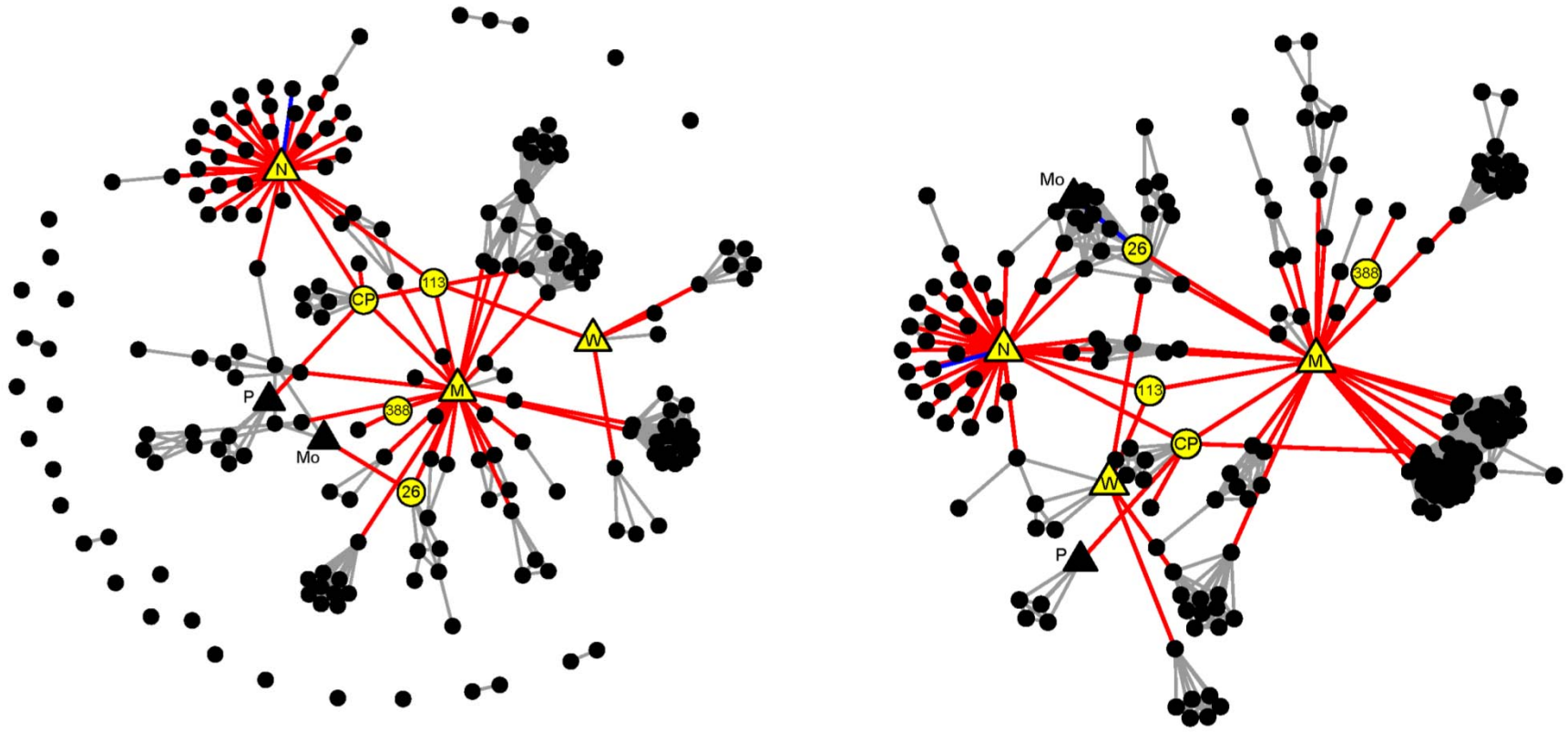


Figure 4.3 Proximity networks representing spatial relationships between premises holding horses infected in the first 10 days of the 2007 equine influenza outbreak in Australia.
(a) Distance cut-off dichotomised at 5 km, (b) 'Maximal network' with distance cut-off of 15.3 km. Graphed in arbitrary space, including isolates. Nodes are coloured yellow if their out-degree ≥1. Horse events where transmission is known to have occurred denoted by triangles. Node labels: Centennial Parklands Equestrian Centre Sydney (CP), Maitland event (M), Moonbi event (Mo), Narrabri event (N), Parkes show (P), Warwick event (W), other labels are unique premises identifiers.



a) **b)**

Figure 4.4 Contact-and-proximity networks representing spatial relationships and infected horse movements between all premises holding horses infected in the first 10 days of the 2007 equine influenza outbreak in Australia.
(a) Distance cut-off dichotomised at 5 km. (b) 'Maximal network' with distance cut-off of 15.3 km, all nodes are incorporated into the same weak component. Graphed in arbitrary space, including isolates. Red links represent the movement of infected horses, grey links represent spatial proximity with distance cut-off set at 5 km, blue links signify the movement of infected horses over a distance less than the cut-off. Nodes are coloured yellow if their out-degree ≥1, triangles denote horse events. Centennial Parklands Equestrian Centre Sydney (CP), Maitland event (M), Moonbi event (Mo), Narrabri event (N), Parkes show (P), Warwick event (W).

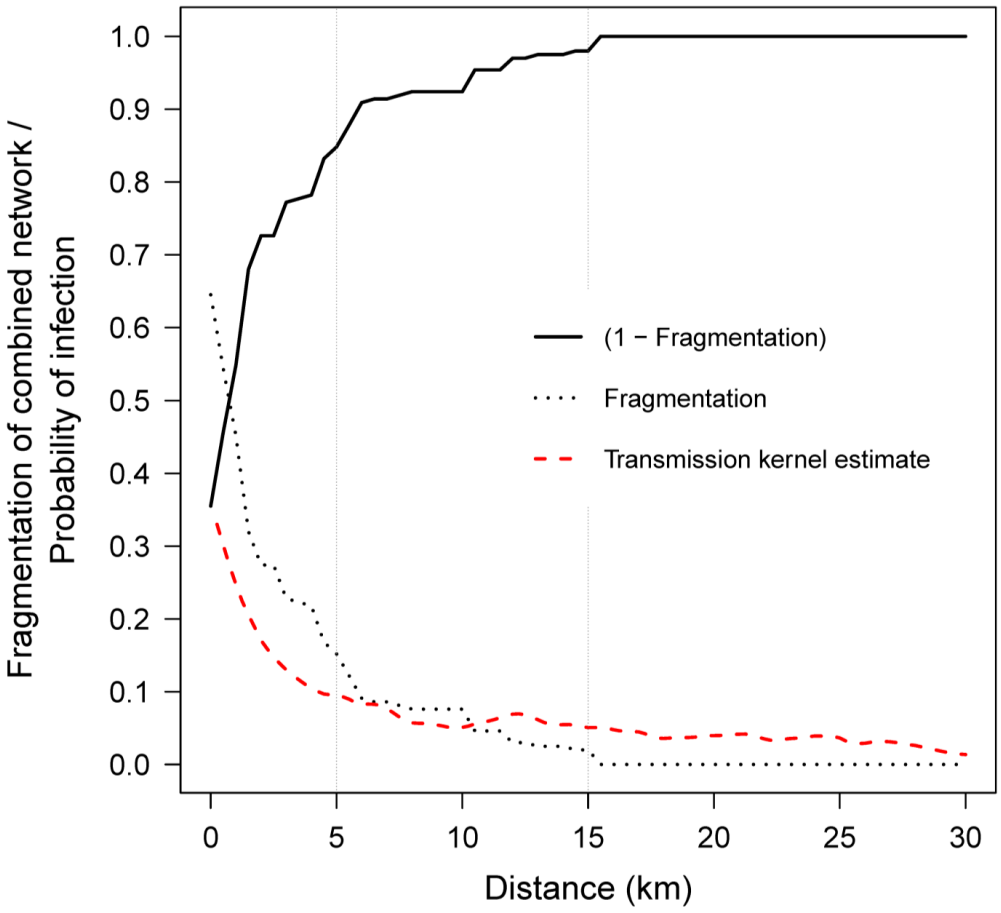


Figure 4.5 Estimates of the effective range of local spread in the first 10 days of the 2007 equine influenza outbreak in Australia.
The inverted fragmentation index represents the proportion of nodes included in the giant weak component of the combined contact-and-proximity network when the distance cut-off is varied. The directly estimated transmission kernel for the same time period is overlaid for comparison. Most local spread occurred within 5 km of an infected premises, with very little local spread beyond 15 km.

4.3.5. Cluster description

Forty-four clusters were identified based on the maximal contact-and-proximity network applying a distance cut-off of 15.3 km (Figure 4.6). Over the entire epidemic, 82.6% of infected premises were contained within the boundaries of these clusters. Of the 70 infected horse movements, only 43 movements introduced infection into a new spatial cluster. Fifteen of these originated from the Maitland event on 19 August 2007, and 21 from the Narrabri event on 26 August 2007. The remaining seven movements occurred between 23 and 25 August 2007, and were all from cutpoints previously identified in the contact network (Centennial Parklands Equestrian Centre, the Warwick event, and nodes '113' and '388').

4.4. Discussion

Social network analysis is a valuable tool for describing epidemics with an underlying contact network structure. By including spatial proximity between infected premises as an additional relationship in the network analysis of the early spread of an actual epidemic of infectious disease, we were able to discriminate network-based from secondary local spatial spread of equine influenza. As the contact and proximity networks were similarly constructed they could be combined to describe contacts between and within clusters, providing a clear picture of how the traced contacts led to the formation of clusters.

The network of infected horse movements during the first 10 days of the 2007 equine influenza outbreak in Australia dictated the spatial extent of the ensuing epidemic. Certain movements introduced infected horses into new regions, thereby initiating new clusters of infection in highly susceptible populations. Over the entire geographic extent of the outbreak, we observed that 85% of those premises infected within the first 10 days were within 5 km of an infected premises described by the contact network; 100% were within 15.3 km. These estimates are highly comparable with the directly estimated transmission kernel and a previous analysis of local spread within a single cluster of this outbreak (Davis et al., 2009) where 91% of newly infected premises were found to be within 2 km of an earlier identified infected premises, and the largest observed spread distance was 13 km.

By constructing a maximal proximity network based on our estimate of 15.3 km for the maximum distance of local spread, spatial clusters of infected premises were identifiable as this network's weak components. Each cluster represents the upper bound of premises that may have been infected through local spread in that small area. Further analysis into the spatial pattern of local spread within individual clusters would be useful for zoning and outbreak management in the event of future outbreaks. As several of the described clusters partly overlapped, these clusters may be coalesced. However, care would be required to retain useful information

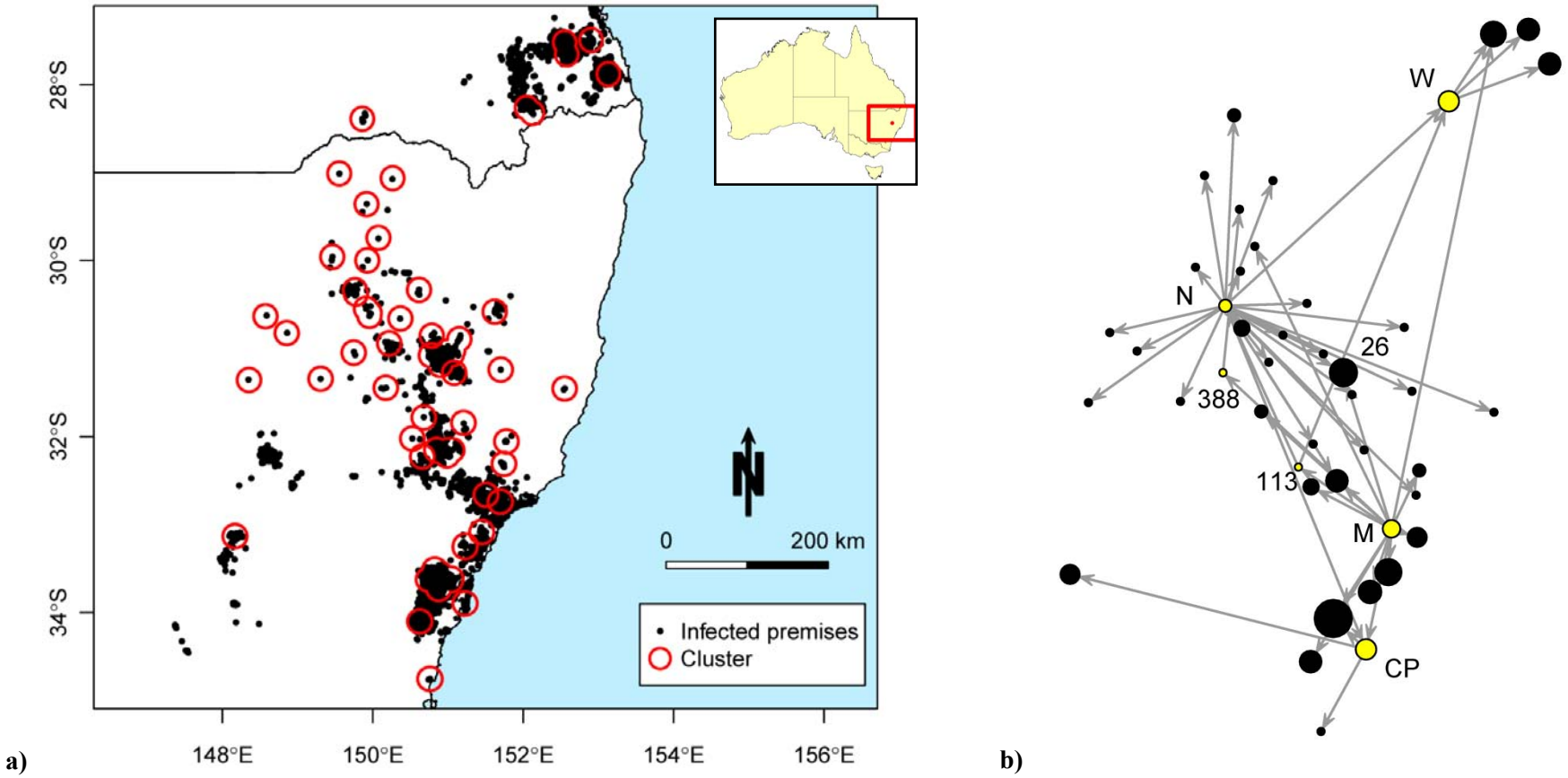


Figure 4.6 Clusters of infected premises described by the 'maximal' contact-and-proximity network (distance cut-off dichotomised at 15.3 km) in the first 10 days of the 2007 equine influenza outbreak in Australia.
(a) Cluster boundaries (red) coalesced from 15.3 km buffers around premises grouped according to the components of the maximal proximity network, and infected premises (closed black circles). (b) Block-modelled network of infected horse movements between clusters with node sizes relative to number of premises per cluster, labelled and coloured yellow if out-degree ≥1.

pertaining to multiple introductions into each coalesced cluster.

The maximal contact-and-proximity network was highly clustered, had short average path length, and possessed a highly skewed degree distribution, the first two being features of networks with 'small-world' properties (Watts and Strogatz, 1998). Each of these characteristics have important implications for infectious disease spread and the design of appropriate control strategies. Through applying bond percolation theory to epidemic modelling (Newman, 2003), it has been shown that the clustering in such networks lowers the threshold at which an infectious agent can infect all nodes, effectively increasing the reproductive ratio of the epidemic. However, this may lead to localised depletion of susceptible individuals within clusters, and fewer individuals overall being infected (Newman, 2003). In the outbreak studied, the combined contact-and-proximity network was highly clustered because of the spatial structure of the data. Yet, prior to the implementation of movement restrictions, the presence of sparse contact links connected distant spatial clusters leading to most nodes being reachable within very few geodesic steps (low average path length). This allowed the epidemic to easily 'percolate' through the population, rapidly infecting a large number of widely dispersed premises. Once movement restrictions were implemented, in effect cutting out the contact links between clusters, the epidemic could be expected to burn out over time in numerous spatially isolated clusters.

It is known from modelling networks with highly skewed degree distributions that disease surveillance and control strategies built upon random sampling will be inefficient compared to those targeted towards certain classes of premises, particularly those with high degree (Callaway et al., 2000). The spread of disease between clusters may be prevented by targeting the links ('bridges') that connect such 'hubs' in different clusters (Cohen et al., 2000). Prior knowledge of the underlying contact-structure of the Australian horse industry would therefore be very useful for appropriately targeting surveillance and interventions in future outbreaks. Several recent SNA studies have done

exactly this for other animal populations, investigating underlying networks with the intention of inferring how certain infectious diseases (such as FMD and highly pathogenic H5N1 avian influenza) might spread if introduced (Christley and French, 2003, Webb, 2005, Natale et al., 2009, Ribbens et al., 2009).

In most infectious disease situations, network and local spatial spread are expected to overlap to some extent. To make decisions on disease control, it is important to assess in broad terms the amount of network versus local spatial spread. If animal movements occur to premises that have already been infected (by local spatial spread), then our network analysis approach would likely over-estimate the proportion of premises infected through the network. However, likelihood-based approaches allow estimation of the probability that infection of individual premises occurred by network or other means (such as fomites or windborne spread). By comparing our network analysis findings to those of a purely spatial likelihood-based approach we were able to identify 14 (95% CI: 9, 20 premises) of the 70 premises in the contact network that were more likely to have been infected by local spatial spread. Furthermore, combining the two approaches provided an explicit description of the contact network and those premises that were most at risk of network spread (in this case 70 of 197 of the premises infected in the first 10 days). It is therefore important to combine this network analysis approach with a likelihood-based approach to avoid over-attributing the contribution of network-associated spread. Such information would be valuable to the decision-maker attempting to control disease spread, since methods of spread can dictate the control strategies implemented.

Spread to around a quarter (28.3%) of those premises infected within the first 10 days of the outbreak could be explained by the contact network structure alone. However, our findings also suggest that over 70% of early spread occurred through local spatial spread. Such local spread may have occurred through direct contact or droplet transmission between horses on adjacent premises, and over short distances through windborne aerosol transmission and spread on fomites (such as riding equipment,

peoples clothes and footwear, or on vehicles). This current analysis was limited in that it was not possible to differentiate the various components of local spread by transmission mechanism. This is an active area of further research utilising stochastic simulation, parameterised in part by the results of this analysis, to further our understanding of the mechanisms of disease flow between premises. Nonetheless, much is known from outbreak investigations, anecdotal accounts, and experimental and observational epidemiological studies. An infected coughing horse can spread the virus for up to 32 m (Miller, 1965), short range transmission on fomites has previously been reported (Guthrie et al., 1999), as has windborne spread of aerosolised virus over distances of 3.2 and 8 km (Huntington, 1990, Dalglish, 1992). Furthermore, Davis et al. (2009) noted that in a single cluster of the 2007 outbreak in Australia, directional spread (within 1–2 km) was consistent with the predominant wind direction over the same time period.

Although equine influenza can survive for around 2 days in soil (Yadav et al., 1993) and on non-porous surfaces (Bean et al., 1982), local spread on fomites is likely to only occur over limited distances because of how rapidly equine influenza is inactivated by sunlight, heat, cold, drying and common disinfectants. Indeed, during this outbreak, on-farm biosecurity measures including disinfecting footwear were found to be associated with a large reduction in odds of infection (Firestone et al., 2011a). Furthermore, only a few rare instances of spread over distances of 20–80 km were noted in which no infected horse movement or other plausible transmission pathway was identified despite thorough investigation (Equine Influenza Epidemiology Support Group, 2008).

We acknowledge that our methods are sensitive to the completeness of the contact-tracing data. Despite considerable resources being devoted to contact-tracing early in this outbreak it is likely that some infected horse movements were missing from the dataset. Clusters whose contact links were missing would be connected to the GWC of the maximal contact-and-proximity network through spatial proximity to infected premises in the next nearest cluster, rather than through a contact link.

Any such missing data might have obscured an important contact network link thereby slightly inflating our estimates of local spread distances.

The date of onset of clinical signs in the first horse affected on a premises was important in the definition of both infected premises and infected horse movements used in this analysis. The study was restricted in time, such that premises were only considered infected if their onset date was within the first 2 weeks of the outbreak. Also, links representing infected horse movements were only added if the date of movement and the onset date on the originating and destination premises were in a logical sequence with respect to the incubation and infectious periods of equine influenza. These criteria were set quite conservatively, including only a 1 day margin of error. These strict criteria explain why some infected premises traced to the events at Maitland and Narrabri, and noted elsewhere to have inaccurate onset dates (Equine Influenza Epidemiology Support Group, 2008), were excluded from this analysis. If these selection criteria were relaxed to further account for such reporting errors this would have increased the likelihood of misclassifying the mechanism of spread to certain infected premises and clusters. We also purposefully excluded from this analysis several records of potential transmission via fomites because routine reporting of such movements between infected premises was not conducted during the outbreak. The inclusion of these records in the contact-network structure might have slightly reduced our estimates of the distance of local spread, but would have introduced a lot of uncertainty.

4.5. Conclusions

This social network analysis provides a complete description of early spread during the 2007 outbreak of equine influenza in Australia, capturing both the spatial relationships and contact patterns between infected premises. Prior to disease detection, around a quarter of early epidemic spread occurred through the movement of infected horses, and these horse movements defined much of the spatial extent of the outbreak. By describing spatial clusters with consideration of contact-tracing network topology, and estimating the shape of the early epidemic's transmission kernel, we found that most local spread occurred within 5 km of an infected premises, with little local spread occurring beyond 15 km. This analysis could be extended through further characterisation of local spread within clusters, and together with research into underlying contact structures of the horse industry be applied to improve targeting of outbreak surveillance and control activities in future similar outbreaks.

Chapter 5: Cluster description using spatial and social network analysis

"*Yo-yos, hula hoops and pogo sticks, …, sweep through schools, and more sporadically leap from school to school, in patterns that differ from a measles epidemic in no serious particular.*

Ten years ago, you could have travelled thousands of miles through the United States and never seen a baseball cap turned back to front. Today, the reverse baseball cap is ubiquitous.

I do not know what the pattern of geographical spread of the reverse baseball cap precisely was, but epidemiology is certainly among the professions primarily qualified to study it."

Richard Dawkins (1991), 'Viruses of the Mind'.

This chapter appears as the following published paper in Preventive Veterinary Medicine:

Firestone, S.M., Ward, M.P., Christley, R.M., Dhand, N.K., 2011. The importance of location in contact networks: describing disease spread using spatial social network analysis. Prev. Vet. Med. 102, 185-195.

5. Cluster description using spatial and social network analysis

5.1. Introduction

Epidemics of emerging infectious diseases are often first observed as several clusters of disease at widespread locations, with initial spread seemingly occurring through a contact or other network (such as a market chain or major transportation route) (Thrusfield et al., 2005b). The spatio-temporal analysis of such epidemics is based on the concept that transmission is more likely to occur between at-risk individuals that are in close proximity, be that in space or time (Pfeiffer et al., 2008). However, empirical research into the spread of H5N1 avian influenza (Small et al., 2007) has shown that when a complex contact network structure underlies an epidemic, traditional approaches may be insufficient to appropriately describe the spatio-temporal pattern of the epidemic spread. We demonstrate the application of social network analysis to integrate the spatial and network dimensions of an epidemic, facilitating differentiation of spatial spread into its network and non-network (local spread) components, thereby providing a means of describing the sequence of cluster formation.

Social network analysis (SNA) has been used to describe the network component of disease transmission during epidemics in animals (Corner et al., 2003, Shirley and Rushton, 2005, Ortiz-Pelaez et al., 2006). The underlying contact networks of animal populations have also been investigated with SNA (Christley and French, 2003, Natale et al., 2009, Ribbens et al., 2009), incorporating spatial (Webb, 2005, Lockhart et al., 2010) and spatio-temporal analyses (Martinez-Lopez et al., 2009a). Depending on the context, a network of nodes can be constructed to represent individual animals or premises, and connections ('links') between nodes may be used to represent a wide range of relationships such as physical contact, movement of animals, or proximity in space. The network can itself be described in terms of its size (number of nodes), cohesion (proportion of clustering or fragmentation) and centrality (the extent to which the network revolves around any single node). Furthermore, the centrality of individual nodes, and thereby their potential importance to

disease flow within the network, can be characterised by calculating node-level attributes such as betweenness and degree. The betweenness of a node is the frequency with which it lies along the shortest path (the 'geodesic') between other nodes in the network, whereas a node's degree is the number of links that are incident upon it (Freeman, 1979).

In this study, we apply SNA methods to describe the formation of clusters during the 2007 equine influenza outbreak in Australia. This was the first reported outbreak in Australia, and at the time only small numbers of imported horses had been recently vaccinated. Equine influenza (A/H3N8) was introduced via infected vaccinated horses from Japan that were quarantined in a facility 40 km west of Sydney (Callinan, 2008). From this facility, the disease spread undetected to local horses and then rapidly through the Australian horse population infecting approximately 67,000 horses on 9359 infected premises (IPs). How the disease escaped from the quarantine facility was not identified, despite thorough investigation. Clinical signs were first observed in the quarantined horses on 17 August 2007, then five days later in two horses kept in a large horse facility in central Sydney (Centennial Parklands Equestrian Centre). Contact-tracing revealed that these horses had attended an equestrian event the previous weekend at Maitland, 160 km north of Sydney. Most of the geographic spread of this epidemic occurred undetected during the first ten days, through movement of infected horses prior to the implementation of movement restrictions (Cowled et al., 2009b). Subsequent spatial spread was believed to be driven by local spread mechanisms (direct contact, fomite and windborne transmission). After four months, equine influenza was eradicated from Australia through the combination of movement restrictions and vaccination (DAFF, 2008).

In this study we apply social network analysis to data from the 2007 outbreak of equine influenza in Australia, to investigate the importance of network and spatial location in disease spread. We also compare clusters of IPs described by spatial-SNA with those detected using the space-time scan statistic, leading to an improved overall description of the dynamics of the outbreak in space and time.

5.2. Materials and methods

5.2.1. The equine influenza dataset

The state governments of New South Wales (NSW) and Queensland (QLD) provided contact-tracing and attribute data on all horse premises investigated during the 2007 outbreak. Premises were defined as infected if they held horses that had been observed with the classical clinical signs of equine influenza (cough, elevated temperature, nasal discharge and lethargy) and if the diagnosis had been confirmed by laboratory testing based on real-time reverse transcription polymerase chain reaction assay (Foord et al., 2009). Contact-tracing records included the date of the movement, and the addresses and unique identifiers of the origin and destination premises. Premises attribute records included address, geocoded coordinates (based on premises centroid), number of horses, and date of onset of clinical signs in the first horse affected ('onset date'), and were linked to the database of laboratory testing records collected at the time of the outbreak.

5.2.2. Exploratory analysis of the outbreak in time and space

The dataset was imported into the R statistical package version 2.12.0 (R Development Core Team, 2011), and an epidemic curve constructed as the count of infected premises reported per day. The spatial coordinates of each premises were converted to the Albers conic equal-area projection based on the Geocentric Datum of Australia 1994 (Geoscience Australia, 2007), using the 'sp' library in R (Pebesma and Bivand, 2005). An estimate of the risk surface was plotted (ArcMap™ 9.3, ESRI Inc., Redland, CA, USA) as the ratio of the intensity of kernel smoothed estimates of infected and non-infected premises. Intensity surfaces were estimated with a Gaussian kernel function using a bandwidth of 20 km. This bandwidth was selected intentionally to be larger than the distance over which direct transmission is considered to occur (Diggle et al., 2007). It was based on our earlier analyses in which we estimated that over 99% of local spread in the first two weeks of the outbreak (through both direct and indirect means) had occurred within 15 km (Firestone et al., 2012a).

Kriging was used to investigate the accuracy of crude spatial modelling (irrespective of known contact network relationships and other variables) in the early outbreak period (30 days) compared to the whole period of the outbreak (130 days). For both periods, iterative reweighted least squares estimation was used to fit anisotropic semivariogram models to empirical semivariograms of differences in onset dates between neighbouring IPs, with weighting determined by the number of point pairs at a series of equal distance intervals (lags), implemented with the R library 'geoR' (Ribeiro and Diggle, 2001). Several lag distances and a number of statistical models were fit to identify optimal semivariogram models for each time period, the parameters of which were then used to prepare the two kriged surfaces. This approach enabled visualisation of the spatio-temporal distribution of the epidemic. Due to the contagious nature of infectious disease spread, spatial aggregation (clustering) of premises with similar onset dates is expected (Ward et al., 2008). If spatial autocorrelation is operating independently of other effects, then semivariance should increase with distance, up to a certain distance ('the practical range') where the differences in onset dates between pairs of IPs are expected to level off to the mean semivariance ('the sill') (Hengl, 2009).

The precision of prediction based on the two kriged surfaces was tested using leave-one-out cross-validation with the 'gstat' library in R. The prediction error ('residual') was estimated at each individual observation point by reproducing a kriged surface without it (based on smoothing over a local neighbourhood with radius 100 km). The accuracy of prediction (R^2) was estimated based on these prediction errors (Hengl, 2009), thus describing the proportion of variability in the date of onset of IPs explained by the crude spatial model alone.

5.2.3. Restriction of the study period

All further spatio-temporal analyses were restricted to the first 30 days of the outbreak to facilitate visualisation of the network, and comparative analyses. Our study period started with the 'Maitland equestrian event' in NSW, the first known spreading event in the epidemic, and ended prior to the

commencement of vaccination and the relaxation of movement controls, encompassing most of the exponential growth phase of this epidemic.

Presently available software is capable of conducting analyses on extremely large networks (>100,000 nodes) but visualisation is poor and prohibitively slow if the network is well-connected (i.e. built on a non-sparse-matrix). The network of the entire outbreak (representing 9360 infected premises as nodes) was very highly connected once spatial relationships were incorporated. The distance matrix alone was >1.2 GB in size, rendering it unsuitable for several analyses. By restricting the study period to 30 days, a more manageable network was constructed with 1740 nodes allowing clear visualisation of links representing spatial and contact relationships between all IPs arising in the first month of the outbreak.

5.2.4. Cluster description using spatial social network analysis

A detailed description of the social network analysis of the first ten days of this outbreak is presented elsewhere (Firestone et al., 2012a), and therefore the methods will be outlined only in brief. Further detail is provided on the method of describing clusters and the application of block-modelling to illustrate the series of infected horse movements and to infer the process of how these clusters formed.

A 'contact network' was constructed to represent all recorded movements of infected horses prior to the implementation of movement standstills (on the tenth day of the outbreak) using the R libraries 'statnet' (Handcock et al., 2003) and 'spatstat' (Baddeley and Turner, 2005). Directed links were added to the contact network representing horse movements originating from infected premises until day 10 of the outbreak. As data on actual numbers of horses moved were unavailable, these links were binary in that movement of a single horse was treated equivalently to several horses moving between two premises. Most movements would have involved only one or two infectious horses being introduced to premises holding less than 10 horses. Epidemiological investigations have remarked on

the exceptional contagiousness of equine influenza in the predominantly naïve Australian horse population and that there was little difference whether one or more horses were introduced onto a premises (Equine Influenza Epidemiology Support Group, 2008).

All of the horse movements were back-traced, with numerous infected horses being moved whilst incubating the disease. Infected horses can shed virus as early as 24 hours after infection and one day before developing clinical signs (Crouch et al., 2004, Soboll et al., 2003). Intra-herd modelling (Garner et al., 2011b) has shown that with several cycles of infection, horses on large premises may remain infectious for up to 21 days. Consequently, movements were only included if they occurred within 21 days of the day before clinical signs were first observed on the originating premises, and if clinical signs were observed at the destination premises within three days of movement. A 1 day margin of error was included in these calculations to allow for inconsistent observation and reporting practices.

A 'proximity network' was constructed based on a distance matrix between all premises holding horses that – based on their date of onset of clinical signs – were likely to have been infected during the first 30 days (17 August–15 September 2007) of this epidemic. The proximity network was transformed into a binary network, with links only between pairs of IPs separated by a Euclidean distance less than a specified distance cut-off. This network was comprised of fragments ('components') representing groups of IPs related in proximity ('spatial clusters'). This network of spatial clusters was then combined with the contact network into a 'contact-and-proximity' network, to demonstrate how spread through the contact network led to the formation of spatially-defined clusters. The contact-and-proximity network was exported to NetDraw version 2.091 (Borgatti, 2002) for graphing.

The distance cut-off can be considered an assumption about the distance over which local spread of

equine influenza was occurring. By varying this cut-off we explored the effect of distance on disease transmission. As we increased the distance cut-off, fragmentation in the contact-and-proximity network reduced, up to the point where nearly all nodes are included in the largest connected component of this network. We considered the distance cut-off at this point (i.e. 15 km) to be an estimate of the maximum distance over which local spread had occurred in the first 30 days of the outbreak.

To describe spatial clusters of IPs whilst also considering network topology, we block-modelled the contact-and-proximity network with distance cut-off equal to the estimated maximal distance of local spread. Block-modelling is a network analysis technique that involves grouping nodes into 'blocks' (also colloquially termed 'clusters') based on some similarity (Wasserman and Faust, 1994). Networks are then reconstructed representing the relationships between these groups. In this analysis, nodes were grouped according to the spatial clusters described by the contact-and-proximity network, then, block-modelled with links representing infected horse movements between groups. Such analysis emphasises movements that introduce infection into new spatial clusters.

IPs whose date of onset were poorly described by their spatial location (with respect to other infected premises in their spatial vicinity) were identified – based on the cross-validation of kriging over the 30 day period – as those with residuals of magnitude greater than three standard deviations from the mean (Hengl, 2009). Their location in space, and importance ('centrality') in the contact-and-proximity network, were further investigated. Attributes of these outliers, including network degree, network (geodesic) distance and Euclidean distance from the index node ('Maitland equestrian event'), betweenness, number of horses and size of premises were compared to estimates for the remainder of the study population.

5.2.5. Cluster detection with the space-time scan statistic

A retrospective analysis for clusters of IPs aggregated in space and time was conducted using the space-time permutation model scan statistic (Kulldorff et al., 2005) in SaTScan version 9.0.1 (www.satscan.org). A maximum scanning window of 50% of the study period (15 days) and a 100 km radius (≈ 10% of the study area) was used to test for observed numbers of IPs within circular spatio-temporal windows that exceeded numbers expected if the IPs were distributed in time and space according to a random process. The spatio-temporal model was run for 999 Monte Carlo simulations to estimate the statistical significance of each observed cluster. Clusters with a *P*-value<0.20 were compared with those described using SNA.

5.3. Results

5.3.1. Exploratory analyses in time and space

The epidemic of equine influenza in Australia lasted for 18 weeks, peaking in the seventh week with over 250 infected premises being reported per day (Figure 5.1). The epidemic curve is characteristic of a propagating epidemic, with an exponential growth phase lasting around 30 days. Over the course of the outbreak, around 70,000 horses were infected on 9360 premises, over a geographic extent of nearly 300,000 square kilometres. Dispersal over most of this spatial extent occurred in the first two weeks (Figure 5.2a). A kernel-density smoothed risk surface (Figure 5.2b) shows that infected premises were tightly aggregated spatially, with several clusters coalescing around the capital cities of Brisbane and Sydney and in the major Thoroughbred horse breeding region of the Hunter Valley.

The semivariogram model that best fit the empirical semivariogram for the first 30 days of the outbreak was a spherical model with lags of 10 km (Figure 5.3a). The mean semivariance between pairs of IPs within 10 km of each other was relatively high; the practical range of spatial autocorrelation was estimated to be 96 km. Leave-one-out cross-validation showed that the corresponding kriged surface (Figure 5.3b) was relatively imprecise (R^2=0.13), with a

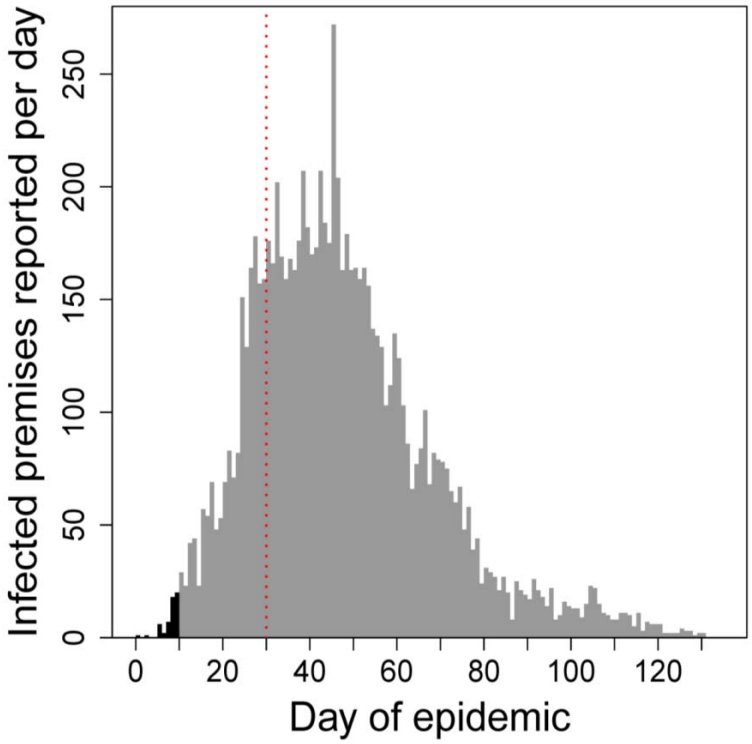


Figure 5.1 Epidemic curve of the 2007 equine influenza epidemic in Australia.
The 10 day period before complete implementation of the movement standstill is shown in black. The exponential phase of the epidemic, used as the study period for spatiotemporal analysis, ends at day 30 (reference line).

negatively skewed distribution of standardised residuals (Figure 5.3c). When kriging with cross-validation was repeated on the entire outbreak dataset, the accuracy of the predicted surface was greatly improved (R^2=0.60), and the distribution of standardised residuals appeared symmetrical.

Nine premises were identified with high magnitude standardised residuals based on the kriged surface of the first 30 days of the outbreak (depicted in yellow in Figure 5.3b). These premises had dates of onset 17–22 days earlier than those predicted by the kriged surface. These outliers were evenly distributed in space amongst the other IPs reported in the first 30 days of the epidemic, and no closer in Euclidean distance to the index case.

5.3.2. Cluster description using spatial social network analysis

The network of recorded movements of infected horses occurring prior to the implementation of

movement standstills (the 'contact network') comprised 70 infected premises (nodes) connected by 70 links (Figure 5.4a). Each link represents the movement of an infected horse(s). This network was dominated by the two equestrian events (at Maitland and Narrabri) where transmission is known to have occurred. Figure 5.4b is a cumulative distribution of premises infected in the first 30 days of the outbreak (n=1740), by distance to the nearest contact-traced infected premises, produced by varying the distance cut-off used to construct the proximity network. Around 89% of the premises infected in the first month of the outbreak were within 5 km of a premises included in the contact network; 98% were within 15 km.

The result of combining the proximity network constructed using a 15 km distance cut-off with the contact network is a graph of the contact and spatial relationship of all premises infected in the first 30 days of the epidemic (Figure 5.5). The contact network forms the core of this combined 'contact-and-proximity network', with infected horse movements connecting spatial clusters of infected premises.

The block-model of the contact-and-proximity network (with distance cut-off set at 15 km) is presented in Figure 5.6. This network comprised 56 spatial clusters of infected premises, the largest comprising 896 infected premises (Table 5.1) and their connection through infected horse movements. The block-model enables identification of the contacts that introduced infection into new spatial clusters, and thereby the progression of cluster formation over time.

The 9 premises identified as outliers based on their kriging residuals were relatively central within the contact-and-proximity network (Figure 5.5), all being within three geodesic steps of the Maitland event and having a higher betweenness relative to the rest of the nodes (being on the shortest path between 350 pairs of nodes compared to mean betweenness of 145 geodesics, P=0.06 from the Wilcoxon rank sum test with continuity correction). These outliers were similar to the rest of the IPs in all other attributes assessed, including: number of horses, size of premises, and network degree.

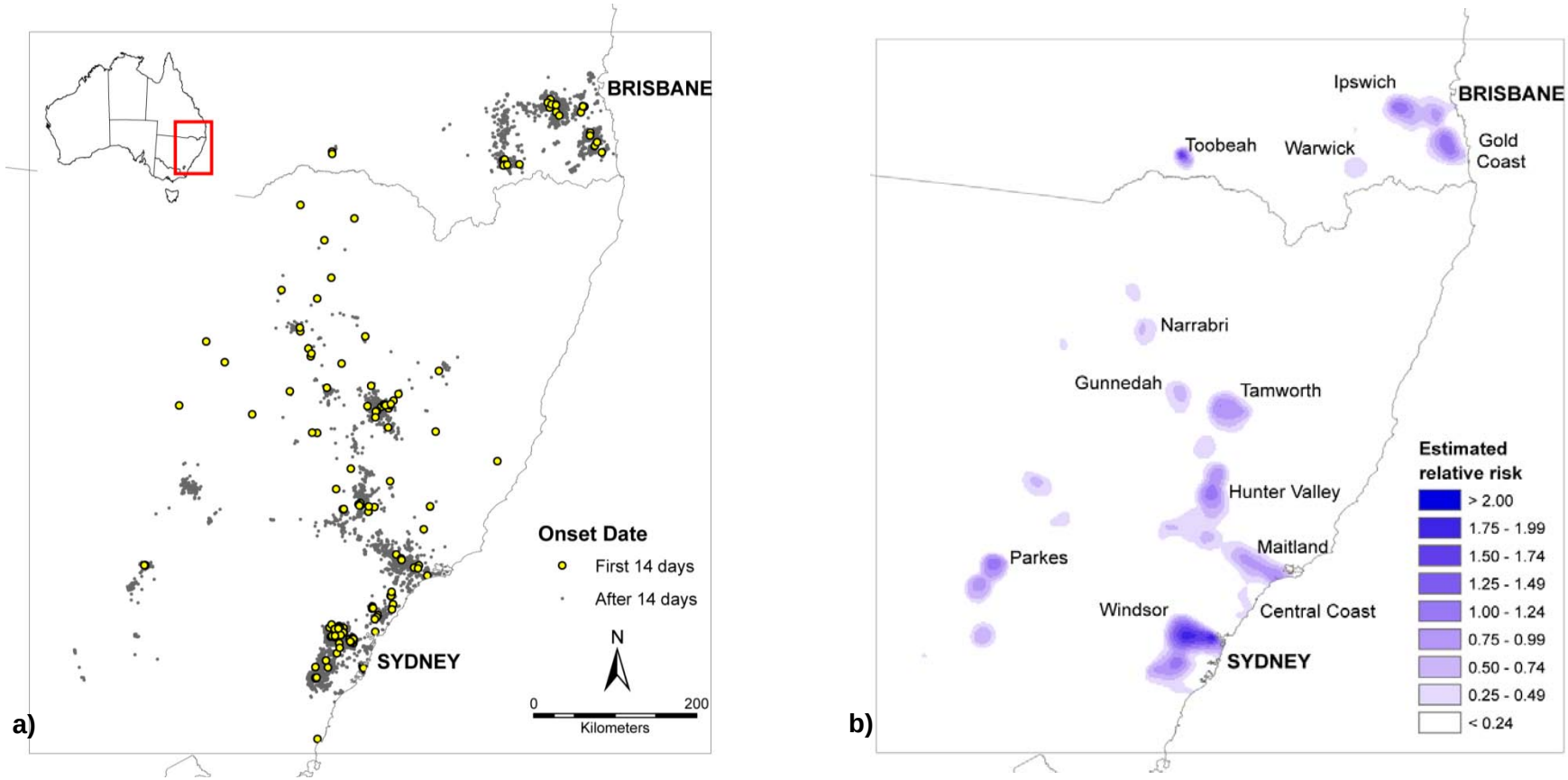


Figure 5.2 Exploratory spatial analyses of the 2007 equine influenza outbreak in Australia.
(a) Geographic extent of the outbreak, with infected premises mapped by reported date of onset of clinical signs in the first horse affected. (b) Estimated risk surface calculated as the ratio of the intensity of Gaussian-kernel smoothed estimates of infected and non-infected premises using a bandwidth of 20 km.

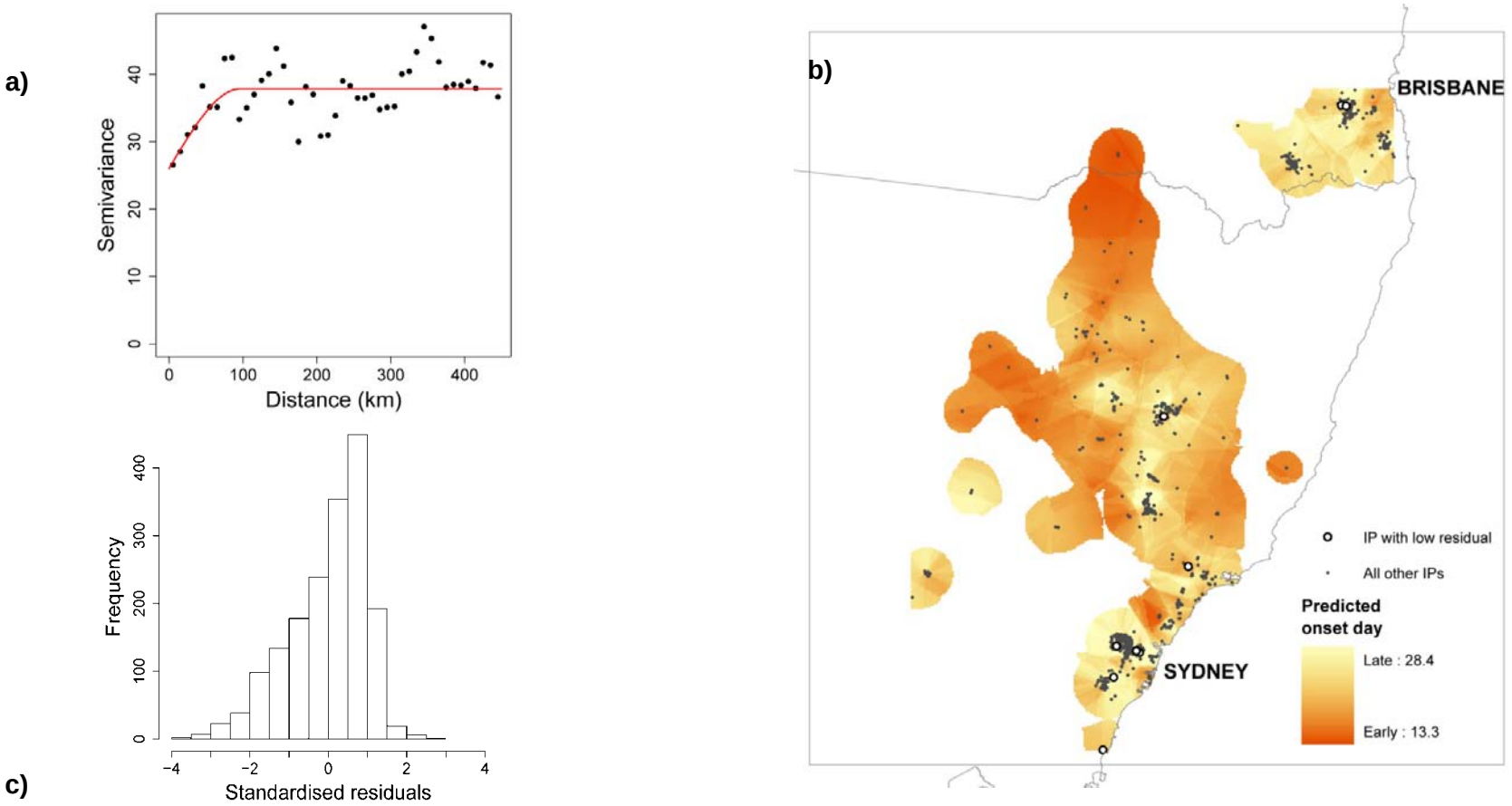


Figure 5.3 Variography and kriging and cross-validation of estimated date of onset data amongst horse premises reported as infected in the first 30 days of the 2007 equine influenza outbreak in Australia.
(a) Empirical semivariogram and fitted spherical model (lag = 10 km) of spatial autocorrelation in estimated onset dates amongst horse premises reported as infected in the first 30 days. (b) Kriged surface of predicted onset dates, highlighting infected premises (IPs) identified as spatiotemporal outliers (residuals of magnitude greater than three standard deviations from the mean). (c) Histogram of standardised residuals estimated using leave-one-out cross-validation of the kriged surface.

