ELSEVIER

Adding the spatial dimension to the social network analysis of an epidemic: Investigation of the 2007 outbreak of equine influenza in Australia

Simon M. Firestone [a,*], Robert M. Christley [b], Michael P. Ward [a], Navneet K. Dhand [a]

[a] Faculty of Veterinary Science, The University of Sydney, 425 Werombi Road, Camden, NSW 2570, Australia
[b] Faculty of Health and Life Sciences, The University of Liverpool, Leahurst Campus, Neston CH64 7TE, United Kingdom

ARTICLE INFO

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. In 2007, in the first reported outbreak in Australia, the virus spread through the horse populations of two states within 4 months. 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. This study applies social network analysis to describe spread of equine influenza between horse premises infected in the early outbreak period, identifying spread occurring through a contact network and secondary local spatial spread.

Social networks were constructed by combining contact-tracing data on horse movements with a distance matrix between all premises holding horses infected within the first 10 days of the outbreak. These networks were analysed to provide a description of the epidemic, identify premises that were central to disease spread and to estimate the relative proportion of premises infected through infected horse movements and through local spatial spread. We then explored the effect of distance on disease spread by estimating the range of local spread (through direct contact, transmission on fomites and windborne transmission) based on the level of fragmentation in the network and also by directly estimating the shape of the outbreak's spatial transmission kernel.

During the first 10 days of this epidemic, 197 horse premises were infected; 70 of these were included in the contact-traced network. Most local spread occurred within 5 km. Local spread was estimated to have occurred up to a distance of 15.3 km – based on the contact-and-proximity network – and at a very low incidence beyond this distance based on the transmission kernel estimate. Of the 70 premises in the contact network, spread to 14 premises (95% CI: 9, 20 premises) was likely to have occurred through local spatial spread from nearby infected premises, suggesting that 28.3% of spread in the early epidemic period was 'network-associated' (95% CI: 25.6, 31.0%). By constructing a 'maximal network' of contact and proximity (based on a distance cut-off of 15.3 km), 44 spatial clusters were described, and the horse movements that initiated infection in these locations were identified.

Characteristics of the combined network, incorporating both spatial and underlying contact relationships between infected premises, explained the high rate of spread, the sequence of cluster formation and the widespread dispersal experienced in the early phase of this epidemic. These results can inform outbreak control planning by guiding the imposition of appropriate control zone diameters around infected premises and the targeting of surveillance and interventions.

* Corresponding author. Tel.: +61 2 9036 7736; fax: +61 2 9351 1693.
E-mail address: simon.firestone@sydney.edu.au (S.M. Firestone).

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 north-east 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 h 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 70 000 horses on over 9000 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., 2009). Clinical signs were first observed in a recently imported horse in quarantine 40 km west of Sydney (Fig. 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., 2009). 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., 2011).

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 (Shirley and Rushton, 2005; Ortiz-Pelaez et al., 2006). These approaches have largely ignored spatial proximity between infected premises. Recently, Green et al. (2008) developed a spatially explicit model to estimate the importance of cattle movements in the spread of bovine tuberculosis, whilst García Álvarez et al. (2011) investigated the importance of animal movements, contacts and spatial relationships between dairy farms in the context of endemic disease spread. 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 (Webb, 2005; Martinez-Lopez et al., 2009).