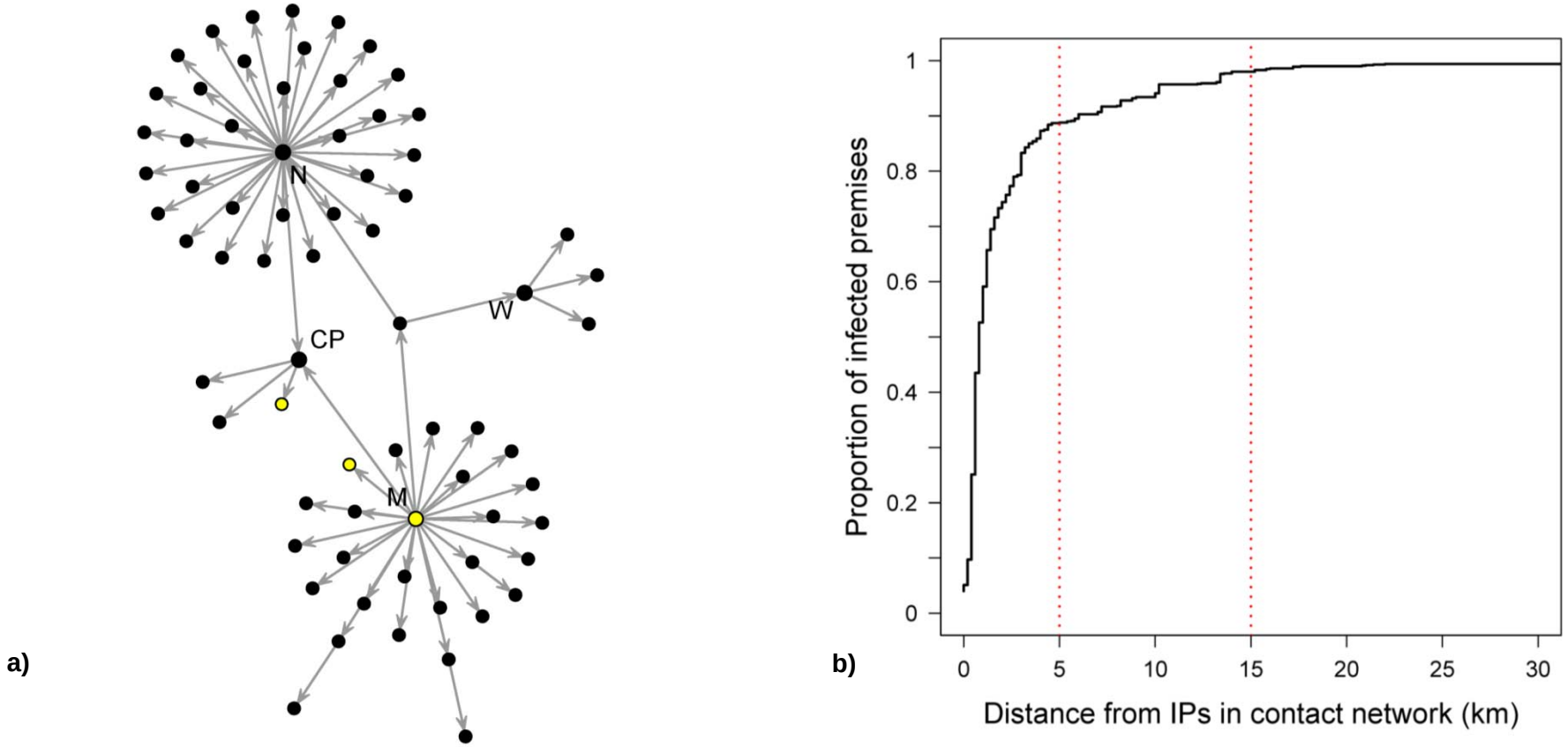


Figure 5.4 The network of recorded movements of infected horses occurring prior to the complete implementation of movement standstills (the 'contact network') on the tenth day of the 2007 equine influenza outbreak in Australia.
(a) Graphed in arbitrary space; nodes in yellow were identified as spatiotemporal outliers through kriging, having onset dates relatively early for their spatial region. CP = Centennial Parklands equestrian centre in central Sydney, M = Maitland equestrian event (the index premises), N = Narrabri equestrian event, W = Warwick equestrian event. (b) Cumulative distribution of infected premises (IPs) with onset in the first 30 days (n=1740), by Euclidean distance from the 70 premises included in contact network.

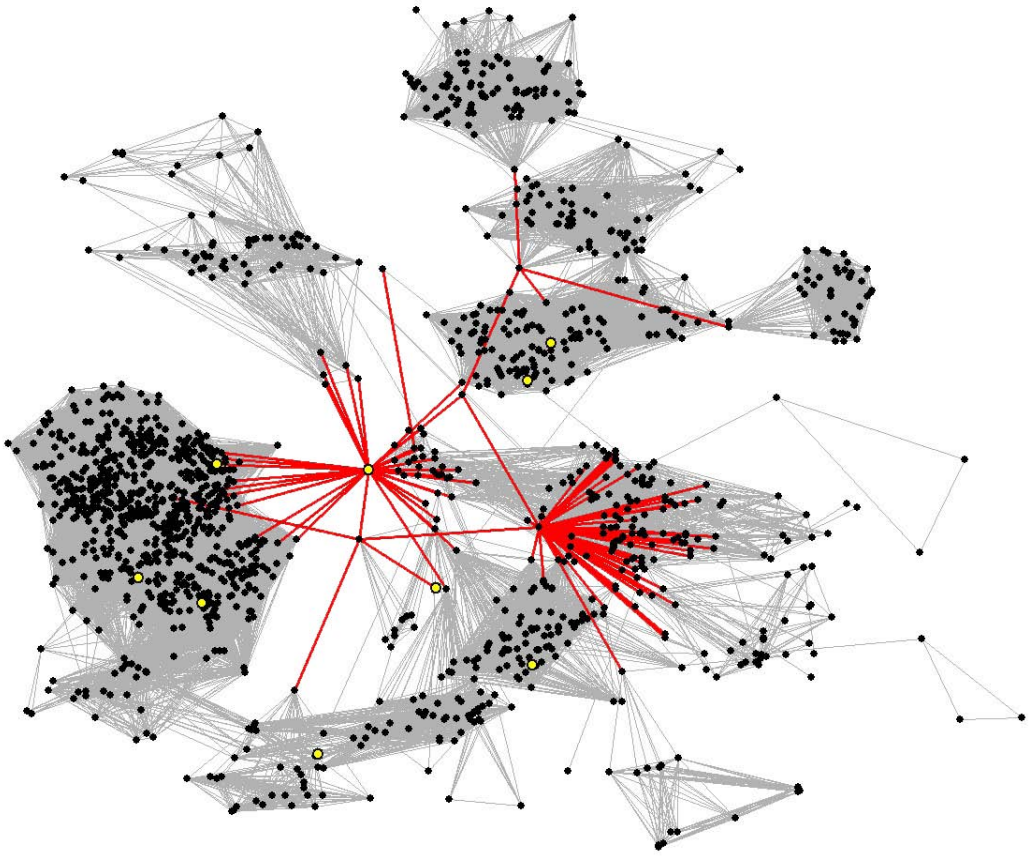


Figure 5.5 Contact-and-proximity network representing spatial and contact relationship between all premises holding horses infected in the first 30 days of the 2007 equine influenza outbreak in Australia.
Spatial relationship represented by grey links, contact relationships (infected horse movements) represented by red links Constructed based on a distance cut-off of 15 km, and graphed in arbitrary space. Nodes in yellow were identified as spatiotemporal outliers through kriging, having onset dates relatively early for their spatial region.

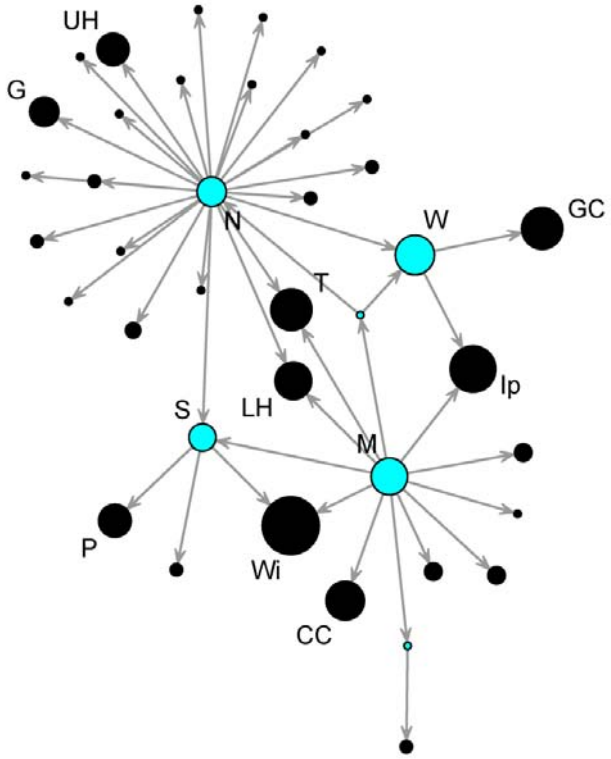


Figure 5.6 Block-model representing infected horse movements between spatial clusters identified as components of the proximity network with distance cut-off 15 km, in the first 10 days of the equine influenza outbreak in Australia.
Graphed in arbitrary space, with nodes sized by the number of infected premises in each cluster. Nodes in blue have out-degree ≥1. CC = Central Coast, G = Gunnedah, GC = Gold Coast (QLD), Ip = Ipswich (QLD), LH = Lower Hunter Valley, M = Maitland, N = Narrabri, P = Parkes, S = Central Sydney, T = Tamworth, UH = Upper Hunter Valley, W = Warwick (QLD), Wi = Windsor.

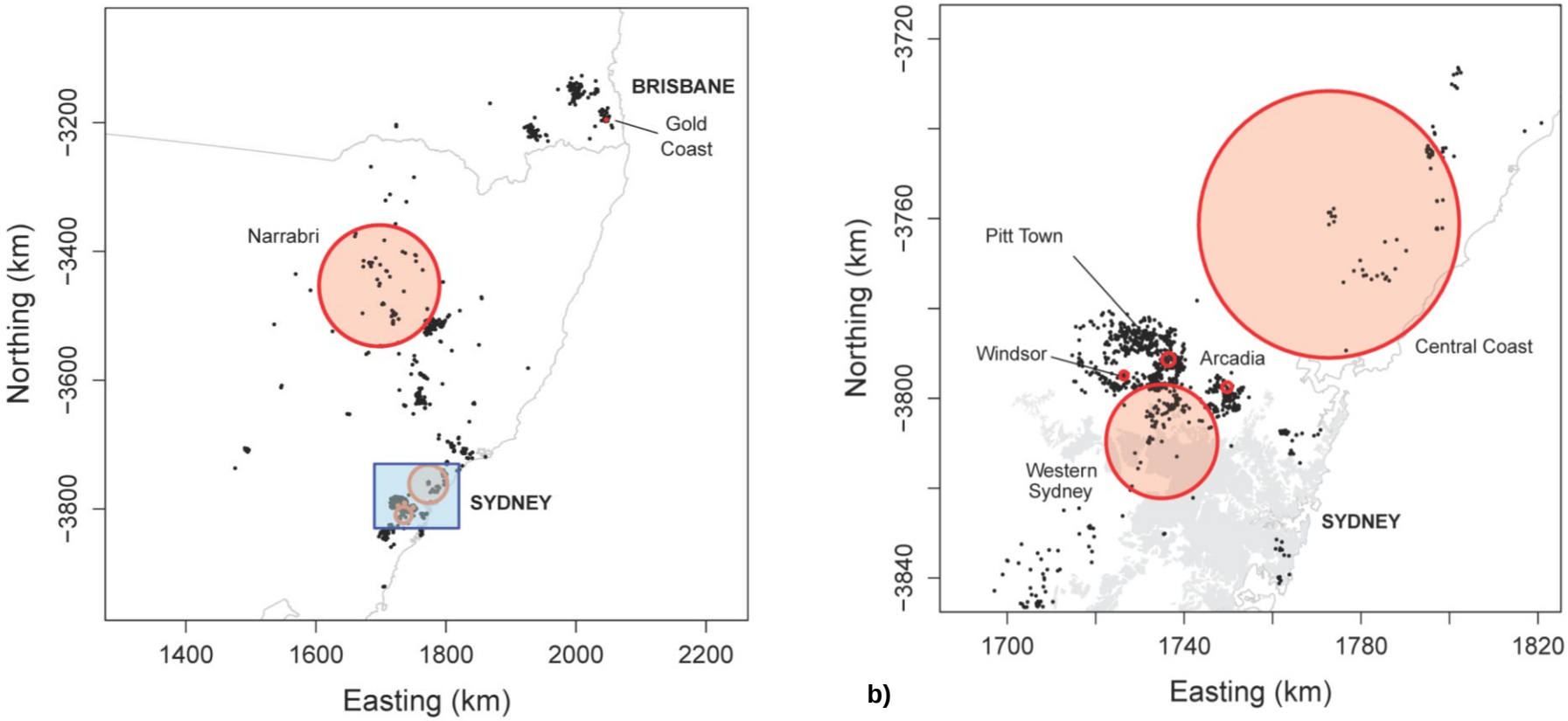


Figure 5.7 Spatiotemporal clusters of infected premises from the first 30 days of the 2007 equine influenza epidemic in Australia, detected using the space-time scan statistic (P<0.20).
(a) Clusters detected across the whole outbreak extent; red circles represent spatial clusters, black points represent infected premises. Blue shaded rectangle denotes area of enlargement in (b) clusters detected in the Greater Sydney region (metropolitan area shaded grey).

Table 5.1 Largest clusters (containing more than 10 infected premises) identified as components of the proximity network with distance cut-off 15 km from the first 30 days of the 2007 equine influenza epidemic in Australia.

Cluster location	Radius (km)	No. of infected premises	Earliest onset date[a]	Earliest infected horse movement source	Earliest infected horse movement date[a]
Windsor	64.2	896	3	Maitland	3
Ipswich (QLD)	28.0	192	6	Maitland	3
Tamworth	27.8	103	6	Maitland	3
Gold Coast (QLD)	20.4	102	9	Warwick (QLD)	9
Warwick (QLD)	26.1	78	9	Maitland[b]	8
Central Coast	39.4	70	6	Maitland	3
Lower Hunter Valley	16.9	56	10	Maitland	3
Maitland	38.9	54	1	Unknown	Unknown
Parkes	4.1	32	8	Central Sydney	8
Upper Hunter Valley	28.7	30	13	Narrabri	10
Gunnedah	10.7	20	12	Narrabri	10
Narrabri	32.3	20	9	Maitland[b]	8
Central Sydney	5.2	15	9	Maitland	3

QLD = Queensland. Central Sydney cluster includes Centennial Parklands Equestrian Centre and subsidiary centres.
[a] Dates presented as epidemic day (i.e. 17 August 2007 is Day 1). [b] Indirect movement from Maitland equestrian event via infected premises in Lower Hunter Valley.

Table 5.2 Clusters of infected premises ($P<0.20$) detected in the first 30 days of the 2007 equine influenza epidemic in Australia, using the space-time permutation model scan statistic with a scanning window of 100 km and 15 days.

Cluster location	Radius (km)	Days of epidemic	Observed number of infected premises	Expected number of infected premises	Maximum likelihood ratio test statistic	*P*-value[b]
Central Coast	29.6	6–12	17	3.2	14.7	**0.001**
Arcadia[a]	1.1	11–15	12	1.8	12.8	**0.010**
Pitt Town[a]	1.5	15–19	17	3.9	12.0	**0.022**
Narrabri	93.6	13–17	26	8.3	12.0	**0.022**
Western Sydney[a]	12.7	28–30	64	32.9	11.7	**0.025**
Gold Coast (QLD)	0.8	18	5	0.2	10.5	0.102
Windsor[a]	0.9	17–18	8	0.9	10.2	0.142

QLD = Queensland.
[a] These four clusters are located in the Greater Windsor region northwest of Sydney. [b] Estimated based on 999 Monte Carlo simulations.

5.3.3. Cluster detection with the space-time scan statistic

The space-time permutation model identified 36 clusters of IPs in the first 30 days of the outbreak. Only five clusters were statistically significant ($P<0.05$); two further clusters had P-values<0.20 (Table 5.2). The primary cluster included 17 IPs located in the Central Coast region (Figure 5.7). Four secondary clusters were detected immediately northwest of Sydney; all were within the Greater Windsor region demarcated on the kernel density map (Figure 5.2b). The largest of these clusters included 64 IPs with onset dates over a three day period. The other three clusters in Greater Windsor covered very small areas with radii <1.5 km, and included a total of 37 IPs reported over a 9 day period.

5.4. Discussion

In the first 30 days of the 2007 outbreak, equine influenza spread rapidly over a very large spatial extent. Kriging was relatively imprecise in describing the pattern of spread during the early phase of this outbreak (explaining only 13% of the variation in date of onset of IPs), because early dissemination was dominated by network-based spread. Movement standstills effectively cut the contact network after only 10 days, when the entire outbreak lasted for 18 weeks. To comprehensively describe spatial spread in the early period of this outbreak, and given that most of the long-distance spread and the seeding of spatial clusters occurred during the first 10 days, a method of spatio-temporal analysis capable of identifying and incorporating the influence of the network of animal movements was required.

Combined analysis of spatial and contact network data demonstrated that local spread emanated outwards from the small number of infected premises in the contact network, up to a distance of around 15 km. Through clarifying the role of network-based spread in the observed spatial pattern of disease it was possible to identify important premises in the early dispersal of this outbreak, and their role in the formation of spatial clusters. Our approach of linking spatial and network analysis facilitated the identification of the specific contacts and premises involved in

introducing disease into new spatial clusters. This enabled estimation of the distribution of the cumulative proportion of local spread occurring over a range of distances. The estimates thus produced were consistent with those from a previous study conducted on a single cluster of this outbreak (Davis et al., 2009). In that case 91% of newly infected premises were found to be within 2 km of an earlier identified IP, and the largest observed local spread distance was 13 km. By studying a single cluster, network-based transmission was restricted out of their analysis. In the current study, our methods enabled us to produce comparable estimates over the entire spatial extent of the outbreak, adjusting for network-based spread.

The combined network of contacts and spatial proximity was constructed based on the estimated maximum distance of local spread. This enabled identification of spatial clusters, each representing the upper bound of premises that may have been infected through local spread following a known introduction into that small area. Within these clusters, local spread may have occurred through a variety of mechanisms, including direct contact or droplet transmission between horses on adjacent premises, windborne transmission as previously reported over distances up to 8 km (Huntington, 1990), or potentially via fomites up to 15 km – the estimated maximum distance of local spread.

The progression of cluster formation can be appreciated when the combined contact-and-proximity network is graphed as a block-model (Figure 5.5). Contacts that introduced infection into new spatial clusters are clearly discernible. Eleven such movements originated from the Maitland equestrian event, initiating infection in a number of large secondary clusters. Indirect movements led to transmission at the equestrian events in Narrabri and Warwick, and horse movements out of these events initiated the formation of the remainder of the observed clusters.

Further investigation of the 9 outliers identified through cross-validation of the kriged surface revealed that they were not discernibly different in spatial location from the rest of the IPs reported in the first 30 days of the outbreak. However, when distance through the contact

network (geodesic distance) was weighted equivalently to the maximum distance of local spread to construct the combined contact-and-proximity network, these premises were found to be closer to the index premises in terms of network location. That is, they were either directly connected to Maitland via a known horse movement, or they were spatially close to one of the 70 premises in the contact network, partially explaining why their onsets were much earlier than predicted using spatial techniques in isolation.

In comparison to the combined spatial-SNA approach, the space-time scan statistic detected relatively few significant clusters, and these were limited to three regions: Central Coast, Windsor and Narrabri. Large clusters were described in each of these regions with spatial-SNA, and they featured prominently in the block-model. However, the scan statistic-derived clusters contained considerably fewer infected premises. This may be related to the temporal aspect of this method of cluster detection: the longest period a significant cluster lasted was 7 days. The relative advantage of using the scan statistic is that it incorporates a hypothesis test, facilitating cluster detection rather than simply cluster description. Conversely, the power of spatio-temporal cluster detection techniques is poorly characterised and may be insufficient to detect small clusters (Ward and Carpenter, 2000a). Small clusters of IPs that are not tightly aggregated in one dimension (time or space) may not attain statistical significance, but may reside in key locations in a contact network and be important in the formation of other secondary clusters. Statistical power might be increased by applying a Bernoulli or Poisson scan statistic model, however the relatively large non-infected population within the study extent (>65,000 premises) precluded use of these models due to computational limitations. To increase power, Ward and Carpenter (2000) recommend using several techniques together when attempting to detect and describe clusters in time and space. By considering the results of spatial SNA together with those of the scan statistic a detailed description of the formation of clusters over time was developed, informed by detail on which clusters were statistically significant.

Whilst Martinez-Lopez et al. (2009a) present an approach for describing clusters of animal

movements incorporating the spatio-temporal scan statistic and SNA, it was not possible to apply their methods to the contact-tracing data from this outbreak. This is due to the purposive nature of sampling which was targeted at infected horse movements, applied over a large geographic extent with finite resources in the midst of an epidemic. Such a dataset is not intended to be representative of all movements (such as horses, vehicles, people and equipment) onto and off all premises (infected or otherwise). Our descriptive approach is suited to such contact-tracing data in that it does not require full knowledge of the population at risk or of all movements within that population. An approach such as that of Martinez-Lopez et al. (2009a) demands a representative, preferably complete, sample of movements. This would be nearly impossible to collect in a wide-spread epidemic, but is well suited to the study of contact patterns within smaller regions with the intention of informing surveillance and disease control initiatives.

The date of onset of clinical signs in the first horse affected on a premises was important in the definition of both premises and infected horse movements in this analysis. Premises were considered for inclusion in the 30 day study period based on their onset date. Links representing infected horse movements were only added if the date of movement and the onset date on the originating and destination premises were in a logical sequence with respect to the incubation and infectious periods of equine influenza. These criteria were set quite conservatively, including only a one day margin of error. Therefore, premises with inaccurately reported onset dates, and horse movements from or to such premises, would have been excluded from this analysis. These strict criteria explain why some infected premises traced to the events at Maitland and Narrabri, noted elsewhere to have inaccurate onset dates (Equine Influenza Epidemiology Support Group, 2008), were excluded. If these selection criteria were relaxed to further account for such reporting errors this would have increased the likelihood of misclassifying the mechanism of spread to certain premises and clusters. We also purposefully excluded several records of potential fomite spread from this analysis as routine reporting of such movements between infected premises was not conducted during the outbreak. The

inclusion of these records may have slightly reduced our estimates of the distance of local spread but would have also introduced a lot of uncertainty.

A potential limitation of this analysis was that the length of the infectious period for a premises was not incorporated into the construction of the proximity network. Therefore, it is possible that some proximity links may emanate from IPs that were no longer infectious. Given the relatively short study period (30 days) compared to the estimated 7–21 days that premises may remain infectious (Garner et al., 2011b), and the tight spatial clustering of IPs, this was considered to have only a minor consequence on the resulting structure of the proximity network. In further research, stochastic modelling and sensitivity analyses could be utilised to ensure that directed proximity links are only added from IPs during the period that they are predicted to be infectious, to produce a directed 'infectious proximity network'.

5.5. Conclusions

This analysis shows the importance of contact network topology in epidemic disease spread. Combined analysis of spatial and contact network data demonstrated that local spread emanated outwards from the small number of infected premises in the contact network, up to a distance of around 15 km. When such a network structure underlies an epidemic, a method of spatial analysis that incorporates the network component of disease spread is required; otherwise the accuracy of estimation of spatial modelling and the estimation of key epidemiological parameters may be compromised. Linking the spatial and network analysis of epidemics facilitates inference about the methods of disease transmission, enabling investigation of spread both through a network and in space and providing a means of describing the formation of spatial clusters.

Chapter 6: Within cluster spatio-temporal analysis

– the influence of meteorological factors

"***Weather.***—*Atmospheric phenomena have always occupied a prominent place among the influences of environment which man regards as most potent for good or for evil so far as his own well-being and that of his livestock was concerned.*

Weather and its effects, the most appreciable of all manifestations of nature, have ever aroused his interests, and the lore which developed around this subject during the age bears abundant testimony of the belief that the state of health or of disease was in a large measure weather born."

Leunis Van Es, Professor of Animal Pathology and Hygiene, The University of Nebraska.

In: Van Es, L. (1932), 'The Principles of Animal Hygiene and Preventive Veterinary Medicine', John Wiley & Sons, Inc. New York. p.183

This chapter appears as the following published paper in the journal Public Library of Science ONE:

Firestone, S.M., Cogger, N., Ward, M.P., Toribio, J.-A.L.M.L., Moloney, B.J., Dhand, N.K., 2012. The Influence of Meteorology on the Spread of Influenza: Survival Analysis of an Equine Influenza (A/H3N8) Outbreak. PLoS ONE 7, e35284.

6. Within cluster spatio-temporal analysis – the influence of meteorological factors

6.1. Introduction

Influenza A viruses are enveloped RNA viruses of the family *Orthomyxoviridae*, and a major cause of morbidity and mortality in both humans and livestock, worldwide (Myers and Wilson, 2006, Alexander, 2007, Belser et al., 2010). Spread may be via direct contact, over short distances on large 'cough' droplets (diameter >10 μm), over longer distances in aerosols of small droplet nuclei (diameter <10 μm) and on fomites (Weber and Stilianakis, 2008, Tellier, 2009). Meteorological variables such as air temperature, relative humidity, rainfall and wind have been suggested as important drivers of the spread and seasonality of influenza in both human (Lowen et al., 2007, Weber and Stilianakis, 2008, Murray and Morse, 2011, Steel et al., 2011) and animal populations (Fang et al., 2008). Recently, Lowen et al. (2007, 2008) described, under laboratory conditions, how relative humidity and ambient temperature combine to influence the transmission of both seasonal (A/H3N2) and pandemic (A/H1N1) human influenza A (Steel et al., 2011). The effects of several other environmental variables (soil pH, sunlight and surface permeability) on the survivability of influenza A viruses were established in earlier laboratory-based experimentation (Bean et al., 1982, Yadav et al., 1993). Analyses of the contribution of wind to the spread of epidemics of influenza, and indeed other infectious diseases, are more limited. Most studies present either circumstantial evidence that the mean direction of epidemic spread coincides with prevailing wind conditions at the time of an outbreak (Davis et al., 2009, Kedmi et al., 2010), analyses of data aggregated to a low temporal or spatial resolution (Yuan et al., 2006, du Prel et al., 2009), or associate spread from a small number of sources with atmospheric dispersal modelling outputs (Gloster et al., 2011). Such research must also overcome the added complexity of movement of individuals within the population at risk.

In their animal model of human influenza A transmission, Lowen et.al. (2007) have shown that dry cool conditions (low relative humidity and cold ambient temperatures) increase the spread of influenza. They suggest that this mechanism is mediated by a complex interaction that affects the survivability of both aerosol droplet nuclei and virus particles. A detailed analysis, at high spatial and temporal resolution, comparing actual influenza outbreak data with concurrent meteorological data is required to validate and provide context to their model outside of controlled laboratory environments, thus furthering our understanding of how meteorological factors truly influence influenza spread. Outbreaks of disease in animal populations present a unique opportunity to study such effects, 'in the field'. Research on detailed animal outbreak datasets has several distinct advantages over comparable research on public health influenza data (Charland et al., 2009, du Prel et al., 2009, Chan et al., 2010, Shoji et al., 2011). Firstly, human populations move about on a daily basis (albeit with some regularity). Implementing a complete human movement standstill ('a 24 hour curfew') to control and contain an outbreak is considered an extraordinary and perhaps unfeasible social distancing measure, reserved for the most severe of human influenza pandemics (Australian Government Department of Health and Ageing, 2010). Conversely, the movements of farm animal populations (such as horses, cattle and sheep) are mostly confined to within single premises, and in the event of an emergency animal disease outbreak, a complete movement ban is often the first control measure to be implemented (Animal Health Australia, 2011). Furthermore, ethical concerns (namely privacy) may constrain the research of human outbreak data, limiting the amount of detailed information that can be collated on the movement of individual people. Given that certain human and animal sub-types of influenza A share generally similar modes and patterns of transmission (Weber and Stilianakis, 2008), research that utilises detailed animal outbreak datasets has the potential to inform our understanding of the complex mechanisms that influence human influenza A spread and seasonality. The 2007 outbreak of equine influenza in Australia presented an excellent opportunity to study the effects of meteorology on the spread of an influenza A virus as it infected a mostly immunologically naïve population, spatially confined (in paddocks).

Equine influenza virus (A/H3N8) is a highly contagious cause of low mortality, high morbidity respiratory disease capable of infecting all members of the horse family (*Equidae*). It is considered endemic to equine populations across most of the world (Myers and Wilson, 2006). The disease is similar in many clinical and epidemiological respects to seasonal human influenza A, and major outbreaks have occurred when novel strains of equine influenza have gained entry into highly susceptible equine populations (Myers and Wilson, 2006). The typical incubation period of equine influenza is 1–3 days (McQueen et al., 1966, Mumford et al., 1994, Paillot et al., 2006), however delayed onset of clinical signs of up to 5 days has been observed after low dose aerosol exposure (Mumford et al., 1990). In 2007, following a breach in the quarantine of infected imported horses (Callinan, 2008), Australia experienced its first ever outbreak of equine influenza. Less than 900 horses are imported annually into Australia from countries that vaccinate for equine influenza (Callinan, 2008), therefore almost the entire horse population was susceptible at the start of this outbreak. Over the course of 4 months, approximately 67,000 horses were infected on 9359 premises in two Australian States—New South Wales (NSW) and Queensland (QLD). Timely and complete implementation of a horse movement ban has been widely credited as the most effective of the control measures that facilitated the rapid eradication of this disease from the Australian horse population (Callinan, 2008). Although vaccination was used to eradicate the disease, its implementation only commenced 6 weeks into the outbreak, well after the peak of reported daily infections (Garner et al., 2011a). Vaccination was initially restricted to disease containment zones and the protection of high value horses (Garner et al., 2011a).

Contact-tracing early in the 2007 outbreak revealed that the disease initially spread through a network of equestrian events, linked by the movement of infected horses prior to detection of the outbreak, producing clusters of infected premises in widespread locations (Callinan, 2008, Firestone et al., 2011b, Firestone et al., 2012a). Epidemiological investigations noted rare instances of presumed windborne spread over short ranges (≤1.5 km, and rarely up to 5 km) based on failure to identify other potential means of transmission (i.e. close contact or fomites)

(Moloney et al., 2011). Previous epidemiological analyses of this outbreak have investigated the spatial and network components of early spread (Cowled et al., 2009b, Firestone et al., 2012a, Firestone et al., 2011b, Moloney et al., 2011), and premises-level risk factors for disease spread such as compliance with advised biosecurity measures (Firestone et al., 2011a). Two further analyses have specifically investigated environmental factors that might have influenced the spread of this outbreak (Davis et al., 2009, East, 2009). In one cluster of 437 infected premises, a relationship was observed between prevailing wind conditions and the global direction of spread (Davis et al., 2009).

In this paper we present a comprehensive analysis of the influence of meteorological variables on time to infection based on an influenza A virus outbreak dataset. This spatio-temporal analysis aims to identify and quantify the association between four meteorological variables (air temperature, relative humidity, rainfall, wind velocity) and time to infection in the largest cluster of the 2007 equine influenza (A/H3N8) outbreak in Australia. We are unaware of any previously published analysis that combines such a large and spatio-temporally detailed influenza outbreak dataset with concurrent daily meteorological data, to allow meaningful estimation of the contribution of such factors in the spread of an influenza A outbreak.

6.2. Materials and methods

6.2.1. The equine influenza dataset

The state government of New South Wales provided contact-tracing and laboratory testing data on all horses investigated during the 2007 outbreak. This dataset was collected at the level of individual horses and aggregated to the premises level for analysis. Study designs that use groups as the unit of interest (such as herds or flocks) rather than individuals, are common in veterinary epidemiological research (Dohoo et al., 2009). Premises attribute records included address, geocoded coordinates (based on premises centroid), number of horses, date of onset of clinical signs in the first horse affected (‘onset date’), vaccination status and date of vaccination.

Premises were defined as infected (IP) if they held horses that had been observed with the classical clinical signs of equine influenza (cough, elevated temperature, nasal discharge and lethargy). This status was confirmed by laboratory testing based on real-time reverse transcription polymerase chain reaction assay (Foord et al., 2009), however, around the peak of the outbreak not all horses were tested due to resource constraints (Moloney et al., 2011). Contact-tracing records included the date of the movement, the addresses and unique identifiers for the origin and destination premises between which horses were moved prior to the horse movement ban.

6.2.2. Study extent: cluster delineation

There was a single 'index' for the 2007 outbreak of equine influenza in Australia: an equestrian event located 160 km north of Sydney, at which transmission was known to have occurred. This analysis focused on local spread within the single largest cluster of the outbreak, centred 60 km northwest of Sydney's city centre (Figure 6.1). To maintain a computationally tractable dataset, premises were selected for inclusion in the study (from the equine influenza dataset) if their centroid was within 15 km of nine contact-traced 'source' premises. All nine contact-traced premises were identified (based on an earlier likelihood-based analysis (Firestone et al., 2012a)) to have been infected in the first week of the outbreak following the movement of infected horses from the 'index'. The 15 km buffer used to delineate the cluster was selected based on a previous analysis in which we identified that 98% of premises infected in the first month of the outbreak were within this distance of a contact-traced 'source' premises (Firestone et al., 2011b). The 'Northwest Sydney' cluster studied was approximately 65 km in diameter, bounded to the North and West by national parks (where horses are prohibited) and to the South and East by metropolitan Sydney.

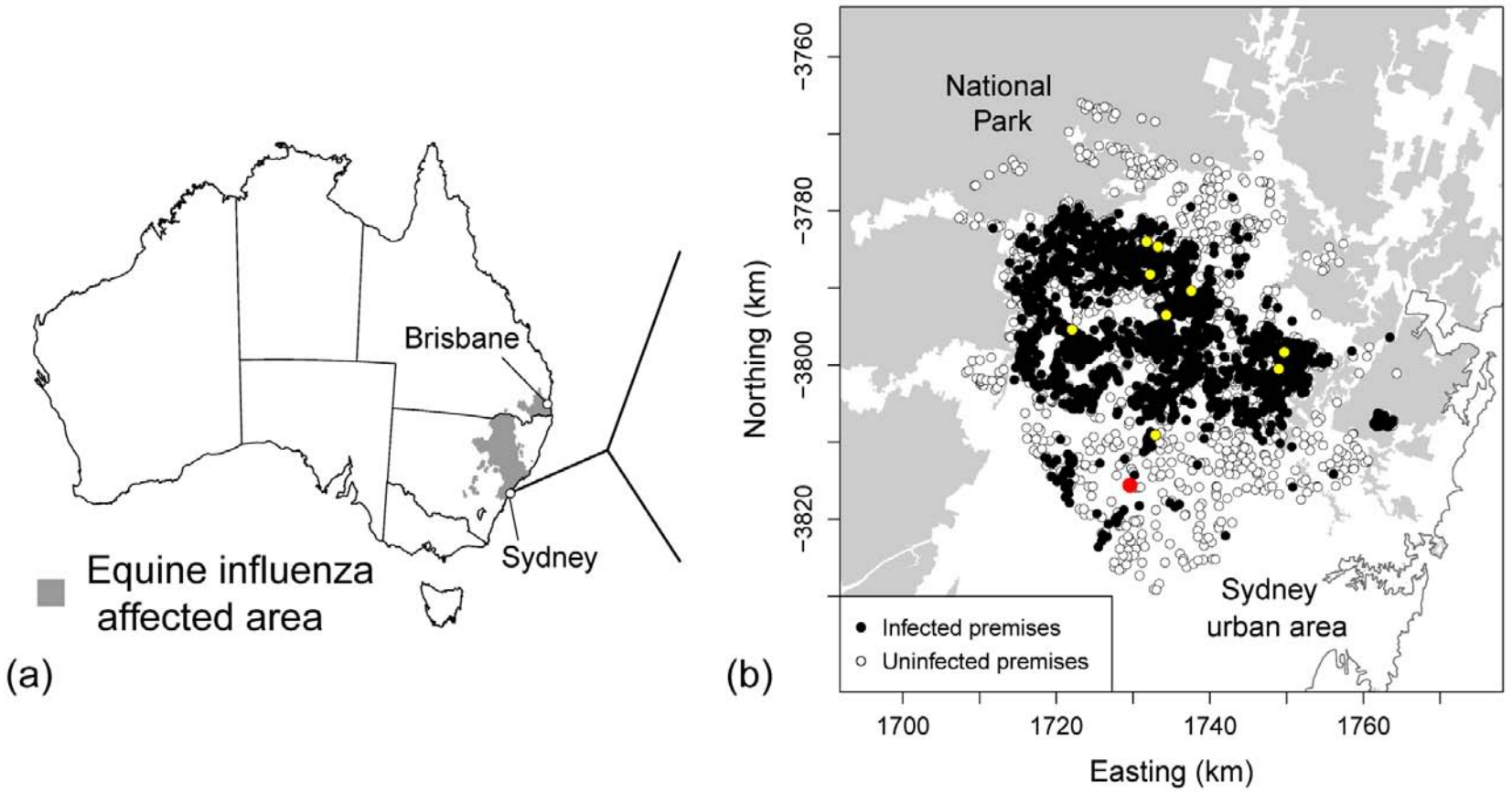


Figure 6.1 Study extent in a survival analysis of the influence of meteorological factors on the spread of the 2007 equine influenza outbreak in Australia. (a) The area affected by the 2007 equine influenza outbreak in Australia. (b) The largest cluster (n = 3624), northwest of Sydney, as defined by a 15 km buffer around the nine earliest infected premises (depicted in yellow) that were contact-traced to events where disease transmission was known to have occurred in the first week of the outbreak. Clinical signs were first observed on 17 August 2007 in a horse in quarantine at Eastern Creek Quarantine Station (red closed circle). The cluster is surrounded by national parks and Sydney urban areas.

6.2.3. Exploratory spatial and temporal analyses

The dataset was imported into the R statistical package version 2.13.0 (R Development Core Team, 2011), and an epidemic curve constructed as the count of infected premises reported per day. The spatial coordinates of each premises were converted to the Australian Albers conic equal-area projection which is based on the Geocentric Datum of Australia 1994 (Geoscience Australia, 2007). Extraction mapping was used to investigate the spatial pattern of risk of infection over time. To identify areas of elevated risk, risk surfaces with upper 95% tolerance contours were estimated as the Gaussian-smoothed kernel density surface of infected horse premises divided by the surface of the population of horse premises at risk in 4-week time periods. A spatially adaptive variable smoothing parameter was used to prepare the risk surfaces (Davies and Hazelton, 2009), with edge effect correction, implemented in R with the 'sparr' library (Davies et al., 2011). The amount of smoothing (bandwidth) applied varied across the study extent in inverse proportion to the population at risk in each time period. To test for directional spread, the mean geographic centre of the outbreak was estimated by week as the mean of the coordinates of the infected premises with dates of onset in each week of the outbreak (Ward and Carpenter, 2000b).

6.2.4. Survival analysis

We applied semi-parametric Cox regression modelling to estimate the association between potential risk factors and the times to infection of individual premises. A geodatabase was compiled in Microsoft Access 2007 (Microsoft Corporation, Redmond, WA, USA) to maintain all premises and meteorological data, with spatial covariates added using ArcMap 9.3 (ESRI, Redlands, CA, USA). The dataset was structured into a daily 'counting process' formulation to enable investigation of the effects of time-varying predictors (Anderson and Gill, 1982), in this case time-lagged premises-level meteorological variables. In this formulation, each premises contributes one observation for every day that it is at risk (until either clinical signs are observed in horses on the premises, or the end of the study period). See the Appendix Tables 6.6 & 6.7

for a sample of the survival dataset used in this analysis. Time-varying covariates and the counting process formulation were arranged using the R statistical package.

In the counting process generalisation of the Cox proportional hazards model, the hazard function depends on time in ways other than only through the baseline hazard function (Hosmer et al., 2008). The proportional hazards assumption does not apply, allowing for inclusion of time-dependent covariates (Allison, 2010). Each subject contributes one observation for every day that it is at risk and each observation contains covariates for the subject at each time point of observation and a start and stop time denoting the interval of risk, i.e. (start, stop] (Therneau and Grambsch, 2000). This enables covariate values for individual subjects to either be time-invariant or to change with time, and to be incorporated into a generalised Cox regression model (Anderson and Gill, 1982) of the form:

$$h_i(t) = h_0(t) \cdot exp\left(\beta_1 x_{i1} + \beta_2 x_{i2(t)} + \ldots + \beta_k x_{ik(t)}\right) \quad (6.1)$$

where $h_i(t)$ is the hazard that an individual, i, from the population yet to experience an event, will experience the event at time t; $h_0(t)$ is the baseline hazard at time t; β_1 and β_2 are the regression coefficients for the time-invariant, x_{i1}, and time-dependent covariates, $x_{i2(t)}$, respectively. The partial likelihood specification for the counting process Cox regression model is described in detail by (Anderson and Gill, 1982), and is estimated including a term for each unique event time, summing over those observations that are still at risk at each actual event time. As there is no overlap in intervals of risk in the set of observations for each subject, the likelihood never involves more than one observation for a subject (Therneau and Grambsch, 2000).

Network and spatial spread in the early outbreak period (the first 14 days of this outbreak) is described in detail elsewhere (Firestone et al., 2012a). To focus this analysis on the

meteorological factors associated with local spread, we excluded any premises that may have been infected in the first 10 days of the outbreak, before the complete implementation of horse movement bans (i.e. any premises with an onset date in the first 14 days of the outbreak), setting the origin of the survival analysis at 30 August 2007. This period ends one typical incubation period (3 days) after movement bans were implemented, with an additional 1 day error margin for delay in observation and reporting (Firestone et al., 2011a). All premises that remained uninfected on the 131st day of the outbreak (25 December 2007, the reported date of onset of the last known infected premises) were right censored on this date.

6.2.4.1. Explanatory variables

Explanatory covariates tested for associations with the time to infection of premises in the Northwest Sydney cluster are listed in Table 6.1. Premises boundaries were extracted from cadastral data provided by the NSW Government Department of Finance and Services. These boundaries were used to generate a continuous variable representing the length of fence that each horse premises shared with any contiguous horse premises in the equine influenza dataset. Premises elevation was extracted from a digital elevation model of Australia (Geoscience Australia, 2008), which is a grid of ground level elevation covering the whole of Australia with a grid spacing of approximately 250 metres, as the mean of all grid cells needed to cover a premises. Distance to the nearest main road was calculated from the premises boundary using vector data of road Classes 1–3 (freeways, highways, primary and arterial roads) (Geoscience Australia, 2006). Human population density, within approximately 1 km of the premises centroid, was estimated based on high resolution gridded population data from 2005 (Center for International Earth Science Information Network et al., 2011), adjusted by 3% for population growth between 2005–2007 (World Bank, 2011).

Table 6.1 Explanatory variables analysed for associations with time to infection of premises in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia.

Variable group	**Variable name**	**Variables** (Units)
Meteorological covariates (time-lagged)	RAIN $t_{-1},..,t_{-5}$	Rainfall (mm day^{-1})[a]
	RH_9AM $t_{-1},..,t_{-5}$	Relative humidity (%) measured daily at 9am[a]
	RH_3PM $t_{-1},..,t_{-5}$	Relative humidity (%) measured daily at 3pm[a]
	TEMP_MAX $t_{-1},..,t_{-5}$	Maximum daily air temperature (°C)[a]
	TEMP_MIN $t_{-1},..,t_{-5}$	Minimum daily air temperature (°C)[a]
	$WIND\_SPD_{undir}$ $t_{-1},..,t_{-5}$	Maximum daily wind speed – undirected (km hour^{-1})[a,b]
	$WIND\_SPD_{dir(k)}$ $t_{-1},..,t_{-5}$	Maximum daily wind speed – directed (km hour^{-1})[a,b]
Premises attributes	AREA	Area (acres)
	HORSE_DENSITY	Horse density (horses acre^{-1})
	HORSES_NUMBER	Number of horses
	SHARED_FENCE	Length of shared fence with other horse premises (m)
	VACC	Vaccination status (1=Yes, 0=No)[a]
	VACC_DAYS	Days since vaccination[a]
Spatial covariates	ELEV	Elevation (m)
	HUMAN_DENS	Human population density within approximately 1 km of the premises (people km^{-2})
	ROAD_DIST	Distance to nearest main road (km)[c]

[a] Time-changing covariate.
[b] Maximum daily wind speed was either based on wind from all directions ('undirected') or wind only from within 45° arcs centred on the direction of the k nearest infected premises for $k = 1,2,3$ (see Figure 6.2 for details) assuming that premises were infectious for 14 days and one of the nearest k infective premises was the source of infection.
[c] Main roads include freeways, highways, primary and arterial roads (Classes 1 to 3).

Estimation of meteorological time-varying predictors

Hourly wind velocity data (wind direction and speed) and daily data for five other meteorological variables (rainfall, minimum and maximum daily air temperature, and relative humidity measured at 9am and 3pm) were obtained from 132 weather stations. All of these weather stations were operated by the Australian Bureau of Meteorology during the study period, and were located either within the cluster or within 20 km of the cluster boundary. Most stations reported only daily rainfall measurements. Ordinary kriging (Matheron, 1963) was used to interpolate daily values at each individual premises location for the meteorological time-varying predictors: maximum wind speed (km hour^{-1}), rainfall (mm), maximum and minimum surface air temperature (°C), and relative humidity (%, measured at 9am and 3pm). Each time-varying meteorological covariate was then time-lagged by 1–5 days to encompass the full range

of incubation periods observed in experimental infection studies (Mumford et al., 1990).

Kriging is a geostatistical smoothing technique that involves modelling the underlying spatial dependency (autocorrelation) in spatially continuous data based on a covariance function (Figure 6.2a) (Matheron, 1963). For each observation point (hour or day), for each meteorological variable, a binned isotropic empirical variogram was plotted that represented covariance (as semivariance) up until half of the maximum pairwise distance between any two weather stations contributing data at that time point, with bin widths (h) of approximately 10% of the average distance between weather stations (Journel and Huijbregts, 1978). A stationary exponential variogram model was then fit to the empirical variogram, using iterative least squares regression, and parameter estimates used to interpolate values at each premises location (Matheron, 1963).

Generation of wind speed covariates

Hourly wind velocity data were available from sixteen of the weather stations, automatically measured on masts at 10 m above the earth's surface. These wind data were supplied in a polar coordinate structure, comprising the average direction of origin of the wind (in degrees from true north) and the maximum wind speed (in km $hour^{-1}$), measured over the 10 minutes leading up to the observation time. To avoid the issue of northerly bearings being split at true north (i.e. true bearings of 1° and 359° seeming distant when they are only 2° apart), prior to variography and kriging, the wind velocity data was converted into a Cartesian coordinate system—defined by two components (Figure 6.2b,c): "u" representing the East-to-West component of the wind velocity, and "v" representing the North-to-South component (Inggs and Lord, 1996). A negative value for the “u” component therefore represents a wind from one of the westerly bearings (i.e. NW, W or SW). Kriging was then conducted on the two wind velocity vector components (Inggs and Lord, 1996). Hourly wind velocity vectors were interpolated for each premises and back-transformed into the original polar coordinates (direction of wind origin and maximum wind speed).

Two approaches were taken to aggregate the hourly wind velocity vectors for each premises into daily maximum wind speed covariates. First, to test the hypothesis that increased wind speed *from any direction* was associated with increased hazard of infection we generated 'undirected' maximum daily wind speed covariates ('WIND_SPD$_{undir}$') without making any directional assumptions, taking the maximum of all hourly wind speed estimates for each premises on each day.

Next, to explore the directionality of wind exposure risk we generated 'directed' maximum daily wind speed covariates ('WIND_SPD$_{dir}$') based only on wind coming from within the direction of the nearest k infected premises (for k=1,2,3) by selecting wind from within 45° arcs centred on the bearing of the nearest k infected premises to each premises on each day. For each premises, on each day of observation, we identified the nearest three infected premises from amongst those infected premises that had a date of onset (of clinical signs in the first horse infected on the premises) within the previous 14 days. Though it is known that individual unvaccinated horses remain infectious for up to 7 days (Myers and Wilson, 2006, Daly et al., 2011), the duration of infectivity may vary on multi-horse premises because of differences in contact rates between individual horses, and individual variability in susceptibility, latency and virus shedding. To infer which premises were holding infectious horses at each time point we assumed that the period of infectivity was 14 days for all premises based on case reports from horse premises of a range of sizes (Dups et al., 2011, Faehrmann et al., 2011, Wong, 2011), intra-herd simulation modelling (Garner et al., 2011b) and that almost the entire population was immunologically naïve to equine influenza at the start of the outbreak.

Finally, these time-varying predictors were lagged by 1–5 days to serve as proxies for wind within the range of incubation periods that have been observed for equine influenza, producing 20 time-lagged explanatory covariates: 'WIND_SPD$_{undir}$ t_{-1}, t_{-2}, . . ., t_{-5}' and 'WIND_SPD$_{dir(k)}$ t_{-1}, t_{-2}, . . ., t_{-5}', for k=1,2,3.