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
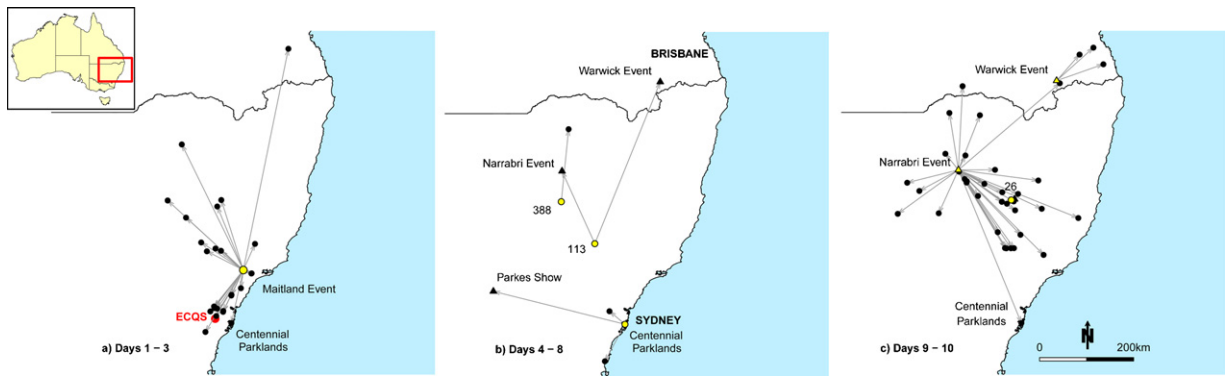
Fig. 1. Infected horse movements during the first 10 days of the equine influenza outbreak of 2007 in Australia. (a) 17–19 August 2007, (b) 20–24 August 2007, (c) 25–26 August 2007. Nodes are coloured yellow and labelled if their out-degree ≥ 1. Horse events where transmission is known to have occurred 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. (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.)

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.

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.

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 (Callinan, 2008; The EI Epidemiology Support Group, 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.

2.2. Data management and preparation

The infected premises dataset had previously been prepared for epidemiological analyses as described elsewhere (Cowled et al., 2009). 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 Foundation for Statistical Computing, Vienna, Austria).

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).

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 h 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.

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 pairwise 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., 2011). 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.

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.

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 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.

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

Table 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	Heterogeneity in a node-level parameters, used to describes the extent that the network revolves around any single
Degree	node. Calculated as a normalised proportion, summing the differences between the largest value in the network and
Reach	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.

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. (2009) 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 naive estimate (ρ) for the early outbreak period, using the following formula (May et al., 2001):

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

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.

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:

$$\pi_{Si} = 1 - \pi_{Ni}$$

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 ($j = 1,...,m$), and the Euclidean distance to each such potential source premises was calculated (d_{ij}). The probability that premises i was infected by each of these potential source premises, $p(d_{ij})$, was estimated using the transmission kernel. The complementary probabilities, $[1 - p(d_{ij})]$, 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:

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

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:

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

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 $var(\pi_N) = var(\pi_S)$.

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.

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 (Fig. 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).

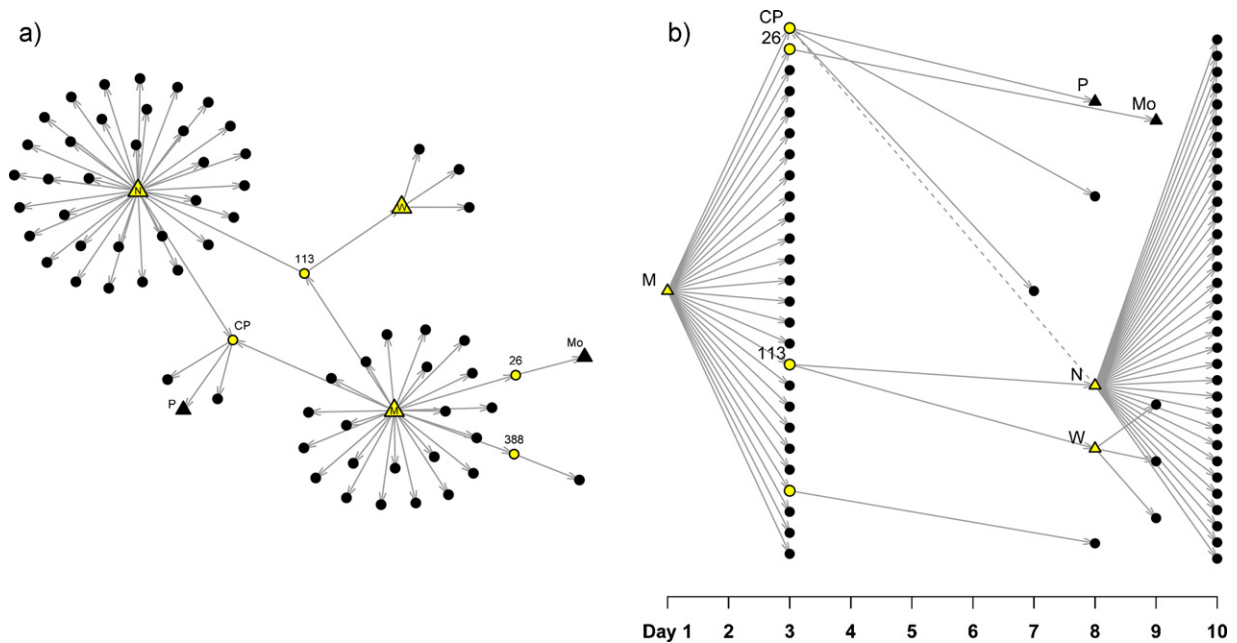### 3.1. The contact network of infected horse movements