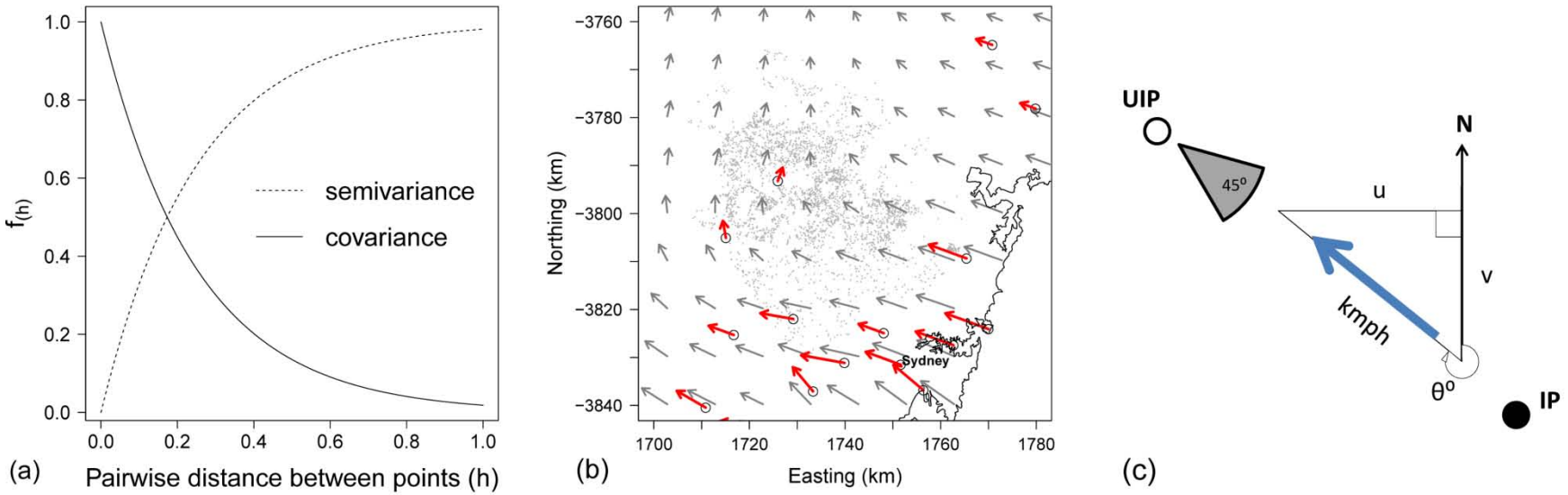


Figure 6.2 Generation of covariates representing premises-level wind exposure risk.
(a) Exponential covariance function (with practical range = 0.25) and its related semivariance function. (b) Hourly wind velocity data from sixteen automated weather stations (open circles) within a 20 km buffer of the cluster's boundary were converted into their East-to-West ('u') and North-to-South ('v') components, and smoothed using kriging to predict hourly wind speed and direction at each premises (small grey dots). (c) For each premises on each day prior to infection or censoring, the ('directed') maximum wind speed originating from within 45^0 arcs centred on the direction of the nearest 1–3 infected premises was estimated for time lags of 1–5 days.

6.2.4.2. Univariable analyses

Instantaneous hazard curves were constructed for each time-invariant covariate with the R library 'epiR' (Stevenson et al., 2010), categorising continuous variables into quartiles. The instantaneous hazard rate, $h(t)$, is the rate at time t, that a randomly-selected individual from the population yet to experience an event, experiences the event at time t (Allison, 2010), and is mathematically defined as:

$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \tag{6.2}$$

where T is the time that an event is experienced. In this study, the unit of interest was the horse premises, and events were defined as the infection of horses with equine influenza virus on a previously uninfected premises.

Univariable Cox models were then constructed and the statistical strength of the association between each variable (categorical or continuous) and the outcome assessed using likelihood ratio tests (Dohoo et al., 2009). The linearity of the relationship between log hazard of infection and each continuous variable was assessed graphically using restricted cubic splines (Harrell 2001) with knots spaced at quintiles in the data. To differentiate linear and nonlinear component terms, partial likelihood ratio tests were conducted comparing a model containing all spline terms to a nested model containing only a single linear term (Harrell 2001). If a highly non-linear relationship was detected, the spline of the continuous variable was retained for multivariable analysis. All continuous covariates were tested for collinearity in pairs by calculating Spearman's rank correlation coefficient (ρ). Intrinsic temporal autocorrelation was expected amongst certain groups of time-lagged time-varying meteorological predictors, such as: 'TEMP_MIN t_{-1}'with 'TEMP_MIN t_{-2}, . . ., t_{-5}' and 'RH_3PM t_{-1}' with 'RH_9AM t_{-1}, t_{-2}, . . ., t_{-5}'. From amongst any pair of highly correlated (ρ >|0.70|) time-invariant covariates, and from amongst intrinsically temporally autocorrelated groups of time-varying predictors, only the

variable with the strongest statistical association with the outcome was retained for further analysis (Armitage et al., 2001).

6.2.4.3. Multivariable analyses

All remaining variables (unconditionally statistically associated with the log hazard of infection at *P*-value <0.25) were entered into a generalised 'counting process' Cox regression model (Anderson and Gill, 1982). Each eligible candidate variable was then individually tested by excluding it from the maximal model and conducting likelihood ratio tests, eliminating any variables with *P*-value ≥0.10. To assess confounding, all eliminated variables were individually added back into the model, retaining any terms that resulted in a >20% change in any regression coefficient. The time-varying predictor representing vaccination status was forced into all multivariable models as it was considered a priori to confound disease spread. The linearity of the relationship between the outcome and each continuous variable still included in the model was assessed again, using restricted cubic splines (Harrell 2001). Finally, tests were conducted for all two-way interactions of terms in the preliminary main effects model.

Goodness of fit of the final model was assessed using 'Martingale' residuals. The influence of every individual observation was tested by omitting it and observing for change in the regression coefficients (Therneau et al., 1990). To test for spatial dependency (autocorrelation) we examined the spatial structure of the residuals of the final model by mapping normalised martingale residuals ('deviance residuals') and plotting an empirical semivariogram (von Klot et al., 2009).

6.3. Results

6.3.1. Exploratory spatial and temporal analysis

The Northwest Sydney cluster of the 2007 equine influenza outbreak in Australia contained 3624 horse premises, of which 1922 were reported to be infected during the 131 day outbreak (cumulative incidence = 53.0%, 95% CI: 51.4, 54.7%).

Surfaces of spatial relative risk by four week period are included as Figure 6.3. In the first 4 weeks of the outbreak there were two areas of elevated spatial risk localised around the nine source premises for this cluster. Over the next 4 weeks, the two areas of elevated risk coalesced and expanded. Between weeks 9–12, the areas of spatial risk dissipated into several smaller pockets of infection. Over the remainder of the outbreak, the spatial risk faded out in isolated pockets of infection.

The mean centre of the outbreak did not move predominantly in any single direction over the study period, moving northwest at 3.0 km week^{-1} in the first 4 weeks, then southwest at 3.9 km week^{-1} for 4 weeks, before moving back to the East at 4.1 km week^{-1} whilst the epidemic faded out.

6.3.2. Survival analysis

The complete survival dataset included 3153 premises containing 1727 events (infections) during the study period. Data on 57 infected horse premises were excluded because their onset dates occurred in the first 14 days of the outbreak (a period when they could possibly have been infected by the movement of infected horses rather than by local spatial spread). Sixty-seven infected premises were missing a date of onset, and 347 premises (71 infected and 276 uninfected premises) were missing data on their number of horses. Once data on these premises (which were evenly distributed across the study extent) had been excluded, data on all variables were complete. The median survival time, the point at which half of the premises in this cluster

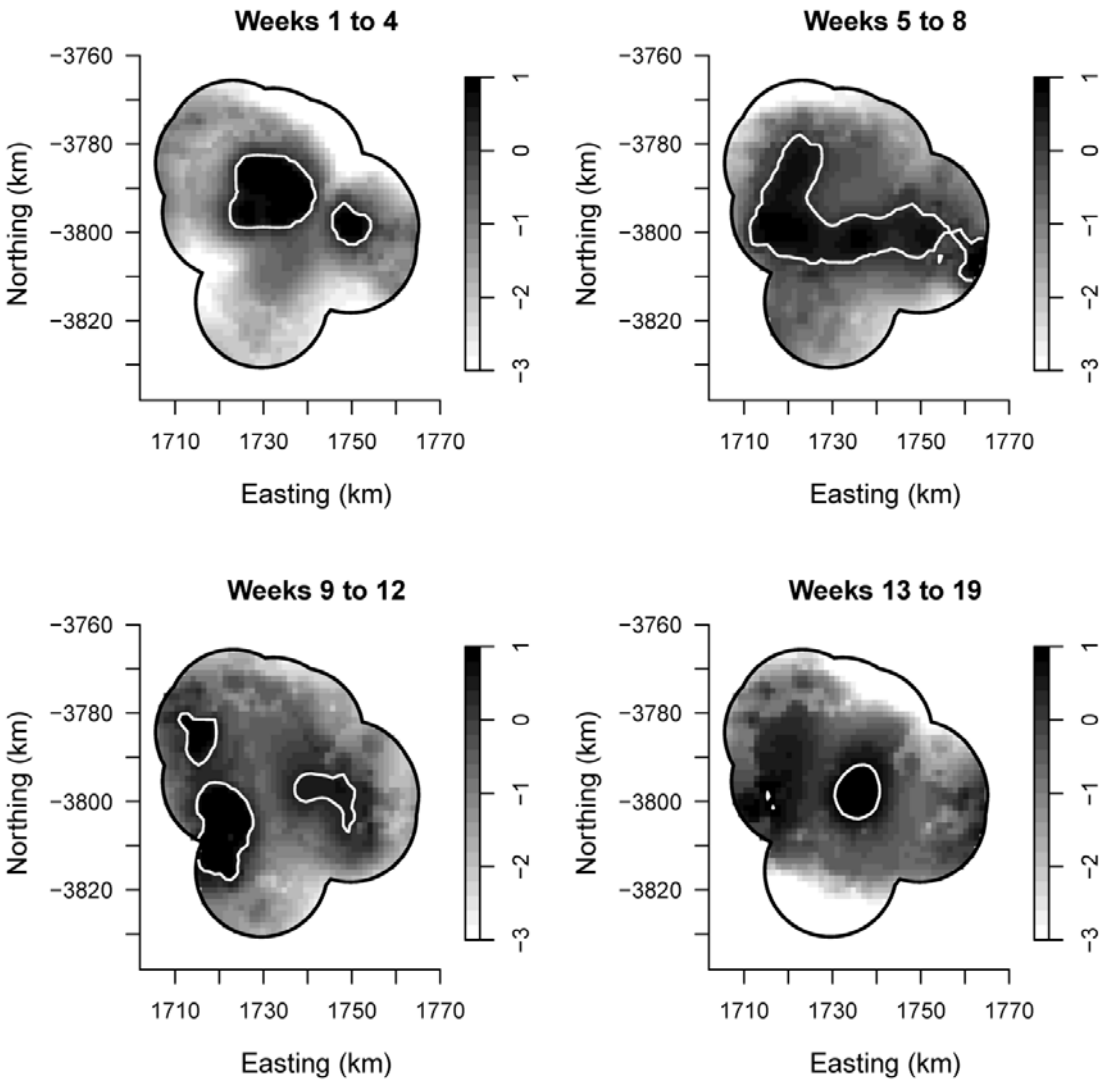


Figure 6.3 Spatial spread of equine influenza in the largest cluster (n = 3624), northwest of Sydney, of the 2007 outbreak in Australia.
Surfaces of log relative risk were estimated in 4-week intervals using adaptive kernel estimation, with upper 95% tolerance contours (solid white lines). With this method the amount of smoothing (bandwidth) is inversely proportional to the density of the population at risk.

were infected, was day 55 of this outbreak (95% CI : 52, 61). The instantaneous hazard, the proportion of infections per day in the population surviving uninfected until that day, peaked on day 28 (Figure 6.4); 92 premises were reported to be infected on this day.

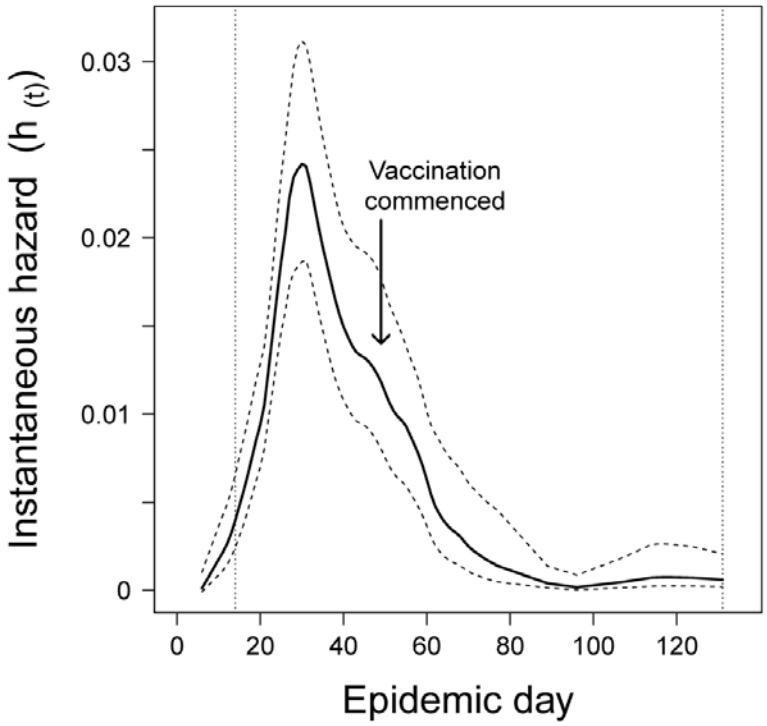


Figure 6.4 Smoothed instantaneous hazard of infection (with 95% confidence intervals) in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia. Horse movement standstills were implemented from day 10, and vaccination commenced in this cluster on day 49. Dotted reference lines denote the survival analysis study period (between days 14 and 131 of the outbreak).

6.3.2.1. Univariable analysis

Meteorological covariates and hazard of infection

Most horse premises were relatively close to a weather station, with the mean distance to the nearest weather station reporting wind data being 11.7 km (SD = 5.4 km, maximum = 27.4 km). For all meteorological data, there was a paucity of weather stations in the northwest corner of the study extent (because this region is bordered by a national park).

Daily rainfall data were available from 127 weather stations in the study extent. Over the study period, the median estimated daily rainfall per premises was 0.1 mm day^{-1} (IQR: 0 to 2.8 mm day^{-1}, maximum = 106.5 mm day^{-1}). No statistically significant associations were detected

between time-lagged rainfall covariates and hazard of infection (Table 6.2). Moderate temporal correlation ($\rho \approx 0.60$) was observed between rainfall data 1 day apart, and between rainfall and relative humidity measurements conducted within 1 day of each other. A detailed correlation matrix of all continuous covariates is provided in the Appendix Table 6.8.

Relative humidity data measured twice daily (at 9am and 3pm) were available from eighteen weather stations (Figure 6.5a). The mean of the estimated 9am and 3pm relative humidity measurements for the horse premises under observation were 70.8% (SD = 17.5%) and 52.9% (SD = 20.3%), respectively. Conditions were drier when measured at the same station at 3pm compared to 9am, on any given day, with paired relative humidity measurements 16.0% on average lower in the afternoon (95% CI: 15.3, 16.8%). Moderate to high temporal autocorrelation ($\rho \approx 0.70$) was observed between 9am and 3pm relative humidity data on the same day and at the same time 1 day apart.

A negative cubic relationship was observed between relative humidity and hazard of infection (Figure 6.5b). Risk of equine influenza infection was highest in dry conditions (<20% relative humidity), decayed rapidly until increasing at intermediate relative humidity (40–60%). Once relative humidity was >80% there was effectively no risk. This relationship was independent of whether relative humidity was measured at 9am or 3pm, and was also independent of the time lag applied (Table 6.2). The strongest statistical association was with the 3pm measurement time-lagged by 5 days, thus, 'RH_3PM t_{-5}' was selected as a proxy for relative humidity, irrespective of diurnal variation or time-lag.

Daily surface air temperature data were available from 21 weather stations (Figure 6.5c). The mean of the estimated daily maximum and minimum temperatures at the 3153 horse premises was 24.0 °C (95% CI: 22.4, 25.7), and 12.6 °C (95% CI: 9.1, 16.1 °C), respectively. There was an increasing trend in temperature across the study period as the season changed from spring to summer, and a low level of correlation ($\rho \approx 0.40$) between maximum and minimum temperature

measured on the same day. Minimum daily temperature data 1–4 days apart were moderately to highly correlated ($0.61 \leq \rho \leq 0.75$), less correlation was observed between maximum daily temperatures 1 day apart ($\rho \approx 0.53$), and a low cross-correlation ($0.30 < \rho < 0.50$) was observed between minimum daily temperature, rainfall and 9am relative humidity data for 1–5 days.

A highly nonlinear relationship was observed between infection and maximum daily air temperature (Figure 6.5d), with risk of infection greatest toward both extremes of the range of observed maximum temperatures (<16 °C and >28 °C). The statistical strength of this association was greatest at a time-lag of 3 days (Table 6.2), however, the shape was consistent across time-lags. Hazard of infection increased linearly as minimum daily temperatures decreased, and the statistical strength of this association was also greatest when a time-lag of 3 days was applied. Combining daily maximum and minimum measurements into a midpoint daily temperature resulted in statistically weaker associations (Table 6.2).

Hourly wind velocity data were available from sixteen weather stations (Figure 6.2b). Wind conditions varied considerably in time with little temporal autocorrelation observed (Appendix Table 6.8). There was no clearly discernible predominant wind pattern over the study period. The median of the maximum daily reported wind speeds estimated for each premises (*from all directions*) was 26.6 km hour^{-1} (IQR: 22.3, 33.3 km hour^{-1}, maximum = 73.3 km hour^{-1}).

The univariate relationship between hazard of infection and wind speed, making no directional assumptions ('undirected'), is presented in Figure 6.6, by time-lag. Maximum daily wind speed, lagged by 3 days, had the strongest statistical association with the outcome (Table 6.2). Increased hazard of infection was observed on days when the maximum daily wind speed was >30 km hour^{-1}. The univariate relationships between hazard of infection and maximum daily wind speed from the direction of the *k* nearest neighbours are presented in Table 6.3, and plots of the restricted cubic splines of these relationships are shown in Figure 6.7 (only for a time-lag of 3 days). The strongest statistical association between any wind speed covariate and hazard of

infection was identified based on ‘directed’ wind speed from the direction of the three nearest neighbours, time-lagged by 3 days.

The following five candidate meteorological variables were consequently selected for multivariable analysis: linear terms for rainfall and minimum daily air temperature, both time-lagged by 3 days, a restricted cubic spline for relative humidity measured at 3pm time-lagged by 5 days, and splines of maximum daily air temperature and maximum daily wind speed from the direction of the three nearest infected premises, both time-lagged by 3 days.

Table 6.2 Univariable analysis of the association between meteorological covariates (time-changing and time-lagged) and time to infection of premises in the largest cluster (n = 3153), northwest of Sydney, during the 2007 equine influenza outbreak in Australia.

Meteorological Factor	Time-lag	*b*	SE(*b*)	LRT	*df*	*P*-value[a]
Rainfall (mm day^{-1})	t_{-1}	0.006	0.033	0.0	1	0.852
	t_{-2}	-0.005	0.028	0.0	1	0.870
	t_{-3}	-0.045	0.037	1.5	1	0.215
	t_{-4}	-0.024	0.027	0.9	1	0.342
	t_{-5}	-0.028	0.031	0.9	1	0.344
Relative humidity, measured daily at 9am (%)	t_{-1}	*nonlinear spline*		29.9	4	<0.001
	t_{-2}	*nonlinear spline*		29.4	4	<0.001
	t_{-3}	*nonlinear spline*		14.3	4	0.006
	t_{-4}	*nonlinear spline*		47.1	4	<0.001
	t_{-5}	*nonlinear spline*		41.9	4	<0.001
Relative humidity, measured daily at 3pm (%)	t_{-1}	*nonlinear spline*		47.7	4	<0.001
	t_{-2}	*nonlinear spline*		39.0	4	<0.001
	t_{-3}	*nonlinear spline*		35.1	4	<0.001
	t_{-4}	*nonlinear spline*		71.6	4	<0.001
	t_{-5}	*nonlinear spline*		81.4	4	<0.001
Maximum daily air temperature (°C)	t_{-1}	*nonlinear spline*		58.4	4	<0.001
	t_{-2}	*nonlinear spline*		44.2	4	<0.001
	t_{-3}	*nonlinear spline*		64.1	4	<0.001
	t_{-4}	*nonlinear spline*		49.5	4	<0.001
	t_{-5}	*nonlinear spline*		47.2	4	<0.001
Midpoint daily air temperature (°C)	t_{-1}	-0.046	0.065	0.5	1	0.478
	t_{-2}	-0.019	0.065	0.1	1	0.775
	t_{-3}	-0.085	0.065	1.7	1	0.190
	t_{-4}	-0.070	0.067	1.1	1	0.297
	t_{-5}	-0.027	0.067	0.2	1	0.684
Minimum daily air temperature (°C)	t_{-1}	-0.040	0.031	1.6	1	0.204
	t_{-2}	-0.033	0.032	1.1	1	0.289
	t_{-3}	-0.068	0.032	4.6	1	0.031
	t_{-4}	-0.052	0.033	2.5	1	0.111
	t_{-5}	-0.027	0.032	0.7	1	0.395
Maximum daily wind speed (km hour^{-1})	t_{-1}	*nonlinear spline*		10.2	4	0.038
	t_{-2}	*nonlinear spline*		19.6	4	<0.001
	t_{-3}	*nonlinear spline*		52.0	4	<0.001
undirected[b]	t_{-4}	*nonlinear spline*		35.5	4	<0.001
	t_{-5}	*nonlinear spline*		14.9	4	0.005

[a] *P*-values derived from likelihood ratio tests (LRT) comparing univariable to null Cox regression models.
[b] Maximum daily wind speed based on wind from all directions ('undirected'), making no assumption concerning nearest infected premises assumption.

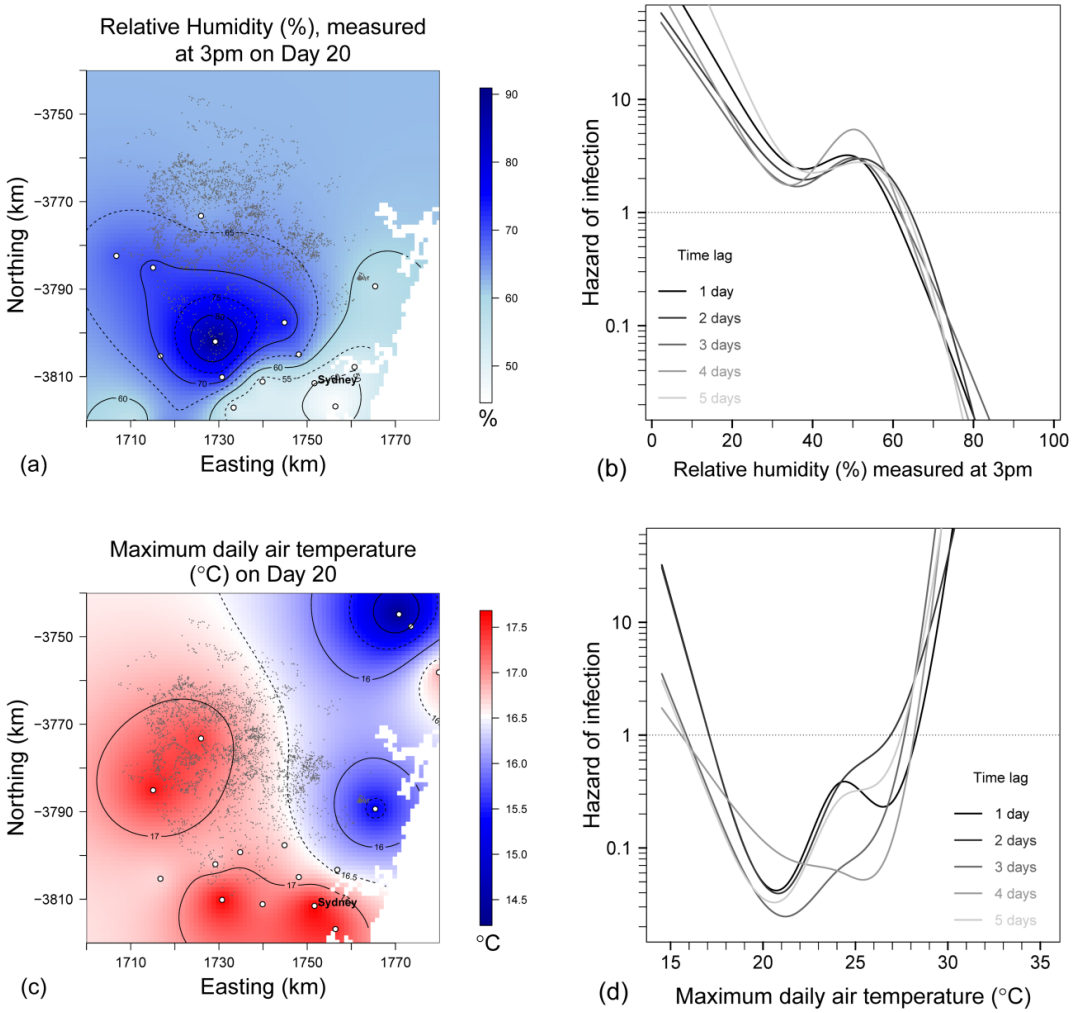


Figure 6.5 Daily meteorological data provided by Australian Bureau of Meteorology weather stations smoothed using kriging, and time-lagged by 1–5 days, to investigate the association with premises-level hazard of infection in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia.
Smoothed estimate of (a) relative humidity measured at 3pm on Day 20 of the outbreak. Small grey dots denote the horse premises. (b) Restricted cubic splines of the crude relationship between hazard of infection and relative humidity (3pm measurement) at time-lags of 1–5 days over the entire study period. (c) Smoothed daily maximum air temperature on Day 20 and (d) the relationship between daily maximum air temperature and hazard of infection. White closed circles represent the location of weather stations for which data was available.

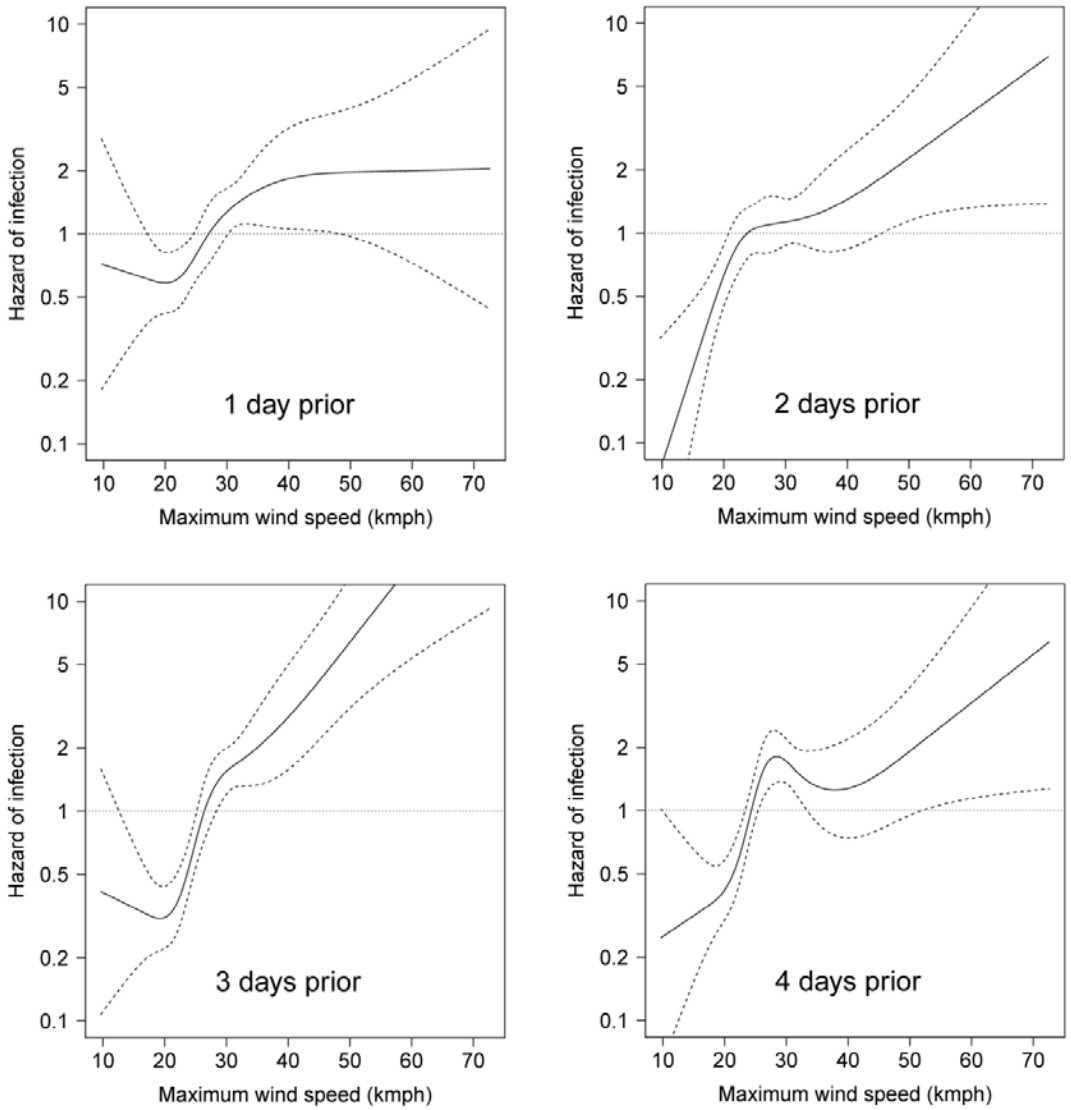


Figure 6.6 The crude relationships between hazard of infection and maximum daily wind speed *from all directions*, time-lagged by 1–4 days, in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia.
Estimates are based on hourly wind data from all directions. Dashed lines represent 95% confidence intervals.

Table 6.3 Univariable analysis of the association between directed wind speed covariates (time-changing and time-lagged) and time to infection of premises in the largest cluster (n = 3153), northwest of Sydney, during the 2007 equine influenza outbreak in Australia.

Meteorological Factor	Time-lag	Term	LRT	*df*	*P*-value[b]
Maximum daily wind speed (km hour^{-1}) *directed* ($k = 1$)[a]	t_{-1}	*nonlinear spline*	3.8	4	0.430
	t_{-2}	*nonlinear spline*	9.1	4	0.058
	t_{-3}	*nonlinear spline*	16.5	4	0.002
	t_{-4}	*nonlinear spline*	6.6	4	0.159
	t_{-5}	*nonlinear spline*	3.4	4	0.499
Maximum daily wind speed (km hour^{-1}) *directed* ($k = 2$)[a]	t_{-1}	*nonlinear spline*	14.0	4	0.007
	t_{-2}	*nonlinear spline*	25.3	4	<0.001
	t_{-3}	*nonlinear spline*	34.5	4	<0.001
	t_{-4}	*nonlinear spline*	8.2	4	0.083
	t_{-5}	*nonlinear spline*	24.6	4	<0.001
Maximum daily wind speed (km hour^{-1}) *directed* ($k = 3$)[a]	t_{-1}	*nonlinear spline*	41.2	4	<0.001
	t_{-2}	*nonlinear spline*	49.5	4	<0.001
	t_{-3}	*nonlinear spline*	75.6	4	<0.001
	t_{-4}	*nonlinear spline*	38.0	4	<0.001
	t_{-5}	*nonlinear spline*	52.3	4	<0.001

[a] Maximum daily wind speed ('directed') based on wind only from within 45° arcs centred on the direction of the k nearest infected premises for $k = 1,2,3$ (see Figure 6.2 for details) assuming that premises were infectious for 14 days and one of the nearest k infective premises was the source of infection.

[b] *P*-values derived from likelihood ratio tests (LRT) comparing univariable to null Cox regression models.

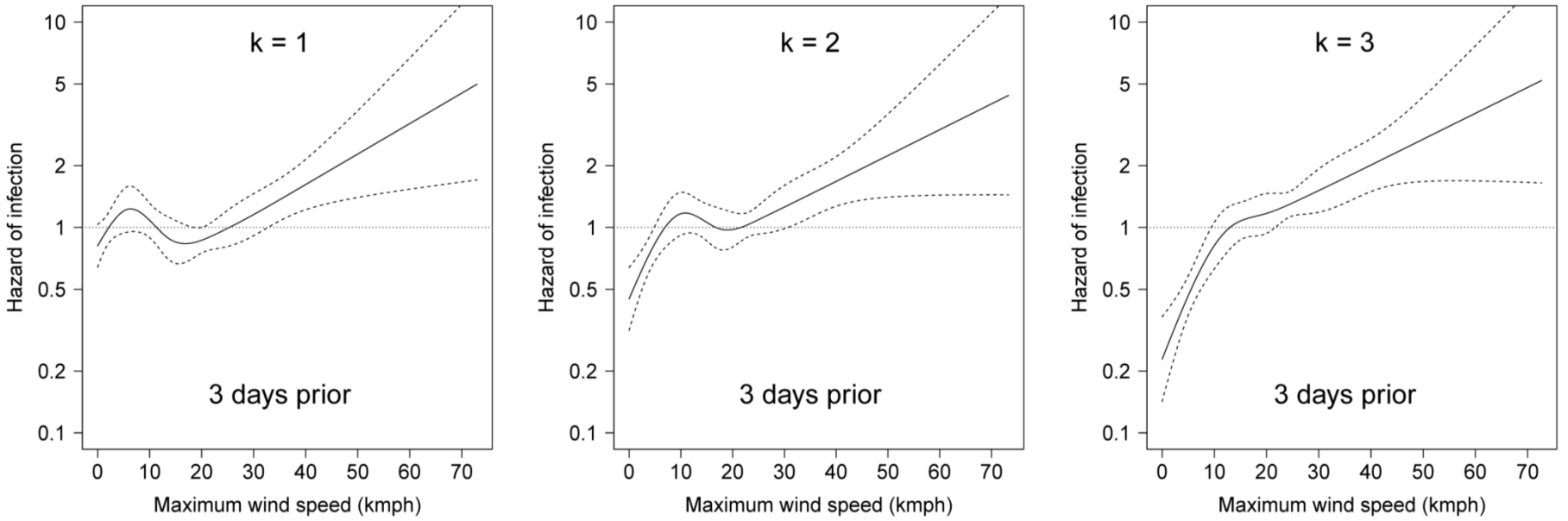


Figure 6.7 The crude relationships between hazard of infection and maximum daily wind speed selected from the direction of the *k* nearest infected premises (*k* = 1,2,3), time-lagged by 3 days, in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia.
Estimates are based only on hourly wind data from within 45^0 arcs centred on the direction of the *k* nearest infected premises (arcs may overlap if nearest infected premises are in the same direction, see Figure 6.2c for details). Dashed lines represent 95% confidence intervals.

Horse premises attributes and hazard of infection

Horse premises in the Northwest Sydney cluster were highly skewed in terms of their land area and the number of horses they held at the time of the outbreak. The median premises held 2 horses (IQR: 1, 5 horses; maximum: 139 horses) on 5.1 acres (IQR: 4.8, 15.2 acres; maximum: 2 225 acres). These variables were log transformed for all further analyses, with results back-transformed for presentation. Highly non-linear crude relationships were observed between hazard of infection and premises area and horse density (Table 6.4 and Figure 6.8a). Medium sized (4.8–15.2 acres) and medium density premises (1–5 acres per horse) were at increased risk of infection, as were horse premises that shared a fence with another horse premises. Hazard also increased with the number of horses held on a premises; this trend was well represented by categorisation based on quartiles.

A trend existed across the study area in terms of premises elevation and surrounding human population density. Hazard of infection was higher on horses premises located at lower elevations (<45 m) and >2.2 km from main roads (Table 6.4). Risk was also higher on horse premises located in peri-urban areas (human population densities between 1–500 people km^{-2}) compared to premises located either away from residential areas (human population density within 1 km = 0) or within urban areas (>500 people km^{-2}) (Figure 6.8b).

Premises area and premises horse density were the only highly correlated pairing (ρ = -0.74), amongst the premises attribute variables. Of these two covariates, premises area was the more strongly associated with the outcome. The following premises attribute variables were therefore included in multivariable analysis: splines of premises area and local human population density, number of horses, length of shared fence with other horse premises, premises elevation and distance to the nearest main road. Vaccination status was retained as it was considered an *a priori* confounder.

Table 6.4 Univariable analysis of non-meteorological covariates with time to infection of premises in the largest cluster (n = 3153), northwest of Sydney, during the 2007 equine influenza outbreak in Australia.

Factor	Category	No.	Hazard ratio	(95% CI)	*P*-value[a]
Premises attributes					
Area (acres)	>15.2	788	0.99	(0.85, 1.15)	<0.001
	5.1–15.2	788	1.94	(1.69, 2.23)	
	4.8–5.1	789	2.09	(1.83, 2.40)	
	<4.8	788	1.00		
Horse density (horses acre^{-1})	>1.00	776	1.50	(1.29, 1.74)	<0.001
	0.40–1.00	799	2.51	(2.18, 2.89)	
	0.20–0.40	787	1.85	(1.63, 2.17)	
	<0.20	791	1.00		
Number of horses	>5	662	3.28	(2.82, 3.82)	<0.001
	3–5	902	2.48	(2.14, 2.88)	
	2	787	2.08	(1.79, 2.43)	
	1	802	1.00		
Length of shared fence with other horse premises (m)	>300	742	1.45	(1.29, 1.63)	<0.001
	1–300	725	1.64	(1.47, 1.84)	
	0	1 686	1.00		
Vaccination status[b]	Yes	490	0.28	(0.04, 2.13)	0.137
	No	2 663	1.00		
Spatial covariates					
Elevation (m)	>115	785	0.72	(0.63, 0.82)	<0.001
	45–115	777	0.66	(0.58, 0.76)	
	25–45	786	1.02	(0.90, 1.15)	
	<25	805	1.00		
Human population density (people km^{-2})	>500	1 059	1.05	(0.94, 1.18)	<0.001
	1–500	954	1.29	(1.48, 1.44)	
	0	1 140	1.00		
Distance to nearest main road (km)	>2.2	787	1.23	(1.08, 1.41)	0.021
	1.1–2.2	789	1.14	(1.00, 1.31)	
	0.4–1.0	788	1.11	(0.97, 1.27)	
	<0.4	786	1.00		

[a] *P*-values derived from likelihood ratio (LRT) tests comparing univariable to null Cox regression models.
[b] Time-changing covariate.

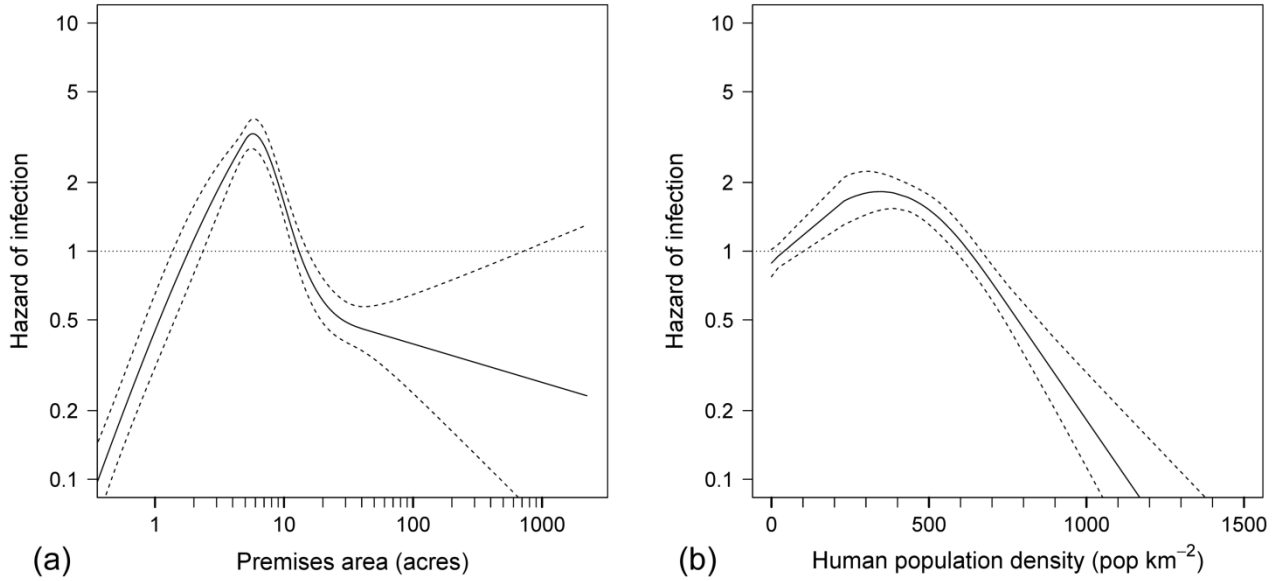


Figure 6.8 The crude relationships between hazard of infection and (a) premises area and (b) local human population density (people residing within approximately 1 km of the horse premises), in the largest cluster, northwest of Sydney, during the 2007 equine influenza outbreak in Australia..
Dashed lines represent 95% confidence intervals

6.3.2.2. Multivariable analysis

The final model is presented in Table 6.5. Two variables were eliminated during multivariable model-building: 'distance to nearest main road' and 'minimum daily air temperature'. No first order interaction terms were significant at $P<0.05$.

The shape of the restricted cubic splines representing the nonlinear relationships between hazard of infection and relative humidity, maximum daily air temperature, maximum daily wind speed (from the direction of the nearest three infected premises), premises area and human population density, were all largely unchanged from their crude forms (as presented in Figures 6.5–6.8). Post-adjustment, rainfall was detected to be weakly protective. The increased hazard amongst premises with higher numbers of horses persisted, as did the reduction in hazard amongst premises at higher elevations, with a 42% reduction in risk for every order of magnitude increase in elevation. Premises that were adjacent to another horse premises were at increased hazard of equine influenza infection.

Model goodness-of-fit and residual analysis

The final model accounted for a quarter of the variability in the data (Schemper and Stare pseudo-R^2 = 25.8%). No issues were identified based on inspection of martingale and deviance residuals, both overall, and when plotted against each variable included in the final model. Residual spatial structure was not evident in the empirical semivariogram of the deviance residuals, suggesting that spatial correlation was not unduly influencing our effect estimates (or their associated standard errors). Influence statistics identified only one important outlying premises, infected 36 days after the vaccination of the 2 horses on the property. These horses did not receive a second vaccination, whilst up to three doses may be required to attain protective immunity.

Table 6.5 Final multivariable Cox regression model for time to infection of premises in the largest cluster (n = 3153), northwest of Sydney, during the 2007 equine influenza outbreak in Australia.

Factor	Category	Hazard ratio	(95% CI)	*P*-value[a]
Meteorological covariates				
Rainfall, t_{-3}[b] (mm day^{-1})	*Linear*	0.91	(0.82, 1.00)	0.055
Relative humidity (%), measured daily at 3pm, t_{-5}[b]	*Nonlinear spline*	—	—	<0.001
Maximum daily air temperature, t_{-3}[b] (°C)	*Nonlinear spline*	—	—	<0.001
Maximum daily wind speed, t_{-3}[b], *directed* (k = 3)[c] (km hour^{-1})	*Nonlinear spline*	—	—	<0.001
Premises attributes				
Area (acres)	*Nonlinear spline*	—	—	<0.001
Number of horses	>5	3.16	(2.70, 3.69)	<0.001
	3–5	2.19	(1.89, 2.55)	
	2	1.93	(1.66, 2.26)	
	1	1.00		
Length of shared fence with other horse premises (m)	>300	1.30	(1.15, 1.48)	<0.001
	1–300	1.27	(1.13, 1.43)	
	0	1.00		
Vaccination status[b]	Yes	0.28	(0.04, 2.09)	0.134
	No	1.00		
Spatial covariates				
log10 (Elevation (m))	*Linear*	0.58	(0.51, 0.67)	<0.001
Human population density (people km^{-2})	*Nonlinear spline*	—	—	<0.001

Number of events = 1727, Log likelihood = −12,847.4; df = 25; $P < 0.001$; R^2 = 25.8%

[a] *P*-values derived from Likelihood ratio tests (LRT).

[b] Time-changing covariate, time-lagged 3 or 5 days as noted.

[c] Maximum daily wind speed ('directed') based on wind only from within 45° arcs centred on the direction of the three nearest infected premises assuming that premises were infectious for 14 days and one of the three nearest infective premises was the source of infection.

6.4. Discussion

To our knowledge, this empirical analysis provides the first estimates of the contribution of humidity, air temperature and wind to the spread of an actual outbreak of influenza ('in the field'). We have demonstrated that it is possible to detect an association between wind velocity and disease spread, and directly estimate the strength of such an association. This advances our understanding of the windborne spread of influenza from purely circumstantial association to a hypothesis statistically-tested with empirical data.

Relative humidity and influenza spread

Our analysis shows that influenza spread in this cluster was highly dependent on relative humidity. Recent reviews (Tellier, 2006, Tellier, 2009, Weber and Stilianakis, 2008) present contradictory results from laboratory trials of influenza A virus survival at intermediate humidities (Hemmes et al., 1960, Schaffer et al., 1976), and disagreement concerning the importance of aerosol transmission. The negative cubic relationship that we observed between hazard of infection and relative humidity provides field validation for some of these laboratory trials. The curve in Figure 6.5b across the whole range of relative humidities observed under natural conditions, exactly complements the results presented by Hemmes et. al. (Hemmes et al., 1960) of inactivation of aerosolised influenza A virus under controlled conditions. Our findings also support the theory presented by Lowen et al. that the relationship between influenza transmission and relative humidity is mediated by both virion and aerosol droplet nuclei stability (Lowen et al., 2007). In cool dry conditions, droplets are desiccated and remain small, which may stabilise influenza aerosols and facilitate longer range transmission, whereas at high relative humidity, the droplets absorb water and settle (Tellier, 2006). The small rise in hazard of infection at intermediate relative humidities (40–60%) is perhaps due to a summation of two effects: as relative humidity increases within this range so too does viral survival (Schaffer et al., 1976), whilst droplet nuclei settle more readily. This rise was most pronounced in the spline of relative humidity time-lagged by 3 days, yet the 5 day time-lagged variable was

the predictor in the group with the strongest statistical association with hazard of infection. Amongst all other groupings of autocorrelated meteorological variables, a time-lag of 3 days was the predictor with the strongest statistical association with hazard of infection, corresponding closely with the typical 1–3 day incubation period of equine influenza.

Recent research has suggested that in certain situations absolute humidity may better represent the relationship between air humidity and influenza A virus survival (McDevitt et al., 2010) and aerosol transmission (Shaman and Kohn, 2009). However, the dependency is perhaps more complex (Steel et al., 2011), because the amount of water vapour that air can hold increases with temperature. Absolute and relative humidity are related metrics for the amount of water vapour in moist air. Absolute humidity is the mass of water vapour per cubic meter of total moist air, whereas, relative humidity is absolute humidity expressed as a percentage of the amount of water vapour needed for saturation at a specific temperature. We used relative humidity rather than absolute humidity because: the relative humidity data were more complete over the study period, corresponding 9am and 3pm air temperature data were not available for all data points so back-transformation of relative humidity measurements into absolute humidity would have resulted in less complete data, and we wanted to ensure we could directly compare our results with the original research describing the dependency between relative humidity, air temperature and influenza virus transmission and survival (Hemmes et al., 1960, Lowen et al., 2007).

Air temperature and influenza spread

The shape of the highly nonlinear relationship that we observed between hazard of equine influenza infection and maximum daily air temperature suggests two mechanisms of influenza transmission. Hazard was lowest on days when the maximum air temperature was between 20–25 °C, and greatly increased on days with lower and higher maximum temperatures. Aerosol transmission of influenza A viruses has been shown to be enhanced in cooler conditions (Lowen et al., 2007), and on days when maximum daily temperature was <20 °C the air temperature

would be expected to remain in the optimal range for aerosol transmission equine influenza for longer. The marked increase in hazard of infection when maximum temperature was >25 °C is also consistent with recent research. Whilst high temperatures block aerosol transmission of influenza A viruses, the success of animal-to-animal contact transmission is unchanged at high temperatures (Lowen et al., 2008) perhaps explaining the spread of influenza in warm tropical environments.

Wind velocity and influenza spread

There is some consensus in the literature that airborne transmission of influenza is at least possible; however, there is strong disagreement about its importance (Weber and Stilianakis, 2008). In the cluster investigated, we observed an association between hazard of infection and increasing wind speed from the direction of nearby potential sources of equine influenza infection. A similar association was found when wind speed covariates were generated without making any directional assumptions, in effect testing the general hypothesis that hazard of infection was increased on days with increased wind speed (*from any direction*). Irrespective of our approach, wind speeds >30 km hour^{-1}, lagged by 3 days, were consistently associated with increased hazard of infection.

In developing proxy covariates for the directed formulation of the daily wind speed covariates ('WIND_SPD$_{dir(k)}$') certain assumptions were required. Wind data was only included if it was within a 45° arc of the nearest *k* infected premises to each uninfected premises on each observation day. The nearest *k* infected premises were assumed to be the only windborne source of equine influenza virus for a susceptible premises, and we assumed the duration of infectivity at the premises level was 14 days for all premises. The statistical strength of association between hazard of infection and wind speed increased as more nearest neighbours were cumulatively incorporated into the method of wind covariate generation, suggesting that the nearest neighbour was not always the only source of windborne infection. It would be computationally intensive to continue incorporating further nearest neighbours into this method,

so we cannot definitively state how far to extend this process. There must exist a point after which adding infected neighbours to the process of generating wind speed covariates results in weaker associations, as the associations were statistically stronger when wind speed covariates were generated with three nearest neighbours than when we made no directional assumptions. We can state that incorporating three nearest infected neighbours is better than only one or two, that our findings are relatively robust to the assumptions that we made whilst generating directed wind speed covariates, and that the detected association between hazard of infection and wind speed appears related to the direction of proximate infected premises.