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 (Fig. 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 Fig. 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' (Fig. 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 2). It consisted of a single component, including 36% of the nodes, and 127 unconnected premises (these isolates are excluded from Fig. 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 $R = 14.6$.

3.2. The proximity of infected premises in space

The proximity network was highly fragmented irrespective of the distance cut-off (Fig. 3). With the distance cut-off dichotomised at 5 km, the proximity network comprised 28 relatively small components and 54 isolated premises (Fig. 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

Fig. 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. (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.)

Table 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. (2011).

[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. (2009).
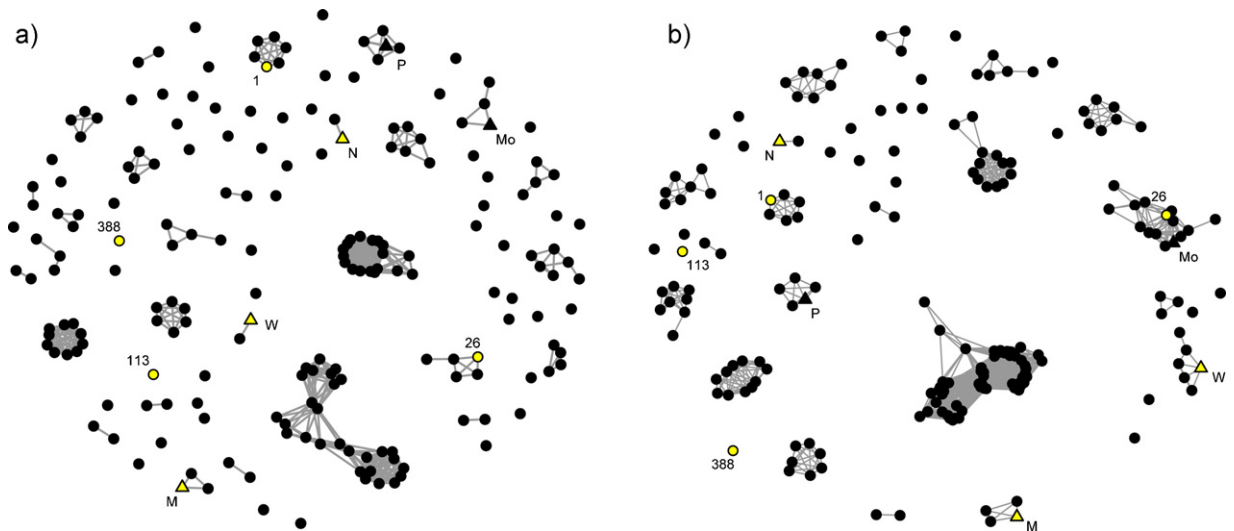
Fig. 3. Proximity networks representing spatial relationships between premises holding horses infected in the first 10 days of the 2007 equine influenza outbreak in Australia. (a) Distance cut-off dichotomised at 5 km, (b) 'Maximal network' with distance cut-off of 15.3 km. Graphed in arbitrary space, 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. (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.)