The associations that we have detected between increasing wind speed and hazard of infection need to be interpreted in the context of our study design. There is potential for ecological fallacy in aggregated data analyses such as this, in which the unit of interest is not an individual animal but a group. Furthermore, it is not possible in such observational epidemiological analyses to definitively identify windborne spread from any other transmission route (direct contact, cough droplet and spread on fomites). Nonetheless, the detected association, presumably representing windborne spread of equine influenza, is biologically plausible, and its increasing strength with increasing wind speed from the direction of nearby infected premises is difficult to explain by spread through other means alone.

At wind speeds of >30 km hour^{-1} an aerosol of influenza droplet nuclei would only need to be stable for minutes to be able to infect horses on nearby premises. Equine influenza viruses have been shown to survive for periods of hours to days in soil and water, even in direct sunlight (Yadav et al., 1993), and infected horses shed large amounts of virus (>10^3 EID_{50}/ml of swab extract) throughout the roughly 7 days that they are infectious (Wood et al., 2007). Infection is more reliably achieved by inhalation of aerosolised virus that intranasal inoculation (Mumford et al., 1990), with a minimum infective dose of 10^2 EID_{50}/ml. We therefore consider it plausible that infected horses on one premises could cough or otherwise produce a sufficient quantity of aerosolised equine influenza virus, which after travelling wind-assisted could constitute an

infectious dose for a horse on a nearby premises (whether inhaled immediately or after surviving a short period on soil or in drinking water).

A recent time-series analysis investigated correlation between the frequency of paediatric influenza A hospital admissions and several meteorological variables including wind velocity (du Prel et al., 2009). A statistically significant univariable association was observed between increasing wind velocity and increased influenza A hospital admissions, in data collected from one hospital and one weather station (du Prel et al., 2009). However, in multivariable analyses no association was observed between wind velocity and influenza A hospitalisations, perhaps due to the level of spatial and temporal data aggregation (across the hospital catchment and into 14-day time intervals). Aggregated analyses of the association of meteorological factors with the spread of the severe acute respiratory syndrome (SARS) in Beijing (Yuan et al., 2006), and hand, foot and mouth disease (HFMD) of humans in Hong Kong (Ma et al., 2010), have found statistically significant associations with increasing wind velocity, albeit at much lower wind velocities.

Atmospheric dispersal modelling of the picornavirus that causes foot-and-mouth disease (FMD) in cloven-hoofed ungulates has consistently found that the virus is likely to be dispersed even in calm conditions (Garner and Cannon, 1995, Gloster et al., 2011, Gloster et al., 2010). The dependency of influenza A virus survival on relative humidity (Hemmes et al., 1960, Lowen et al., 2007) is completely different to that of poliovirus (Hemmes et al., 1960), HFMD (McGeady et al., 1979) and FMD virus (Garner and Cannon, 1995) (all much smaller non-enveloped RNA viruses of the family *Picornaviridae*). Therefore, it is perhaps not unexpected that the survival of aerosols of influenza A viruses, which are enveloped, and smaller picornaviruses could depend on different wind conditions.

When interpolating meteorological covariates and estimating nearest neighbour distances we used centroids to reduce the complexity of the analytical methods. For >99% of the premises in

our dataset we estimate that the maximum distance between the centroid and premises boundary was <500 m. When the largest 1% of premises (in area) were excluded from the final model, the only regression coefficients to change by >20% were the two highest order spline components for relative humidity and maximum daily temperature, and these changes were not discernible in post-adjustment plots. We therefore consider our findings to be insensitive to measurement bias introduced by representing premises by their centroids.

Environmental variables capable of influencing airborne disease spread (such as local horse density, tree density or terrain undulation) vary considerably in the different regions and clusters of premises infected during the 2007 equine influenza outbreak in Australia. A potential limitation of this analysis was that we focussed on only one cluster (the largest and most dense cluster in terms of population at risk) from a very large outbreak. There were two considered reasons for our detailed focus: counting process survival analysis involves analysing a very large dataset (204,909 observations on 3153 premises); and owing to a wide variance in local environmental characteristics and potential for differences in disease transmission dynamics, mixing clusters in the same analysis might dilute any meaningful results. Before generalising our findings to the whole outbreak, or indeed other outbreaks, follow-up research to assess the importance of the risk factors investigated in broadly dissimilar environments, is therefore required.

A classical geostatistical approach (Matheron, 1963) (kriging based on a least squares fit of an empirical variogram) was applied to interpolate premises-level meteorological covariates from weather station data. A more sophisticated model-based geostatistical approach (Diggle and Ribeiro, 2007) (maximum-likelihood based model fitting that does not rely on an empirical variogram) may be more appropriate. However, we considered that the numerous fine adjustments required to undertake a model-based approach to be impractical for fitting the 5616 separate models (117 days × 24 hours × 2 wind component vectors) that were required to produce hourly wind vector estimates at each of the 3153 individual premises locations across

the entire study period. It was also not possible to assess the assumptions of stationarity or isotropy for each of the thousands of semivariogram models required to generate all of the daily meteorological covariates for each premises. These assumptions appeared justified based on semivariograms of the mean conditions for each meteorological covariate over the entire study period. Our interpolation approach could have been refined by incorporating elevation (regression kriging), a spatial trend or even anisotropy into the method. The study extent covered the northern half of the Sydney basin, which is relatively flat, and bounded by a plateau of national parks where horses are prohibited. These refinements would be recommended when conducting further similar research of clusters located in more varied terrain.

In the cluster of infection investigated, disease did not appear to spread predominantly in any single direction. We purposefully focussed on this cluster rather than other large clusters in which a single global direction of spread has been noted (Davis et al., 2009), with the intention of estimating the typical contribution of wind to disease spread rather than circumstantially associating prevailing wind with the global direction of disease spread. In any cluster in which an overall direction of spread is detected, an important further research question remains: What proportion of this anisotropic spread is directly attributable to windborne disease spread? Our methods provide a means to answer this research question, and to retrospectively investigate the contribution of windborne aerosol spread to local disease spread during outbreaks such as the foot-and-mouth disease outbreak in the United Kingdom in 2001.

By restricting this analysis to a study period after the horse movement ban was put in place, we focussed this study on factors influencing the local spread of equine influenza. We also adjusted for a number of relevant confounders of the meteorological associations we aimed to estimate: vaccination status of horses on the premises, premises size (in terms of area and number of horses), whether premises were adjacent to another premises holding horses, and local human population density. A small misclassification bias is known to be present in the equine influenza dataset, due to under-reporting of infected premises by owners either attempting to avoid

movement restrictions or who failed to detect infection (Dhand and Sergeant, 2011a). A previous analysis found <1% under-reporting occurred in this region, suggesting that <13 infected premises were misclassified as uninfected (Dhand and Sergeant, 2011a); we considered this bias negligible.

6.5. Conclusions

By combining influenza outbreak and concurrent meteorological data, we have shown how relative humidity, air temperature and wind velocity combined to influence the spread of an actual influenza outbreak. Hazard of equine influenza infection was higher when relative humidity was <60% and lowest on days when daily maximum air temperature was 20–25 °C. Wind speeds >30 km hour^{-1} from the direction of nearby infected premises were associated with increased hazard of infection. Our analysis supports, and extends, the findings of studies into influenza A transmission conducted under controlled conditions. The relationships described are of direct importance for managing disease risk during influenza outbreaks in horses, and more generally, advance our understanding of the transmission of influenza A viruses under natural conditions.

6.6. Appendix

Table 6.6 Example of survival dataset formulations for (a) time-independent and (b) time-dependent ('counting formulation') Cox regression modelling of factors associated with time to infection of premises in the largest cluster of the 2007 equine influenza outbreak in Australia.

INFECTED PREMISES (unvaccinated)

(a) ID	IP	DAY	VACC	VACC_DAYS	HORSE_N	AREA	CENT_DIST	CENT_DIR	ROAD_DIST	ELEV	WIND_ST
0001	1	55	0	–	11	84.3	21.0	36.0	0.6	26	8

(b) ID	START	STOP	IP	VACC_TD	HORSE_N	AREA	IP_DIST	IP_DIR	WIND_SPD (KMPH)			TEMP_MIN (°C)		
									t_{-1}	t_{-2}	t_{-3}	t_{-1}	t_{-2}	t_{-3}
0001	14	15	0	0	11	84.3	8.45	102.4	0	17	0	13.4	10.9	9.1
0001	15	16	0	0	11	84.3	8.45	102.4	8	0	17	14.1	13.4	10.9
0001	16	17	0	0	11	84.3	7.74	77.9	17	8	0	8.7	14.1	13.4
0001	17	18	0	0	11	84.3	7.74	77.9	11	17	8	12.2	8.7	14.1
0001	18	19	0	0	11	84.3	7.74	77.9	15	11	17	12.6	12.2	8.7
….	….	….	….	….	….	….	….	….	….	….	….	….	….	….
0001	50	51	0	0	11	84.3	0.18	299.1	9	17	46	8.0	13.1	10.9
0001	51	52	0	0	11	84.3	0.18	299.1	31	9	17	11.0	8.0	13.1
0001	52	53	0	0	11	84.3	0.18	299.1	21	31	9	12.6	11.0	8.0
0001	53	54	0	0	11	84.3	0.18	299.1	0	21	31	12.4	12.6	11.0
0001	54	55	1	0	11	84.3	0.18	299.1	5	0	21	11.6	12.4	12.6

IP = premises infection status, DAY = date of onset, VACC = vaccination status (Yes =1), VACC_DAY = day of vaccination, VACC_TD = vaccination status (time-dependent variable), HORSE_N = number of horses, CENT_DIST/DIR = distance (km) and direction (degrees) from cluster centre, IP_DIST/DIR = distance (km) and direction (degrees) from nearest potential source premises, ROAD_DIST = distance (km) from nearest main road, WIND_ST = ID of nearest weather station, WIND_SPD = time-lagged maximum wind speed from within 45° of the direction of the nearest infected premises, TEMP_MIN = time-lagged minimum surface air temperature.

Table 6.7 Example of survival dataset formulations for (a) time-independent and (b) time-dependent ('counting formulation') Cox regression modelling of factors associated with time to infection of premises in the largest cluster of the 2007 equine influenza outbreak in Australia.

UNINFECTED PREMISES (vaccinated)

(a) ID	IP	DAY	VACC	VACC_DAYS	HORSE_N	AREA	CENT_DIST	CENT_DIR	ROAD_DIST	ELEV	WIND_ST
0002	0	131	1	77	2	41.5	21.5	79.0	1.8	199	8

(b) ID	START	STOP	IP	VACC_TD	HORSE_N	AREA	IP_DIST	IP_DIR	WIND_SPD (KMPH)			TEMP_MIN (°C)		
									t_{-1}	t_{-2}	t_{-3}	t_{-1}	t_{-2}	t_{-3}
0002	14	15	0	0	2	41.5	10.96	59.8	13	5	11	14.1	7.5	6.5
0002	15	16	0	0	2	41.5	10.96	59.8	0	13	5	10.7	14.1	7.5
0002	16	17	0	0	2	41.5	11.52	67.2	17	8	11	6.7	10.7	14.1
0002	17	18	0	0	2	41.5	10.39	65.5	13	17	8	13.2	6.7	10.7
….	….	….	….	….	….	….	….	….	….	….	….	….	….	….
0002	75	76	0	0	2	41.5	5.44	111.5	21	0	0	11.0	19.0	13.2
0002	76	77	0	0	2	41.5	5.44	111.5	15	21	0	13.8	11.0	19.0
0002	77	78	0	1	2	41.5	5.44	111.5	22	15	21	14.7	13.8	11.0
….	….	….	….	….	….	….	….	….	….	….	….	….	….	….
0002	127	128	0	1	2	41.5	3.85	70.6	24	8	17	19.4	17.8	16.4
0002	128	129	0	1	2	41.5	3.85	70.6	18	24	8	22.1	19.4	17.8
0002	129	130	0	1	2	41.5	3.85	70.6	26	18	24	12.1	22.1	19.4
0002	130	131	0	1	2	41.5	30.01	97.7	18	26	0	16.0	12.1	22.1

IP = premises infection status, DAY = date of onset, VACC = vaccination status (Yes =1), VACC_DAY = day of vaccination, VACC_TD = vaccination status (time-dependent variable), HORSE_N = number of horses, CENT_DIST/DIR = distance (km) and direction (degrees) from cluster centre, IP_DIST/DIR = distance (km) and direction (degrees) from nearest potential source premises, ROAD_DIST = distance (km) from nearest main road, WIND_ST = ID of nearest weather station, WIND_SPD = time-lagged maximum wind speed from within 45° of the direction of the nearest infected premises, TEMP_MIN = time-lagged minimum surface air temperature.

Table 6.8 Correlations between continuous explanatory variables (time-independent and time-changing) analysed in Cox regression for association with time to infection of premises in the largest cluster of the 2007 equine influenza outbreak in Australia.

Covariate	Lag	Spearmans Correlation Coefficients																															
		High correlation ρ > |0.70|																															
		Moderate correlation |0.50| < ρ ≤ |0.70|																															
		Low correlation |0.30| < ρ ≤ |0.50|																															
RAIN	-1	1.00																															
	-2	0.61	1.00																														
	-3	0.47	0.63	1.00																													
	-4	0.40	0.48	0.62	1.00																												
	-5	0.31	0.39	0.48	0.63	1.00																											
RH_9AM	-1	0.61	0.51	0.43	0.36	0.31	1.00																										
	-2	0.60	0.58	0.49	0.43	0.36	0.64	1.00																									
	-3	0.46	0.60	0.59	0.50	0.43	0.42	0.62	1.00																								
	-4	0.33	0.47	0.59	0.58	0.51	0.36	0.42	0.64	1.00																							
	-5	0.17	0.29	0.42	0.56	0.57	0.30	0.37	0.40	0.61	1.00																						
RH_3PM	-1	0.61	0.47	0.35	0.28	0.21	0.75	0.48	0.33	0.21	0.17	1.00																					
	-2	0.69	0.58	0.45	0.34	0.28	0.72	0.75	0.48	0.32	0.23	0.62	1.00																				
	-3	0.57	0.69	0.59	0.46	0.34	0.52	0.69	0.75	0.48	0.28	0.47	0.60	1.00																			
	-4	0.46	0.59	0.69	0.59	0.45	0.41	0.50	0.70	0.74	0.43	0.35	0.45	0.61	1.00																		
	-5	0.33	0.47	0.59	0.69	0.59	0.34	0.39	0.50	0.70	0.70	0.19	0.33	0.46	0.62	1.00																	
TEMP_MAX	-1	-0.30	-0.21	-0.11	-0.01	0.05	-0.18	-0.10	-0.05	0.03	0.11	-0.47	-0.22	-0.15	-0.11	0.04	1.00																
	-2	-0.27	-0.28	-0.20	-0.10	-0.01	-0.22	-0.19	-0.09	-0.05	0.02	-0.18	-0.47	-0.21	-0.13	-0.10	0.54	1.00															
	-3	-0.17	-0.28	-0.29	-0.20	-0.10	-0.15	-0.21	-0.19	-0.10	-0.03	-0.11	-0.18	-0.48	-0.22	-0.14	0.34	0.53	1.00														
	-4	-0.11	-0.19	-0.29	-0.30	-0.20	-0.06	-0.13	-0.20	-0.20	-0.08	-0.04	-0.10	-0.19	-0.49	-0.23	0.35	0.33	0.54	1.00													
	-5	-0.04	-0.10	-0.19	-0.31	-0.30	-0.03	-0.06	-0.12	-0.21	-0.21	0.12	-0.04	-0.09	-0.20	-0.50	0.23	0.35	0.32	0.52	1.00												
TEMP_MIN	-1	0.31	0.21	0.19	0.18	0.21	0.44	0.44	0.29	0.23	0.32	0.40	0.36	0.31	0.22	0.20	0.39	0.47	0.38	0.38	0.41	1.00											
	-2	0.37	0.32	0.21	0.19	0.19	0.38	0.43	0.45	0.29	0.19	0.39	0.39	0.38	0.31	0.22	0.20	0.39	0.46	0.36	0.38	0.73	1.00										
	-3	0.38	0.38	0.34	0.22	0.19	0.37	0.36	0.43	0.46	0.25	0.38	0.37	0.40	0.39	0.32	0.20	0.21	0.37	0.44	0.36	0.61	0.73	1.00									
	-4	0.39	0.40	0.39	0.33	0.23	0.41	0.35	0.38	0.43	0.41	0.42	0.36	0.39	0.41	0.40	0.23	0.21	0.19	0.34	0.42	0.65	0.61	0.75	1.00								
	-5	0.39	0.41	0.41	0.39	0.34	0.47	0.39	0.37	0.38	0.39	0.45	0.40	0.38	0.40	0.42	0.21	0.24	0.20	0.16	0.32	0.67	0.65	0.62	0.74	1.00							
WIND_SPD	-1	-0.03	-0.04	-0.03	-0.02	0.00	-0.08	-0.04	-0.04	-0.04	-0.02	-0.11	-0.10	-0.05	-0.06	-0.04	0.13	0.15	0.06	0.05	0.05	0.04	0.02	0.00	-0.02	-0.02	1.00						
	-2	-0.03	-0.03	-0.03	-0.02	-0.02	-0.06	-0.08	-0.04	-0.04	-0.04	-0.04	-0.11	-0.10	-0.05	-0.05	0.02	0.13	0.15	0.05	0.05	-0.01	0.04	0.01	-0.01	-0.02	0.24	1.00					
	-3	-0.03	-0.03	-0.02	-0.04	-0.02	-0.04	-0.07	-0.08	-0.04	-0.04	-0.05	-0.04	-0.11	-0.10	-0.05	0.05	0.01	0.12	0.13	0.04	-0.01	-0.02	0.03	0.01	-0.01	0.15	0.23	1.00				
	-4	-0.03	-0.02	-0.03	-0.03	-0.03	-0.04	-0.04	-0.06	-0.08	-0.05	-0.01	-0.06	-0.04	-0.10	-0.09	0.03	0.05	0.01	0.11	0.13	0.01	-0.01	-0.02	0.03	0.01	0.19	0.15	0.24	1.00			
	-5	0.01	-0.01	-0.01	-0.01	-0.03	0.01	-0.04	-0.04	-0.05	-0.09	0.02	-0.01	-0.04	-0.03	-0.09	0.02	0.04	0.04	0.00	0.11	0.03	0.03	0.01	0.00	0.05	0.09	0.19	0.16	0.25	1.00		
	Lag	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5	-1	-2	-3	-4	-5		
	Covariate	RAIN					RH_9AM					RH_3PM					TEMP_MAX					TEMP_MIN					WIND_SPD						

AREA	1.000						
HORSE_DENSITY	-0.740	1.000					
HORSES_NUMBER	0.260	0.340	1.000				
ELEV	0.010	-0.060	-0.090	1.000			
HUMAN_DENS	-0.450	0.370	-0.110	0.090	1.000		
ROAD_DIST	0.180	-0.160	0.060	-0.020	-0.070	1.000	
	AREA	HORSE_DENSITY	HORSES_NUMBER	ELEV	HUMAN_DENS	ROAD_DIST	

AREA = premises area, HORSE_DENSITY = horses acre^{-1}, HORSE_NUMBER = number of horses, ELEV = premises elevation, HUMAN_DENS = population within 1 km of premises centroid, ROAD_DIST = distance from nearest main road, RAIN = rainfall, RH_9AM/RH_3PM = Relative humidity measured at 9am/3pm daily, TEMP_MAX/TEMP_MIN = maximum/minimum air temperature, WIND_SPD = maximum wind speed from within 45° of the direction of the nearest infected premises.

Chapter 7: Development of a dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia

"*Essentially, all models are wrong, but some are useful.*"

George E. P. Box, Professor of Statistics at the University of Wisconsin.

In: Box, G.E.P, Draper N. (1987), 'Empirical Model-Building and Response Surfaces', John Wiley & Sons Canada, Ltd. p.424.

This chapter is a manuscript in preparation for submission for peer-review:

Firestone, S.M., Jewell, C.P., Garner, M.G., Moloney, B.J., Ward, M.P., Dhand, N.K., *unpublished manuscript*. Development of a dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia.

7. Development of a dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia

7.1. Introduction

When a transboundary animal disease enters a highly susceptible population, the course and eventual outcome of the ensuing outbreak depends critically on the timeliness and appropriateness of outbreak management interventions. At the time of disease detection, predictive models have been used to augment epidemiological investigations and aid the development of appropriate outbreak control strategies (Kao, 2002, Jewell et al., 2009a). To be useful, these decision-support tools need to accommodate chance events and large uncertainties in the data available early in such outbreaks, present realistic and biologically plausible predictions and be validated against data not used in their construction (Kitching et al., 2006, Taylor, 2003). Otherwise, as occurred during the 2001 outbreak of foot-and-mouth disease (FMD) in the United Kingdom, controversy will surround the usefulness of these models and the appropriateness of any decisions informed by their predictions (Kao, 2002, Kitching et al., 2006).

Equine influenza is a highly contagious respiratory disease of horses, characterised in naïve horses by a harsh dry cough, pyrexia, lethargy, anorexia and occasionally a nasal discharge (Myers and Wilson, 2006). Although most horses recover uneventfully, large outbreaks have occurred when highly susceptible horse populations have been exposed to novel viral strains (Daly et al., 2011), causing substantial economic and social impacts (Callinan, 2008). In August 2007, following a breach in the quarantine of sub-clinically infected imported horses (Callinan, 2008), Australia experienced its first ever outbreak of equine influenza. Only small numbers of horses imported for breeding had been recently vaccinated, leaving almost the entire horse population susceptible. Following this incursion, an outbreak spread rapidly via the movement of infected horses to and from equestrian events prior to disease detection, and through local

spread within 5–15 km of infected premises (Firestone et al., 2012a, Firestone et al., 2011a, Cowled et al., 2009b, Davis et al., 2009). Eradication was achieved in less than 5 months, through a combination of horse movement restrictions, on-farm biosecurity measures and targeted vaccination (Callinan, 2008, Glanville and Christie, 2011), but not before around 67,000 horses were infected on 9359 premises.

Animal health authorities responsible for decision-making during the 2007 outbreak of equine influenza in Australia have acknowledged that the size and extent of the outbreak was by far in excess of initial estimates (Glanville et al., 2009). Early in the response, when important strategic decisions were being made, data on the population at risk was extremely limited and authorities were only aware of three infected premises, when it has been estimated that over 80 premises had already been infected (Garner et al., 2011b). These uncertainties undermined confidence during critical stages of the containment and eradication program (Glanville et al., 2009).

Simulation models of epidemic processes in domesticated animal populations rely on three types of data: contact-tracing data on animal movements, population at risk data and probability distributions of key parameters used to represent the epidemic process. Epidemiological investigations in the early epidemic period target the known uncertainties in such data in an attempt to answer key questions such as: where has infection come from; where have infected animals moved; where might the disease spread; what processes are involved and which groups are most at risk? A key report into the application of models to inform decision-making during the 2001 FMD outbreak in the UK cautioned against the use of predictive models in decision-support (Taylor, 2003), recommending that the focus should be on retrospective analysis of epidemics. Where such models were used in the face of an outbreak, it was suggested that they should only be used to augment (rather than replace) detailed epidemiological investigations by providing comparisons of observed and modelled outcomes for the purpose of alerting epidemiologists to unexpected events, and aiding the targeting of interventions (Taylor, 2003).

Recent research into real-time Bayesian statistical inference of epidemics has provided novel methods for predicting undetected infections (Jewell et al., 2009a) and risk of infection (Jewell et al., 2009b). The benefit of this approach over previous predictive models is that in the face of an outbreak such models can utilise field data from an ongoing outbreak, and incorporate these with prior estimates, expert opinion, contact-tracing and other animal premises data into a single analytical framework which can be used to develop predictions on the course of the epidemic, including description of the level of uncertainty in these estimates (Jewell et al., 2009c).

Previous modelling of the 2007 equine influenza outbreak in Australia has been conducted, and validated, at the premises level with the specific strategic purpose of recreating the spatial and temporal patterns of spread to enable retrospective evaluation of different control strategies (Garner et al., 2011b). Outbreaks of equine influenza at horse racing facilities in the United States, United Kingdom and Japan have been modelled at the individual horse level, to specifically describe intra-premises epidemiological processes and the effects of vaccination on disease spread (Satou and Nishiura, 2006, Glass et al., 2002, de la Rua-Domenech et al., 2000), with extensions to incorporate the influences of seasonal changes in the population at risk (Park et al., 2003), and antigenic drift in vaccinated horse populations (Park et al., 2004). A larger more general metapopulation model of between yard spread of equine influenza has also been developed based on the 2003 outbreak at the Newmarket facility in the United Kingdom (Baguelin et al., 2010).

The aim of this analysis was to develop a stochastic epidemic model of equine influenza in Australia to provide a dynamic framework for the rapid assessment of future outbreaks of equine influenza in Australia. The purpose of the model was not to replicate the 2007 outbreak exactly, but rather to develop informative prior probability distributions to inform future Bayesian predictive risk-mapping, thereby making most use of the type of uncertain data anticipated at the time of detection of such an emergency animal disease event. Inference of key

epidemiological parameters was also intended to advance general understanding of the process underlying the spread of equine influenza under Australia's unique environmental conditions.

7.2. Materials and methods

7.2.1. Study area

This study was conducted on three highly affected regions of the 2007 equine influenza outbreak in Australia (Figure 7.1); two primary study regions (Greater Sydney and the Hunter Valley) were used for Bayesian inference and simulation, and a third region (Tamworth) was used only for validation purposes. These regions were selected because the peri-urban Greater Sydney region was the most dense in terms of horse premises at the time of the 2007 outbreak and included nearly a third of all the premises infected during the 2007 outbreak (Table 7.1). Once horse movement bans were completely implemented (on 26 August 2007, day 10 of the outbreak), this region was effectively isolated, being bounded by National parks (where horses are prohibited) on three sides and by metropolitan Sydney to the East. The semi-rural Hunter Valley is a major centre of Thoroughbred horse breeding in Australia, and although of similar spatial extent to the Greater Sydney region, this cluster included a smaller number of horse premises than in the Greater Sydney region; these premises were on average larger in area and held greater numbers of horses at the time of the 2007 outbreak (Table 7.1). Tamworth is a rural city with many surrounding recreational and commercial horse premises, similar in area to those in the Hunter Valley, but on average holding fewer horses.

7.2.2. Data collection

The State governments of New South Wales (NSW) and Queensland (QLD) provided premises-level covariate and contact-tracing data collected at the time of the 2007 equine influenza outbreak. The collation and cleaning of these data in preparation for epidemiological analysis is described in detail elsewhere (Cowled et al., 2009b, Garner et al., 2011b, Firestone et al., 2012a). These data are considered a quasi-census of the population at risk in outbreak affected

areas in 2007 (Garner et al., 2011b, Moloney et al., 2011), and contained the following covariates: address, geocoded coordinates (based on property centroid), number of horses, premises area, vaccination status (and date of vaccination for NSW data). Infected premises data also included the estimated date of onset of first clinical signs of the first horse affected ('onset date'), and were linked to laboratory testing records.

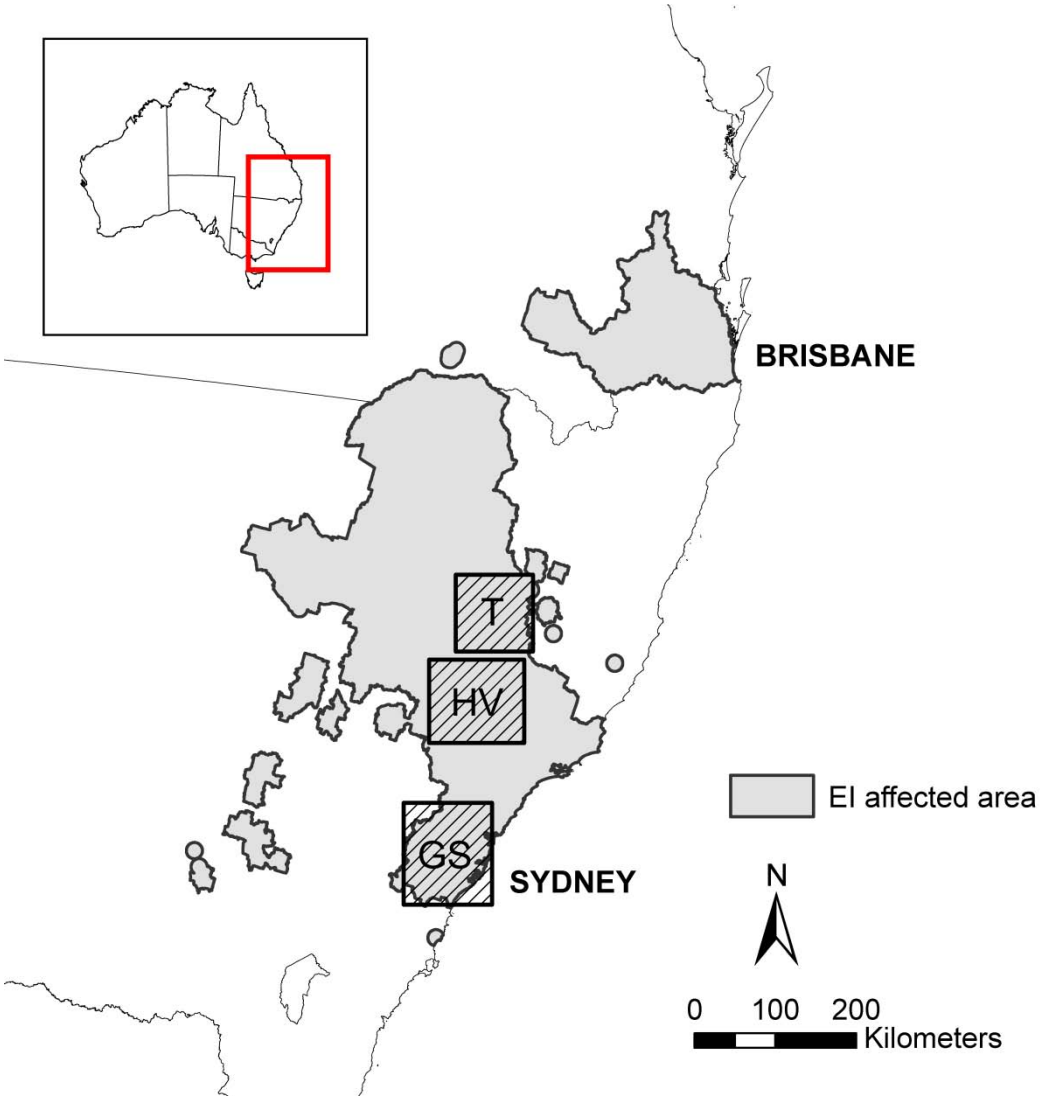


Figure 7.1 Regions modelled in a Bayesian SIR simulation of equine influenza outbreaks in Australia. Parameter distributions were developed by a Bayesian reversible-jump Markov chain Monte-Carlo algorithm trained on two highly affected regions in the 2007 outbreak.
GS = Greater Sydney region (n=7 670 horse premises) and HV = Hunter Valley region (n=1 540 premises). Forward simulations for were conducted on these regions, and a further highly affected region for validity testing: T = Tamworth region (n= 1931 premises).

Premises were classified as infected if they held horses that had been observed with the classical clinical signs of equine influenza (cough, elevated temperature, nasal discharge and lethargy), and (for the first 4 weeks of the outbreak only) laboratory confirmation based on the results of real-time reverse transcription polymerase chain reaction assays, antigen-detecting enzyme-linked immunosorbent assays or haemaglutinin inhibition testing (Kirkland et al., 2011, Moloney et al., 2011). The contact-tracing dataset included 1034 horse movements onto and off premises investigated by government veterinary services during the outbreak. Each movement record included the date of the movement, and the address and unique identifier of both the origin and destination premises.

Table 7.1 Descriptive statistics for highly affected regions in the Australian 2007 equine influenza outbreak, used as study sites in the development and validation of a real-time Bayesian model of equine influenza in Australia.

	Greater Sydney	**Hunter Valley**	**Tamworth**	**All outbreak affected areas**
Number of horse premises	7 670	1 540	1 931	65 005
Land area (km^2)	9 000	12 000	7 200	168 000
Horse premises density (premises km^{-2})	0.85	0.13	0.27	0.39
Number of infected premises	2 895	403	529	9 359
Cumulative incidence (%) (95% CI)	37.7 (36.7, 38.8)	26.2 (24.0, 28.4)	27.4 (25.4, 29.4)	14.4 (14.1, 14.7)
Horses per premises median (range)	2 (1, 380)	4 (1, 741)	3 (1, 200)	2 (1, 1200)
Horse premises area (acres) median (range)	5 (1, 17 178)	94 (1, 6 273)	41 (1, 5 225)	36 (1, 74 894)

All data were collated into a relational Microsoft Access 2007 database (Microsoft Corporation, Redmond WA , USA). Bayesian inference and simulation runs were implemented in purpose built software applications coded in the computer language C++, and all further statistical analyses were conducted in the R statistical package version 2.13.0 (R Development Core Team, 2011).

7.2.3. Bayesian inference

7.2.3.1. Model structure

A Bayesian susceptible-infected-recovered (SIR) approach was applied to stochastically model equine influenza transmission between horse premises, separately in each of the two primary study regions (Greater Sydney and the Hunter Valley), as an adaptation of the method described in detail by Jewell et al. (2009a). This method involves utilising a Bayesian reversible jump Markov chain Monte Carlo (RJ-MCMC) algorithm (Neal and Roberts, 2004, O'Neill and Roberts, 1999) to infer probability distributions for transmission rates and unknown parameters for disease transmission.

The unit of interest in this analysis was the horse premises. Each premises was assumed to exist at any time in only one of three mutually exclusive states (Figure 7.2): susceptible (*S*), infected (*I*) or recovered (*R*). Each infected premises i was associated with event times I_i, O_i, and R_i representing the time of infection, onset of clinical signs, and time of recovery, respectively.

The infection time for each infected premises i was assumed to be distributed according to a continuous time inhomogeneous Poisson (Jewell et al., 2009b). Since infection times are not directly observable – but onset times of clinical signs are – the incubation period $(O_i - I_i)$ was imputed by sampling from a Gamma distribution with a fixed shape parameter $\alpha = 4$, and a rate parameter $\lambda = 2$. These parameters were selected to return a distribution that reflects the typical 1–3 day incubation period of equine influenza (Paillot et al., 2006), with a small probability of

longer incubation periods (4–6 days) as observed in ponies experimentally infected with low doses of virus (Mumford et al., 1990), thus:

$$(O_i - I_i) \sim Gamma(4, 2) \quad (7.1)$$

The recovery times of infected premises were unobserved in the data unlike the case in many SIR-type model setups. In order to estimate these, a deterministic susceptible-latent-infected-

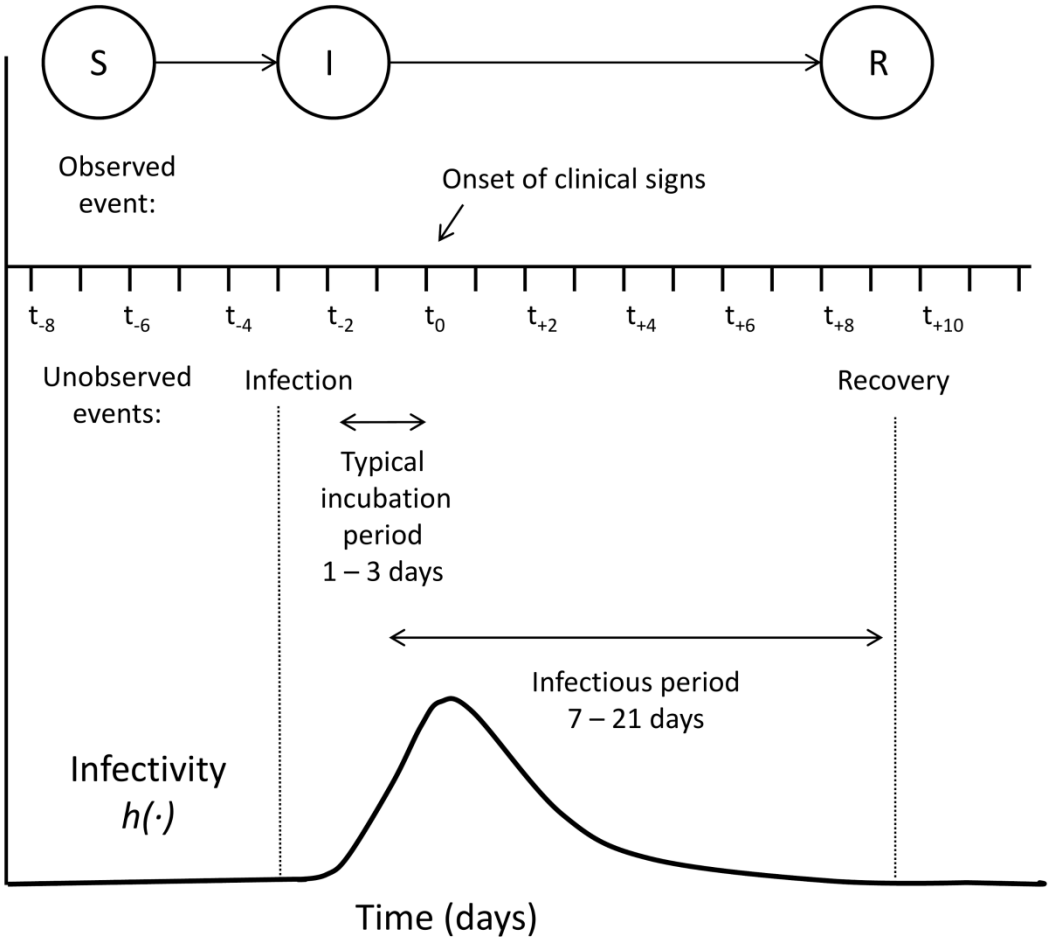


Figure 7.2 Diagrammatic representation of the Bayesian SIR model developed for real-time prediction of equine influenza outbreaks in Australia, with smooth infectivity function, h(·), showing observed and unobserved events, after Jewell et al. (2009a).
S = Susceptible premises, holding horses that are able to be infected, and no horses that are infected by equine influenza. I = Infected premises, holding ≥1 infected horse(s) capable of infecting horses on susceptible premises. R = Recovered premises, all of the horses on a premises have acquired immunity either through recovering from recent infection or vaccination, and are no longer infectious or susceptible to infection. The smooth infectivity function representing intra-premises epidemic processes was based on a standard deterministic SLIR model.

recovered (SLIR) model was used (based on the number of horses data for each premises) assuming that a single index case at each infected premises was infected at time I_i. The estimated proportion of horses infected on the premises at time t was used as a time-continuous measure of premises-level infectivity (see Figure 7.2), a more biologically plausible approach than the standard SIR-type step changes in infectivity (Jewell et al., 2009b). Although a stochastic intra-premises SLIR model might be considered preferable, the additional complexity and computational requirement that this would introduce into the RJ-MCMC algorithm was considered to be not worthwhile for what was anticipated as only a modest further gain in biological plausibility.

Transmission rates

The disease transmission rate $\tau_{j(t)}$ was defined as the sum of the individual infection pressures, $\beta_{ij(t)}$, acting on any susceptible premises j (in the set $\boldsymbol{J}_{(t)}$) from all infective premises (in the set $\boldsymbol{I}_{(t)}$), at time t, as:

$$\tau_{j(t)} = \varepsilon_{(t)} + \sum \beta_{ij(t)} \qquad i \in \boldsymbol{I}_{(t)},\ j \in \boldsymbol{S}_{(t)} \tag{7.2}$$

where, $\varepsilon_{(t)}$ was the time-dependant background transmission rate used to represent transmission other than by explicitly modelled local spread in three separate time periods: ε_1 prior to the horse movement ban (t_1 <10 days), ε_2 during the horse movement ban and prior to the commencement of vaccination (10≤ t_2 <44 days), and ε_3 once vaccination had commenced and horse movement restrictions were relaxed in certain areas (t_3 ≥44 days). Non-explicitly modelled transmission may include movements of infected horses separate to the explicitly modelled primary disease incursion into a region, or rare instances of long distance transmission on fomites.

Following Jewell et at. 2009a, and considering the available premises-level covariate data and the results of previous epidemiological analyses (Firestone et al., 2012a, Firestone et al., 2012b,

Firestone et al., 2011a, Firestone et al., 2011b) the pair-wise infection pressure between individual infected and susceptible premises, $\beta_{ij(t)}$, was formulated as the product of three functions representing: the infectivity of *i*, the susceptibility of *j*, and the decay of infection pressure with increasing Euclidean distance between *i* and *j*, assuming that:

$$\beta_{ij(t)} = \mu \cdot q_{i(t)} \cdot s_{j(t)} \cdot K(d_{ij}) \quad (7.3)$$

where μ is the baseline transmission rate, $q_{i(t)}$ is the infectivity of premises *i* at time *t*, $s_{j(t)}$ is the susceptibility of premises *j* at time *t*, and $K(d_{ij})$ is a spatial kernel describing how disease transmission rate decays with distance between *i* and *j*.

The infectivity of premises *i* at time *t* was defined as:

$$q_{i(t)} = n_i a_i^{\xi} h(t - I_i) \quad (7.3.1)$$

where n_i and a_i are the number of horses on, and area of premises *I,* respectively, and ξ is an unknown parameter that allows for nonlinear effects of premises area. $h(t - I_i)$ returns the proportion of infected horses on premises *I* at time *t*, obtained directly from the solution of the SLIR system of differential equations.

The susceptibility of premises *j* at time *t* was defined as:

$$s_{j(t)} = n_j a_j^{\zeta} \theta^{\mathbf{1}[j \in \boldsymbol{V}_{(t)}]} \quad (7.3.2)$$

where n_j and a_j are the number of horses on, and area of premises *j,* respectively, ζ is an unknown parameter and θ is the effect of vaccination. The indicator function $\mathbf{1}[j \in \boldsymbol{V}(t)]$ returns 1 if *j* is a member of the set $\boldsymbol{V}_{(t)}$ of vaccinated premises at time *t*, and 0 otherwise.

Finally, the spatial kernel $K(d_{ij})$ was assumed to take the Cauchy form:

$$K(d_{ij}) = \frac{\delta}{\delta^2 + \rho_{ij}^2} \tag{7.3.3}$$

where ρ_{ij}^2 is the squared Euclidean (straight line) distance between the centroids of premises *i* and *j*, and δ is an unknown scaling parameter. The Cauchy form was selected for the spatial transmission kernel because it is computationally efficient and also because it may assume a heavier tail than exponential or geometric kernels (allowing for increased likelihood of long-range transmission) without having to include an extra parameter (Kypraios, 2007).

Intra-premises deterministic SLIR process

The SLIR system of differential equations were formulated based on three transmission rates (β_{intra}, ψ_{intra}, γ_{intra}) between individual horses within the infected premises *i*, at time *t*, assuming density-independent transmission and homogenous random mixing within each premises, such that:

$$h(t - I_i) = \begin{cases} \frac{n_{(t)}^I}{n_i} & \text{if } t \geq I_i \\ 0 & \text{otherwise} \end{cases} \tag{7.4}$$

$$\frac{dn_{(t)}^S}{dt} = -\frac{\beta_{intra} n_{(t)}^S n_{(t)}^I}{n_i} \tag{7.4.1}$$

$$\frac{dn_{(t)}^L}{dt} = \frac{\beta_{intra} n_{(t)}^S n_{(t)}^I}{n_i} - \psi_{intra} n_{(t)}^L \tag{7.4.2}$$

$$\frac{dn_{(t)}^I}{dt} = \psi_{intra} n_{(t)}^L - \gamma_{intra} n_{(t)}^I \tag{7.4.3}$$

$$\frac{dn_{(t)}^R}{dt} = \gamma_{intra} n_{(t)}^I \tag{7.4.4}$$

where $n_{(t)}^S$, $n_{(t)}^L$, $n_{(t)}^I$ and $n_{(t)}^R$ are the numbers of susceptible, latent, infective and recovered

horses, respectively, on premises *i* at time *t*. Time was represented quasi-continuously in the intra-premises SLIR model (in contrast to the continuous inter-premises processes) by small time steps (dt = 0.1 days), to enable the computationally efficient return of $h(t - I_i)$. Therefore, as an example, if premises *i* held 6 horses, 4.2 horses could be modelled as infected 7.8 days after the index horse was infected on the premises, in which case: $h(t - I_i)$ = 4.2/6 = 0.70.

The intra-premises transmission rate for individual horses moving between susceptible and infected states, β_{intra}, was modelled as an unknown parameter. The latent period of individual horses, 1/ ψ_{intra}, was fixed at 1 day, and the recovery rate, 1/γ_{intra}, fixed to 6 days based on experimental evidence that horses infected for the first time commence shedding equine influenza virus from 1 day after infection and then for 5–7 days depending on dose and method of infection (Daly et al., 2011, Mumford et al., 1990). The basic reproductive ratio of an epidemic (R_0), may be defined as the average number of secondary infections produced by a typical infective case assuming homogenous random mixing (Anderson and May, 1991). At the individual horse level, R_0 was estimated using the ratio:

$$R_0 = \beta_{intra} \,/\, \gamma_{intra} \qquad (7.5)$$

7.2.3.2. Model fitting and parameter estimation

The reliability of predictions produced by a model critically depends on how reasonably it was parameterised. In general, the shapes of the key probability distributions that drive stochastic epidemic models have been parameterised based on either the review of past outbreaks, experimental and observational epidemiological studies, the elicitation of 'sensible' parameters by consulting experts, or, more recently, by using Bayesian frameworks to statistically infer prior transmission parameter distributions whilst accounting for unobserved information (Jewell et al., 2009b).

For each primary study region, our model was fitted to spatial and covariate data from the equine influenza dataset, and a Bayesian RJ-MCMC algorithm (Jewell et al., 2009c) was used to infer the joint posterior probability distribution for the nine unknown parameters (ε_1, ε_2, ε_3, μ, ξ, ζ, δ, θ, β_{intra}). Vague priors were placed on all unknown parameters as detailed in Table 7.2. The unknown parameters ξ and ζ were structured based on earlier observation of a highly nonlinear association between premises area and risk of infection (Firestone et al., 2012b, Firestone et al., 2011a).

Table 7.2 Prior probability distributions placed on unknown parameters in Bayesian reversible-jump Monte-Carlo Markov Chain inference of data from highly affected regions of the 2007 equine influenza outbreak in Australian.

Unknown parameter	Prior probability distribution	Description (reference)
Time-dependent background transmission rates[a]	$\varepsilon_{(t)} \sim$ Gamma(1,1)	Vague priors with support (allowed values) ≥0
Baseline transmission rate	$\mu \sim$ Gamma(1,1)	
Decay parameter for spatial transmission kernel	$\delta \sim$ Gamma(1,1)	
Intra-premises (horse-to-horse) transmission rate	$\beta_{intra} \sim$ Gamma(1,1)	
Index of effect of premises area on infectivity	$\xi \sim$ Normal(0,0.1)	Vague priors with support in the range (-∞,+∞) centred on zero, to suggest that premises area had no effect on the risk of horses being infected by or infecting horses on another premises.
Index of effect of premises area on susceptibility	$\zeta \sim$ Normal(0,0.1)	
Protective effect of vaccination	$\theta \sim$ Beta(2,2)	Vague prior with support in the range [0,1] based on the results of previous survival analysis (Firestone et al., 2012b).

[a] Three background transmission rates (ε_1, ε_2, ε_3) were used to represent transmission by non-explicitly modelled routes in the period prior to the horse movement ban (t_1<10 days), during the horse movement ban and prior to the commencement of vaccination (10≤ t_2<44 days), and once vaccination had commenced and horse movement restrictions were relaxed in certain areas (t_3 ≥44 days), respectively.

For the Greater Sydney region, the first 2000 iterations of the RJ-MCMC were discarded ('burn-in') to allow the chain to equilibrate and a further 18,000 iterations were used for estimation of the joint posterior distribution. The RJ-MCMC algorithm took longer to converge for the Hunter Valley region, so the first 10,000 iterations were discarded as burn-in and a further 90,000 iterations were used for estimation of the joint posterior distribution.

7.2.3.3. Treatment of missing data

Data were missing on the number of horses held by 3099 of 65,005 premises (4.8%) in the equine influenza affected area. Horse premises with missing number of horses data were evenly distributed across the studied areas. In the Bayesian RJ-MCMC analysis, such missing observations may be treated as latent variables and inferred, but this comes at a high computational cost.