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).

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 (Fig. 4). With the distance cut-off set at 5 km, this network contained 6 components and 19 isolated nodes (Fig. 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.
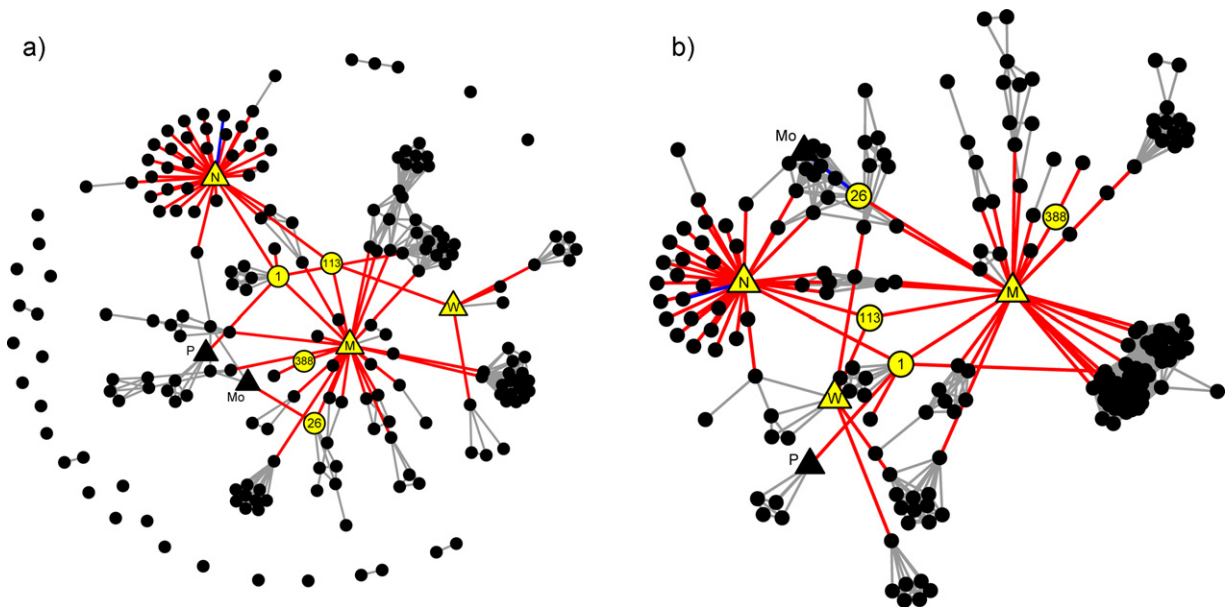
Overall, this contact-and-proximity network (with distance cut-off set at 5 km) inherited a relatively high out-degree centralisation from the contact network, reflecting the relative importance of several nodes with high out-degree in the contact network to the combined network structure. This network 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.
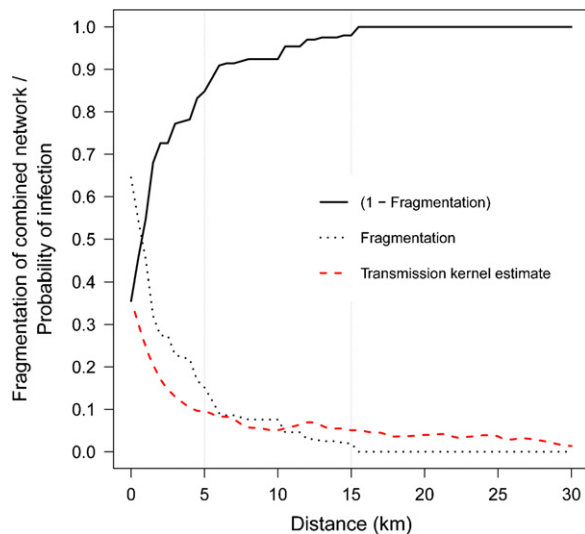
3.4. Estimates of the range of local spatial spread

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 (Fig. 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 Fig. 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 (Fig. 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). Figs. 3b and 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 (Fig. 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

Fig. 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), other labels are unique premises identifiers. (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.)

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%).

Fig. 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.

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 (Fig. 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. 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
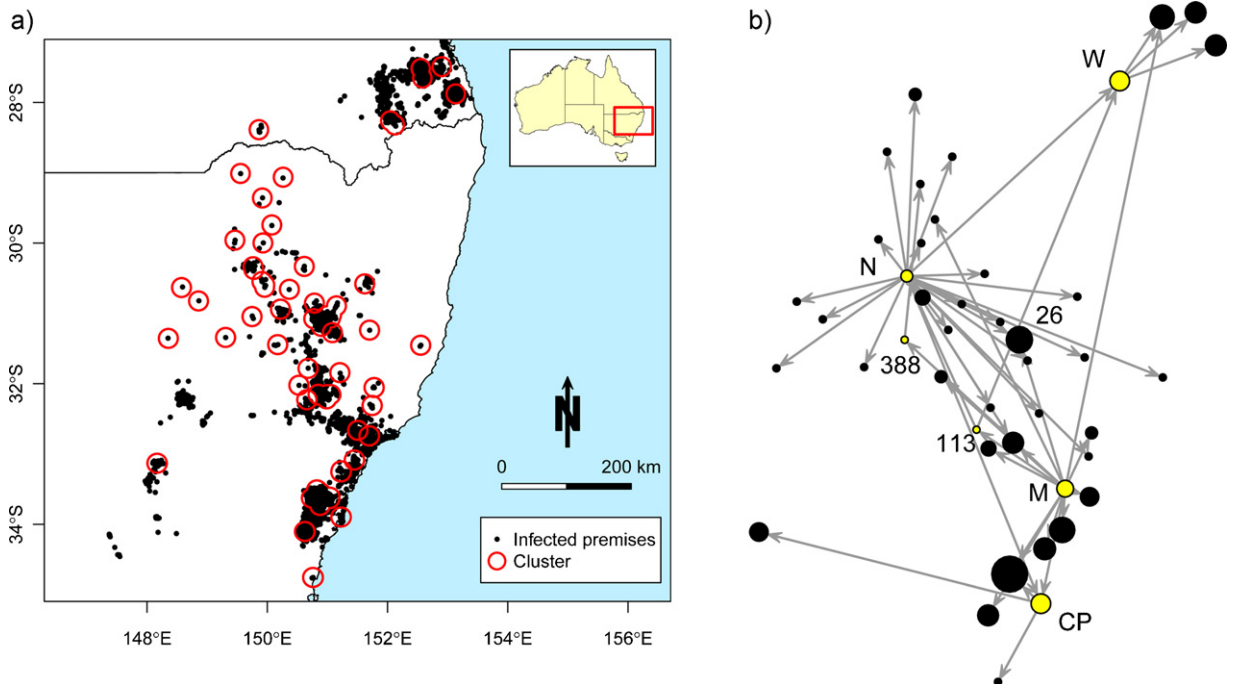
Fig. 6. Clusters of infected premises described by the 'maximal' contact-and-proximity network (distance cut-off dichotomised at 15.3 km) in the first 10 days of the 2007 equine influenza outbreak in Australia. (a) Cluster boundaries (red) coalesced from 15.3 km buffers around premises grouped according 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. (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.)

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 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 a premises that has 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., 2011). 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 (The EI 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 (The EI 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.

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.

Conflicts of interest

None.