These data could not be considered as missing completely at random (MCAR) because the probability that number of horses data were missing for a premises was likely to be correlated with other non-missing values in the data (such as the infection status of a premises) and may even have depended on the unobserved data itself. The probability of a premises having missing number of horses data may have depended on the number of horses it held if owners of premises holding small numbers of horses were less inclined to self-report than owners of premises holding large numbers of horses. These data were therefore considered missing not at random (MNAR).

Methods of reducing biases related to MNAR data are an area of active research. Certain Bayesian and multiple imputation methods are considered able to reduce the unknown biases caused by MNAR data (Dohoo et al., 2009). Rather than using data-augmentation methods for all of the missing data on premises horse numbers within the Bayesian RJ-MCMC algorithm, these data were imputed, considering that the gain in accuracy through data-augmentation on this variable would be far outweighed by the computational cost. Multiple imputation were

conducted using Gibbs sampling of chained equations (van Buuren and Groothuis-Oudshoorn, in press) on missing horse number data, prior to Bayesian inference, stratified by region and taking into account observed data on premises area and whether or not they were reported as infected. The missing horse data were imputed three times. For each imputation, the Bayesian RJ-MCMC algorithm was repeated, producing a joint probability distribution of unknown parameters. As a component of the formal sensitivity analysis of the model, simulation model outputs were statistically compared based on the inferred joint parameter distributions of each imputation.

7.2.4. Stochastic simulations

Continuous time stochastic simulation (Bailey, 1975, Gillespie, 1977) was conducted based on the epidemic model specified above, running 300 simulations for each primary study region. A *post hoc* analysis was conducted to assess whether this number of stochastic simulation runs (effective sample size) was sufficient. The change in the coefficient of variation (CV) of each of four key output parameters was monitored to ascertain when little further gain in information was occurring with extra iterations: the number of premises infected by seeding through the horse movement network, final epidemic size, the timing (in days) and the magnitude (in numbers of premises infected per day of the epidemic peak) (Cowled et al., 2012).

Each simulation run was parameterised using a single sample from the joint posterior distributions produced by the Bayesian analysis. The resulting collection of 300 simulation results were used as an estimation of the Bayesian predictive distribution of the epidemic, and were used to quantify model fit (Bernardo and Smith, 1994).

In each run, infection was seeded at the single known index premises, k, the location of the equestrian event where disease transmission was first known to occur in the 2007 outbreak. Prior to the implementation of the horse movement ban (t <10), only the movement of infected

horses was simulated. The infection of premises at widely dispersed spatial locations across the population at risk was stochastically modelled based on contact-tracing data of the network of actual horse movements originating from the event at premises *k*. In our previous likelihood-based analysis of the contact-tracing dataset (Firestone et al., 2012a), we estimated that the mean probability of infection at a susceptible premises that received a horse movement from an infected premises in the first 10 days of the 2007 outbreak was 18.7% (95% CI: 16.7, 20.3%), after accounting for risk of infection through local spatial spread. So for any susceptible premises, *j*, that received a horse movement from any premises *i* holding infected horses in the first 10 days of the simulated outbreak, the probability that premises *j* was infected by this movement, π_{ij}, was randomly sampled from a Bernoulli distribution centred on 0.187.

Once an infection occurred in the modelled study region, the simulated spread by the movement of horses was continued (which ceased when t=10 days), and concurrent simulation of local spatial spread commenced. The standard stochastic simulation method for local spatial spread was adapted to incorporate the smooth infectivity function by retrospective sampling (Jewell, 2009) to appropriately thin the Poisson process (Karr, 1991) thus maintaining local time-homogeneity. Further detail on these methods which enable computationally efficient simulation based on the Exponential distribution is provided in the appendix to this chapter (see equation 7.A1). The derivation of these methods is described in detail elsewhere (Jewell, 2009).

7.2.5. Model validation

There are numerous challenges when attempting to validate epidemic models (Reeves et al., 2011). In the case of equine influenza in Australia, there has only ever been one outbreak, and the dataset, upon which the model's validity depends, contains known differential ascertainment biases of the population at risk due to under-reporting (Dhand and Sergeant, 2011a) and incentives to clear infection in certain areas, missing data, misclassification and case definition changes (Moloney et al., 2011). The Australian 2007 equine influenza dataset is very large, and

considerable effort and resources were devoted to collecting, collating and cleaning this surveillance data. However, enabling epidemiological analyses and modelling were not the main considerations when amassing this dataset; outbreak surveillance was conducted to directly inform disease control and monitoring activities.

A range of subjective and objective techniques were used to validate our model. Firstly, to assess the 'face validity' of the model concept and to ensure that once developed it would be potentially useful for decision-makers the model formulation was discussed at a meeting of veterinary epidemiologists and government officials who had been involved in the actual outbreak response and follow-up research activities. Outputs from simulation runs based on a pilot RJ-MCMC analysis were animated and analysed to verify that the formulation and coding were producing the desired outputs.

To objectively test the 'operational validity' of the model, final model outputs were compared with empirical estimates from the actual 2007 outbreak. The model was also pseudo-validated by populating it with data not previously used in model training (from the Tamworth region), then simulated model outputs were compared to the observed realisation in this region during the 2007 outbreak. For these iterations, the simulation model was parameterised using the posterior probability distributions from the immediately adjacent Hunter Valley study region, given similarities in the area and spatial aggregation of horse premises across the two regions (Table 7.1), noting that the horse industry differed considerably in these two regions. For completeness, a second validation trial was run populated with population at risk data from the Tamworth region and the inferred posterior probability distributions from the Greater Sydney region.

Finally, model outputs were compared against those produced by the less general simulation model developed and validated by Garner et al. (2011b) for a different purpose: to replicate the 2007 outbreak and test the effectiveness of different vaccination strategies. The previous model

was originally populated based on data from the whole spatial extent of the 2007 outbreak, but restricted in time to the period after the outbreak was detected until immediately prior to vaccination commencing ($10 \leq t < 44$ days). For the purposes of this inter-model comparison, the previous model was populated with the dataset for the Greater Sydney region only, seeding each simulation with the status of the actual outbreak at $t = 9$ days.

7.2.6. Sensitivity analysis

A univariable sensitivity analysis of 12 key parameters (π_{ij}, ε_1, ε_2, ε_3, μ, ξ, ζ, θ, δ, ψ_{intra}, β_{intra}, γ_{intra}) was conducted to test the importance of uncertainty in input parameters on modelled estimates (Garner and Hamilton, 2011) and elucidate which transmission rates had the most influence on simulated epidemic outcomes in the two primary study regions (Reeves et al., 2011). For each parameter of interest, the change in model outputs was investigated after conducting 300 model runs with the parameter of interest doubled, and then halved (Cowled et al., 2012, Garner et al., 2011b).

Peak immunity occurs around 14 days after a primary vaccination course of the canarypox-vectored vaccine that was used in the 2007 outbreak, reducing clinical signs and shedding in subsequently infected animals (Edlund Toulemonde et al., 2005). For parsimony, the model formulation did not include an unknown parameter representing delay in onset of immunity following vaccination. To test whether our model was structurally sensitive to this assumption, it was reformulated setting a fixed delay in onset of immunity following vaccination of 3,5,7,14 and 21 days after the recorded date of the primary vaccination, and a further 300 iterations run with each scenario.

To further explore the sensitivity of our model to the deterministic intra-premises SLIR process, the unknown intra-premises transmission rate, β_{intra}, was fixed at values equivalent to those estimated in two previous modelling studies that reported the basic reproductive ratio: $R_0 = 3.5$

after Satou and Nishiura (2006), and R_0 = 10 after Glass et al. (2002). Simulations were then repeated for each of the primary study regions, and the change in model outputs investigated.

7.2.7. Statistical analyses

Model validation and sensitivity analysis involved comparing model outputs against the single actual realisation observed in 2007, and inter-model comparisons. These comparisons were conducted on the basis of final epidemic size and temporal and spatial cross-correlation of model outputs. For each region, the 95% credible interval (CrI) of the predicted final epidemic size was compared directly against the 2007 data. Pair-wise inter-model comparisons of final epidemic size were conducted using the Wilcoxon rank sum test to statistically test for differences in the distributions of model outputs, without assuming them to follow the normal distribution.

Temporal cross-correlations of model outputs against 2007 observations were conducted using Spearman's rank correlation coefficient (ρ). The number of infected premises predicted per day in each model iteration were compared against those actually observed, reported with 95% credible intervals. The same method was applied for pair-wise inter-model comparison using the median predicted number of infected premises per day, across all iterations.

A grid-based method was applied to evaluate the level of spatial correlation in the location of predicted infected premises between models and against observed data (Dubé et al., 2007, Garner et al., 2011b). Each study region was divided into regular grids of cells 10 km × 10 km in size (Garner et al., 2011b), producing a total of 132 cells in the Greater Sydney region, 143 cells in the Hunter Valley region and 81 cells in the Tamworth region. The predicted number of infected premises per cell was counted for each iteration, and the median count per grid cell determined. Cross-correlation between simulated counts per cell and observed data were conducted using Spearman's rank correlation coefficient (ρ_{grid}), reporting 95% credible

intervals. Pair-wise inter-model comparisons were conducted similarly, comparing simulated counts of each iteration of one model, against the median count per grid cell of the reference model.

7.3. Results

7.3.1. Bayesian inference

Model convergence was achieved for all unknown parameters after discarding the first 2,000 samples as burn-in for the Greater Sydney region. Autocorrelation was low (serial autocorrelation function, ACF<0.5) after thinning the joint posterior distribution samples by 60 iterations, to ultimately produce an effective sample size of 300 iterations (see Appendix Figure 7.9). The Bayesian RJ-MCMC algorithm took longer to converge in the analysis of the Hunter Valley region. Model convergence was achieved after discarding the first 10,000 samples as burn-in, and low autocorrelation (ACF<0.5) was achieved after thinning by 300 iterations, to again produce an effective sample size of 300 iterations (see Appendix Figure 7.10).

Density plots of each inferred parameter from the thinned joint posterior distribution, by region are presented in Figure 7.3. Except for ε_1 in the Hunter Valley region, all marginal posteriors diverged from the priors. The earliest onset dates reported in the Hunter Valley region during the 2007 outbreak were for 3 IPs with onset on day 10 (these were therefore included in time period *t2*). This lack of data caused the posterior probability distribution for ε_1 in the Hunter Valley to closely match that of its vague $Gamma(1,1)$ prior.

Background and baseline transmission rates

In the Greater Sydney region, the posterior mean of the baseline transmission rate, μ, was 0.28 infections per 1000 premises per day (95% highest posterior density region (HPD): 0.25, 0.31). The baseline transmission rate in the Hunter Valley region was an order of magnitude lower (posterior mean of μ = 0.02 infections per 1000 premises per day, 95% HPD: 0.01, 0.04).

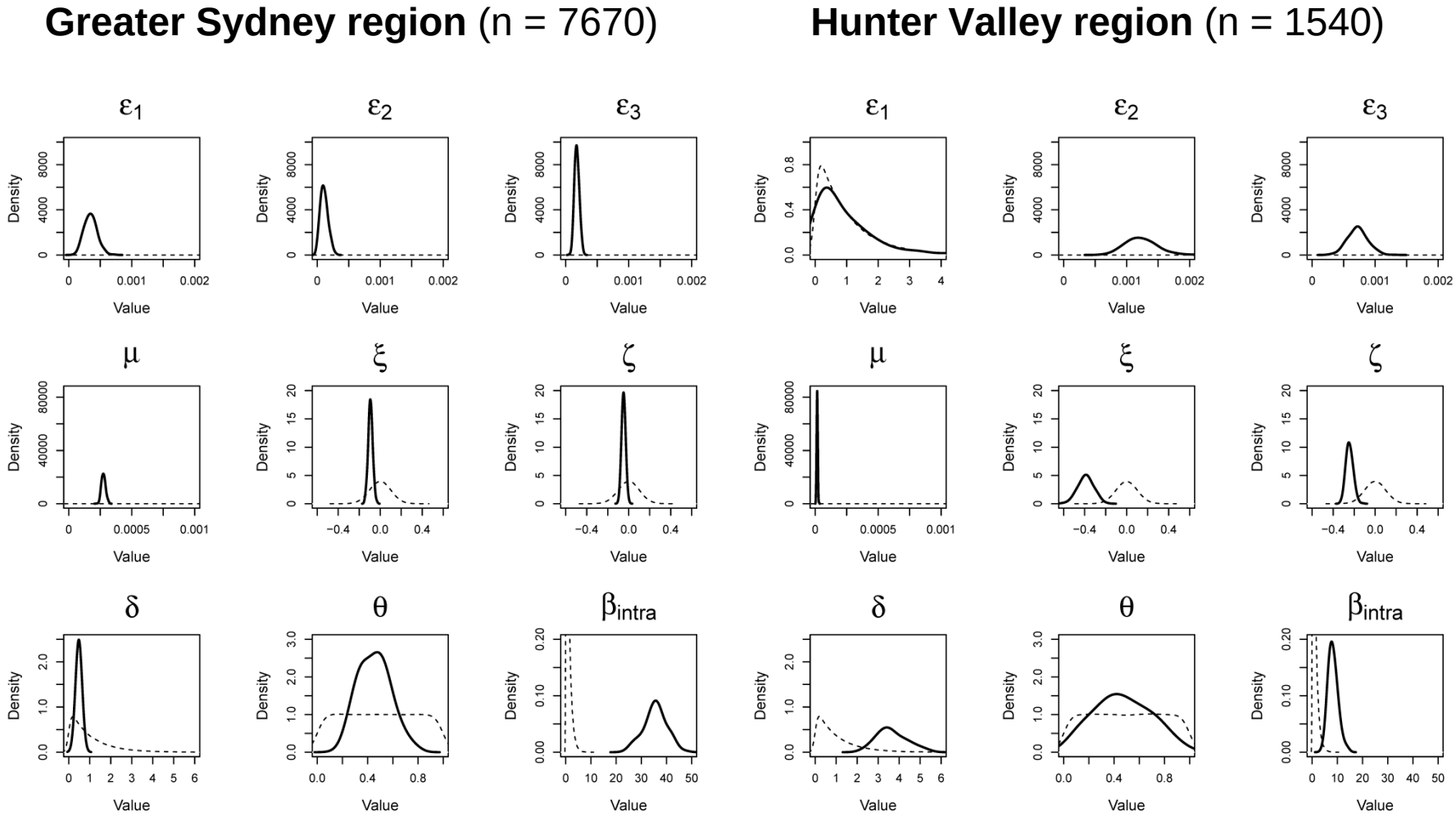


Figure 7.3 Probability density functions, by region, for inferred parameters from the joint posterior parameter distribution after discarding burn-in, and thinning to reduce auto-correlation in the sample used to populate the simulation model.
Dotted lines indicate vague prior distributions used for each individual parameter in the reversible-jump Markov Chain Monte Carlo algorithm.

The background transmission rate in Greater Sydney during the first 10 days of the outbreak (posterior mean of ε_1 = 0.35 infections per 1000 premises per day, 95% HPD: 0.17, 0.56) was more than three times the background rate immediately after implementation of the horse movement ban (posterior mean of ε_2 = 0.11 infections per 1000 premises per day, 95% HPD: 0.01, 0.24). In all three time periods, the time-dependent background transmission rates were higher in the Hunter Valley than in the Greater Sydney region (Figure 7.3).

Shape of the spatial transmission kernel

Estimates of the transmission kernels for each region are presented in Figure 7.4. The posterior mean of the distance decay parameter (δ) in the Greater Sydney region was 0.50 (95% HPD: 0.42, 0.54), where 50% of local spatial spread was estimated to have occurred within 500 m, and 95% of local spatial spread within 2.2 km (95% HPD: 1.8, 2.4 km) of premises holding infectious horses. Local spread was observed over much further distances in the Hunter Valley region, where the posterior mean of the distance decay parameter (δ) was 3.65 (95% HPD: 2.36, 5.25). In this major horse-breeding region with fewer, larger horse premises, 50% of local spatial spread occurred within 3.65 km, and 95% of local spatial spread within 15.9 km (95% HPD: 10.2, 22.9 km).

The effect of premises area on local spread

The posterior probability distributions for the parameters ξ and ζ, representing the effect of premises area on premises-level infectivity and susceptibility, respectively, covered negative ranges and did not cross zero for either region. This suggests that horses on larger premises were both less likely to be infected and less likely to infect horses on another premises. In the Greater Sydney region, the posterior mean values for ζ of -0.05 and ξ of -0.09, suggest that as premises area doubled, risk of infection decreased by 3.3% (95% HPD: 1.6, 4.6%) and the risk of infecting horses on another premise decreased 6.2% (95% HPD: 3.6, 8.9%). Stronger diluting effects were observed in the Hunter Valley region (the posterior mean values for ζ and ξ were -0.24 and -0.39, respectively), where as premises area doubled, risk of infection decreased by

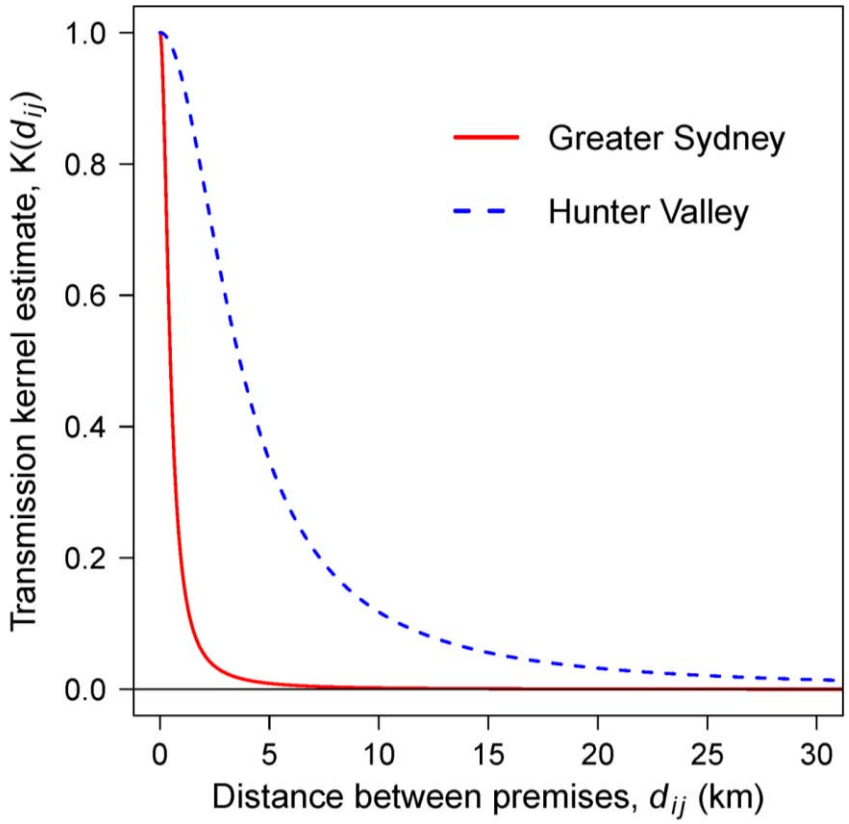


Figure 7.4 The estimated decay in transmission by distance from infected premises for each of the primary study regions.
Transmission kernels were estimated using the posterior mean values of δ, the Cauchy distribution scaling parameter.

15.6% (95% HPD: 11.4, 19.1%) and the risk of infecting horses on another premise decreased 23.7% (95% HPD: 15.9, 31.0%).

The effect of vaccination

After vaccination had occurred, clinically observable infection was 54.5% less likely to be reported on a premises in the Greater Sydney region (95% HPD: 27.3, 77.0%). A similar protective effect was observed in the Hunter Valley, however the marginal posterior for θ, the parameter representing the effect of vaccination, did not diverge markedly from its prior.

The intra-premises transmission rate

The posterior mean for β_{intra} in the Greater Sydney region was 35.5 (95% HPD: 26.1, 44.2). The

intra-premises transmission rate β_{intra} was much lower in the Hunter Valley region (posterior mean = 8.3, 95% HPD: 5.6, 12.8). The largest premises in the Greater Sydney region held 380 horses, its duration of premises-level infectiousness was estimated to be 13.8 days (95% HPD: 13.7, 14.0 days) if infected. The largest premises in the Hunter Valley held 741 horses, if infected its duration of premises-level infectiousness was estimated to be 16.4 days (95% HPD: 15.7, 17.3 days).

7.3.2. Stochastic simulations

Trace plots of four key output parameters stabilised within 300 simulation runs for both the Greater Sydney and Hunter Valley regions (see Appendix Figure 7.11). The simulated epidemics in the Greater Sydney and Hunter Valley regions are presented in Figures 7.5 and 7.6 in comparison with observed data from the 2007 outbreak. The observed final epidemic size was within the intervals predicted by the model for both regions, and the simulated and observed epidemic curves in these two regions were very highly cross-correlated (Table 7.3). Spatial outputs were also highly cross-correlated in both regions. The simulated epidemic in the Hunter Valley failed to capture a late peak after movement restrictions were eased to accommodate the Thoroughbred breeding season ($t \geq 44$ days). The median number of premises predicted to have been infected through the movement of infected horses prior to the horse movement ban was 8 premises in Greater Sydney (95% CrI: 3, 13 premises), and 3 premises in the Hunter Valley (95% CrI: 1, 7 premises). During the actual 2007 outbreak we previously estimated that 8 premises in the Greater Sydney region and 5 in the Hunter Valley were infected in this manner.

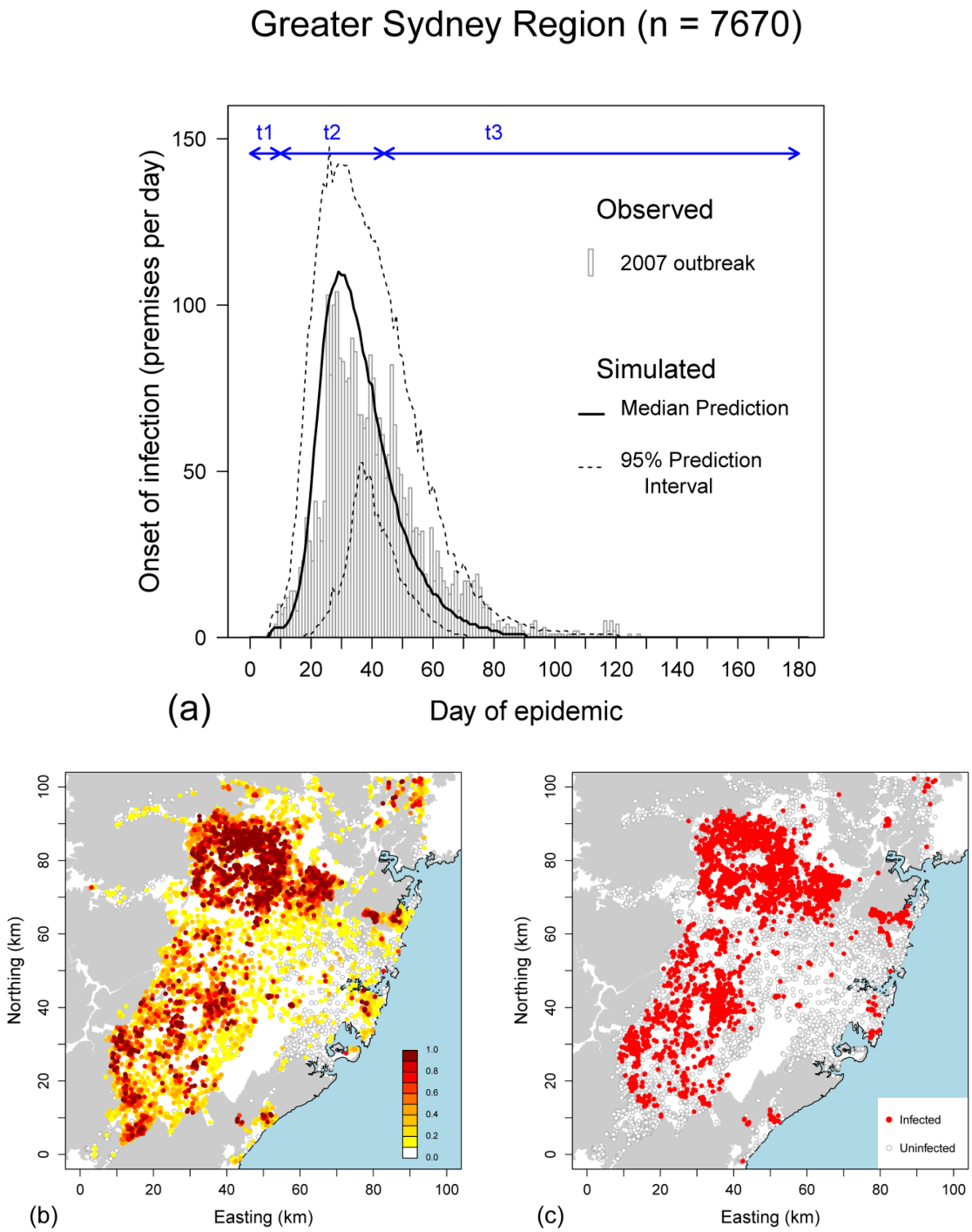


Figure 7.5 Comparison of simulated and observed outbreaks for the Greater Sydney region. (a) Simulated epidemic curve, with 95% credible interval, compared to the realisation observed during the 2007 equine influenza outbreak in Australia. Time periods: $t1$<10 days; 10≤ $t2$<44 days; and $t3$ ≥44 days. (b) Proportion of runs in which a premises was infected based on 300 simulation runs, compared to (c) the single observed realisation of the 2007 outbreak.

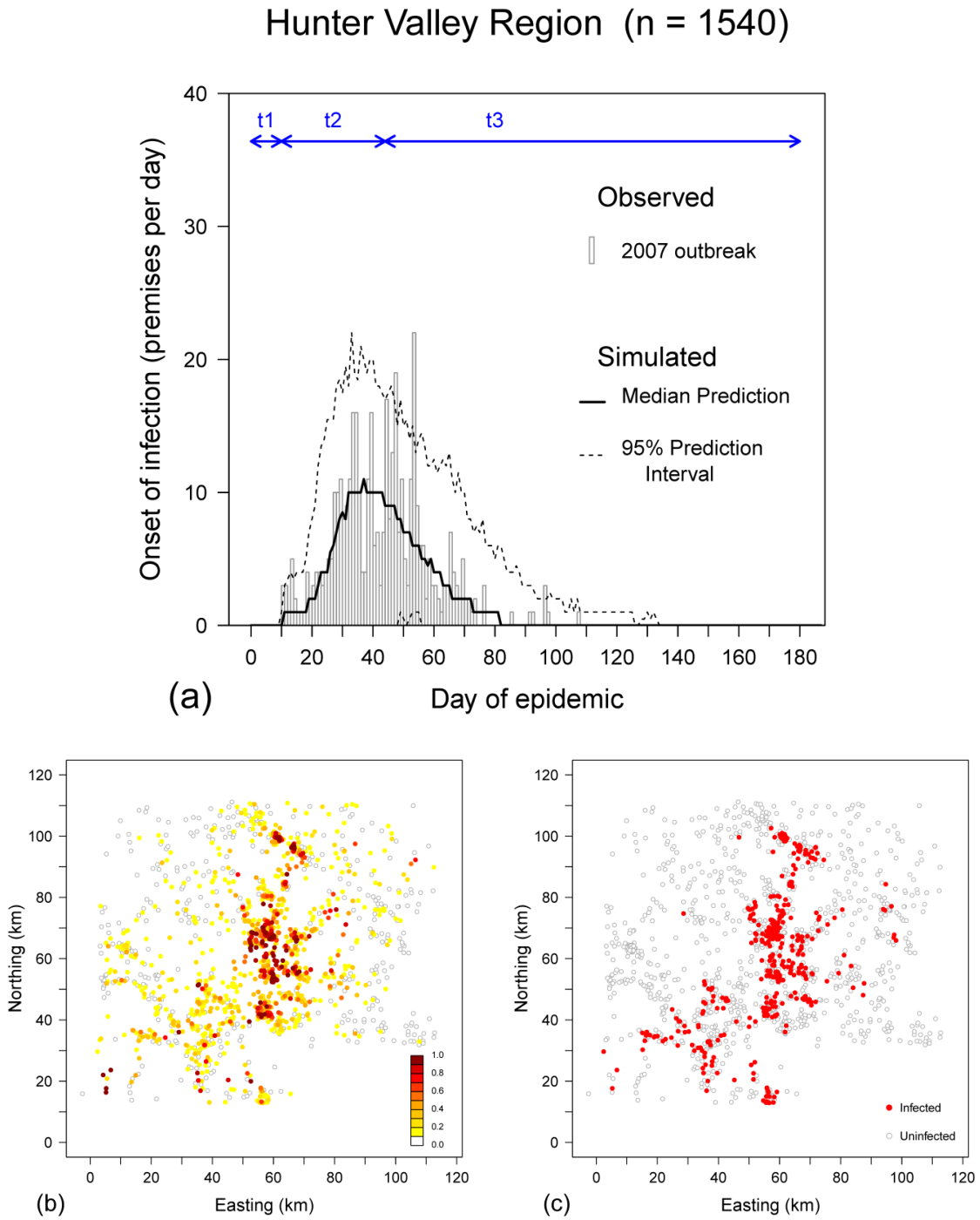


Figure 7.6 Comparison of simulated and observed outbreaks for the Hunter Valley region. (a) Simulated epidemic curve, with 95% credible interval, compared to the realisation observed during the 2007 equine influenza outbreak in Australia. Time periods: $t1<10$ days; $10\leq t2<44$ days; and $t3 \geq 44$ days. (b) Proportion of runs in which a premises was infected based on 300 simulation runs, compared to (c) the single observed realisation of the 2007 outbreak.

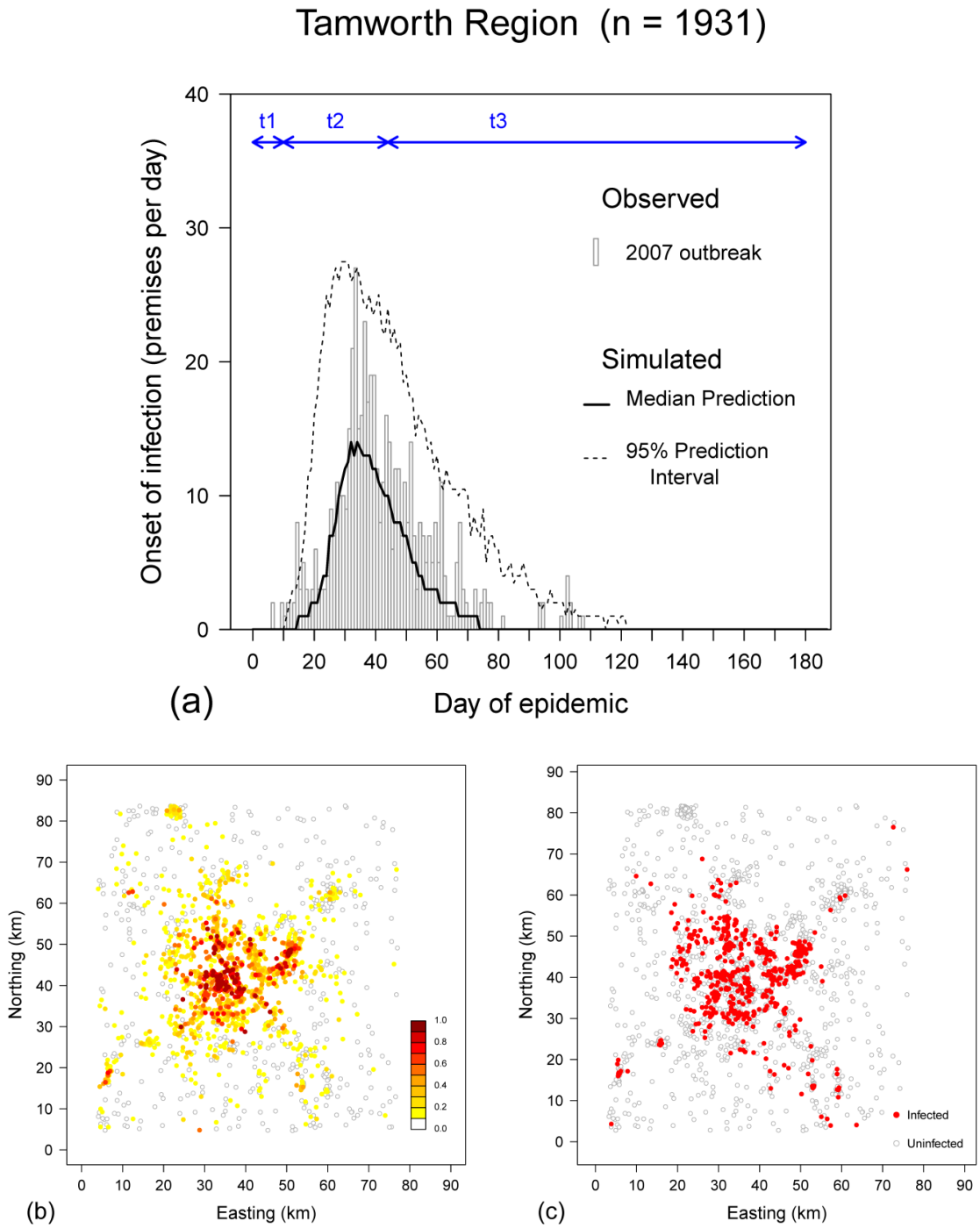


Figure 7.7 Validation against data not used in model-building based on the Tamworth region. (a) Simulated epidemic curve, with 95% credible interval, compared to the realisation observed in the Tamworth region during the 2007 equine influenza outbreak in Australia. The model was parameterised with population at risk in the Tamworth region, and unknown parameters distributions inferred for the Hunter Valley region. Time periods: $t1$<10 days; 10≤ $t2$<44 days; and $t3$ ≥44 days. (b) Proportion of runs in which a premises was infected based on 300 simulation runs, compared to (c) the single observed realisation of the 2007 outbreak.

Table 7.3 Comparison of stochastic simulation outputs against observations from the 2007 outbreak of equine influenza in Australia, by study region.

Region	Final epidemic size (number of infected premises)		Cross-correlation of simulated and observed epidemic curves ρ (95% CrI)	Cross-correlation of simulated and observed spatial outputs ρ_{grid} (95% CrI)
	Observed	Simulated (95% CrI)		
Greater Sydney	2 895	2 933 (2 803, 3 078)	0.916 (0.843, 0.937)	0.885 (0.867, 0.903)
Hunter Valley	403	386 (299, 460)	0.831 (0.189, 0.888)	0.748 (0.670, 0.816)
Tamworth[a]				
Validation trial 1: **HV**	529	429 (2, 535)	0.825 (0.059, 0.880)	0.766 (0.204, 0.848)
Validation trial 2: **GS**	529	388 (1, 475)	0.663 (-0.075, 0.870)	0.738 (0.175, 0.854)

[a] In both validation trials the model was parameterised with population at risk data from the Tamworth region (not used in model building). Unknown parameters were based on the joint posterior probability distribution of the Hunter Valley region (HV) in trial 1, and the Greater Sydney region (GS) in trial 2.

7.3.3. Model validation

7.3.3.1. Simulation in a new study population

Modelling outputs from the validation trial conducted on the population at risk in the Tamworth region using the joint posterior parameter distribution from the adjacent Hunter Valley region are presented in Figure 7.7. The model marginally under-predicted the final epidemic size in this trial (Table 7.3, trial 1), however, the simulated and observed epidemics were highly cross-correlated temporally and spatially. When the epidemic in the Tamworth region was simulated using the joint parameter distribution from the Greater Sydney region under-prediction was exacerbated, and cross-correlation of the simulated and observed epidemics was lower (Table 7.3, trial 2).

7.3.3.2. Inter-model comparison

The previous model was only run on the Greater Sydney region for the second time period ($10 \leq t2 < 44$ days). In this time period, the epidemic curves simulated by this specific epidemic model closely followed the actual epidemic curve observed in the 2007 outbreak (see Appendix Figure 7.12). Median predictions of the previous model were highly cross-correlated with the outputs of 300 iterations of our model, both temporally (ρ = 0.818, 95% CrI: 0.482, 0.902) and spatially (ρ_{grid} = 0.919, 95% CrI: 0.863, 0.937). The previous model predicted an epidemic size of 2145 premises between days 10 and 43 of the outbreak (95% CrI: 1 954, 2 304), five less than our model (difference of medians 95% CrI: -27.0, 39.0).

7.3.4. Sensitivity analysis

The results of varying 12 key parameters on the simulated final epidemic size in the Greater Sydney and Hunter Valley regions are presented in Figure 7.8. In both regions, simulation modelling outputs were highly sensitive to input values for parameters representing the baseline transmission rate (μ), the shape parameter for the spatial transmission kernel (δ), the nonlinear diluting effects of area (ξ and ζ), and the period that individual horses remain infectious (γ_{intra}).

When the baseline transmission rate was halved, the final epidemic size decreased by 51% in the Greater Sydney region, and 48% in the Hunter Valley. Increasing δ, and thereby extending the effective range of local spread, had more of an effect on increasing the size of predicted epidemics in the Greater Sydney region than in the Hunter Valley region (35% versus 19% increase in final epidemic size). On the other hand, model outputs for the Hunter Valley were more sensitive to doubling and halving ξ and ζ, than model outputs from the Greater Sydney region. All of these changes reduced cross-correlation of model outputs with observed data, both spatially and temporally (all *P*-values<0.001).

Halving the recovery rate of infected horses (γ_{intra}) increased final epidemic size by 53% in Greater Sydney and 60% in the Hunter Valley. In both primary study regions, this resulted in increased temporal cross-correlation of modelled outputs against observed data (*P*-values<0.001), with little difference observed in the spatial cross-correlation of outcomes. Halving the latent period of individual horses (ψ_{intra}) led to modest improvements in the temporal correlation of model outputs when compared to the observed outbreak data. Increasing ψ_{intra} reduced epidemic size in the Greater Sydney region by 1%.

Halving and doubling the intra-premises transmission rate (β_{intra}) did not result in a statistically significant difference in modelling outcomes. Fixing β_{intra} at a much lower value so that the intra-premises R_0 = 10 resulted in a 2% increase in temporal cross-correlation of the modelled and observed epidemic curves for the Hunter Valley region (*P*<0.001), and no change to other modelling outcomes. Fixing β_{intra} so that the intra-premises R_0 = 3.5 resulted in a 7% reduction in the final epidemic size in Greater Sydney and a 4% reduction in the Hunter Valley. Temporal and spatial model outputs for both regions were less well correlated with observed data when R_0 = 3.5 (*P*-values<0.001).

Reducing seeding of infection through contact-traced horse movements (halving π_{ij}) only effected the final size of simulated outbreaks in the Hunter Valley. Although this led to smaller

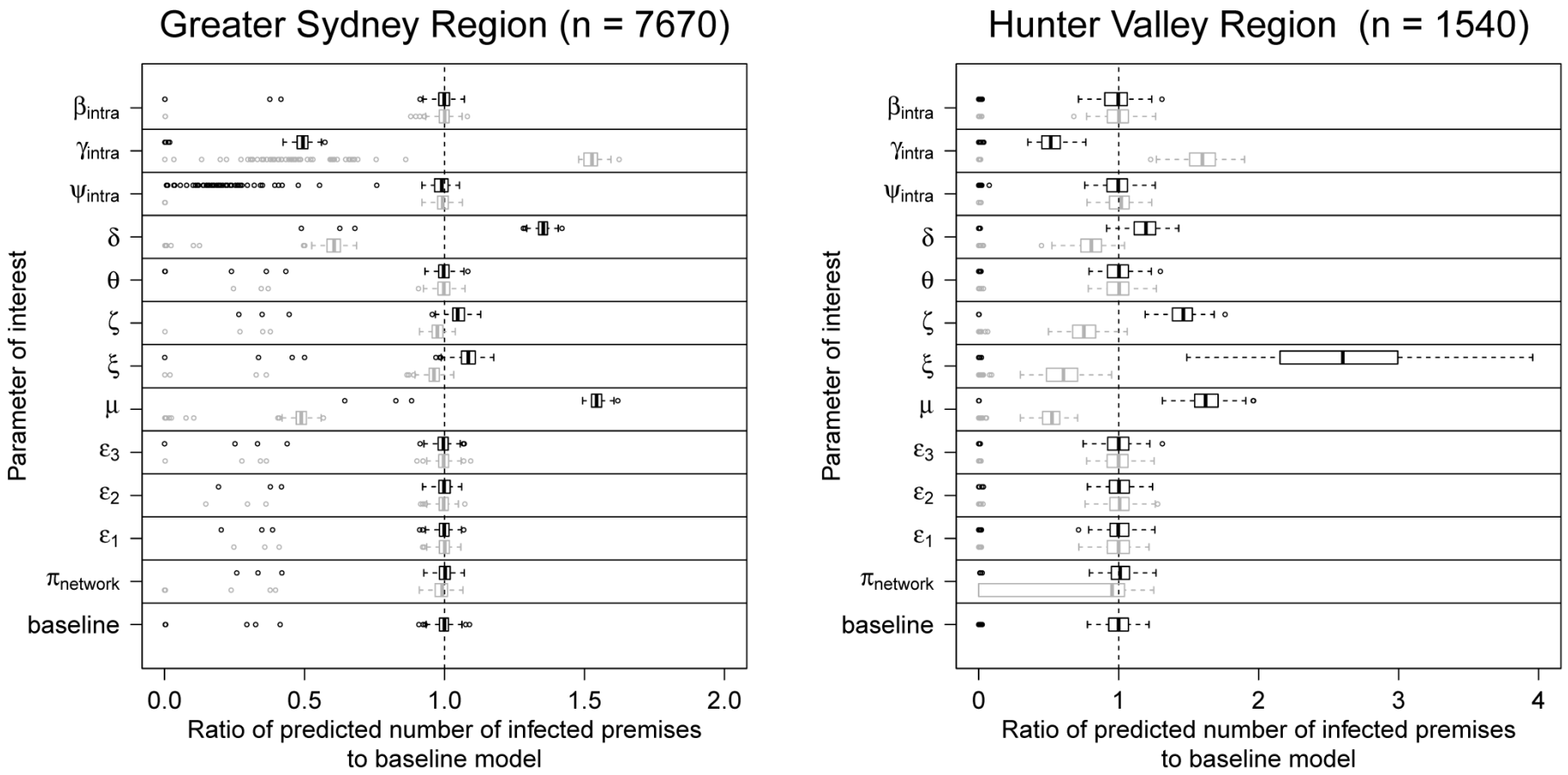


Figure 7.8 Sensitivity analysis of the Bayesian dynamic modelling framework for the rapid assessment of future outbreaks of equine influenza in Australia. Ratios of final epidemic size (number of infected premises) produced from 300 iterations of the baseline model, and from 300 iterations when the parameter of interest is doubled (black) and halved (grey), by region.

outbreaks, it did not result in an overall reduction in correlation of temporal or spatial model outputs with observed data. Modelling outputs were also insensitive to large variations in the background transmission rates (ε_1, ε_2, ε_3) and the parameter representing the effect of vaccination (θ). Although there was a marginal increase in the numbers of infected premises when the model was reformulated to incorporate a delayed onset of vaccine-induced immunity of up to 21 days, these changes were not statistically significant. Nor was there any significant difference in simulation modelling outputs when input parameter distributions were derived from RJ-MCMC Bayesian inference of different imputations of missing data on number of horses in the population at risk.

7.4. Discussion

In this study a rigorous Bayesian statistical analysis of data from the 2007 Australian equine influenza outbreak was conducted to infer key epidemiological parameters for two very different regions where horses are kept. A stochastic simulation model was then populated and validated, tailored to the types of data and uncertainties commonly experienced early in animal disease emergencies. In addition to the parameter estimates presented, the output of this study is a general stochastic modelling framework capable of being utilised for rapid assessment of future outbreaks of equine influenza, and decision-support in regions of Australia with comparable horse industries to either of the primary study regions.

Model validation involves assessing whether the model and its outputs are a sufficiently accurate representation of the modelled system to be useful for the intended purpose (Garner and Hamilton, 2011, Reeves et al., 2011). The purpose of the model developed in this analysis was to estimate key epidemiological parameters, thereby enabling meaningful prediction during future outbreaks of equine influenza in Australia. The model was not intended to replicate the 2007 outbreak in these regions exactly, but rather to develop probability distributions that could inform Bayesian predictive risk-mapping for decision-support during future outbreaks. As the

intention was to present a set of realistic scenarios, including the single actual realisation observed in the 2007 outbreak, simulated outputs would not be expected to exactly match those of the empirical estimates; rather to include the actual realisation, as just one in a set of plausible simulated scenarios. Nor would simulated outputs be expected to be identical to those of the less general model developed by Garner et al. (2011b), as their modelling objective was different: these authors attempted to replicate the 2007 outbreak so as to evaluate different vaccination scenarios. Nevertheless, the modelling outputs produced were highly comparable to the single actual observed realisation in the 2007 equine influenza outbreak in Australia, and to those of the previous model.

The model was pseudo-validated by populating it with population at risk data not used in its development. This data was derived from a region with broad similarities in geography and spatial aggregation of horse premises to one of the primary study areas, and known differences in that the horse industry in the Tamworth region includes relatively less commercial horse breeding premises than in the Hunter Valley region. Nevertheless, model outputs were highly comparable to the single realisation observed in 2007.

In the event of a future outbreak of equine influenza affecting a region comparable to either of the primary study regions, the Bayesian framework could be populated with updated premises-level data, and informative prior parameters based on the posterior probability distributions of this analysis (from the comparable primary study region). Predictive simulations could then be run, estimating the risk of infection on all premises in the newly affected region. As the outbreak progresses, the model would be updated with data collected during epidemiologic investigations and parameter learning used to improve the accuracy of predictions. This approach has been demonstrated in a simulation study of highly pathogenic avian influenza (Jewell et al., 2009b) and in an analysis of the 2007 foot-and-mouth disease outbreak in the United Kingdom (Jewell et al., 2009a). Assessment of the usefulness of such a Bayesian framework requires application during a large-scale emergency animal disease outbreak. In such

a setting, obtaining detailed and high quality contact-tracing and outbreaks surveillance data is a non-trivial exercise. Delays in accessing data from the field, and the quality of such data in the early stages of a large outbreak impact on the ability to develop meaningful predictions of the course of the outbreak.

A large degree of regional heterogeneity was observed in the underlying epidemic process during the 2007 equine influenza outbreak, which suggests that different outbreak control strategies might be considered in different regions. In the major horse-breeding region, where larger area horse premises were less densely distributed than in the peri-urban region, local spread occurred over an increased range (seven times lower rate of decay of transmission with distance) and the underlying epidemic process was more stochastic (higher background and lower baseline transmission rates). The background transmission rates represented non-explicitly modelled 'random' spread. In contrast, the baseline transmission rates were used to represent local spatial spread, centred around infected premises according to the shape of the inferred spatial transmission kernel.

The increased background transmission rates in the Hunter Valley region suggest that the epidemic process was less predictable based on a simple spatial transmission kernel in this major horse-breeding area. Direct contact and short-range aerosol transmission of equine influenza between horses on adjoining and nearby premises would be less likely to occur in a sparsely populated region with larger area horse premises. Therefore, the focus in such areas needs to be on preventing spread through other transmission routes (i.e. on fomites and horse movements). The greater range of local spread observed in the Hunter Valley region also suggests increased importance of transmission on fomites and perhaps airborne transmission, relative to direct and close contact, in major horse breeding areas.

The inferred parameter estimates are highly consistent with the findings of previous analyses of the 2007 equine influenza outbreak in Australia. In a peri-urban cluster of infection in south-east

Queensland, Davis et al. (2009) found that following the horse movement standstill 91% of newly infected premises were within 2 km of an earlier infected premises, and that the largest observed local spread distance was 13 km. Our own previous estimates, across the whole outbreak extent, are that after adjusting for animal movements, 89% of local spread occurred within 5 km, and 98% within 15 km of an infected premises (Firestone et al., 2012a).

The present Australian policy for halting the local spread of equine influenza is targeted vaccination and heightened surveillance within 10 km of infected premises, together with ring vaccination inwards from an outer buffer zone at least 20 km from infected premises (Animal Health Australia, 2011). This appears appropriate for high density peri-urban regions, such as Greater Sydney, where local disease spread occurred almost entirely within 5 km. However, the inference of data from the Hunter Valley presented in this study, and field experience from south-east Queensland during the 2007 outbreak (The EI Epidemiology Support Group, 2008), suggests that there was a non-negligible probability of local spread rapidly escaping a 20 km buffer in major horse-breeding areas. Where resources permit, an extended buffer should be considered to keep ahead of infection in such areas.

In both of the primary study regions, horses on larger land-parcels were less likely to infect or be infected. The effect was more pronounced in the Hunter Valley region where horse premises are generally larger in area (Table 7.2). In earlier research, a similar nonlinear relationship was identified between premises area and risk of infection with equine influenza during the 2007 outbreak in Australia (Firestone et al., 2012b, Firestone et al., 2011a). A biologically plausible explanation may be that horses would be more likely to be further from premises fence-lines on larger premises. Different patterns of contact and reduced opportunities for transmission on larger area horse premises has also been suggested as a reason underlying an observed association between herd sizes and attack rates on infected premises (Dhand and Sergeant, 2011b).

In both of the highly affected study regions, following vaccination of horses on a premises the likelihood of clinically observable infection being reported was inferred to have halved. A lack of data limited divergence of the marginal posteriors for the unknown variable representing the effect of vaccination (θ) from the specified priors. A similar yet non-significant protective effect of vaccination was observed in a previous survival analysis of a subset of the available data for the Greater Sydney region (Firestone et al., 2012b). Vaccination was only implemented well after the peak of the outbreak, due to supply and regulatory delays (Callinan, 2008, Garner et al., 2011b). Although the efficacy of the administered (canarypox-vectored) vaccine has been demonstrated against the 2007 Australian outbreak strain (Bryant et al., 2010), several factors may impact on the effectiveness of vaccination in the face of such a large-scale outbreak. It is unlikely that all vaccinated animals would achieve the same level of immune response as horses in experimental vaccination trials and in an outbreak setting horses would be exposed to virus at different points in their vaccination schedules (Bryant et al., 2010). During the 2007 outbreak, many horses received only one primary dose before being exposed to virus (Kannegieter et al., 2011) and some horses were inadvertently vaccinated whilst incubating disease or when they were at very high risk of infection (Perkins et al., 2011).

Reformulating the model to incorporate a more biologically plausible delay in onset of immunity resulted in little to no improvement in correlation of model outputs against the data observed during the 2007 outbreak. If field data were available from an outbreak of equine influenza in a similarly highly susceptible population, where vaccination was implemented earlier, then it may be possible to provide inference on the delay in onset of immunity, based on an additional unknown parameter.

Our model did not capture a late peak of reported cases in the Hunter Valley region epidemic during the period when movement restrictions were purposefully relaxed in this region to minimise disruption to the Thoroughbred horse breeding season (Figure 7.6). If complete horse movement data were available characterisation of the epidemic process could be improved by

explicitly modelling such contact data, rather than only the known transmission network, as has been done for a simulation study of highly pathogenic avian influenza in the United Kingdom (Jewell et al., 2009b).

In both of the primary study regions, the inferred estimates of intra-premises basic reproductive ratio (R_0) were considerably higher than those reported in two earlier modelling studies (Glass et al., 2002, Satou and Nishiura, 2006). This is not unexpected considering that apart from a small number of recently imported horses (that would have been vaccinated in pre-export quarantine) the Australian horse population was immunologically naïve to equine influenza at the commencement of the 2007 outbreak. The disparity may also in part relate to differences in the population at risk datasets on which each study was based and differences in the modelling approach. The two previous modelling studies were populated with data on a small number of racetrack premises, each premises holding >100 horses. The dataset used in this analysis contained many more horse premises holding on average far fewer horses (92% of premises in Greater Sydney region held ≤10 horses at the time of the outbreak, whereas 80% of premises in the Hunter Valley held ≤10 horses). Furthermore, in contrast to the two previous studies, the unit of interest in this study was the horse premises not the individual horse. In such a highly susceptible population, all horses on a premises are treated as a unit for quarantine, vaccination and other outbreak control measures. Thus, the modelling approach was less focussed on estimation of intra-premises transmission rates *per se*, rather, the intra-premises SLIR model (and inference of β_{intra}) had two mechanistic purposes unrelated to inference of R_0: to provide a means of modelling a smooth premises-level infectivity function, and to facilitate a tractable method of thinning the stochastic Poisson process to maintain local time-homogeneity whilst simulating epidemics.

Sensitivity analysis showed that our model was insensitive to very large changes in β_{intra}. Interestingly, simulation outputs were relatively stable when we fixed β_{intra} so that $R_0 = 10$ after Glass et al. (2002); temporal outputs even improved marginally for the Hunter Valley, which

has a greater number of larger breeding premises. Conversely, modelling outputs were less well correlated with observed data when we fixed β_{intra} so that R_0 = 3.5 after Satou and Nishiura (2006), which suggests that 3.5 is too low a value for the R_0 of equine influenza virus in Australia. Indeed the findings of our analysis, and epidemiological reports of all horses being infected within 8–14 days of infection on premises holding several hundred horses during the 2007 outbreak, support an estimate of $R_0 \geq 10$.

Our model was highly sensitive to four inferred parameters (μ, δ, ξ and ζ) and the fixed intra-premises transmission parameter γ_{intra} which represents the recovery rate of individual horses. In each of the primary regions there appeared to be sufficient data to develop highly accurate inferences of μ, δ, ξ and ζ, with known uncertainty. The value of γ_{intra} was fixed based on substantial experimental evidence (control animals in equine influenza vaccination trials) that the duration of virus shedding of previously unexposed horses is 5–7 days (Daly et al., 2011, Mumford et al., 1990). In further research, our model could be improved by inferring two intra-premises transmission parameters γ_{intra} and β_{intra}.

The methods of sensitivity analysis could be improved by conducting global rather than local approaches that are either model independent or only assume monotonicity rather than linearity. Local methods are not strictly appropriate for non-linear and non-additive models, where independence and linearity can't be assumed. In further research, global approaches as reviewed by (Saltelli et al., 2004) will be applied to assess the presented model.

The method applied to assess spatial cross-correlation of the location of predicted infected premises between models and with observed data was subjective in that the grid-based approach (ρ_{grid}) requires a decision on the grid cell size. Conducting these comparisons at a higher spatial resolution may lead to lower estimates of spatial cross-correlation of modelling outputs. Inter-model spatial agreement has been assessed by applying Fleiss' Kappa with a similar grid-based approach (Dubé et al., 2007), and applying Cohen's Kappa to compare model outputs to

observed data (Garner et al., 2011b). These approaches involve categorisation of the cumulative incidence per cell, requiring yet another subjective decision. A kernel density estimation approach to graphical inter-model comparison has recently been implemented (Sanson et al., 2011b), however this method is dependent on the amount of smoothing applied.

Modelling the data by region reduced the considerable computational resource requirements involved in the Bayesian inference. The observed regional heterogeneity in inferred parameters is one reason why this is a desirable approach when modelling large multi-centric outbreaks. This research could be extended by inferring unknown parameters for all spatial clusters in the 2007 outbreak of equine influenza in Australia, within a single Bayesian RJ-MCMC modelling framework. Although computationally demanding, such an analysis would allow comparison of the joint parameter distributions both within and between groups of comparable regions, and increase the generalisability of the dynamic modelling framework to a broader range of regions.

The estimates in this study are conditional on the model formulation. A parsimonious model was intentionally formulated, with respect to the premises-level covariates required in its parameterisation. Model fit to the 2007 outbreak data may be improved by reformulating the model and forward simulating to test for improvements in correlation of the simulated model outputs with observed outbreak data. However, it must be considered that specifically replicating the 2007 outbreak was not the purpose of this modelling study.

On-farm management practices, meteorological and other known risk factors for equine influenza spread could be incorporated into the modelling approach, but such complexity increases the amount of data and number of assumptions required to parameterise such a model, thereby reducing its usefulness in real-time applications. For this reason, the modelling approach was tailored to the type of premises-level data anticipated in the event of a future outbreak of equine influenza in Australia.

7.5. Conclusions

In developing the model of equine influenza in Australia presented in this paper, probability distributions of key epidemiological parameters were inferred for two regions with very different horse industry structures, horse premises densities and geographic conditions. Spread dynamics were observed to differ markedly in a major horse-breeding and a peri-urban region highly affected by the 2007 outbreak of equine influenza in Australia, suggesting that different outbreak control procedures need to be prioritised in different areas during future outbreaks. Pseudo-validation of the model against data not used in its development, provided a demonstration of how our model may be applied to develop predictions during future outbreaks affecting horse populations in generally comparable regions to those studied. The inferred parameter distributions will be useful as informative priors in the Bayesian modelling framework developed for rapid assessment of future outbreaks of equine influenza in Australia.

7.6. Appendix

Stochastic simulation of local spread incorporating a smooth infectivity function

The following adaption of the standard stochastic simulation method for local spatial spread allows the incorporation of the smooth infectivity function by retrospective sampling (Jewell, 2009), ensuring that the Poisson process is appropriately thinned so as to maintain local time-homogeneity (Karr, 1991). A detailed derivation of these methods, which enable computationally efficient simulation based on the Exponential distribution, is provided by Jewell (2009).

Once a simulated infection had been stochastically seeded in the study region, based on the contact-traced network of horse movements forwards from the index premises, k, realised during the 2007 outbreak of equine influenza in Australia, time to the next infection was proposed, t^*, based on a Poisson Process with rate equal to the total infection pressure in the system at time t:

$$t^* = t + Exponential\left(\varepsilon_{(t)} + \sum_{i \epsilon I_{(t)}} \sum_{j \epsilon S_{(t)}} \beta_{ij(t)}\right) \quad (7.A1)$$

The next event was chosen depending on whether t^* was earlier than the minimum of the set of all other possible events (onset or removal time of an infected premises).

If t^* was the earliest time, then an infection was proposed as the next event and the proposed infectee, j^*, was stochastically selected from all of the susceptible premises, using an empirical distribution function (after Jewell, 2009) to weight the selection based on the amount of infection pressure exerted on each susceptible premises at time t. To maintain local time-homogeneity in the Poisson Process, so that inter-event times (as proposed) could be modelled using the Exponential distribution, retrospective sampling of the proposed infector, i^*, was conducted to appropriately thin the Poisson process (Karr, 1991). Thus, i^* was stochastically

selected from all of the infected premises, again using an empirical distribution function, this time weighting the selection based on the amount of infection pressure that each infected premises is exerting on j^*, the proposed infectee. Finally, the proportion of infected horses on the proposed infector premises i^* was treated as a probability, $P(h(t\text{-}I_i))$, and a number, u, randomly sampled from the range [0,1]. The proposed infection on j^* was only accepted to have occurred if $x<P(h(t:i))$, otherwise, the Poisson process was thinned by advancing time without an event occurring.

When an infection was accepted to occur, the state of j^* was changed from *S* to *I*. The time of onset of clinical signs in the first horse infected on j^* (*O*) was stochastically sampled from the distribution Gamma(α, λ) and was stored, and the standard deterministic within-premises SLIR was run (see equations 4.1 – 4.4), to produce infection times for each individual horse on the newly infected premises. These were also stored as these changes modulated the infection pressure in the system through the function $h(\cdot)$ in equation 3. The time of recovery (R_i) for the newly infected premises was estimated using the intra-premises SLIR and stored. Finally, time was advanced, infection pressures updated and a new infection proposed.

If the next event was not a proposed infection, rather a change of state of an infected premises (from *I* to *R*), the status of that premises was updated, time advanced accordingly, infection pressures recalculated, and a new infection proposed.

This process was repeated until either there were no infected premises remaining in the system, or for a maximum of 200 days, whichever occurred first. This study period was selected considering that the observed 2007 outbreak of equine influenza in Australia lasted 131 days.

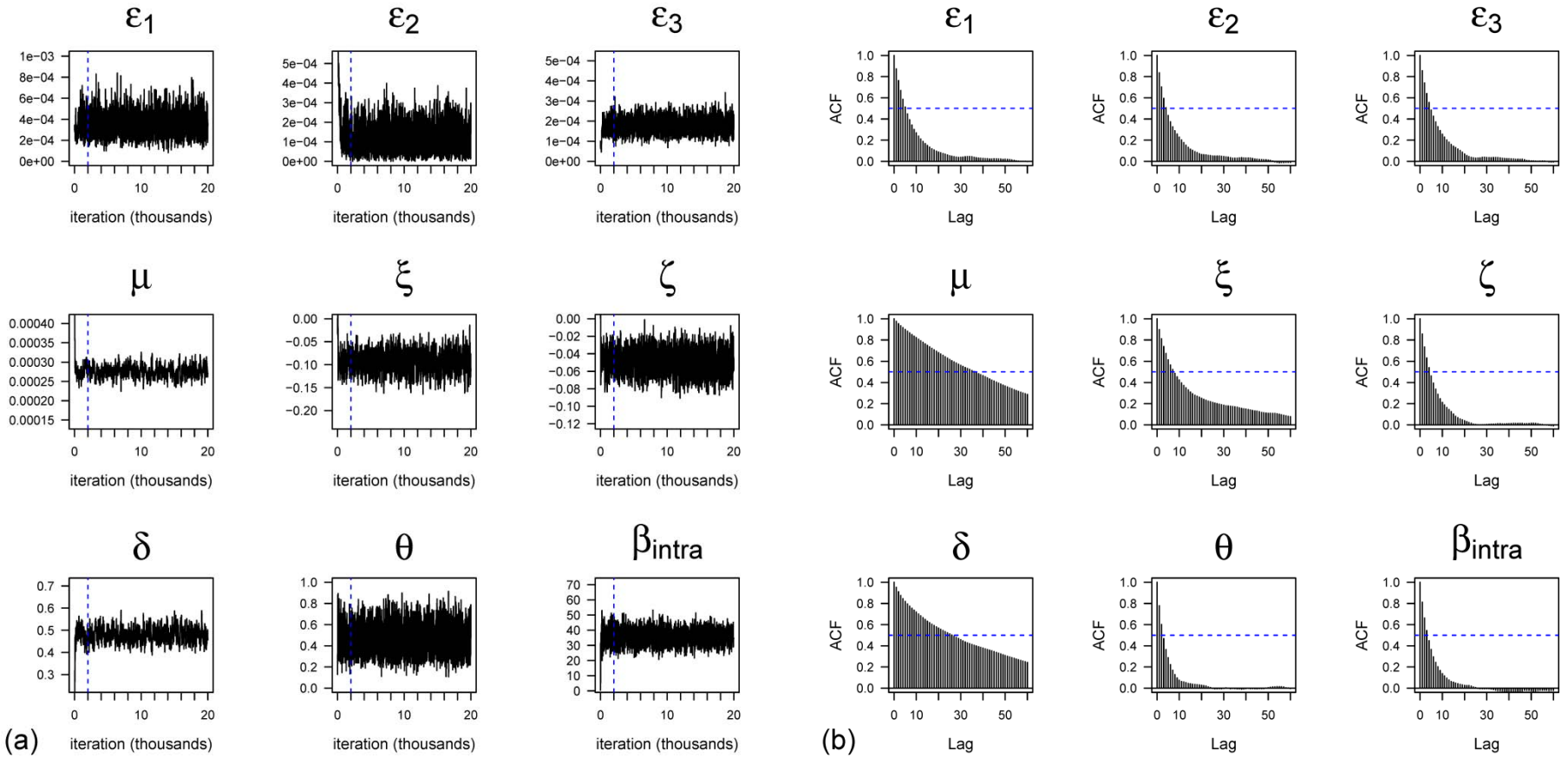


Figure 7.9 Convergence and thinning of the joint posterior probability distributions for the Greater Sydney region.
(a) Trace plots showing mixing and convergence of the Bayesian reversible-jump Markov Chain Monte Carlo algorithm. Reference line depicts the applied burn-in of 2000 iterations. (b) Auto-correlation function (ACF) plots, showed that auto-correlation was low (<0.5) after thinning by 60 iterations.

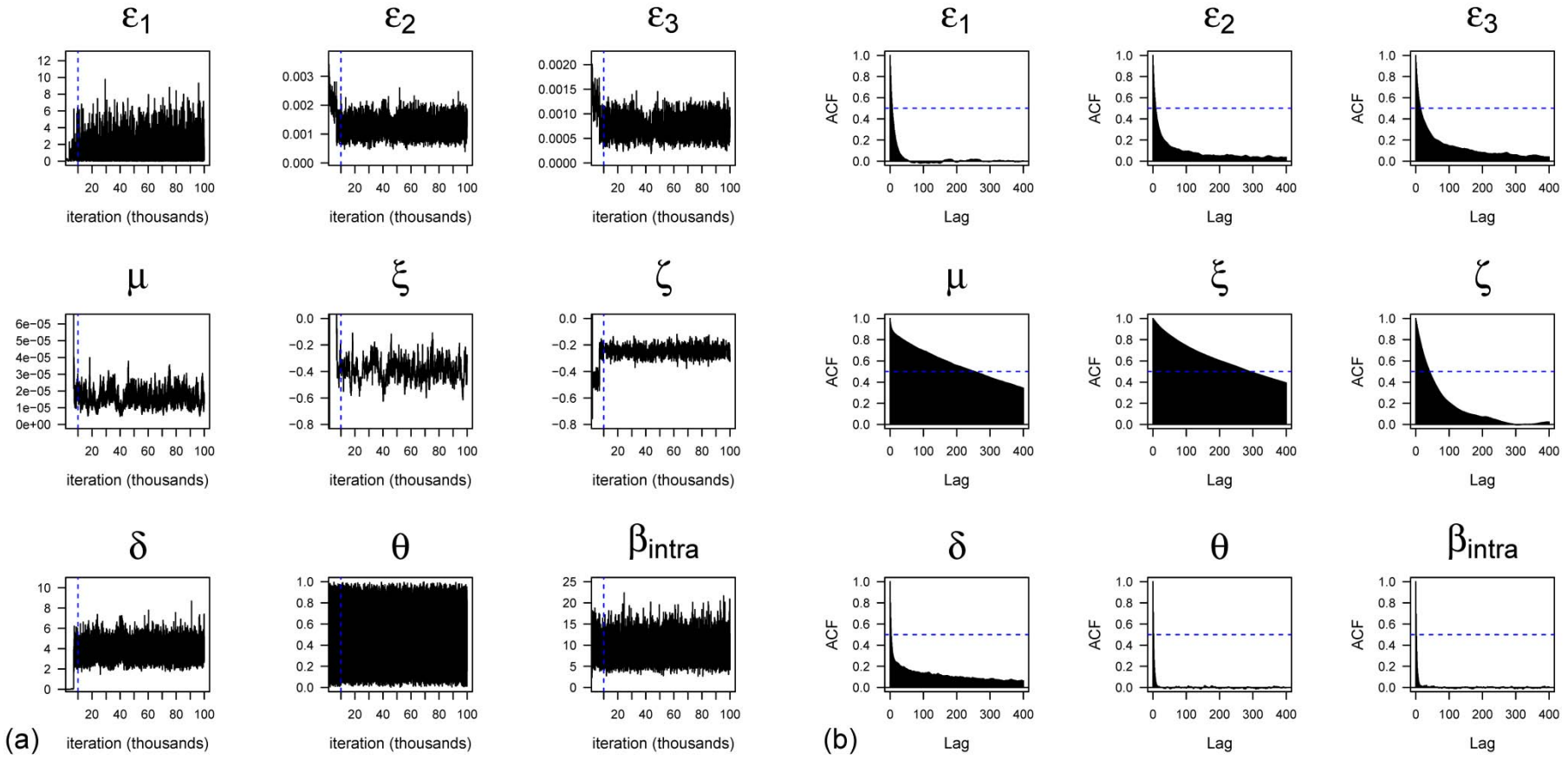


Figure 7.10 Convergence and thinning of the joint posterior probability distributions for the Hunter Valley region.
(a) Trace plots showing mixing and convergence of the Bayesian reversible-jump Markov Chain Monte Carlo algorithm. Reference line depicts the applied burn-in of 10,000 iterations. (b) Auto-correlation function (ACF) plots, showed that auto-correlation was low (<0.5) after thinning by 300 iterations.

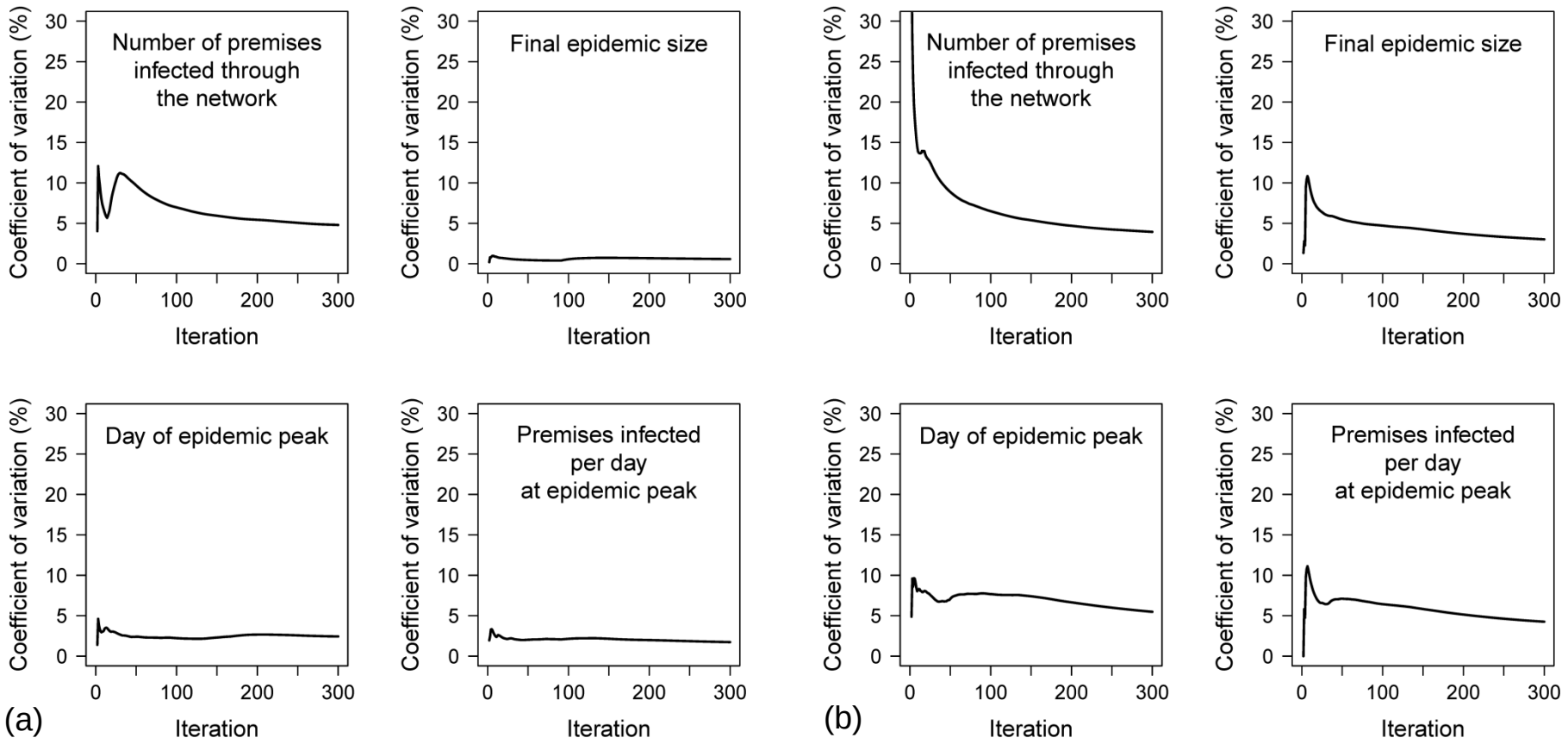


Figure 7.11 *Post hoc* assessment of the number of iterations required to achieve stable predictions for stochastic simulation. (a) The Greater Sydney region, and (b) the Hunter Valley region.

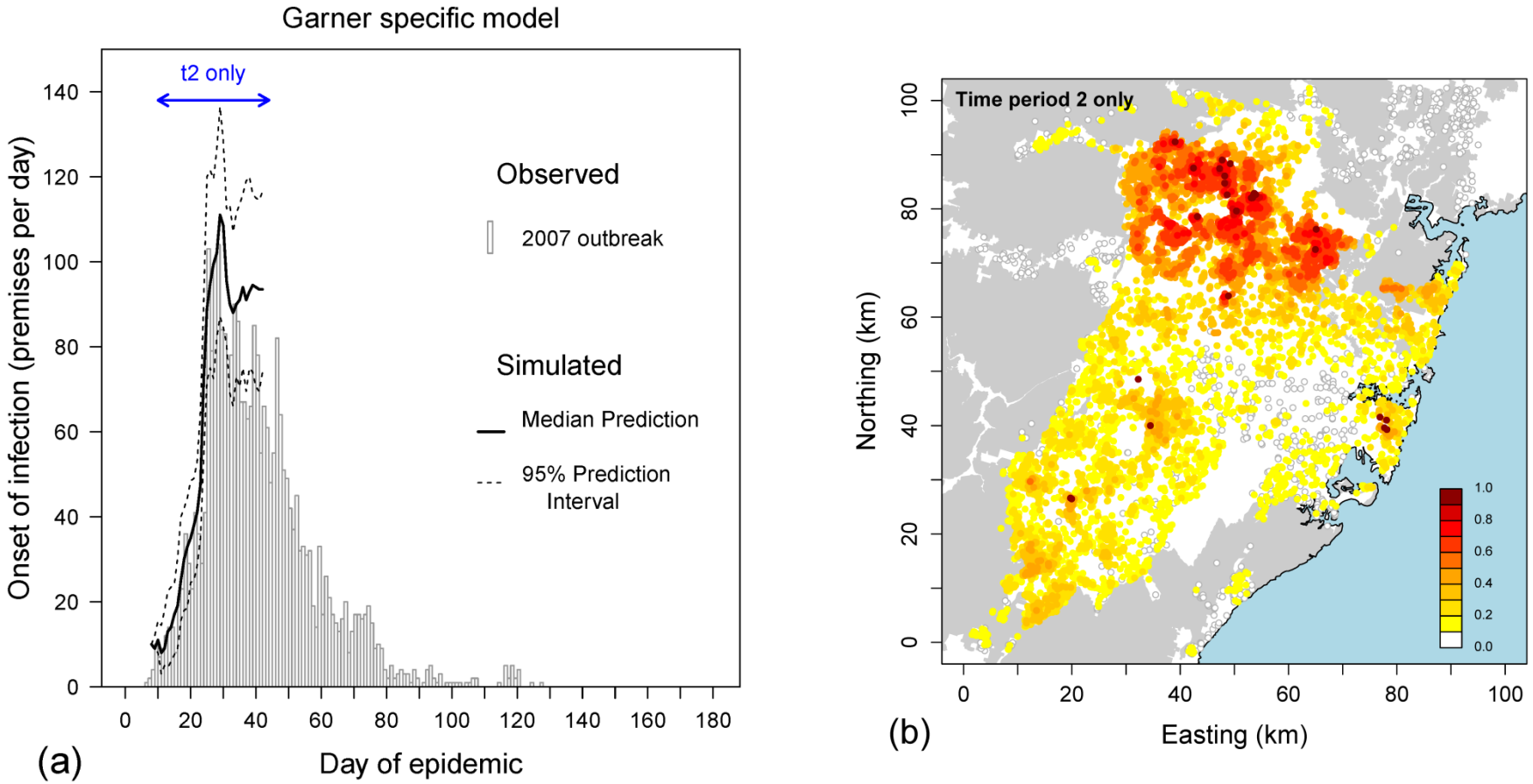


Figure 7.12 Simulation outputs of the baseline model by Garner et al. (2011b).
(a) Simulated epidemic curve produced by 100 iterations with 95% credible intervals, compared to the actual realisation observed in the Greater Sydney region during the 2007 equine influenza outbreak in Australia. The Garner model was seeded based on the actual outbreak data at day 9, and only run during time period 2 ($10 \leq t < 44$), between the blue arrow. (b) Proportion of runs in which a premises was infected.

Chapter 8: General discussion

8. General discussion

The 2007 outbreak of equine influenza was the largest emergency animal disease outbreak in Australia's history. The horse industry in affected areas suffered substantial hardships and losses owing to the outbreak and the response measures required to control it. A clear understanding of the mechanisms underlying the rapid spread of this outbreak, and how each of the response measures contributed to the control and eradication of equine influenza from the Australian horse herd is important for informing future outbreak responses. A large amount of epidemiological investigation and analysis was conducted during and immediately after the outbreak, however many questions remained unanswered. The series of epidemiological studies that comprise this thesis were therefore conducted to further investigate the factors that facilitated or prevented the spread of equine influenza in Australia; thereby contributing to our understanding of how similar rapidly-spreading outbreaks could be better controlled and more rapidly contained. The principal significance of each of the key findings of this thesis will be discussed in turn, identifying areas requiring further research and recommendations for the control of future outbreaks of equine influenza. Many of the recommendations are applicable to other outbreaks of infectious diseases in animals.

8.1. Biosecurity compliance and other premises-level risk factors

Equine influenza virus is known to spread over short ranges by direct contact, in cough droplets or aerosols (Hannant and Mumford, 1996), and also by the movement of infected animals and on fomites (Guthrie, 2006, Guthrie et al., 1999, Gildea et al., 2011). Prior to this thesis no objective information was available about whether the adoption of biosecurity measures could in fact prevent the spread of equine influenza infection through these routes. Biosecurity measures such as personal hygiene, equipment hygiene and access control measures were recommended to horse owners, however, there were conflicting anecdotal reports about their effectiveness during the 2007 outbreak (Arthur and Suann, 2011, Frazer et al., 2011, Kung et al., 2011).

The case-control study of premises-level risk factors (Chapter 3) quantified the association between compliance with advised on-farm biosecurity measures and prevention of spread of equine influenza infection onto horse premises in highly affected areas. Compliance with measures aimed at reducing the risk of transmission on fomites was associated with a nearly four-fold reduction in odds of infection. These findings were consistent with the experience of the horse owners and managers interviewed, most of whom (>80%) considered that the recommended on-farm biosecurity measures would be very effective in the event of a future outbreak (Schemann et al., 2012). Furthermore, an analysis (Gildea et al., 2011) published around the same time as the case-control study found that biosecurity measures were effective in preventing the spread of equine influenza in Ireland, where equine influenza is endemic.

There was a high amount of correlation amongst several variables representing personal 'barrier hygiene' biosecurity measures (hand-washing, changing clothes and shoes, and having a footbath in place), all however, were generally protective. Further research is required to establish the effectiveness of each individual measure; this is complicated by their non-independence because practicing one biosecurity measure was found to be associated with the practicing of others. Nevertheless, the findings of the case-control study showed that vigilant compliance with on-farm biosecurity can prevent infection, and that these measures should be recommended to horse owners to limit spread in any future outbreak.

A consistent finding across the two risk factor studies (Chapters 3 & 6) was that premises holding higher numbers of horses were at increased risk of infection, and that there was less risk of infection on larger area premises. Bayesian inference of data from two regions (Chapter 7) complemented these findings. Horses kept in high density housing have previously been identified to be at increased risk of equine influenza infection (Gildea et al., 2011, Morley et al., 2000b). Under-reporting of infection on premises holding fewer horses or on larger area premises may be an alternative explanation for these findings. A study into the level of under-

reporting during the 2007 outbreak found that the proportion of under-reporting was highest in rural clusters (Dhand and Sergeant, 2011a), where horse premises are larger in area. The regional differences in under-reporting did not appear to be related to the number of horses held on a premises, and the overall level of under-reporting was estimated to be very low (1.2%) (Dhand and Sergeant, 2011a). Therefore it is recommended that during future outbreaks surveillance and strict quarantine should be targeted at medium to higher density horse premises, such as those holding many horses in small areas or in stable complexes.

A measurement bias may have been introduced during the process of cleaning the data to consolidate duplicated uninfected premises records based on cadastral property boundaries (Chapter 2) leading to under-estimation of herd size on some multi-owner horse premises. It was not possible to assess the frequency or magnitude of this bias because records were de-identified for privacy reasons. It would only have affected estimates in the survival analysis (Chapter 6) and Bayesian inference (Chapter 7), and even then only for a proportion of the premises that had multiple owners.

Once a horse was infected on a premises, infection rapidly spread to almost all horses on the premises irrespective of individual animal risk factors. Little association was detected in the case-control study (Chapter 3) between infection status and the age and sex structure of herds. This is understandable considering the immunologically naïve status of almost the entire Australian horse population prior to this outbreak. As very few horses had a history of vaccination, individual animal-level risk factors were relatively less important in determining which horses were infected. This is consistent with the very high proportions of infected horses ('attack rates') observed on infected premises (>96%) (Dhand and Sergeant, 2011b), the high intra-premises reproductive ratios (R_0 >10) estimated in the Bayesian inference of two regions (Chapter 7), and detailed descriptions from confined horse populations infected during the outbreak (Wong, 2011, Britton, 2007, Equine Influenza Epidemiology Support Group, 2008, Kung et al., 2011, Morton et al., 2011, Ryan, 2011). In countries where equine influenza is

endemic, and vaccination is practiced, the age, sex and vaccination history of individual horses have previously been identified as important risk factors for equine influenza transmission (Gildea et al., 2011, Barquero et al., 2007). Considering that vaccination for equine influenza is currently not practiced in Australia (in the absence of infection), and that any future outbreak strain is likely to be antigenically distinct from the 2007 outbreak strain, the age and sex of individual horses are unlikely to be any more important as risk factors for equine influenza infection in a future outbreak in Australia.

The case-control study was intentionally designed to minimise the potential for poor recall, including the use of on-farm interviews, open discussion leading up to questioning of memorable anchor points, referring to available documentation and veterinary records, and focussing questions on horse management practices during a specific week prior to a memorable event. Respondents from case and control premises were similarly precise in their ability to recall key dates and events, and recall bias was considered to have had a negligible effect on the observed results. Cases and controls were sourced from a laboratory testing database to ensure that misclassification of outcome status did not occur and that control premises had a memorable event to structure interviews around (the testing of horses on the premises).

Confounding is a major issue in risk factor studies due to the non-random allocation of subjects to study groups. The case-control study period was restricted to the period prior to the implementation of vaccination and the staged relaxation of movement restrictions, to eliminate confounding associated with these variables. Data were collected at interview on all other factors considered *a priori* as likely confounders, and multivariable modelling was used to test and adjust for potential confounding in the case-control study analysis. A decision was made not to match, although this was considered when designing the case-control study. Reasons for not matching included: uncertainty over exactly which factors would be controlled for by matching on location (potential 'overmatching') and that matching may introduce selection bias into case-control studies (Dohoo et al., 2009). A term representing the proximity of case and control

premises to the nearest infected premises was included in multivariable analyses, this allowed the infection risk at a range of distances from infected premises to be adjusted for (and quantified) when estimating the protective effect of biosecurity practices. Clustering of observations from the same geographic region was identified as potentially important, *a priori*. A region-level random effect term was therefore included in multivariable analysis; this both confirmed and adjusted for clustering at this level.

8.2. Windborne spread of equine influenza and other meteorological influences

The role of wind in the spread of equine influenza in Australia was a major knowledge gap prior to this thesis. Research into windborne spread of diseases cannot be easily accomplished in a laboratory, and evidence for windborne spread of equine influenza from observational epidemiological studies was weak. Although there was consensus in the literature that airborne transmission of influenza viruses was at least possible, there was strong disagreement about its importance (Weber and Stilianakis, 2008). Many instances of possible windborne spread were noted during the 2007 outbreak of equine influenza in Australia (Moloney et al., 2011), mostly where no other potential means of transmission could be identified. One study presented circumstantial evidence that the direction of prevailing winds was similar to the direction that statistically significant clusters arose in a highly affected region of Queensland (Davis et al., 2009).

In order to statistically test for an association between risk of equine influenza infection and premises-level wind conditions, a method of generating variables to represent daily windborne disease exposure risk was required that could be implemented on a dataset containing >1000 infected premises.

A variety of methods have been used to investigate the windborne spread of outbreaks. Previous analyses have presented: circumstantial evidence that the mean direction of spread of an outbreak has coincided with prevailing wind conditions at the time (Davis et al., 2009, Kedmi et

al., 2010), analysed data aggregated to a low temporal or spatial resolution (du Prel et al., 2009, Yuan et al., 2006), or used atmospheric dispersal modelling to investigate spread from a small number of sources (Gloster et al., 2011, Ssematimba et al., 2012). A small yet innovative case-control study (Johnson et al., 2005) investigated the spread of an outbreak of infectious laryngotracheitis (ILT) between commercial poultry flocks by using plume modelling from a single weather station to estimate a dichotomous covariate representing whether subjects were downwind of other infected premises on each day of the study period. Johnson et al (2005) then conducted hypothesis-testing on the association between the binary variable representing windborne exposure risk and ILT infection for the 18 cases and 122 control premises, and found that case premises were nine times more likely than controls to have been located within the estimated wind vector of another case premises. None of these methods appeared appropriate for studying a dataset as large as the equine influenza outbreak dataset.

Through the combined application of geostatistics and survival analysis, an approach was developed to generate premises-level covariates to represent the maximum daily wind speed from the direction of nearby infected premises, based on spatially smoothed hourly meteorological data (Chapter 6). Using generalised Cox regression modelling, it was then possible to statistically test the hypothesis that hazard of infection was increased on individual premises on days when wind speeds were higher from the direction of nearby infected premises. An association was detected between hazard of infection and wind speed from the direction of proximate infected premises. Sensitivity analysis showed this finding to be robust to the assumptions made whilst generating directional wind speed covariates. The results of the survival analysis (Chapter 6) therefore advanced understanding of the windborne spread of equine influenza from purely circumstantial association and anecdote to a hypothesis statistically-tested with empirical data. The implication of these findings is that no matter how stringent movement restrictions and on-farm biosecurity practices are, some local spread of equine influenza is going to occur by windborne aerosol. Further research is however required, to ascertain whether similar estimates of windborne spread of equine influenza are observable in

broadly dissimilar environments. This could be implemented by analysing other clusters of the 2007 equine influenza outbreak. In any cluster in which an overall direction of spread is detected, it will be important to ascertain the proportion of anisotropic spread that is directly attributable to windborne disease spread.

Two other meteorological variables were identified as important risk factors for equine influenza infection: air temperature and relative humidity. The findings of the Cox regression modelling of time to infection with equine influenza in the largest cluster of the 2007 outbreak in Australia (Chapter 6) were highly comparable with those of experimental research. Hazard of equine influenza infection was highest when relative humidity was <60% and lowest on days when daily maximum air temperature was 20–25 °C.

The dependence of human influenza virus transmission on relative humidity and temperature has been the focus of considerable research under controlled laboratory conditions (Lowen et al., 2007, Lowen et al., 2008, Steel et al., 2011, Shaman and Kohn, 2009, McDevitt et al., 2010). Virus survival and aerosol droplet nuclei stability have been used to explain differences in the aerosol transmission of influenza A viruses in different conditions: droplets are desiccated in dry conditions and remain small, which may stabilise influenza aerosols and facilitate longer range transmission, whereas at high relative humidity, the droplets absorb water and settle (Lowen et al., 2007, Tellier, 2006, Tellier, 2009). Lowen et al. (2007) have demonstrated in experimental infection trials that the aerosol transmission of influenza A viruses between guinea pigs is favoured by cooler conditions. This may be related to increased survival of these enveloped viruses, and the longer duration of peak viral shedding observed when infected animals were kept in cold conditions (5 °C) (Lowen et al., 2007).

The increase in hazard of infection at air temperatures over 25 °C (Chapter 6) is not easily explained and requires further research. Lowen et al. (2008) demonstrated in a further set of transmission experiments that whilst higher temperatures block aerosol spread of influenza A,

contact transmission was unaffected by air temperatures >30 °C. The survival analysis was conducted on a dense cluster of infected premises in the Special Restricted Area (purple zone). Movement restrictions were relaxed in this area (NSW DPI, 2007b) from the 36th day of the outbreak (21 September 2007), consequently it is possible that the finding of increased hazard at higher temperatures was confounded by increased movement of infected horses later in the outbreak, when conditions were warmer. Restricting the survival analysis study period to the 26 days when there were presumably no movements at all (between days 11–36) would exclude the peak in this cluster. Repeating the survival analysis in additional clusters (in the red zone, where movement restrictions were not relaxed until much later) is therefore required to further establish the influence of the meteorological factors on equine influenza spread. Any such research will need to be conducted on another cluster of infected premises for which meteorological data can be obtained from a sufficient number of nearby weather stations.

As the 2007 outbreak of equine influenza progressed, the change of season and resulting changes in air temperature and humidity may have affected rates of spread. Whilst contact transmission would have continued unabated in all conditions, whenever humidity or air temperature rose, there would have been less potential for aerosol transmission between premises (Lowen et al., 2008). This has to do both with the enveloped nature of influenza viruses, making them viable for shorter times in warmer conditions, and factors affecting the survival of the vehicle of their airborne transmission (drop nuclei). Importantly, certain fomites may protect the virus from environmental desiccation increasing the period of viability even in warm conditions. The implication of these findings is that disease spread will be harder to contain in cooler conditions, however, when humidity or air temperature are high, the priority should be on measures that target direct contact and transmission on fomites (i.e. reducing contact between horses on adjoining premises and disinfection of potential fomites). If sufficiently detailed and high quality meteorological data are available on air temperature, relative humidity and wind velocity, then these meteorological factors are worth considering for inclusion in future stochastic simulation models of equine influenza outbreaks. Improved

understanding of the importance of certain modes of transmission given the weather conditions at the time of an outbreak will be highly informative to those responsible for managing the outbreak response.

8.3. Characterisation of disease spread through contact networks and space

Considered at the premises-level, equine influenza can gain entry to a herd of uninfected animals by the introduction of an infected horse(s), direct or close range contact with horses on contiguous premises, aerosol (or airborne) transmission from nearby premises, or over variable distances on fomites (people, equipment or vehicles). The premises-level covariate most strongly associated with infection status in the case-control study (Chapter 3) was proximity to the nearest infected premises. Two effects were observed, presumably representing local spread and long distance spread (prior to complete implementation of the horse movement ban). Cases were more likely than controls to be within 5 km of an infected premises. This finding was consistent with earlier studies of local spread during the 2007 outbreak (Cowled et al., 2009b, Davis et al., 2009), and previous estimates from outbreaks of equine influenza overseas (Dalglish, 1992, Huntington, 1990). Longer distance spread (>10 km from an infected premises) was associated with the movement of infected horses, presumably incubating equine influenza whilst returning from events where transmission is known to have occurred.

Although local and long distance spread appeared distinct, it was not possible in the case-control study to identify the mechanism of transmission of equine influenza to individual premises, especially when a premises had received horses from a known infected premises around the same time that other infected premises were reported nearby. Identification of the most likely means of spread to such premises was required to estimate the importance of animal movements in the spread of the outbreak.

Early in the 2007 equine influenza outbreak, animal movement networks ('contact-networks') seeded clusters of cases in widespread locations (Chapter 4), dictating the spatial extent of the

ensuing outbreak. To improve the characterisation of disease spread, given that a contact-traced network was known to underlie early spread, spatial and social network analysis were applied in combination to explicitly describe the transmission network of infected animal movements and spatial relationships between newly infected premises (Chapters 4 & 5). Social network analysis is suited to handling several complex relationships between epidemiological units of interest (in this case, horse premises) and has recently been applied to analyse spatial and contact relationships between animal holdings in studies of bovine tuberculosis (Green et al., 2008) and endemic disease spread (García Álvarez et al., 2011), and to develop hypothetical models of how diseases might spread through animal populations (Webb, 2005, Martinez-Lopez et al., 2009a, Lockhart et al., 2010).

The combined application of spatial and social network analyses enabled estimation of the risk of local spread by spatial distance from infected premises ('the spatial transmission kernel') whilst accounting for the underlying contact network of infected horse movements (Chapters 4 & 5). Extension of these methods enabled likelihood-based estimation that 28.3% of early disease spread (prior to the implementation of horse movement restrictions) was through the movement of infected horses (Chapter 4). Premises infected by this route were central to the seeding of infection in widespread sites (Chapter 5). Using the likelihood-based approach the 56 horse premises most likely to have been infected through horse movements were identified, as were the 44 distinct spatial clusters seeded by these movements. These findings extend those of Cowled et al. (2009), who estimated that whilst only 2.5% of the total number of infected premises had onset dates in the first 2 weeks of the outbreak, these premises were dispersed across 83% of eventual extent of the outbreak-affected area. Clearly, if the implementation of horse movement restrictions had been delayed, further movements of horses incubating equine influenza would have spread disease to new areas. This would have made disease containment and eradication more difficult (perhaps impossible) to achieve.

These methods were extended to develop a mixed transmission model of the sequence of cluster

formation, incorporating spatial relationships between infected premises and contact network topology. Describing the underlying contact network structure was important when delineating clusters (Chapter 5), and the seeding of stochastic simulations based on the contact network data was very effective in modelling disease spread during the early dissemination phase of the outbreak (Chapter 7). These analyses demonstrated how best to use contact-tracing data to identify which premises are most important to the widespread dissemination of an outbreak in a highly susceptible population. In future outbreaks of equine influenza or other emergency animal diseases, similar methods could be applied to inform the risk-based targeting of surveillance and control measures (such as the 'predictive' vaccination of animals on premises identified as highly likely to be central to disease spread).

Although the contact network dictated where clusters were seeded, local spatial spread was the most common means of infection throughout the entire period of the outbreak. Spatial proximity to an infected premises was critically important, so a considerable focus of the research in this thesis involved describing the risk of infection by distance from infected premises. Consistent estimates of the range of local spread were obtained in the case-control study (Chapter 3), the two spatial social network analyses (Chapters 4 & 5), and through Bayesian inference of the shape of the transmission kernel based on estimation of the distance decay parameter (Chapter 7). In all cases, local spread was found to occur most commonly within 5 km of an infected premises, however, a non-negligible risk of local spread existed out to around 20 km. Regional heterogeneity in these estimates was observed in the Bayesian inference in two regions with markedly different horse industry characteristics. Local spread within a high density peri-urban region mostly occurred within 3 km of infected premises, whilst the effective range of local spread in a major Thoroughbred horse breeding region was 16 km.

The present Australian policy for halting the spread of any further equine influenza outbreaks is to implement immediate movement restrictions, and then target vaccination and surveillance within 10 km of infected premises, together with ring vaccination inwards from an outer buffer

zone at least 20 km from infected premises (Animal Health Australia, 2011). This strategy appears appropriate for high density peri-urban regions, however, there is a non-negligible probability that local spread will rapidly move beyond a 20 km buffer in major horse-breeding and rural areas. This has been a consistent finding across the thesis, and is supported by accounts of field experiences in south-east Queensland during the 2007 outbreak (The EI Epidemiology Support Group, 2008). Therefore, extended buffer may be required to contain infection in such areas during any future outbreaks of equine influenza in Australia.

The network analyses were limited by the completeness of the provided contact-tracing data. Despite considerable resources being devoted to investigation and contact-tracing early in this outbreak some infected horse movements were no doubt missing from the dataset, which may have obscured important contact network links and led to under-estimation of the importance of the movement of infected animals based on the combined spatial and social network analysis approach.

When attempting to map the progression of future EAD outbreaks, the network analysis methods presented in this thesis could be rapidly applied to contact-tracing and population at risk data to describe the spatial and contact relationships between all newly identified infected and suspect premises, delineate clusters and estimate the range of local spread in different regions. Such analyses would greatly inform disease containment and control efforts.

8.4. Development of a dynamic modelling framework for rapid assessment of future outbreaks

Equine influenza virus is a highly unpredictable agent (Daly et al., 2011). Recent outbreak experience has demonstrated that it may mutate and present novel variants to which even vaccinated populations are highly susceptible (Newton et al., 2006). A future incursion of equine influenza in Australia is possible through a range of routes, and such an outbreak may

not present in a similar form to the last, complicating the decision-making required in outbreak response and control.

Drawing on the findings of the epidemiological studies presented in Chapters 3 to 6, a spatially-explicit stochastic epidemic model of equine influenza transmission was developed to estimate key epidemiological parameters that could be used to inform predictive modelling of the course of future outbreaks of equine influenza in Australia. The Bayesian reversible jump Markov chain Monte Carlo algorithms used in Chapter 7 have only recently been applied to infer probability distributions for transmission rates in epidemics (Neal and Roberts, 2004, O'Neill and Roberts, 1999, Jewell et al., 2009a, Minh et al., 2011, Jewell et al., 2009b). The model was intentionally formulated in a parsimonious manner and tailored to the type of premises-level and contact-tracing data anticipated to be available early in an outbreak. The entire 2007 equine influenza outbreak dataset was used to train the model in the Greater Sydney and Hunter Valley regions. An important research question still to be completely addressed is at what stage into a future outbreak will it be possible to make reasonable predictions based on biologically plausible priors (rather than analysing based on entire outbreak datasets). An attempt to answer this question has been made by assessing parameter learning based on simulated epidemic data (Jewell et al., 2009b), and it would be possible to extend this research based on the equine influenza outbreak dataset.

The simulation modelling outputs produced were highly comparable to the single actual realisation observed in the three study regions during the 2007 equine influenza outbreak. In any future equine influenza outbreaks in Australia, the fitted probability distributions for key epidemiological parameters from the two primary study regions can be utilised by incorporating them as informative priors into Bayesian predictive modelling of the course of the outbreak. Assessment of the usefulness of such a Bayesian framework requires application during a large-scale emergency animal disease outbreak. In such a setting, obtaining detailed and high quality contact-tracing and outbreaks surveillance data is a non-trivial exercise. Delays in accessing

data from the field, and the quality of such data in the early stages of a large outbreak impact on the ability to develop meaningful predictions of the course of the outbreak. During the 2007 foot-and-mouth disease outbreak in the United Kingdom, this approach was used to predict risk and identify infected but as yet undetected premises based on parameters fitted to data from the 2001 outbreak (Jewell et al., 2009a). As the Bayesian inference was only conducted in two regions, further research is required to increase the generalisability of the dynamic modelling framework by estimating probability distributions for the required model parameters using data from several further highly-affected regions in the 2007 outbreak.