Acknowledgements

References

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford.

Bean, B., Moore, B.M., Sterner, B., Peterson, L.R., Gerding, D.N., Balfour, H.H., 1982. Survival of influenza viruses on environmental surfaces. J. Infect. Dis. 146, 47–51.

Callaway, D.S., Newman, M.E.J., Strogatz, S.H., Watts, D.J., 2000. Network robustness and fragility: percolation on random graphs. Phys. Rev. Lett. 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.

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, 586–589.

Cohen, M.L., 2000. Changing patterns of infectious disease. Nature 406, 762–767.

Cohen, R., Erez, K., ben-Avraham, D., Havlin, S., 2000. Resilience of the Internet to random breakdowns. Phys. Rev. Lett. 85, 4626–4628.

Cowled, B., Ward, M.P., Hamilton, S., Garner, G., 2009. The equine influenza epidemic in Australia: spatial and temporal descriptive analyses of a large propagating epidemic. Prev. Vet. Med. 92, 60–70.

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, pp. 37–53.

Davis, J., Garner, M.G., East, I.J., 2009. Analysis of local spread of equine influenza in the Park Ridge Region of Queensland. Transbound. Emerg. Dis. 56, 31–38.

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. J. Appl. Prob. 35, 448–462.

Erdős, P., Rényi, A., 1961. On the strength of connectedness of a random graph. Acta Math. Hung. 12, 261–267.

Firestone, S.M., Schemann, K.A., Toribio, J.-A.L.M.L., 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.

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). Vet. Microbiol. 137, 1–9.

Freeman, L.C., 1979. Centrality in social networks conceptual clarification. Soc. Netw. 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. Epidemiol. Infect., 1–12.

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. Vet. Rec. 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.

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. Proc. R. Soc. B 275, 1001–1005.

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, 179–185.

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.

Hanneman, R.A., Riddle, M., 2005. Introduction to Social Network Methods. University of California, Riverside, CA, Available at: http://faculty.ucr.edu/~hanneman/.

Huntington, P.J., 1990. Equine influenza – the disease and its control. Technical Report Series No. 184. Department of Agriculture and Rural Affairs, Victoria.

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.

Martinez-Lopez, B., Perez, A.M., Sanchez-Vizcaino, J.M., 2009. Combined application of social network and cluster detection analyses for temporal–spatial characterization of animal movements in Salamanca, Spain. Prev. Vet. Med. 91, 29–38.

Matthews, L., Woolhouse, M., 2005. New approaches to quantifying the spread of infection. Nat. Rev. Microbiol. 3, 529–536.

May, R.M., Gupta, S., McLean, A.R., 2001. Infectious disease dynamics: what characterizes a successful invader? Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci. 356, 901–910.

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. Int. J. Tuberc. Lung Dis. 7, S486–S493.

Miller, W.C., 1965. Equine influenza – further observations on "coughing" outbreak 1965. Vet. Rec. 77, 455–456.

Myers, C., Wilson, W.D., 2006. Equine influenza virus. Clin. Tech. Equine Pract. 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. Prev. Vet. Med. 92, 341–350.

Newman, M.E.J., 2003. Properties of Highly Clustered Networks. PhRvE 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, 185–192.

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. Prev. Vet. Med. 76, 40–55.

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.

Ribbens, S., Dewulf, J., Koenen, F., Mintiens, K., De Kruif, A., Maes, D., 2009. Type and frequency of contacts between Belgian pig herds. Prev. Vet. Med. 88, 57–66.

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. Sex. Transm. Dis. 34, 280–287.

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, 1023–1032.

Small, M., Walker, D.M., Tse, C.K., 2007. Scale-free distribution of avian influenza outbreaks. Phys. Rev. Lett., 99.

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.

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.

Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press Cambridge, New York.

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. Prev. Vet. Med. 68, 3–17.

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. Int. J. Anim. Sci. 8, 93–98.