Regional heterogeneity in the underlying epidemic process was observed in the studied regions. The importance of this finding is that different outbreak control procedures may need to be prioritised in different areas. Transmission over longer distances in rural and major Thoroughbred horse breeding areas after horse movements had ceased (presumably on fomites), suggests that limiting human contact with susceptible horses, and implementing strict on-farm biosecurity will help to prevent spread in these areas.

The Bayesian modelling presented in Chapter 7 was limited in that it did not explicitly incorporate contact relationships between premises. Rather, contact-tracing data was used only for seeding simulations in disparate clusters, based on likelihood-based estimates from the network analysis (Chapter 4) of the proportion of horse movements out of infected premises that resulted in infection at the destination premises. This approach worked effectively for the limited purpose of the simulation model developed in Chapter 7, but is a crude method of seeding infections in the study regions and is not predictive of the range of scenarios that may develop *prior* to outbreak detection and the implementation of a horse movement standstill. Contact and spatial relationships between premises have been explicitly incorporated into a similarly structured Bayesian modelling framework (Jewell et al., 2009b), however that study was based on a simulated H5N1 avian influenza outbreak dataset, rather than actual contact-tracing data. Modelling the entire contact network in a population (rather than only the mixed

transmission network emanating from infected premises) requires a representative sample of all contacts. Contact-tracing conducted during a large outbreak is not going to produce such data. Data from animal identification and movement-tracking systems are required, such as the National Livestock Identification System (NLIS) for cattle and sheep movements in Australia. Possessing such high quality data on daily movements of susceptible species enables prediction of the set of outbreak scenarios of where disease might spread prior to detection, based on seeding at different locations and different times. Most Australian States have recently implemented mandatory registration of premises holding horses under the Property Identification Code (PIC) system. The feasibility of recording horse movements under the NLIS is a topic for future research.

An improved understanding of the distribution of the Australian horse population, horse movements and other contacts between horse premises are required to better inform risk prediction of the course of a future outbreak of equine influenza (or another EAD affecting horses such as West Nile /Kunjin virus or African horse sickness). Registration of horse ownership (online mechanisms and collaboration with racing, breeding and equestrian associations) and horse movements (Travelling horse statements) were important means of gaining this data in the 2007 outbreak. Following the eradication of equine influenza from Australia, the systematic collection and collation of data on the movements of horses was discontinued. Although most States now require the registration of all properties that keep horses (along with all other classes of livestock), this is not compulsory in all States. A consistent national approach to the collection of data on the distribution and demographics of the Australian horse population is required to enable rapid identification of the population at risk in any future outbreaks of emergency animal diseases affecting horses. Any further research attempting to extend the modelling approach taken in this thesis by explicitly incorporating contact relationships will also require more representative data on the distribution of horses in Australia and the contacts between horse premises.

8.5. Conclusions and recommendations

The research included in this thesis has elucidated the key factors underlying the spread of the 2007 equine influenza outbreak in Australia, and presented new methods of describing such rapidly spreading epidemics. The movement of infected horses, meteorological variables (air temperature, humidity and wind speed), on-farm biosecurity measures and intrinsic features of horse premises (proximity to other infected premises, numbers of horses held and premises area) were all important variables that influenced the spread of infection onto horse premises. The following conclusions can be made:

- On-farm biosecurity measures were effective in preventing equine influenza spread onto horse premises in highly affected areas. Compliance with advised on-farm biosecurity measures by horse owners, managers, veterinarians and others visiting horse premises in highly affected areas was associated with a four-fold reduction in odds of infection;
- Risk of infection was higher on premises holding increased numbers of horses;
- Horses on larger area premises were less likely to be infected, and less likely to infect horses on nearby premises once infectious, however this finding may be confounded by under-reporting in rural areas;
- The contact network of infected horse movements prior to disease detection described the sequence of cluster formation and the geographic extent of spread;
- Purely spatial methods of modelling epidemic spread performed poorly in the early epidemic phase, when a contact network was underlying the widespread dissemination of infection;
- A likelihood-based approach linking spatial and social network analyses allowed inference on the mode of transmission (infected animal movements versus local spatial spread) to infected premises in the early epidemic period. Subsequently, 28% of disease spread prior to the implementation of horse movement restrictions in the 2007 outbreak of equine influenza was identified to be associated with the movement of infected horses;

- Most local spread occurred within 5 km of infected premises, however there was regional heterogeneity in the risk of infection by distance from infected premises and there was a non-negligible risk of infection (presumably on fomites) up to 20 km from infected premises in rural and major Thoroughbred horse breeding areas; and
- Local spatial spread of equine influenza was dependent on relative humidity, air temperature and wind velocity. Wind speeds >30 km hour^{-1} from the direction of nearby infected premises were associated with increased hazard of infection. The findings of experiments conducted under controlled conditions were supported, in that hazard of equine influenza infection was higher when relative humidity was <60% and lowest on days when daily maximum air temperature was 20–25°C.

Based on the research presented in this thesis the following recommendations are made to agencies responsible for managing future outbreaks of equine influenza in Australia:

- Compliance with the advised on-farm biosecurity measures prevented spread of infection during the 2007 outbreak, therefore such measures should again be recommended to horse owners to limit spread in any future outbreak of equine influenza in Australia;
- Likelihood-based approaches linking spatial and social network analyses should be applied to describe the sequence of cluster formation, estimate the effective range of local spread, and infer whether newly detected infected premises were infected by local spatial spread or through other means (movements of infected animals or transmission on fomites);
- Windborne spread of equine influenza was found to be important over short ranges, therefore movement restrictions and on-farm biosecurity can never be expected to completely contain disease spread;

- Meteorological factors had an important influence on the spread of this outbreak and should be considered for inclusion in stochastic simulation modelling of future outbreaks of equine influenza, if sufficient high quality data is available;
- The implementation of control zones and vaccination buffers should consider regional heterogeneity in effective range of local spread. Zones may need to be based on 20 km buffers in rural regions and major Thoroughbred breeding areas;
- Bayesian inference and risk prediction should be implemented to produce rapid assessments of the course of future outbreaks of equine influenza in Australia. Risk predictions may be made based on the prior probability distributions estimated in this thesis, from data on the 2007 equine influenza outbreak, and updated for individual premises as the outbreak progresses. The meaningfulness of the predictions produced will depend on the timeliness and quality of outbreak surveillance, contact-tracing and population at risk data available early in the outbreak, and how comparable the regions are to those already studied;
- A consistent national approach is required for the collection and collation of data on the distribution and demographics of the Australian horse population, and horse movements, to enable the rapid identification of the population at risk and prediction of the range of scenarios of where disease may spread prior to detection in future outbreaks of emergency animal diseases affecting horses; and
- An approach to automated geocoding and data checking against cadastral property boundaries should be considered when designing surveillance databases for such emergency animal disease events, considering the amounts of duplication identified in the equine influenza outbreak dataset, even after intensive data cleaning efforts.

Further research could be conducted into the following aspects of the epidemiology of the 2007 outbreak of equine influenza in Australia:

- Estimation of the effectiveness of each of the non-independent on-farm biosecurity measures using graphical Bayesian network modelling or a similar statistical approach on the case-control study data;
- Repeating the cohort study into factors associated with infection at the first equestrian event where transmission was known to occur, with a revised outcome measure, now that the premises most likely to have been infected through the movement of infected horses have been identified;
- Further establishing the influence of meteorological factors on equine influenza spread by conducting survival analysis in additional clusters;
- Atmospheric dispersal modelling in several small rural clusters of infected premises;
- Estimating the intra-premises basic reproductive ratio for each of the large closed populations investigated early in the outbreak; and
- Further generalising the dynamic modelling framework by Bayesian inference on data from several additional highly-affected regions.

Bibliography

AHMED, S. S. U., ERSBOLL, A. K., BISWAS, P. K. & CHRISTENSEN, J. P. (2010) The space-time clustering of highly pathogenic avian influenza (HPAI) H5N1 outbreaks in Bangladesh. *Epidemiology and Infection,* 138**,** 843-852.

ALEXANDER, D. J. (2007) An overview of the epidemiology of avian influenza. *Vaccine,* 25**,** 5637-5644.

ALEXANDERSEN, S., BROTHERHOOD, I. & DONALDSON, A. I. (2002) Natural aerosol transmission of foot-and-mouth disease virus to pigs: minimal infectious dose for strain O-1 Lausanne. *Epidemiology and Infection,* 128**,** 301-312.

ALFORD, P., GELLER, S., RICHARDSON, B., SLATER, M., HONNAS, C., FOREMAN, J., ROBINSON, J., MESSER, M., ROBERTS, M., GOBLE, D., HOOD, D. & CHAFFIN, M. (2001) A multicenter, matched case-control study of risk factors for equine laminitis. *Preventive Veterinary Medicine,* 49**,** 209-222.

ALLISON, P. D. (2010) *Survival Analysis Using SAS: A Practical Guide. Second Edition,* Cary, NC: SAS Institute Inc.

ANDERSON, P. K. & GILL, R. D. (1982) Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics,* 10**,** 1100-1120.

ANDERSON, R. M. & MAY, R. M. (1991) *Infectious Diseases of Humans: Dynamics and Control,* Oxford, Oxford University Press.

ANIMAL HEALTH AUSTRALIA (2007) *Disease strategy: Equine influenza (Version 3.0). Australian Veterinary Emergency Plan (AUSVETPLAN), Edition 3, Primary Industries Ministerial Council, Canberra, ACT.*

ANIMAL HEALTH AUSTRALIA (2008) *Summary Document (Version 3.1). Australian Veterinary Emergency Plan (AUSVETPLAN), Edition 3, Primary Industries Ministerial Council, Canberra, ACT.*

ANIMAL HEALTH AUSTRALIA (2011) *Disease strategy: Equine influenza (Version 3.1). Australian Veterinary Emergency Plan (AUSVETPLAN), Edition 3, Primary Industries Ministerial Council, Canberra, ACT.*

ANTHONY, N. D. (2011) Clinical impression of equine influenza at Morgan Park and the western region of Brisbane, Queensland, Australia. *Australian Veterinary Journal,* 89**,** 16-17.

ARCHER, D. C., PINCHBECK, G. L., FRENCH, N. P. & PROUDMAN, C. J. (2008) Risk factors for epiploic foramen entrapment colic in a UK horse population: A prospective case-control study. *Equine Veterinary Journal,* 40**,** 405-410.

ARMITAGE, P., BERRY, G. & MATTHEWS, J. N. S. (2001) *Statistical Methods in Medical*

Research, 4th Edition, Wiley-Blackwell. pp 356-360.

ARROYO, M., PEREZ, A. M. & RODRIGUEZ, L. L. (2011) Characterization of the temporal and spatial distribution and reproductive ratio of vesicular stomatitis outbreaks in Mexico in 2008. *American Journal of Veterinary Research,* 72**,** 233-238.

ARTHUR, R. J. & SUANN, C. J. (2011) Biosecurity and vaccination strategies to minimise the effect of an equine influenza outbreak on racing and breeding. *Australian Veterinary Journal,* 89**,** 109-113.

AUSTRALIAN ANIMAL HEALTH COUNCIL LIMITED (2001) *Emergency Animal Disease Response Agreement: Government and livestock industry cost sharing deed in respect of emergency animal disease responses. Variation No. 10/02 – 26/10/10,* Kingston, ACT.

AUSTRALIAN GOVERNMENT DEPARTMENT OF HEALTH AND AGEING (2010) *Australian Health Management Plan for Pandemic Influenza (December 2009 update)*. Canberra, ACT.

AXON, J. E., RUSSELL, C. M., TYNER, G. A., BURBURY, M. E. & CARRICK, J. B. (2008) Clinical observations of equine influenza in a naive foal population. *Journal of Veterinary Internal Medicine,* 22**,** 819-819.

BADDELEY, A. & TURNER, R. (2005) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software,* 12**,** 1-42.

BAGUELIN, M., NEWTON, J. R., DEMIRIS, N., DALY, J., MUMFORD, J. A. & WOOD, J. L. N. (2010) Control of equine influenza: scenario testing using a realistic metapopulation model of spread. *Journal of the Royal Society Interface,* 7**,** 67-79.

BAILEY, N. T. J. (1975) *The Mathematical Theory of Infectious Diseases and its Applications, second edition.*, Hafner Press, Macmillan Publishing Co., Inc. New York.

BARQUERO, N., DALY, J. M. & NEWTON, J. R. (2007) Risk factors for influenza infection in vaccinated racehorses: Lessons from an outbreak in Newmarket, UK in 2003. *Vaccine,* 25**,** 7520-7529.

BARTON, D. E., DAVID, F. N. & MERRINGTON, M. (1965) A criterion for testing contagion in time and space. *Annals of Human Genetics,* 29**,** 97-102.

BATES, T. W., THURMOND, M. C. & CARPENTER, T. E. (2003) Description of an epidemic simulation model for use in evaluating strategies to control an outbreak of foot-and-mouth disease. *American Journal of Veterinary Research,* 64**,** 195-204.

BEAN, B., MOORE, B. M., STERNER, B., PETERSON, L. R., GERDING, D. N. & BALFOUR, H. H. (1982) Survival of influenza viruses on environmental surfaces. *Journal of Infectious Diseases,* 146**,** 47-51.

BEELER, E. (2009) Influenza in Dogs and Cats. *Veterinary Clinics of North America. Small Animal Practice,* 39**,** 251-264.

BEGG, A. P., REECE, R. L., HUM, S., TOWNSEND, W., GORDON, A. & CARRICK, J. (2011) Pathological changes in horses dying with equine influenza in Australia, 2007. *Australian Veterinary Journal,* 89**,** 19-22.

BELSER, J. A., MAINES, T. R., TUMPEY, T. M. & KATZ, J. M. (2010) Influenza A virus transmission: contributing factors and clinical implications. *Expert Reviews in Molecular Medicine,* 12.

BIGRAS-POULIN, M., BARFOD, K., MORTENSEN, S. & GREINER, M. (2007) Relationship of trade patterns of the Danish swine industry animal movements network to potential disease spread. *Preventive Veterinary Medicine,* 80**,** 143-165.

BIGRAS-POULIN, M., THOMPSON, R. A., CHRIEL, M., MORTENSEN, S. & GREINER, M. (2006) Network analysis of Danish cattle industry trade patterns as an evaluation of risk potential for disease spread. *Preventive Veterinary Medicine,* 76**,** 11-39.

BIOSECURITY AUSTRALIA (2010) *Import risk analysis report for horses from approved countries: final report.,* Canberra.

BOENDER, G. J., HAGENAARS, T. J., BOUMA, A., NODELIJK, G., ELBERS, A. R. W., DE JONG, M. C. M. & VAN BOVEN, M. (2007) Risk maps for the spread of highly pathogenic avian influenza in poultry. *Plos Computational Biology,* 3**,** 704-712.

BORGATTI, S. P. (2002) Netdraw Network Visualization. Analytic Technologies: Harvard, MA.

BOWMAN, A. W. & AZZALINI, A. (1997) *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-PLUS Illustrations.,* Oxford, Oxford University Press,.

BRENNAN, M. L., KEMP, R. & CHRISTLEY, R. M. (2008) Direct and indirect contacts between cattle farms in north-west England. *Prev Vet Med,* 84**,** 242-60.

BRITTON, A. (2007) Australian equine influenza outbreak 2007: an epidemiological assessment between 3-25 August 2007 within NSW. NSW Department of Primary Industries report submitted to the Callinan inquiry. Accessed online 15/4/2010 at http://www.equineinfluenzainquiry.gov.au.

BRITTON, A. L., MAJOR, D. A., PERRY, G. H. & READ, A. J. (2011) Spatial association and clinical development of equine influenza in horses yarded overnight at an equestrian event at Maitland prior to propagating the 2007 epidemic in Australia. *Australian Veterinary Journal,* 89**,** 68-69.

BRYANT, N. A., PAILLOT, R., RASH, A. S., MEDCALF, E., MONTESSO, F., ROSS, J., WATSON, J., JEGGO, M., LEWIS, N. S., NEWTON, J. R. & ELTON, D. M. (2010) Comparison of two modern vaccines and previous influenza infection against challenge with an equine influenza virus from the Australian 2007 outbreak. *Veterinary Research,* 41.

BRYANT, N. A., RASH, A. S., WOODWARD, A. L., MEDCALF, E., HELWEGEN, M.,

WOHLFENDER, F., CRUZ, F., HERRMANN, C., BORCHERS, K., TIWARI, A., CHAMBERS, T. M., NEWTON, J. R., MUMFORD, J. A. & ELTON, D. M. (2011) Isolation and characterisation of equine influenza viruses (H3N8) from Europe and North America from 2008 to 2009. *Veterinary Microbiology,* 147**,** 19-27.

CALLAWAY, D. S., NEWMAN, M. E. J., STROGATZ, S. H. & WATTS, D. J. (2000) Network robustness and fragility: percolation on random graphs. *Physical Review Letters,* 85**,** 5468-5471.

CALLINAN, I. (2008) *Equine influenza - the August 2007 outbreak in Australia. Report of the Equine Influenza Inquiry. The Hon. Ian Callinan AC*.

CARPENTER, T. E. (2011) Stochastic, spatially-explicit epidemic models. *Revue Scientifique Et Technique-Office International Des Epizooties,* 30**,** 417-424.

CENTER FOR INTERNATIONAL EARTH SCIENCE INFORMATION NETWORK, INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE, THE WORLD BANK & CENTRO INTERNACIONAL DE AGRICULTURA TROPICAL (2011) Global Rural-Urban Mapping Project (GRUMPv3): Urban Extents Data Collection, Alpha Version 3. Columbia University, New York, NY.

CENTRE FOR INTERNATIONAL ECONOMICS (2007) *Estimating horse industry GVP and horse numbers. Briefing paper prepared for Animal Health Australia,* Canberra, Australia.

CHAN, T. C., KING, C. C., YEN, M. Y., CHIANG, P. H., HUANG, C. S. & HSIAO, C. K. (2010) Probabilistic Daily ILI Syndromic Surveillance with a Spatio-Temporal Bayesian Hierarchical Model. *PLoS ONE,* 5.

CHARLAND, K. M. L., BUCKERIDGE, D. L., STURTEVANT, J. L., MELTON, F., REIS, B. Y., MANDL, K. D. & BROWNSTEIN, J. S. (2009) Effect of environmental factors on the spatio-temporal patterns of influenza spread. *Epidemiology and Infection,* 137**,** 1377-1387.

CHRISTLEY, R. M. & FRENCH, N. P. (2003) Small-world topology of UK racing: the potential for rapid spread of infectious agents. *Equine Vet J,* 35**,** 586-9.

CHRISTLEY, R. M., HODGSON, D. R., ROSE, R. J., WOOD, J. L., REIDS, S. W., WHITEAR, K. G. & HODGSON, J. L. (2001) A case-control study of respiratory disease in Thoroughbred racehorses in Sydney, Australia. *Equine Vet J,* 33**,** 256-64.

CHRISTLEY, R. M., PINCHBECK, G. L., BOWERS, R. G., CLANCY, D., FRENCH, N. P., BENNETT, R. & TURNER, J. (2005) Infection in social networks: using network analysis to identify high-risk individuals. *Am J Epidemiol,* 162**,** 1024-31.

CLARK, P. J. & EVANS, F. C. (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology,* 35**,** 445-453.

COGGER, N. (2006) Epidemiology of musculoskeletal injuries in two- and three-year-old

Australian Thoroughbred racehorses. Ph.D. thesis. Faculty of Veterinary Science, University of Sydney, Sydney, Australia.

COHEN, M. L. (2000) Changing patterns of infectious disease. *Nature,* 406**,** 762-767.

COHEN, N. D., MACKAY, R. J., TOBY, E., ANDREWS, F. M., BARR, B. S., BEECH, J., BERNARD, W. V., CLARK, C. K., DIVERS, T. J., FURR, M. O., KOHN, C. W., LEVY, M., REED, S. M., SEAHORN, T. L. & SIOVIS, N. M. (2007) A multicenter case-control study of risk factors for equine protozoal myeloencephalitis. *Journal of the American Veterinary Medical Association,* 231**,** 1857-1863.

COHEN, R., EREZ, K., BEN-AVRAHAM, D. & HAVLIN, S. (2000) Resilience of the Internet to random breakdowns. *Physical Review Letters,* 85**,** 4626-4628.

COOK, A. (2008) Summary of the Australian equine influenza outbreak. *Veterinary Record,* 163**,** 378-378.

CORNER, L. A. L., PFEIFFER, D. U. & MORRIS, R. S. (2003) Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (Trichosurus vulpecula). *Preventive Veterinary Medicine,* 59**,** 147-167.

COWLED, B., GARNER, G. & MOLONEY, B. (2009a) *The cleaning and collation of the NSW and QLD equine influenza datasets in preparation for epidemiological analysis. Unpublished manuscript.,* Canberra, Australia.

COWLED, B., WARD, M. P., HAMILTON, S. & GARNER, G. (2009b) The equine influenza epidemic in Australia: spatial and temporal descriptive analyses of a large propagating epidemic. *Preventive Veterinary Medicine,* 92**,** 60-70.

COWLED, B. D., GARNER, M. G., NEGUS, K. & WARD, M. P. (2012) Controlling disease outbreaks in wildlife using limited culling: modelling classical swine fever incursions in wild pigs in Australia. *Vet Res,* 43**,** 3.

CRAWFORD, P. C., DUBOVI, E. J., CASTLEMAN, W. L., STEPHENSON, I., GIBBS, E. P. J., CHEN, L., SMITH, C., HILL, R. C., FERRO, P., POMPEY, J., BRIGHT, R. A., MEDINA, M. J., JOHNSON, C. M., OLSEN, C. W., COX, N. J., KLIMOV, A. I., KATZ, J. M. & DONIS, R. O. (2005) Transmission of equine influenza virus to dogs. *Science,* 310**,** 482-485.

CRISPE, E., FINLAISON, D. S., HURT, A. C. & KIRKLAND, P. D. (2011) Infection of dogs with equine influenza virus: evidence for transmission from horses during the Australian outbreak. *Australian Veterinary Journal,* 89**,** 27-28.

CROUCH, C. F., DALY, J., HANNANT, D., WILKINS, J. & FRANCIS, M. J. (2004) Immune responses and protective efficacy in ponies immunised with an equine influenza ISCOM vaccine containing an 'American lineage' H3N8 virus. *Vaccine,* 23**,** 418-425.

DAFF (2008) Australian OIE declaration of freedom from equine influenza. Australian Government Department of Agriculture Fisheries and Forestry (DAFF). Canberra.

DALGLISH, R. A. (1992) The international movement of horses – the current infectious disease situation. IN SHORT, C. R. (Ed.) *Proceedings of the 9th International Conference of Racing Analysts and Veterinarians.* Louisiana State University, New Orleans.

DALY, J., NEWTON, J. R., SMITH, K. C. & MUMFORD, J. (2006) Epidemiology of Equine Influenza Viruses: Pathogenicity and Transmissibility. *Rad 496. Medicinske znanosti (Croatian translated),* 30**,** 87-94.

DALY, J. M., BLUNDEN, A. S., MACRAE, S., MILLER, J., BOWMAN, S. J., KOLODZIEJEK, J., NOWOTNY, N. & SMITH, K. C. (2008) Transmission of equine influenza virus to English foxhounds. *Emerging Infectious Diseases,* 14**,** 461-4.

DALY, J. M., LAI, A. C. K., BINNS, M. M., CHAMBERS, T. M., BARRANDEGUY, M. & MUMFORD, J. A. (1996) Antigenic and genetic evolution of equine H3N8 influenza A viruses. *Journal of General Virology,* 77**,** 661-671.

DALY, J. M., MACRAE, S., NEWTON, J. R., WATTRANG, E. & ELTON, D. M. (2011) Equine influenza: A review of an unpredictable virus. *Veterinary Journal,* 189**,** 7-14.

DALY, J. M., NEWTON, J. R. & MUMFORD, J. A. (2004) Current perspectives on control of equine influenza. *Vet Res,* 35**,** 411-23.

DANON, L., FORD, A. P., HOUSE, T., JEWELL, C. P., KEELING, M. J., ROBERTS, G. O., ROSS, J. V. & VERNON, M. C. (2011) Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases,* 2011**,** 1-28.

DAVIES, T. M. & HAZELTON, M. L. (2009) Adaptive kernel estimation of spatial relative risk. *Statistics in Medicine,* 29**,** 2423-2437.

DAVIES, T. M., HAZELTON, M. L. & MARSHALL, J. C. (2011) sparr: Analyzing Spatial Relative Risk Using Fixed and Adaptive Kernel Density Estimation in R. *Journal of Statistical Software,* 39**,** 1-14. Available at: http://www.jstatsoft.org/v39/i01/.

DAVIS, J., GARNER, M. G. & EAST, I. J. (2009) Analysis of local spread of equine influenza in the Park Ridge Region of Queensland. *Transboundary and Emerging Diseases,* 56**,** 31-38.

DE LA RUA-DOMENECH, R., REID, S. W. J., GONZALEZ-ZARIQUIEY, A. E., WOOD, J. L. N. & GETTINBY, G. (2000) Modelling the spread of a viral infection in equine populations managed in Thoroughbred racehorse training yards. *Preventive Veterinary Medicine,* 47**,** 61-77.

DENT, J. E., KAO, R. R., KISS, I. Z., HYDER, K. & ARNOLD, M. (2008) Contact structures in the poultry industry in Great Britain: exploring transmission routes for a potential avian influenza virus epidemic. *BMC Veterinary Research,* 4.

DEPARTMENT OF AGRICULTURE FORESTRY AND FISHERIES (2008) *Recovery of EI Country Free Status: Australian report to the World Organisation for Animal Health*

(OIE), Canberra.

DHAND, N. K. & SERGEANT, E. S. G. (2011a) Assessment of the proportion of under-reporting during the 2007 equine influenza outbreak in New South Wales, Australia. *Australian Veterinary Journal,* 89**,** 73-74.

DHAND, N. K. & SERGEANT, E. S. G. (2011b) Attack risk on infected properties during the 2007 equine influenza outbreak in New South Wales, Australia. *Australian Veterinary Journal,* 89**,** 70-72.

DIEKMANN, O., DE JONG, M. C. M. & METZ, J. A. J. (1998) A deterministic epidemic model taking account of repeated contacts between the same individuals. *Journal of Applied Probability,* 35**,** 448-462.

DIGGLE, P. J. (1983) *Statistical Analysis of Spatial Point Patterns,* London, Academic Press.

DIGGLE, P. J. (2006) Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research,* 15**,** 325-336.

DIGGLE, P. J. & CHETWYND, A. G. (1991) Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics,* 47**,** 1155-1163.

DIGGLE, P. J., CHETWYND, A. G., HAGGKVIST, R. & MORRIS, S. E. (1995) Second-order analysis of space-time clustering. *Stat Methods Med Res,* 4**,** 124-36.

DIGGLE, P. J., GOMEZ-RUBIO, V., BROWN, P. E., CHETWYND, A. G. & GOODING, S. (2007) Second-order analysis of inhomogeneous spatial point processes using case-control data. *Biometrics,* 63**,** 550-557.

DIGGLE, P. J. & RIBEIRO, P. J. (2007) *Model-based Geostatistics. Springer Series in Statistics.,* Springer, New York, NY.

DOHOO, I., MARTIN, W. & STRYHN, H. (2009) *Veterinary epidemiologic research, 2nd Edition.*, VER Inc, Charlottetown, Prince Edward Island, Canada.

DOMMERGUES, L., RAUTUREAU, S., PETIT, E. & DUFOUR, B. (2011) Network of Contacts between Cattle Herds in a French Area Affected by Bovine Tuberculosis in 2010. *Transboundary and Emerging Diseases*.

DU PREL, J. B., PUPPE, W., GRONDAHL, B., KNUF, M., WEIGL, J. A. I., SCHAAFF, F. & SCHMITT, H. J. (2009) Are Meteorological Parameters Associated with Acute Respiratory Tract Infections? *Clinical Infectious Diseases,* 49**,** 861-868.

DUBÉ, C., RIBBLE, C., KELTON, D. & MCNAB, B. (2009) A Review of Network Analysis Terminology and its Application to Foot-and-Mouth Disease Modelling and Policy Development. *Transboundary and Emerging Diseases,* 56**,** 73-85.

DUBÉ, C., RIBBLE, C., KELTON, D. & MCNAB, B. (2011) Introduction to network analysis and its implications for animal disease modelling. In: Models in the management of animal diseases (P. Willeberg, ed.). *Revue Scientifique Et Technique De L Office International Des Epizooties,* 30**,** 425-436.

DUBÉ, C., STEVENSON, M. A., GARNER, M. G., SANSON, R. L., CORSO, B. A., HARVEY, N., GRIFFIN, J., WILESMITH, J. W. & ESTRADA, C. (2007) A comparison of predictions made by three simulation models of foot-and-mouth disease. *New Zealand Veterinary Journal,* 55**,** 280-288.

DUPS, J. N., MORTON, J. M., ANTHONY, N. D. & DWYER, J. F. (2011) Clinical signs of equine influenza in a closed population of horses at a 3-day event in southern Queensland, Australia. *Australian Veterinary Journal,* 89**,** 17-18.

EAST, I. J. (2009) The Role of Land Use Patterns in Limiting the Spread of Equine Influenza in Queensland During the 2007 Epidemic. *Transboundary and Emerging Diseases,* 56**,** 292-302.

EDLUND TOULEMONDE, C., DALY, J., SINDLE, T., GUIGAL, P. M., AUDONNET, J. C. & MINKE, J. M. (2005) Efficacy of a recombinant equine influenza vaccine against challenge with an American lineage H3N8 influenza virus responsible for the 2003 outbreak in the United Kingdom. *Veterinary Record,* 156**,** 367-371.

EL-HAGE, C. M., SAVAGE, C. J., MINKE, J. M., FICORILLI, N. P. & GILKERSON, J. R. (2009) An accelerated vaccination schedule for use in an equine influenza emergency response. IN WHITE, N. A. (Ed.) *Proceedings of the 55th Annual Convention of the American Association of Equine Practitioners.* Las Vegas, NV, USA.

EQUINE INFLUENZA EPIDEMIOLOGY SUPPORT GROUP (2008) *Equine influenza 2007. The Australian Experience: Report to the Consultative Committee on Emergency Animal Disease.,* Canberra, Department of Agriculture Forestry and Fisheries.

ERDŐS, P. & RÉNYI, A. (1961) On the strength of connectedness of a random graph. *Acta Mathematica Hungarica,* 12**,** 261-267.

FAEHRMANN, P., RIDDELL, K. & READ, A. J. (2011) Longitudinal study describing the clinical signs observed in horses naturally infected with equine influenza. *Australian Veterinary Journal,* 89**,** 22-23.

FANG, L.-Q., DE VLAS, S. J., LIANG, S., LOOMAN, C. W. N., GONG, P., XU, B., YAN, L., YANG, H., RICHARDUS, J. H. & CAO, W.-C. (2008) Environmental Factors Contributing to the Spread of H5N1 Avian Influenza in Mainland China. *PLoS ONE,* 3**,** e2268.

FENTON, S. E., CLOUGH, H. E., DIGGLE, P. J., EVANS, S. J., DAVISON, H. C., VINK, W. D. & FRENCH, N. P. (2009) Spatial and spatio-temporal analysis of Salmonella infection in dairy herds in England and Wales. *Epidemiology and Infection,* 137**,** 847-857.

FERGUSON, N. M., DONNELLY, C. A. & ANDERSON, R. M. (2001a) The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science,* 292**,** 1155-1160.

FERGUSON, N. M., DONNELLY, C. A. & ANDERSON, R. M. (2001b) Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature,* 413**,** 542-548.

FERNS, L., DOHOO, I. & DONALD, A. (1991) A case-control study of Nocardia mastitis in Nova Scotia dairy herds. *canadian Veterinary Journal,* 32**,** 673-677.

FIRESTONE, S. M., CHRISTLEY, R. M., WARD, M. P. & DHAND, N. K. (2012a) Adding the spatial dimension to the social network analysis of an epidemic: Investigation of the 2007 outbreak of equine influenza in Australia. *Preventive Veterinary Medicine,* 106**,** 123-135.

FIRESTONE, S. M., COGGER, N., WARD, M. P., TORIBIO, J.-A. L. M. L., MOLONEY, B. J. & DHAND, N. K. (2012b) The Influence of Meteorology on the Spread of Influenza: Survival Analysis of an Equine Influenza (A/H3N8) Outbreak. *PLoS ONE,* 7**,** e35284.

FIRESTONE, S. M., SCHEMANN, K. A., TORIBIO, J.-A. L. M. L., WARD, M. P. & DHAND, N. K. (2011a) A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. *Preventive Veterinary Medicine,* 100**,** 53-63.

FIRESTONE, S. M., WARD, M. P., CHRISTLEY, R. M. & DHAND, N. K. (2011b) The importance of location in contact networks: Describing early epidemic spread using spatial social network analysis. *Preventive Veterinary Medicine,* 102**,** 185-195.

FOORD, A. J., SELLECK, P., COLLING, A., KLIPPEL, J., MIDDLETON, D. & HEINE, H. G. (2009) Real-time RT-PCR for detection of equine influenza and evaluation using samples from horses infected with A/equine/Sydney/2007 (H3N8). *Veterinary Microbiology,* 137**,** 1-9.

FRAZER, J. L., PERKINS, N. R. & PITT, D. (2011) Role of personal decontamination in preventing the spread of equine influenza. *Australian Veterinary Journal,* 89**,** 120-124.

FREEMAN, L. C. (1979) Centrality in social networks conceptual clarification. *Social Networks,* 1**,** 215-239.

GARCÍA ÁLVAREZ, L., WEBB, C. R. & HOLMES, M. A. (2011) A novel field-based approach to validate the use of network models for disease spread between dairy herds. *Epidemiology and Infection***,** 1-12.

GARNER, M. G. & BECKETT, S. D. (2005) Modelling the spread of foot-and-mouth disease in Australia. *Australian Veterinary Journal,* 83**,** 758-766.

GARNER, M. G. & CANNON, R. M. (1995) *Potential for wind-borne spread of foot-and-mouth disease virus in Australia,* Bureau of Resource Sciences, Department of Primary Industries and Energy. Canberra, ACT.

GARNER, M. G., COWLED, B., EAST, I. J., MOLONEY, B. J. & KUNG, N. (2011a) Evaluating the effectiveness of the response to equine influenza in the Australian

outbreak and the potential role of early vaccination. *Australian Veterinary Journal,* 89**,** 143-145.

GARNER, M. G., COWLED, B., EAST, I. J., MOLONEY, B. J. & KUNG, N. Y. (2011b) Evaluating the effectiveness of early vaccination in the control and eradication of equine influenza - A modelling approach. *Preventive Veterinary Medicine,* 99**,** 15-27.

GARNER, M. G. & HAMILTON, S. A. (2011) Principles of epidemiological modelling. In: Models in the management of animal diseases (P. Willeberg, ed.). *Revue Scientifique Et Technique De L Office International Des Epizooties,* 30**,** 407-416.

GARNER, M. G., SCANLAN, W. A., COWLED, B. D. & CARROLL, A. (2011c) Regaining Australia's equine influenza-free status: a national perspective. *Australian Veterinary Journal,* 89**,** 169-173.

GEOSCIENCE AUSTRALIA (2006) Geodata TOPO-250K series 3. Australian Government Department of Resources, Energy and Tourism, Canberra, ACT.

GEOSCIENCE AUSTRALIA (2007) Geocentric Datum of Australia. http://www.ga.gov.au/geodesy/datums/gda.jsp (accessed on 20.04.10).

GEOSCIENCE AUSTRALIA (2008) GEODATA 9 Second Digital Elevation Model, Version 3 and flow direction grid. Australian Government Department of Resources, Energy and Tourism, Canberra, ACT.

GIBBENS, J. C. & WILESMITH, J. W. (2002) Temporal and geographical distribution of cases of foot-and-mouth disease during the early weeks of the 2001 epidemic in Great Britain. *Veterinary Record,* 151**,** 407-412.

GIEBULTOWICZ, S., ALI, M., YUNUS, M. & EMCH, M. (2011) A comparison of spatial and social clustering of cholera in Matlab, Bangladesh. *Health & Place,* 17**,** 490-497.

GILCHRIST, P. & SERGEANT, E. S. G. (2011) Risk of an equine influenza virus reservoir establishing in wild horses in New South Wales during the Australian epidemic. *Australian Veterinary Journal,* 89**,** 75-78.

GILDEA, S., ARKINS, S. & CULLINANE, A. (2011) Management and environmental factors involved in equine influenza outbreaks in Ireland 2007-2010. *Equine Veterinary Journal,* 43**,** 608-617.

GILKERSON, J. R. (2011) Equine influenza in Australia: a clinical overview. *Australian Veterinary Journal,* 89**,** 11-13.

GILLESPIE, D. T. (1977) Exact stochastic simulation of coupled chemical-reactions. *Journal of Physical Chemistry,* 81**,** 2340-2361.

GLANVILLE, R. J. & CHRISTIE, B. (2011) High-level coordination and strategy in the 2007 equine influenza outbreak response. *Australian Veterinary Journal,* 89**,** 97-100.

GLANVILLE, R. J., CHRISTIE, B. M., WEBSTER, W. R., ROTH, I. J., DUNN, S. E. & CROOK, A. (2009) Eradication of Equine Influenza (EI) from Australia. *Proceedings*

of the 12th International Symposium on Veterinary Epidemiology and Economics (ISVEE), 2009. Durban, South Africa.

GLASS, K., WOOD, J. L. N., MUMFORD, J. A., JESSET, D. & GRENFELL, B. T. (2002) Modelling equine influenza 1: a stochastic model of within-yard epidemics. *Epidemiology and Infection,* 128**,** 491-502.

GLOSTER, J., BURGIN, L., JONES, A. & SANSON, R. (2011) Atmospheric dispersion models and their use in the assessment of disease transmission. In: Models in the management of animal diseases (P. Willeberg, ed.). *Revue Scientifique Et Technique De L Office International Des Epizooties,* 30**,** 457-465.

GLOSTER, J., JONES, A., REDINGTON, A., BURGIN, L., SORENSEN, J. H., TURNER, R., DILLON, M., HULLINGER, P., SIMPSON, M., ASTRUP, P., GARNER, G., STEWART, P., D'AMOURS, R., SELLERS, R. & PATON, D. (2010) Airborne spread of foot-and-mouth disease - Model intercomparison. *Veterinary Journal,* 183**,** 278-286.

GOULD, W. W., PITBLADO, J. & SRIBNEY, W. M. (2006) *Maximum Likelihood Estimation with Stata. 3rd ed.*, College Station, TX: Stata Press.

GREEN, D. M., KISS, I. Z. & KAO, R. R. (2006) Modelling the initial spread of foot-and-mouth disease through animal movements. *Proc. Biol. Sci.,* 273.

GREEN, D. M., KISS, I. Z., MITCHELL, A. P. & KAO, R. R. (2008) Estimates for local and movement-based transmission of bovine tuberculosis in British cattle. *Proceedings of the Royal Society B-Biological Sciences,* 275**,** 1001-1005.

GUO, Y., WANG, M., ZHANG, G. S., LI, W. K., KAWAOKA, Y. & WEBSTER, R. G. (1995) Seroepidemiological and molecular evidence for the presence of two H3N8 equine influenza viruses in China in 1993-1994. *J. Gen. Virol.,* 76**,** 2009-2014.

GUO, Y. J., WANG, M., KAWAOKA, Y., GORMAN, O., ITO, T., SAITO, T. & WEBSTER, R. G. (1992) Characterization of a new avian-like influenza-A virus from horses in China. *Virology,* 188**,** 245-255.

GUTHRIE, A. J. (2006) Equine influenza in South Africa, 2003 outbreak. *In: Proceedings of the 9th World Equine Veterinary Association Congress.* Marrakech, Morocco, 22–26 January 2006.

GUTHRIE, A. J., STEVENS, K. B. & BOSMAN, P. P. (1999) The circumstances surrounding the outbreak and spread of equine influenza in South Africa. *Rev Sci Tech,* 18**,** 179-85.

HAMMELL, K. L. & DOHOO, I. R. (2005) Risk factors associated with mortalities attributed to infectious salmon anaemia virus in New Brunswick, Canada. *Journal of Fish Diseases,* 28**,** 651-661.

HANDCOCK, M. S., HUNTER, D. R., BUTTS, C. T., GOODREAU, S. M. & MORRIS, M. (2003) Software Tools for the Statistical Modeling of Network Data. Version 2.1-1.

HANNANT, D. & MUMFORD, J. A. (1996) Equine influenza. IN STUDDERT, M. (Ed.) *Virus*

Infections of Equines. Amsterdam, Elsevier Science, BV.

HANNANT, D., MUMFORD, J. A. & JESSETT, D. M. (1988) Duration of circulating antibody and immunity following infection with equine influenza virus. *Veterinary Record,* 122**,** 125-128.

HANNEMAN, R. A. & RIDDLE, M. (2005) Introduction to social network methods., University of California, Riverside, CA, Available at: http://faculty.ucr.edu/~hanneman/.

HAPPOLD, J. & RUBIRA, R. (2011) Equine influenza: patterns of disease and seroprevalence in Thoroughbred studs and implications for vaccination. *Australian Veterinary Journal,* 89**,** 135-137.

HARRELL , F. E. (2001) *Regression Modeling Strategies With Applications to Linear Models Logistic Regression, and Survival Analysis,* New York, Springer.

HARVEY, N., REEVES, A., SCHOENBAUM, M. A., ZAGMUTT-VERGARA, F. J., DUB, C., HILL, A. E., CORSO, B. A., MCNAB, W. B., CARTWRIGHT, C. I. & SALMAN, M. D. (2007) The North American Animal Disease Spread Model: A simulation model to assist decision making in evaluating animal disease incursions. *Preventive Veterinary Medicine,* 82**,** 176-197.

HAYDON, D. T., CHASE-TOPPING, M., SHAW, D. J., MATTHEWS, L., FRIAR, J. K., WILESMITH, J. & WOOLHOUSE, M. E. J. (2003) The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. *Proceedings of the Royal Society of London Series B-Biological Sciences,* 270**,** 121-127.

HEATH, M. F., VERNON, M. C. & WEBB, C. R. (2008) Construction of networks with intrinsic temporal structure from UK cattle movement data. *BMC Veterinary Research,* 4**,** 1-11.

HEINE, H. G., TRINIDAD, L., SELLECK, P. & LOWTHER, S. (2007) Rapid detection of highly pathogenic avian influenza H5N1 virus by TaqMan reverse transcriptase-polymerase chain reaction. *Avian Diseases,* 51**,** 370-372.

HEMMES, J. H., WINKLER, K. C. & KOOL, S. M. (1960) Virus survival as a seasonal factor in influenza and poliomyelitis. *Nature,* 188**,** 430-431.

HENDERSON, R., SHIMAKURA, S. & GORST, D. (2002) Modeling spatial variation in leukemia survival data. *Journal of the American Statistical Association,* 97**,** 965-972.

HENGL, T. (2009) *A Practical Guide to Geostatistical Mapping.,* Amsterdam. http://spatial-analyst.net/book/.

HENNING, K. A., HENNING, J., MORTON, J., LONG, N. T., HA, N. T. & MEERS, J. (2009) Farm- and flock-level risk factors associated with Highly Pathogenic Avian Influenza outbreaks on small holder duck and chicken farms in the Mekong Delta of Viet Nam. *Preventive Veterinary Medicine,* 91**,** 179-188.

HOSMER, D. W. & LEMESHOW, S. (2000) *Applied Logistic Regression, 2nd edition,* New York, USA, John Wiley & Sons, Inc.

HOSMER, D. W., LEMESHOW, S. & MAY, S. (2008) *Applied survival analysis: Regression modeling of time-to-event data, 2nd ed.,* Hoboken, N.J., Wiley.

HOX, J. J. (1994) Hierarchical regression-models for interviewer and respondent effects. *Sociological Methods & Research,* 22**,** 300-318.

HUNTINGTON, P. J. (1990) *Equine Influenza – The Disease and its Control. Technical Report Series No. 184.*, Department of Agriculture and Rural Affairs, Victoria.

INGGS, M. R. & LORD, R. T. (1996) Interpolating satellite derived wind field data using Ordinary Kriging, with application to the nadir gap. *IEEE Transactions on Geoscience and Remote Sensing,* 34**,** 250-256.

ISMAIL, T. M., SAMI, A. M., YOUSSEF, H. M. & ZAID, A. A. A. (1990) An outbreak of equine influenza type 1 in Egypt in 1989. *Veterinary Medical Journal of Giza,* 38**,** 195-206.

ITO, M., NAGAI, M., HAYAKAWA, Y., KOMAE, H., MURAKAMI, N., YOTSUYA, S., ASAKURA, S., SAKODA, Y. & KIDA, H. (2008) Genetic Analyses of an H3N8 Influenza Virus Isolate, Causative Strain of the Outbreak of Equine Influenza at the Kanazawa Racecourse in Japan in 2007. *Journal of Veterinary Medical Science,* 70**,** 899-906.

JACKSON, A. E. (2011) In this issue: Equine influenza in Australia in 2007-the definitive story. *Australian Veterinary Journal,* 89**,** 1-2.

JEGGO, M., HAMMOND, J. M. & KIRKLAND, P. D. (2008) The initial laboratory diagnosis of equine influenza in Australia in 2007. *Microbiol. Aust.,* 29**,** 80-82.

JENSEN, E. S., LUNDBYE-CHRISTENSEN, S., SAMUELSSON, S., SORENSEN, H. T. & SCHONHEYDER, H. C. (2004) A 20-year ecological study of the temporal association between influenza and meningococcal disease. *European Journal of Epidemiology,* 19**,** 181-187.

JEWELL, C. P. (2009) Real-time inference and risk-prediction for notifiable diseases of animals. Ph.D. thesis. Department of Mathematics and Statistics, Lancaster University, Lancaster.

JEWELL, C. P., KEELING, M. J. & ROBERTS, G. O. (2009a) Predicting undetected infections during the 2007 foot-and-mouth disease outbreak. *Journal of the Royal Society Interface,* 6**,** 1145-1151.

JEWELL, C. P., KYPRAIOS, T., CHRISTLEY, R. M. & ROBERTS, G. O. (2009b) A novel approach to real-time risk prediction for emerging infectious diseases: A case study in Avian Influenza H5N1. *Preventive Veterinary Medicine,* 91**,** 19-28.

JEWELL, C. P., KYPRAIOS, T., NEAL, P. & ROBERTS, G. O. (2009c) Bayesian Analysis for

Emerging Infectious Diseases. *Bayesian Analysis,* 4**,** 465-496.

JOHNSON, Y. J., GEDAMU, N., COLBY, M. M., MYINT, M. S., STEELE, S. E., SALEM, M. & TABLANTE, N. L. (2005) Wind-Borne Transmission of Infectious Laryngotracheitis Between Commercial Poultry Operations. *International Journal of Poultry Science,* 4**,** 263-267.

JOURNEL, A. G. & HUIJBREGTS, C. (1978) *Mining geostatistics,* London, Academic Press.

KANNEGIETER, N. J., FROGLEY, A., CRISPE, E. & KIRKLAND, P. D. (2011) Clinical outcomes and virology of equine influenza in a naive population and in horses infected soon after receiving one dose of vaccine. *Australian Veterinary Journal,* 89**,** 139-142.

KAO, R. R. (2002) The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. *Trends in Microbiology,* 10**,** 279-286.

KAO, R. R., GREEN, D. M., JOHNSON, J. & KISS, I. Z. (2007) Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK. *Journal of the Royal Society Interface,* 4**,** 907-916.

KARR, A. F. (1991) *Point processes and their statistical inference.,* CRC Press. ISBN:9780824785321.

KEDMI, M., HERZIGER, Y., GALON, N., COHEN, R. M., PEREL, M., BATTEN, C., BRAVERMAN, Y., GOTTLIEB, Y., SHPIGEL, N. & KLEMENT, E. (2010) The association of winds with the spread of EHDV in dairy cattle in Israel during an outbreak in 2006. *Preventive Veterinary Medicine,* 96**,** 152-160.

KEELING, M. J. (1999) The effects of local spatial structure on epidemiological invasions. *Proceeding of the Royal Society of London Series B Biological Sciences,* 266**,** 859-867.

KEELING, M. J. & EAMES, K. T. D. (2005) Networks and epidemic models. *Journal of the Royal Society Interface,* 2**,** 295-307.

KEELING, M. J., WOOLHOUSE, M. E. J., SHAW, D. J., MATTHEWS, L., CHASE-TOPPING, M., HAYDON, D. T., CORNELL, S. J., KAPPEY, J., WILESMITH, J. & GRENFELL, B. T. (2001) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science,* 294**,** 813-817.

KELSALL, J. E. & DIGGLE, P. J. (1995) Non-parametric estimation of spatial variation in relative risk. *Statistics in Medicine,* 14**,** 2335-2342.

KIRKLAND, P. D. (2011) Role of the diagnostic laboratories during the 2007 equine influenza outbreak in Australia. *Australian Veterinary Journal,* 89**,** 29-32.

KIRKLAND, P. D., DAVIS, R. J., WONG, D., RYAN, D., HART, K., CORNEY, B., HEWITSON, G., COOPER, K., BIDDLE, A., EASTWOOD, S., SLATTERY, S., RAYWARD, D., EVERS, M., WRIGHT, T., HALPIN, K., SELLECK, P. & WATSON, J. (2011) The first five days: field and laboratory investigations during the early stages of the equine influenza outbreak in Australia, 2007. *Australian Veterinary*

Journal, 89**,** 6-10.

KIRKLAND, P. D. & DELBRIDGE, G. (2011) Use of a blocking ELISA for antibodies to equine influenza virus as a test to distinguish between naturally infected and vaccinated horses: proof of concept studies. *Australian Veterinary Journal,* 89**,** 45-46.

KIRKLAND, P. D., FINLAISON, D. S., CRISPE, E. & HURT, A. C. (2010) Influenza virus transmission from horses to dogs, Australia. *Emerg Infect Dis,* 16**,** 699-702.

KISS, I. Z., GREEN, D. M. & KAO, R. R. (2005) Disease contact tracing in random and clustered networks. *Proceedings of the Royal Society B-Biological Sciences,* 272**,** 1407-1414.

KISS, I. Z., GREEN, D. M. & KAO, R. R. (2006a) Infectious disease control using contact tracing in random and scale-free networks. *Journal of the Royal Society Interface,* 3**,** 55-62.

KISS, I. Z., GREEN, D. M. & KAO, R. R. (2006b) The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread. *Journal of the Royal Society Interface,* 3**,** 669-677.

KISS, I. Z., GREEN, D. M. & KAO, R. R. (2008) The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *Journal of the Royal Society Interface,* 5**,** 791-799.

KISS, I. Z., SIMON, P. L. & KAO, R. R. (2009) A Contact-Network-Based Formulation of a Preferential Mixing Model. *Bulletin of Mathematical Biology,* 71**,** 888-905.

KITCHING, R. P., THRUSFIELD, M. & TAYLOR, N. M. (2006) Use and abuse of mathematical models: an illustration from the 2001 foot and mouth disease epidemic in the United Kingdom. *Revue Scientifique Et Technique De L Office International Des Epizooties,* 25**,** 293-311.

KLOVDAHL, A. S. (1985) Social networks and the spread of infectious diseases - the AIDS example. *Social Science & Medicine,* 21**,** 1203-1216.

KLOVDAHL, A. S., DHOFIER, Z., ODDY, G., O'HARA, J., STOUTJESDIJK, S. & WHISH, A. (1977) Social networks in an urban area: first Canberra study. *Aust. N. Z. J. Sociol.,* 13.

KLOVDAHL, A. S., POTTERAT, J. J., WOODHOUSE, D. E., MUTH, J. B., MUTH, S. Q. & DARROW, W. W. (1994) Social networks and infectious disease - the Colorado-Springs study. *Social Science & Medicine,* 38**,** 79-88.

KNOX, E. G. (1964) The detection of space-time interactions. *The Royal Statistical Society Series C - Applied Statistics,* 13**,** 25-29.

KNOX, E. G. (1989) Detection of clusters. IN ELLIOT, P. (Ed.) *Methodology of enquiries into disease clustering.* London, Small Area Health Statistics Unit.

KULLDORFF, M., HEFFERNAN, R., HARTMAN, J., ASSUNCAO, R. & MOSTASHARI, F.

(2005) A space-time permutation scan statistic for disease outbreak detection. *Plos Medicine,* 2**,** 216-224.

KULLDORFF, M. & NAGARWALLA, N. (1995) Spatial disease clusters - Detection and inference. *Statistics in Medicine,* 14**,** 799-810.

KUNG, N., MACKENZIE, S., PITT, D., ROBINSON, B. & PERKINS, N. R. (2011) Significant features of the epidemiology of equine influenza in Queensland, Australia, 2007. *Australian Veterinary Journal,* 89**,** 78-85.

KUNG, N. Y., MORRIS, R. S., PERKINS, N. R., SIMS, L. D., ELLIS, T. M., BISSETT, L., CHOW, M., SHORTRIDGE, K. F., GUAN, Y. & PEIRIS, M. J. S. (2007) Risk for infection with highly pathogenic influenza A virus (H5N1) in chickens, Hong Kong, 2002. *Emerging Infectious Diseases,* 13**,** 412-418.

KYPRAIOS, T. (2007) Efficient Bayesian inference for partially observed stochastic epidemics and a new class of semi-parametric time series models. Ph.D. thesis. Department of Mathematics and Statistics, Lancaster University, Lancaster.

LAI, A. C. K., CHAMBERS, T. M., HOLLAND, R. E., MORLEY, P. S., HAINES, D. M., TOWNSEND, H. G. G. & BARRANDEGUY, M. (2001) Diverged evolution of recent equine-2 influenza (H3N8) viruses in the Western Hemisphere. *Archives of Virology,* 146**,** 1063-1074.

LAWSON, A. B. & ZHOU, H. (2005) Spatial statistical modeling of disease outbreaks with particular reference to the UK foot and mouth disease (FMD) epidemic of 2001. *Preventive Veterinary Medicine,* 71**,** 141-156.

LEE, K. I. & KOVAL, J. J. (1997) Determination of the best significance level in forward stepwise logistic regression. *Communications in Statistics - Simulation and Computation,* 26**,** 559 - 575.

LOCKHART, C. Y., STEVENSON, M. A., RAWDON, T. G., GERBER, N. & FRENCH, N. P. (2010) Patterns of contact within the New Zealand poultry industry. *Preventive Veterinary Medicine,* 95**,** 258-266.

LOWEN, A. C., MUBAREKA, S., STEEL, J. & PALESE, P. (2007) Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathogens,* 3**,** 1470-1476.

LOWEN, A. C., STEEL, J., MUBAREKA, S. & PALESE, P. (2008) High temperature (30 degrees C) blocks aerosol but not contact transmission of influenza virus. *Journal of Virology,* 82**,** 5650-5652.

MA, E., LAM, T., WONG, C. & CHUANG, S. K. (2010) Is hand, foot and mouth disease associated with meteorological parameters? *Epidemiology and Infection,* 138**,** 1779-1788.

MAJOR, D. A. & JONES, B. (2011) Behaviour of equine influenza virus in a naive population:

a practitioner's perspective. *Australian Veterinary Journal,* 89**,** 13-14.

MANTEL, N. (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research,* 27**,** 209-220.

MARDONES, F. O., PEREZ, A. M. & CARPENTER, T. E. (2009) Epidemiologic investigation of the re-emergence of infectious salmon anemia virus in Chile. *Diseases of Aquatic Organisms,* 84**,** 105-114.

MARDONES, F. O., PEREZ, A. M., VALDES-DONOSO, P. & CARPENTER, T. E. (2011) Farm-level reproduction number during an epidemic of infectious salmon anemia virus in southern Chile in 2007-2009. *Preventive Veterinary Medicine,* 102**,** 175-184.

MARQUETOUX, N., PAUL, M., WONGNARKPET, S., POOLKHET, C., THANAPONGTHARM, W., ROGER, F., DUCROT, C. & CHALVET-MONFRAY, K. (2012) Estimating spatial and temporal variations of the reproduction number for highly pathogenic avian influenza H5N1 epidemic in Thailand. *Preventive Veterinary Medicine,* DOI: 10.1016/j.prevetmed.2012.01.021.

MARTINEZ-LOPEZ, B., PEREZ, A. M. & SANCHEZ-VIZCAINO, J. M. (2009a) Combined application of social network and cluster detection analyses for temporal-spatial characterization of animal movements in Salamanca, Spain. *Preventive Veterinary Medicine,* 91**,** 29-38.

MARTINEZ-LOPEZ, B., PEREZ, A. M. & SANCHEZ-VIZCAINO, J. M. (2009b) Social Network Analysis. Review of General Concepts and Use in Preventive Veterinary Medicine. *Transboundary and Emerging Diseases,* 56**,** 109-120.

MATHERON, G. (1963) Principles of geostatistics. *Economic Geology and the Bulletin of the Society of Economic Geologists,* 58**,** 1246-1266.

MATTHEWS, L. & WOOLHOUSE, M. (2005) New approaches to quantifying the spread of infection. *Nat Rev Microbiol,* 3**,** 529-36.

MAY, R. M., GUPTA, S. & MCLEAN, A. R. (2001) Infectious disease dynamics: what characterizes a successful invader? *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences,* 356**,** 901-910.

MCDEVITT, J., RUDNICK, S., FIRST, M. & SPENGLER, J. (2010) Role of Absolute Humidity in the Inactivation of Influenza Viruses on Stainless Steel Surfaces at Elevated Temperatures. *Applied and Environmental Microbiology,* 76**,** 3943-3947.

MCELROY, R. D., ROTHENBERG, R. B., VARGHESE, R., WOODRUFF, R., MINNS, G. O., MUTH, S. Q., LAMBERT, L. A. & RIDZON, R. (2003) A network-informed approach to investigating a tuberculosis outbreak: implications for enhancing contact investigations. *International Journal of Tuberculosis and Lung Disease,* 7**,** S486-S493.

MCGEADY, M. L., SIAK, J. S. & CROWELL, R. L. (1979) Survival of Coxsackievirus B3 under Diverse Environmental Conditions. *Applied and Environmental Microbiology,*

37**,** 972-977.

MCKINLEY, T. J. (2007) Spatial survival analysis of infectious animal diseases. Ph.D. thesis. Exeter, University of Exeter.

MCQUEEN, J. L., DAVENPORT, F. M., KEERAN, R. J. & DAWSON, H. A. (1966) Studies on equine influenza in Michigan, 1963. II. Epizootiology. *American Journal of Epidemiology,* 83**,** 280-286.

MIETTINE, O. S. (1970) Matching and design efficiency in retrospective studies. *American Journal of Epidemiology,* 91**,** 111-118.

MIKKELSEN, T., ALEXANDERSEN, S., ASTRUP, P., CHAMPION, H. J., DONALDSON, A. I., DUNKERLEY, F. N., GLOSTER, J., SORENSEN, J. H. & THYKIER-NIELSEN, S. (2003) Investigation of airborne foot-and-mouth disease virus transmission during low-wind conditions in the early phase of the UK 2001 epidemic. *Atmospheric Chemistry and Physics,* 3**,** 2101-2110.

MILLER, W. C. (1965) Equine influenza - further observations on "coughing" outbreak 1965. *Veterinary Record,* 77**,** 455-456.

MINH, P. Q., STEVENSON, M. A., JEWELL, C., FRENCH, N. & SCHAUER, B. (2011) Spatio-temporal analyses of highly pathogenic avian influenza H5N1 outbreaks in the Mekong River Delta, Vietnam, 2009. *Spatial and Spatio-temporal Epidemiology,* 2**,** 49-57.

MINKE, J. M., EL-HAGE, C. M., TAZAWA, P., HOMER, D., LEMAITRE, L., COZETTE, V., GILKERSON, J. R. & KIRKLAND, P. D. (2011) Evaluation of the response to an accelerated immunisation schedule using a canarypox-vectored equine influenza vaccine, shortened interdose intervals and vaccination of young foals. *Australian Veterinary Journal,* 89**,** 137-139.

MINKE, J. M., TOULEMONDE, C. E., COUPIER, H., GUIGAL, P. M., DINIC, S., SINDLE, T., JESSETT, D., BLACK, L., BUBLOT, M., PARDO, M. C. & AUDONNET, J. C. (2007) Efficacy of a canarypox-vectored recombinant vaccine expressing the hemagglutinin gene of equine influenza H3N8 virus in the protection of ponies from viral challenge. *American Journal of Veterinary Research,* 68**,** 213-219.

MOLONEY, B., SERGEANT, E. S. G., CHRISTIE, B. M. & WRIGHT, T. M. (2008) Epidemiology of equine influenza in NSW. *Australian College of Veterinary Scientists - Epidemiology Chapter Meeting - Science Week Conference. Accessed online 12/7/2010 at* http://epicentre.massey.ac.nz/acvsc/html/events_2008.html.

MOLONEY, B., SERGEANT, E. S. G., TARAGEL, C. & BUCKLEY, P. (2011) Significant features of the epidemiology of equine influenza in New South Wales, Australia, 2007. *Australian Veterinary Journal,* 89**,** 56-63.

MOLONEY, B. J. (2011) Overview of the epidemiology of equine influenza in the Australian

outbreak. *Australian Veterinary Journal,* 89**,** 50-56.

MORLEY, P. S., TOWNSEND, H. G. G., BOGDAN, J. R. & HAINES, D. M. (2000a) Descriptive epidemiologic study of disease associated with influenza virus infections during three epidemics in horses. *Journal of the American Veterinary Medical Association,* 216**,** 535-544.

MORLEY, P. S., TOWNSEND, H. G. G., BOGDAN, J. R. & HAINES, D. M. (2000b) Risk factors for disease associated with influenza virus infections during three epidemics in horses. *Journal of the American Veterinary Medical Association,* 216**,** 545-550.

MORRIS, R. S., WILESMITH, J. W., STERN, M. W., SANSON, R. L. & STEVENSON, M. A. (2001) Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain, 2001. *Veterinary Record,* 149**,** 137-+.

MORTON, J. M., DUPS, J. N., ANTHONY, N. D. & DWYER, J. F. (2011) Epidemic curve and hazard function for occurrence of clinical equine influenza in a closed population of horses at a 3-day event in southern Queensland, Australia, 2007. *Australian Veterinary Journal,* 89**,** 86-88.

MULATTI, P., BOS, M. E. H., BUSANI, L., NIELEN, M. & MARANGON, S. (2010) Evaluation of interventions and vaccination strategies for low pathogenicity avian influenza: spatial and space-time analyses and quantification of the spread of infection. *Epidemiology and Infection,* 138**,** 813-824.

MUMFORD, J. A., HANNANT, D. & JESSETT, D. M. (1990) Experimental infection of ponies with equine influenza (H3N8) viruses by intranasal inoculation or exposure to aerosols. *Equine Veterinary Journal,* 22**,** 93-98.

MUMFORD, J. A., WILSON, H., HANNANT, D. & JESSETT, D. M. (1994) Antigenicity and immunogenicity of equine influenza vaccines containing a Carbomer adjuvant. *Epidemiology and Infection,* 112**,** 421-437.

MURCIA, P. R., WOOD, J. L. N. & HOLMES, E. C. (2011) Genome-Scale Evolution and Phylodynamics of Equine H3N8 Influenza A Virus. *Journal of Virology,* 85**,** 5312-5322.

MURRAY, E. J. & MORSE, S. S. (2011) Seasonal Oscillation of Human Infection with Influenza A/H5N1 in Egypt and Indonesia. *PLoS ONE,* 6**,** 1-8.

MYERS, C. & WILSON, W. D. (2006) Equine influenza virus. *Clinical Techniques in Equine Practice,* 5**,** 187-196.

NATALE, F., GIOVANNINI, A., SAVINI, L., PALMA, D., POSSENTI, L., FIORE, G. & CALISTRI, P. (2009) Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. *Preventive Veterinary Medicine,* 92**,** 341-350.

NEAL, P. J. & ROBERTS, G. O. (2004) Statistical inference and model selection for the 1861 Hagelloch measles epidemic. *Biostatistics,* 5**,** 249-261.

NEWMAN, M. E. J. (2003) Properties of Highly Clustered Networks. *Physical Review E,* 68.

NEWTON, J. R., DALY, J. M., SPENCER, L. & MUMFORD, J. A. (2006) Description of the outbreak of equine influenza (H3N8) in the United Kingdom in 2003, during which recently vaccinated horses in Newmarket developed respiratory disease. *Vet Rec,* 158**,** 185-92.

NEWTON, J. R., WOOD, J. L. N. & CHANTER, N. (2003) A case control study of factors and infections associated with clinically apparent respiratory disease in UK Thoroughbred racehorses. *Prev Vet Med,* 60**,** 107-32.

NEWTON, R., COOKE, A., ELTON, D., BRYANT, N., RASH, A., BOWMAN, S., BLUNDEN, T., MILLER, J., HAMMOND, T. A., CAMM, I. & DAY, M. (2007) Canine influenza virus: cross-species transmission from horses. *Veterinary Record,* 161**,** 142-143.

NODTVEDT, A., BERGVALL, K., SALLANDER, M., EGENVALL, A., EMANUELSON, U. & HEDHAMMAR, A. (2007) A case-control study of risk factors for canine atopic dermatitis among boxer, bullterrier and West Highland white terrier dogs in Sweden. *Veterinary dermatology,* 18**,** 309-15.

NÖREMARK, M., HÅKANSSON, N., STERNBERG LEWERIN, S., LINDBERG, A. & JONSSON, A. (2011) Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk based surveillance. *Preventive Veterinary Medicine,* 99**,** 78-90.

NSW DPI (2007a) *Equine influenza outbreak : information for horse owners. Accessed online on 20/4/2010 at http://pandora.nla.gov.au/tep/76322,* [Orange, N.S.W], NSW Dept of Primary Industries.

NSW DPI (2007b) *Equine influenza protection plan,* Orange, NSW, Australia, NSW Dept of Primary Industries.

NSW LAND AND PROPERTY MANAGEMENT AUTHORITY (2009) NSW Digital Cadastral Database (DCDB). Available at: http://www.lpi.nsw.gov.au/mapping_and_imagery/spatial_data/cadastral_data. Last accessed: 21 March 2012.

O'NEILL, P. D. & ROBERTS, G. O. (1999) Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 162**,** 121-129.

OIE (2009) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2009.

ORTIZ-PELAEZ, A., PFEIFFER, D. U., SOARES-MAGALHAES, R. J. & GUITIAN, F. J. (2006) Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Preventive Veterinary Medicine,* 76**,** 40-55.

PAILLOT, R., HANNANT, D., KYDD, J. H. & DALY, J. M. (2006) Vaccination against equine influenza: quid novi? *Vaccine,* 24**,** 4047-61.

PARK, A. W., WOOD, J. L. N., DALY, J. M., NEWTON, J. R., GLASS, K., HENLEY, W., MUMFORD, J. A. & GRENFELL, B. T. (2004) The effects of strain heterology on the epidemiology of equine influenza in a vaccinated population. *Proceedings of the Royal Society B-Biological Sciences,* 271**,** 1547-55.

PARK, A. W., WOOD, J. L. N., NEWTON, J. R., DALY, J., MUMFORD, J. A. & GRENFELL, B. T. (2003) Optimising vaccination strategies in equine influenza. *Vaccine,* 21**,** 2862-70.

PAYUNGPORN, S., CRAWFORD, P. C., KOUO, T. S., CHEN, L. M., POMPEY, J., CASTLEMAN, W. L., DUBOVI, E. J., KATZ, J. M. & DONIS, R. O. (2008) Influenza a virus (H3N8) in dogs with respiratory disease, Florida. *Emerging Infectious Diseases,* 14**,** 902-908.

PEBESMA, E. J. & BIVAND, R. S. (2005) Classes and methods for spatial data in R. *R News,* 5.

PERKINS, N. R., WEBSTER, W. R., WRIGHT, T., DENNEY, I. & LINKS, I. (2011) Vaccination program in the response to the 2007 equine influenza outbreak in Australia. *Australian Veterinary Journal,* 89**,** 126-134.

PFEIFFER, D. U., ROBINSON, T. P., STEVENSON, M., STEVENS, K. B., ROGERS, D. J. & CLEMENTS, A. C. A. (2008) *Spatial Analysis in Epidemiology,* Oxford, Oxford University Press.

PORPHYRE, T., STEVENSON, M., JACKSON, R. & MCKENZIE, J. (2008) Influence of contact heterogeneity on TB reproduction ratio R0 in a free-living brushtail possum Trichosurus vulpecula population. *Vet Res,* 39**,** 31.

POWELL, D. G., WATKINS, K. L., LI, P. H. & SHORTRIDGE, K. F. (1995) Outbreak of equine influenza among horses in Hong Kong during 1992. *Vet Rec,* 136**,** 531-6.

QLD DEPARTMENT OF ENVIRONMENT AND RESOURCE MANAGEMENT (2009) QLD Digital Cadastral Database (DCDB). Available at: http://www.derm.qld.gov.au/products/access_pricing/dig_data/cadastral_data_info.html. Last accessed: 21 March 2012.

R DEVELOPMENT CORE TEAM (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Available at: http://www.R-project.org/ Accessed 20 March 2012.

READ, A. J., FINLAISON, D. S., GU, X., DAVIS, R. J., ARZEY, K. E. & KIRKLAND, P. D. (2011) Application of real-time PCR and ELISA assays for equine influenza virus to determine the duration of viral RNA shedding and onset of antibody response in naturally infected horses. *Australian Veterinary Journal,* 89**,** 42-43.

REEVES, A., SALMAN, M. D. & HILL, A. E. (2011) Approaches for evaluating veterinary epidemiological models: verification, validation and limitations. In: Models in the management of animal diseases (P. Willeberg, ed.). *Revue Scientifique Et Technique De L Office International Des Epizooties,* 30**,** 499-512.

RIBBENS, S., DEWULF, J., KOENEN, F., MINTIENS, K., DE KRUIF, A. & MAES, D. (2009) Type and frequency of contacts between Belgian pig herds. *Preventive Veterinary Medicine,* 88**,** 57-66.

RIBEIRO, P. J. & DIGGLE, P. J. (2001) geoR: a package for geostatistical analysis. *R News,* 1**,** 15-18.

ROBINSON, S. E. & CHRISTLEY, R. M. (2007) Exploring the role of auction markets in cattle movements within Great Britain. *Preventive Veterinary Medicine,* 81**,** 21-37.

ROBINSON, S. E., EVERETT, M. G. & CHRISTLEY, R. M. (2007) Recent network evolution increases the potential for large epidemics in the British cattle population. *Journal of the Royal Society Interface,* 4**,** 669-674.

RORRES, C., PELLETIER, S. T. K., KEELING, M. J. & SMITH, G. (2010) Estimating the kernel parameters of premises-based stochastic models of farmed animal infectious disease epidemics using limited, incomplete, or ongoing data. *Theoretical Population Biology,* 78**,** 46-53.

RYAN, D. (2011) Determining the endpoint of an outbreak of equine influenza in a large population of racing Thoroughbreds. *Australian Veterinary Journal,* 89**,** 25-27.

SALTELLI, A., TARANTOLA, S. & CAMPOLONGO, F. (2004) *Sensitivity analysis in practice: a guide to assessing scientific models,* West Sussex, England, John Wiley & Sons, Ltd.

SANSON, R. L. (1994) The development of a decision support system for an animal disease emergency. Ph.D. thesis. Massey University, Palmerston North, New Zealand.

SANSON, R. L., GLOSTER, J. & BURGIN, L. (2011a) Reanalysis of the start of the UK 1967 to 1968 foot-and-mouth disease epidemic to calculate airborne transmission probabilities. *Veterinary Record,* 169**,** 336-U44.

SANSON, R. L., HARVEY, N., GARNER, M. G., STEVENSON, M. A., DAVIES, T. M., HAZELTON, M. L., O'CONNOR, J., DUBÉ, C., FORDE-FOLLE, K. N. & OWEN, K. (2011b) Foot and mouth disease model verification and 'relative validation' through a formal model comparison. In: Models in the management of animal diseases (P. Willeberg, ed.). *Revue Scientifique Et Technique De L Office International Des Epizooties,* 30**,** 527-540.

SARAMÄKI, J. & KASKI, K. (2005) Modelling development of epidemics with dynamic small-world networks. *Journal of Theoretical Biology,* 234.

SATOU, K. & NISHIURA, H. (2006) Basic reproduction number for equine-2 influenza virus

A (H3N8) epidemic in racehorse facilities in Japan, 1971. *Journal of Equine Veterinary Science,* 26**,** 310-316.

SAVILL, N. J., SHAW, D. J., DEARDON, R., TILDESLEY, M. J., KEELING, M. J., WOOLHOUSE, M. E., BROOKS, S. P. & GRENFELL, B. T. (2006) Topographic determinants of foot and mouth disease transmission in the UK 2001 epidemic. *BMC Veterinary Research,* 2**,** 3.

SAWABE, K., TANABAYASHI, K., HOTTA, A., HOSHINO, K., ISAWA, H., SASAKI, T., YAMADA, A., KURAHASHI, H., SHUDO, C. & KOBAYASHI, M. (2009) Survival of Avian H5N1 Influenza A Viruses in Calliphora nigribarbis (Diptera: Calliphoridae). *Journal of Medical Entomology,* 46**,** 852-855.

SCHAFFER, F. L., SOERGEL, M. E. & STRAUBE, D. C. (1976) Survival of airborne influenza virus: effects of propagating host, relative humidity, and composition of spray fluids. *Archives of Virology,* 51**,** 263-273.

SCHEMANN, K., FIRESTONE, S. M., TAYLOR, M. R., TORIBIO, J. A. L. M. L., WARD, M. P. & DHAND, N. K. (2012) Horse owners'/managers' perceptions about effectiveness of biosecurity measures based on their experiences during the 2007 equine influenza outbreak in Australia. *Preventive Veterinary Medicine,* DOI:10.1016/j.prevetmed.2012.01.013.

SCHLEY, D., BURGIN, L. & GLOSTER, J. (2009) Predicting infection risk of airborne foot-and-mouth disease. *Journal of the Royal Society Interface,* 6**,** 455-462.

SCHOENBAUM, M. A. & DISNEY, W. T. (2003) Modeling alternative mitigation strategies for a hypothetical outbreak of foot-and-mouth disease in the United States. *Preventive Veterinary Medicine,* 58**,** 25-52.

SCHOLTENS, R. G., STEELE, J. H., DOWDLE, W. R., YARBROUGH, W. B. & ROBINSON, R. Q. (1964) United States epizootic of equine influenza, 1963. *Public Health Reports,* 79**,** 393-402.

SENA, A. C., MUTH, S. Q., HEFFELFINGER, J. D., O'DOWD, J. O., FOUST, E. & LEONE, P. (2007) Factors and the sociosexual network associated with a syphilis outbreak in rural North Carolina. *Sexually Transmitted Diseases,* 34**,** 280-287.

SERGEANT, E. S. G., COWLED, B. D. & BINGHAM, P. (2011a) Diagnostic specificity of an equine influenza blocking ELISA estimated from New South Wales field data from the Australian epidemic in 2007. *Australian Veterinary Journal,* 89**,** 43-45.

SERGEANT, E. S. G., KIRKLAND, P. D. & COWLED, B. D. (2009) Field evaluation of an equine influenza ELISA used in New South Wales during the 2007 Australian outbreak response. *Preventive Veterinary Medicine,* 92**,** 382-385.

SERGEANT, E. S. G., STONE, M., MOLONEY, B. J. & ARTHUR, R. (2011b) Quantitative analysis of the risk of spread of equine influenza associated with movements of

vaccinated horses from infected areas during the Australian outbreak. *Australian Veterinary Journal,* 89**,** 103-108.

SERGEANT, E. S. G. & WILSON, G. (2011) Demonstrating freedom from equine influenza in New South Wales, Australia, following the 2007 outbreak. *Australian Veterinary Journal,* 89**,** 164-169.

SHAMAN, J. & KOHN, M. (2009) Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 3243-3248.

SHARKEY, K. J., BOWERS, R. G., MORGAN, K. L., ROBINSON, S. E. & CHRISTLEY, R. M. (2008) Epidemiological consequences of an incursion of highly pathogenic H5N1 avian influenza into the British poultry flock. *Proceedings of the Royal Society B-Biological Sciences,* 275**,** 19-28.

SHIRLEY, M. D. F. & RUSHTON, S. P. (2005) Where diseases and networks collide: lessons to be learnt from a study of the 2001 foot-and-mouth disease epidemic. *Epidemiol Infect,* 133**,** 1023-32.

SHOJI, M., KATAYAMA, K. & SANO, K. (2011) Absolute Humidity as a Deterministic Factor Affecting Seasonal Influenza Epidemics in Japan. *Tohoku Journal of Experimental Medicine,* 224**,** 251-256.

SMALL, M., WALKER, D. M. & TSE, C. K. (2007) Scale-free distribution of avian influenza outbreaks. *Physical Review Letters,* 99.

SMITH, K. C., DALY, J. M., BLUNDEN, A. S. & LAURENCE, C. J. (2005) Canine influenza virus. *Veterinary Record,* 157**,** 599-599.

SMYTH, G. B. & DAGLEY, K. (2011) Internet-based survey of horse owners for mortality and morbidity related to equine influenza in the 2007 Australian epidemic. *Australian Veterinary Journal,* 89**,** 23-25.

SOBOLL, G., HOROHOV, D. W., ALDRIDGE, B. M., OLSEN, C. W., MCGREGOR, M. W., DRAPE, R. J., MACKLIN, M. D., SWAIN, W. F. & LUNN, D. P. (2003) Regional antibody and cellular immune responses to equine influenza virus infection, and particle mediated DNA vaccination. *Veterinary Immunology and Immunopathology,* 94**,** 47-62.

SONG, D., KANG, B., LEE, C., JUNG, K., HA, G., KANG, D., PARK, S., PARK, B. & OH, J. (2008) Transmission of avian influenza virus (H3N2) to dogs. *Emerging Infectious Diseases,* 14**,** 741-6.

SONG, D., LEE, C., KANG, B., JUNG, K., OH, T., KIM, H., PARK, B. & OH, J. (2009) Experimental infection of dogs with avian-origin canine influenza A virus (H3N2). *Emerging Infectious Diseases,* 15**,** 56-8.

SOVINOVÁ, O., TUMOVÁ, B., POUSKA, F. & NEMEC, J. (1958) Isolation of a virus causing respiratory disease in horses. *Acta Virologica,* 2**,** 51-61.

SPICKLER, A. R. (2009) Influenza - Technical Factsheet., The Center for Food Security and Public Health (CFSPH), Iowa State University. Available at: http://www.cfsph.iastate.edu/DiseaseInfo/factsheets.php.

SPOKES, P. J., MARICH, A. J. N., MUSTO, J. A., WARD, K. A., CRAIG, A. T. & MCANULTY, J. M. (2009) Investigation of equine influenza transmission in NSW: walk, wind or wing? *NSW Public Health Bulletin,* 20**,** 152-156.

SSEMATIMBA, A., HAGENAARS, T. J. & DE JONG, M. C. M. (2012) Modelling the wind-borne spread of highly pathogenic avian influenza virus between farms. *PLoS ONE,* 7**,** 1-9.

STEEL, J., PALESE, P. & LOWEN, A. C. (2011) Transmission of a 2009 Pandemic Influenza Virus Shows a Sensitivity to Temperature and Humidity Similar to That of an H3N2 Seasonal Strain. *Journal of Virology,* 85**,** 1400-1402.

STEVENSON, M. (2009) Investigation of Spatial Patterns of Animal Disease - Notes for MVS course 195.821 Advanced Analysis and Interpretation of Animal Health Data., EpiCentre, IVABS Massey University, Palmerston North, New Zealand.

STEVENSON, M., SANCHEZ, J. & THORNTON, R. (2010) epiR: Functions for analysing epidemiological data. R package version 0.9-27. Available at: http://CRAN.R-project.org/package=epiR Accessed: 20 March 2012.

STEVENSON, M. A., BENARD, H., BOLGER, P. & MORRIS, R. S. (2005a) Spatial epidemiology of the Asian honey bee mite (Varroa destructor) in the North Island of New Zealand. *Preventive Veterinary Medicine,* 71**,** 241-252.

STEVENSON, M. A., MORRIS, R. S., LAWSON, A. B., WILESMITH, J. W., RYAN, J. B. M. & JACKSON, R. (2005b) Area-level risks for BSE in British cattle before and after the July 1988 meat and bone meal feed ban. *Preventive Veterinary Medicine,* 69**,** 129-144.

STEVENSON, M. A., WILESMITH, J. W., RYAN, J. B. M., MORRIS, R. S., LOCKHART, J. W., LIN, D. & JACKSON, R. (2000) Temporal aspects of the epidemic of bovine spongiform encephalopathy in Great Britain: individual animal-associated risk factors for the disease. *Veterinary Record,* 147**,** 349-354.

TAKAHASHI, K., KULLDORFF, M., TANGO, T. & YIH, K. (2008) A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics,* 7.

TAUBENBERGER, J. K. & KASH, J. C. (2010) Influenza Virus Evolution, Host Adaptation, and Pandemic Formation. *Cell Host & Microbe,* 7**,** 440-451.

TAYLOR, M. R., AGHO, K. E., STEVENS, G. J. & RAPHAEL, B. (2008) Factors influencing psychological distress during a disease epidemic: Data from Australia's first outbreak of equine influenza. *Bmc Public Health,* 8.

TAYLOR, N. (2003) Review of the use of models in informing disease control policy

development and adjustment. A report for DEFRA. Reading, UK, Veterinary Epidemiology and Economics Research Unit (VEERU), University of Reading.

TAYLOR, N. M., HONHOLD, N., PATERSON, A. D. & MANSLEY, L. M. (2004) Risk of foot-and-mouth disease associated with proximity in space and time to infected premises and the implications for control policy during the 2001 epidemic in Cumbria. *Veterinary Record,* 154**,** 617-626.

TELLIER, R. (2006) Review of aerosol transmission of influenza A virus. *Emerging Infectious Diseases,* 12**,** 1657-1662.

TELLIER, R. (2009) Aerosol transmission of influenza A virus: a review of new studies. *Journal of the Royal Society Interface,* 6**,** S783-S790.

THE EI EPIDEMIOLOGY SUPPORT GROUP (2008) Equine influenza 2007. The Australian Experience: Report from the EI Epidemiology Support Group to the
Consultative Committee on Emergency Animal Disease.

THERNEAU, T. M. & GRAMBSCH, P. M. (2000) *Modeling Survival Data: Extending the Cox Model.,* Berlin, Springer.

THERNEAU, T. M., GRAMBSCH, P. M. & FLEMING, T. R. (1990) Martingale-based residuals for survival models. *Biometrika,* 77**,** 147-160.

THRUSFIELD, M., MANSLEY, L., DUNLOP, P., PAWSON, A. & TAYLOR, J. (2005a) The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 2: Serosurveillance, and efficiency and effectiveness of control procedures after the national ban on animal movements. *Veterinary Record,* 156**,** 269-278.

THRUSFIELD, M., MANSLEY, L., DUNLOP, P., TAYLOR, J., PAWSON, A. & STRINGER, L. (2005b) The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 1: Characteristics and control. *Veterinary Record,* 156**,** 229-252.

THRUSFIELD, M., ORTEGA, C., DE BLAS, I., NOORDHUIZEN, J. P. & FRANKENA, K. (2001) WIN EPISCOPE 2.0: improved epidemiological software for veterinary medicine. *Veterinary Record,* 148**,** 567-572.

THRUSFIELD, M. V. (2005) *Veterinary epidemiology, 3rd edition.,* Oxford, UK, Blackwell Science Ltd.

TILDESLEY, M. J., SMITH, G. & KEELING, M. J. (2011) Modeling the spread and control of foot-and-mouth disease in Pennsylvania following its discovery and options for control. *Preventive veterinary medicine,* 104**,** 224-39.

TRAVERS, J. & MILGRAM, S. (1969) An Experimental Study of the Small World Problem. *Sociometry,* 32**,** 425-443.

TURNER, J., BOWERS, R. G., CLANCY, D., BEHNKE, M. C. & CHRISTLEY, R. M. (2008) A network model of E. coli O157 transmission within a typical UK dairy herd: the effect of heterogeneity and clustering on the prevalence of infection. *J Theor Biol,* 254**,**

45-54.

UPPAL, P. K. & YADAV, M. P. (1987) Outbreak of equine influenza in India. *Veterinary Record,* 121**,** 569-570.

VAN BUUREN, S. & GROOTHUIS-OUDSHOORN, C. G. M. (in press) MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Computation and Simulation*.

VELING, J., WILPSHAAR, H., FRANKENA, K., BARTELS, C. & BARKEMA, H. W. (2002) Risk factors for clinical Salmonella enterica subsp enterica serovar Typhimurium infection on Dutch dairy farms. *Preventive Veterinary Medicine,* 54**,** 157-168.

VON KLOT, S., GRYPARIS, A., TONNE, C., YANOSKY, J., COULL, B. A., GOLDBERG, R. J., LESSARD, D., MELLY, S. J., SUH, H. H. & SCHWARTZ, J. (2009) Elemental Carbon Exposure at Residence and Survival After Acute Myocardial Infarction. *Epidemiology,* 20**,** 547-554.

WARD, M. P. & CARPENTER, T. E. (2000a) Analysis of time-space clustering in veterinary epidemiology. *Preventive Veterinary Medicine,* 43**,** 225-237.

WARD, M. P. & CARPENTER, T. E. (2000b) Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Preventive Veterinary Medicine,* 45**,** 257-284.

WARD, M. P. & FARNSWORTH, M. L. (2009) An Evaluation Of The Space-Time Permutation Test For Detecting Disease Clusters. *Proceedings of the 12th International Symposium on Veterinary Epidemiology and Economics (ISVEE), 2009.* Durban, South Africa.

WARD, M. P., HIGHFIELD, L. D., VONGSENG, P. & GARNER, M. G. (2009a) Simulation of foot-and-mouth disease spread within an integrated livestock system in Texas, USA. *Preventive Veterinary Medicine,* 88**,** 286-297.

WARD, M. P., MAFTEI, D., APOSTU, C. & SURU, A. (2008) Geostatistical visualisation and spatial statistics for evaluation of the dispersion of epidemic highly pathogenic avian influenza subtype H5N1. *Veterinary Research,* 39**,** 22.

WARD, M. P., MAFTEI, D., APOSTU, C. & SURU, A. (2009b) Estimation of the basic reproductive number (R-0) for epidemic, highly pathogenic avian influenza subtype H5N1 spread. *Epidemiology and Infection,* 137**,** 219-226.

WASSERMAN, S. & FAUST, K. (1994) *Social Network Analysis: Methods and Applications,* Cambridge, New York, Cambridge University Press.

WATSON, J., HALPIN, K., SELLECK, P., AXELL, A., BRUCE, K., HANSSON, E., HAMMOND, J., DANIELS, P. & JEGGO, M. (2011a) Isolation and characterisation of an H3N8 equine influenza virus in Australia, 2007. *Australian Veterinary Journal,* 89**,** 35-37.

WATSON, J., SELLECK, P., AXELL, A., BRUCE, K., TAYLOR, T., HEINE, H., DANIELS, P. & JEGGO, M. (2011b) Diagnosis of equine influenza virus infections in quarantine stations in Australia, 2007. *Australian Veterinary Journal,* 89**,** 4-6.

WATTS, D. J. & STROGATZ, S. H. (1998) Collective dynamics of 'small-world' networks. *Nature,* 393**,** 440-442.

WEBB, C. R. (2005) Farm animal networks: unraveling the contact structure of the British sheep population. *Preventive Veterinary Medicine,* 68**,** 3-17.

WEBB, C. R. (2006) Investigating the potential spread of infectious diseases of sheep via agricultural shows in Great Britain. *Epidemiology and Infection,* 134**,** 31-40.

WEBER, T. P. & STILIANAKIS, N. I. (2008) Inactivation of influenza A viruses in the environment and modes of transmission: A critical review. *Journal of Infection,* 57**,** 361-373.

WEBSTER, R. G. & GUO, Y. J. (1991) New influenza-virus in horses. *Nature,* 351**,** 527-527.

WHITE, P. J., WINDSOR, P. A., DHAND, N. K. & TORIBIO, J. (2010) Risk factors for congenital chondrodystrophy of unknown origin in beef cattle herds in south-eastern Australia. *Preventive Veterinary Medicine,* 96**,** 36-48.

WILESMITH, J. W., RYAN, J. B. M., STEVENSON, M. A., MORRIS, R. S., PFEIFFER, D. U., LIN, D., JACKSON, R. & SANSON, R. L. (2000) Temporal aspects of the epidemic of bovine spongiform encephalopathy in Great Britain: holding-associated risk factors for the disease. *Veterinary Record,* 147**,** 319-325.

WILESMITH, J. W., STEVENSON, M. A., KING, C. B. & MORRIS, R. S. (2003) Spatio-temporal epidemiology of foot-and-mouth disease in two counties of Great Britain in 2001. *Preventive Veterinary Medicine,* 61**,** 157-170.

WILSON, G., COOPER, K., WILLIAMS, J., EASTWOOD, S. & PEAKE, C. (2011) Equine influenza immunity in the Special Restricted Area (Purple Zone) of New South Wales, Australia. *Australian Veterinary Journal,* 89**,** 116-120.

WILSON, W. D. (1993) Equine influenza. *Veterinary Clinics of North America. Equine Practice,* 9**,** 257-82.

WONG, D. (2011) Equine influenza: a clinical perspective in Centennial Parklands Equestrian Centre. *Australian Veterinary Journal,* 89**,** 15-16.

WOOD, J., SMITH, K. C., DALY, J. M. & NEWTON, J. R. (2007) Viral infections of the equine respiratory tract. IN MCGORUM, B. C., DIXON, P. M., ROBINSON, N. E. & SCHUMACHER, J. (Eds.) *Equine respiratory medicine and surgery.* Philadelphia, PA, Saunders Elsevier.

WOODRUFF, R. E., GUEST, C. S., GARNER, M. G., BECKER, N., LINDESAY, J., CARVAN, T. & EBI, K. (2002) Predicting Ross River virus epidemics from regional weather data. *Epidemiology,* 13**,** 384-393.

WOODWARD, M. (2005) *Epidemiology: study design and data analysis. Second edition.,* Boca Raton, Florida, Chapman & Hall/CRC.

WOOLHOUSE, M. E. J., SHAW, D. J., LIU, L., MELLOR, D. J. & THOMAS, M. R. (2005) Epidemiological implications of the contact network structure for cattle farms and the 20–80 rule. *Biol. Lett.,* 1.

WORLD BANK (2011) World Development Indicators online data catalogue. Available at: http://data.worldbank.org/ Accessed 20 March 2012.

WORLD HEALTH ORGANIZATION (2003) Severe Acute Respiratory Syndrome (SARS): Status of the Outbreak and Lessons for the Immediate Future. World Health Organization, Geneva, Available online at: http://www.who.int/csr/sars/resources/en/index.html.

WORLD HEALTH ORGANIZATION (2009) New influenza A (H1N1) virus infections: global surveillance summary, May 2009. *Wkly Epidemiol. Rec.,* 84**,** 173-184.

YADAV, M. P., UPPAL, P. K. & MUMFORD, J. A. (1993) Physico-chemical and biological characterization of A/Equi-2 virus isolated from 1987 equine influenza epidemic in India. *International Journal of Animal Sciences,* 8**,** 93-98.

YAMANAKA, T., NIWA, H., TSUJIMURA, K., KONDO, T. & MATSUMURA, T. (2008) Epidemic of equine influenza among vaccinated racehorses in Japan in 2007. *Journal of Veterinary Medical Science,* 70**,** 623-625.

YUAN, J. S., YUN, H. M., LAN, W., WANG, W., SULLIVAN, S. G., JIA, S. W. & BITTLES, A. H. (2006) A climatologic investigation of the SARS-CoV outbreak in Beijing, China. *American Journal of Infection Control,* 34**,** 234-236.

Supporting letters from co-authors of published papers

Faculty of Veterinary Science
The University of Sydney

12 June, 2012

To Whom It May Concern,

We the undersigned are writing this letter to stipulate the role of Simon Firestone in the preparation and submission of the following journal article:

> *Firestone, S.M., Schemann, K.A., Toribio, J.A., Ward, M.P., Dhand, N.K., 2011. A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. Prev. Vet. Med. 100, 53-63.*

Simon Firestone, during his PhD candidature, was the primary author responsible for formulating the research questions, designing and conducting the study and statistical analyses, drafting the manuscript, responding to reviewers' reports and coordinating submission and publication of the original research paper.

Kathrin Schemann assisted with piloting the questionnaire, conducted half of the interviews and entered data. All co-authors were involved in planning the study, and commented on the questionnaire design, manuscript for submission, response to reviewers' reports and final version of the original research paper. Further individual contributions to compilation of the dataset used to construct the sampling frame and provision of comments on study design are specifically acknowledged in the published paper.

Sincerely,

Ms Kathrin Schemann

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1780
Fax: +61 2 9531 1693

A/Prof Jenny-Ann Toribio

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1609
Fax: +61 2 9531 1618

Prof Michael Ward

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1607
Fax: +61 2 9531 1618

Dr Navneet Dhand

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1669
Fax: +61 2 9531 1693

Faculty of Veterinary Science
The University of Sydney

12 June, 2012

To Whom It May Concern,

We the undersigned are writing this letter to stipulate the role of Simon Firestone in the preparation and submission of the following journal article:

> *Firestone, S.M., Christley, R.M., Ward, M.P., Dhand, N.K., 2012. Adding the spatial dimension to the social network analysis of an epidemic: investigation of the 2007 outbreak of equine influenza in Australia. Prev. Vet. Med. DOI: 10.1016/j.prevetmed.2012.01.020.*

Simon Firestone, during his PhD candidature, was the primary author responsible for formulating the research questions, preparing the datasets for analysis, designing and conducting the study and statistical analyses, writing the draft manuscript, responding to reviewer's reports and coordinating submission and publication of the original research paper.

Robert Christley assisted with refining the research methods and implementing the social network analyses. Peter Thomson advised on the development of the likelihood-based approaches presented in the published paper. All co-authors reviewed and commented on the design of the study, the manuscript for submission, response to reviewers' reports and the final version of the original research paper. Further individual contributions to compilation of the contact-tracing and spatial datasets and provision of comments on study design are specifically acknowledged in the published paper.

Sincerely,



Dr Robert Christley

Faculty of Health and Life Sciences,
The University of Liverpool,
Leahurst Campus, Neston,
The United Kingdom CH64 7TE
Tel: + 44 151 794 6170
Fax: + 44 151 794 6028

Prof Michael Ward

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1607
Fax: +61 2 9531 1618

Dr Navneet Dhand

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1669
Fax: +61 2 9531 1693

Faculty of Veterinary Science
The University of Sydney

12 June, 2012

To Whom It May Concern,

We the undersigned are writing this letter to stipulate the role of Simon Firestone in the preparation and submission of the following journal article:

> *Firestone, S.M., Ward, M.P., Christley, R.M., Dhand, N.K., 2011. The importance of location in contact networks: describing disease spread using spatial social network analysis. Prev. Vet. Med. 102, 185-195.*

Simon Firestone, during his PhD candidature, was the primary author responsible for formulating the research questions, designing and conducting the study and statistical analyses, writing the draft manuscript, responding to reviewer's reports and coordinating submission and publication of the original research paper.

Michael Ward provided specific guidance on the geostatistical approach and spatial cluster detection methods. All co-authors reviewed and commented on the design of the study, the manuscript for submission, response to reviewers' reports and the final version of the original research paper. Further individual contributions to compilation of the contact-tracing and spatial datasets and provision of comments on study design are specifically acknowledged in the published paper.

Sincerely,



Prof Michael Ward

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1607
Fax: +61 2 9531 1618

Dr Robert Christley

Faculty of Health and Life Sciences,
The University of Liverpool,
Leahurst Campus, Neston,
The United Kingdom CH64 7TE
Tel: + 44 151 794 6170
Fax: + 44 151 794 6028

Dr Navneet Dhand

Faculty of Veterinary Science,
The University of Sydney
425 Werombi Road, Camden
NSW 2570
Tel: + 61 2 9531 1669
Fax: +61 2 9531 1693

Faculty of Veterinary Science
The University of Sydney

12 June, 2012

To Whom It May Concern,

We the undersigned are writing this letter to stipulate the role of Simon Firestone in the preparation and submission of the following journal article:

> *Firestone, S.M., Cogger, N., Ward, M.P., Toribio, J.-A.L.M.L., Moloney, B.J., Dhand, N.K., 2012. The Influence of Meteorology on the Spread of Influenza: Survival Analysis of an Equine Influenza (A/H3N8) Outbreak. PLoS ONE 7, e35284.*

Simon Firestone, during his PhD candidature, was the primary author responsible for refining the research questions, designing and conducting the study and statistical analyses, writing the draft manuscript, responding to reviewer's reports and coordinating submission and publication of the original research paper.

Naomi Cogger originally proposed that a survival analysis be conducted on the equine influenza dataset and provided ongoing assistance in the formulation of the survival analysis dataset. All co-authors reviewed and commented on the design of the study, the manuscript for submission, response to reviewers' reports and the final version of the original research paper. Further individual contributions to compilation of the equine influenza datasets and provision of comments on study design are specifically acknowledged in the published paper.

Sincerely,

Dr Naomi Cogger

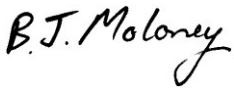
The EpiCentre, Institute of Veterinary Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand
Tel: + 64 6 350 5553
Fax: + 64 6 350 5716

Prof Michael Ward

Faculty of Veterinary Science, The University of Sydney
425 Werombi Road, Camden NSW 2570
Tel: + 61 2 9531 1607
Fax: +61 2 9531 1618

A/Prof Jenny-Ann Toribio

Faculty of Veterinary Science, The University of Sydney
425 Werombi Road, Camden NSW 2570
Tel: + 61 2 9531 1609
Fax: +61 2 9531 1618



Dr Barbara Moloney

New South Wales Government Department of Primary Industries, Orange, New South Wales, Australia
Tel: + 61 2 6391 3687
Fax: + 61 2 6361 9976

Dr Navneet Dhand

Faculty of Veterinary Science, The University of Sydney
425 Werombi Road, Camden NSW 2570
Tel: + 61 2 9531 1669
Fax: +61 2 9531 1693