Animal (2012), **6:7**, pp 1152–1162 © The Animal Consortium 2012
doi:10.1017/S1751731111002631

Structural vulnerability of the French swine industry trade network to the spread of infectious diseases

S. Rautureau[1†], B. Dufour[2] and B. Durand[1]

[1]*Epidemiology unit (EPI), French Agency for Food, Environmental and Occupational Health & Safety (ANSES), 97 700 Maisons-Alfort, France;* [2]*Epidemiology Unit (EPIMAI), Alfort National Veterinary School (ENVA), 97 700 Maisons-Alfort, France*

(Received 1 August 2011; Accepted 28 November 2011; First published online 4 January 2012)

The networks generated by live animal movements are the principal vector for the propagation of infectious agents between farms, and their topology strongly affects how fast a disease may spread. The structural characteristics of networks may thus provide indicators of network vulnerability to the spread of infectious disease. This study applied social network analysis methods to describe the French swine trade network. Initial analysis involved calculating several parameters to characterize networks and then identifying high-risk subgroups of holdings for different time scales. Holding-specific centrality measurements ('degree', 'betweenness' and 'ingoing infection chain'), which summarize the place and the role of holdings in the network, were compared according to the production type. In addition, network components and communities, areas where connectedness is particularly high and could influence the speed and the extent of a disease, were identified and analysed. Dealer holdings stood out because of their high centrality values suggesting that these holdings may control the flow of animals in part of the network. Herds with growing units had higher values for degree and betweenness centrality, representing central positions for both spreading and receiving disease, whereas herds with finishing units had higher values for in-degree and ingoing infection chain centrality values and appeared more vulnerable with many contacts through live animal movements and thus at potentially higher risk for introduction of contagious diseases. This reflects the dynamics of the swine trade with downward movements along the production chain. But, the significant heterogeneity of farms with several production units did not reveal any particular type of production for targeting disease surveillance or control. Besides, no giant strong connected component was observed, the network being rather organized according to communities of small or medium size (<20% of network size). Because of this fragmentation, the swine trade network appeared less structurally vulnerable than ruminant trade networks. This fragmentation is explained by the hierarchical structure, which thus limits the structural vulnerability of the global trade network. However, inside communities, the hierarchical structure of the swine production system would favour the spread of an infectious agent (especially if introduced in breeding herds).

Keywords: network analysis, swine, trade, France

Introduction

Many diseases are susceptible to being spread by movements between holdings, and past events such as foot-and-mouth disease in the United Kingdom (2001) or classical swine fever in the Netherlands (2003) have clearly shown that live animal movements represent a major risk for the transboundary spread of disease (Fèvre *et al.*, 2006). European Union legislation requires identification and registration for cattle, pigs, sheep and goats. Since Regulation (EC) no. 1760/2000 of the European parliament, France, like most Member States, has considerably enhanced cattle industry traceability and extended this measure to all animal productions.

The mandatory system encompasses birth to retail with large amounts of data collected. Traceability becomes a tool that may be used to carry out surveillance, prevention or control of animal disease (Ammendrup and Barcos, 2006). For example, national cattle trade data have been widely studied using network analysis methods (social network analysis (SNA)) for the characterization of trade organization and the detection of communities to qualify their vulnerability structure and to target surveillance (Christley *et al.*, 2005; Bigras-Poulin *et al.*, 2006; Natale *et al.*, 2009; Rautureau *et al.*, 2011; Nöremark *et al.*, 2011). Fewer studies

† E-mail: severine.rautureau@anses.fr; sev_rautureau@yahoo.fr

were dedicated to the pig trade network (Bigras-Poulin *et al.*, 2007; Martinez-Lopez *et al.*, 2009a; Lentz *et al.*, 2011; Nöremark *et al.*, 2011).

SNA based on the study of the relationships among social entities (Wasserman and Faust, 1994) has developed indicators of centrality for nodes (units of interest) or for their relationship (called links) and methods for the identification of cohesive structures (groups or individual clusters). This approach has been introduced in veterinary epidemiological studies only recently (reviewed by Dube *et al.* (2009) or Martinez-Lopez *et al.* (2009b)).

Livestock movements have been analysed thus in several studies to evaluate the potential transmission of an infectious disease through the contact network or to assess the impact of livestock movement regulations. The live animal trade can be represented by a network in which the holdings are represented by nodes and the animal movements between these holdings are represented by links (namely arcs as direction is taken account). Animal movements along the network links can be considered as paths for the diffusion of a disease.

Disease spread is therefore dependent on the structure of the network and network topology may reveal the existence of holdings or groups particularly exposed to infection, defined as 'vulnerable' entities (Bell *et al.*, 1999). Within a network, the vulnerability would be a measure of the likelihood for an individual to become infected. By identifying individuals with high vulnerability, SNA can enhance targeted surveillance and prevention programmes.

This node-level definition may be generalized at the network level, considering that a trade network is vulnerable to the spread of infectious diseases if (i) live animal movements through this network may lead the disease to spread very fast inside a country (in particular before the identification of the first case) and (ii) human and material resources would quickly become insufficient to control the epidemic with the usual control measures. In this context, network connectedness determines how fast and widely an infection may spread, and the existence of large connected components (beyond a certain size) is the basis of network vulnerability.

Swine traceability records have recently become as systematic and detailed as cattle data. While cattle movement registration is individual, swine movement records are collective, for example, by epidemiological unit or animal batch. Moreover, swine husbandry differs from cattle husbandry. The pig industry has a pyramidal structure with movements predominantly going downward through the system; from nucleus herds to multiplier herds, from multiplier herds to production herds and then from breeding units to growing units and from growing to finishing units.

One question of interest is whether these structural differences between cattle and swine trade networks can be measured. We computed current network analysis parameters as a means of addressing this question. The network type commonly studied for network analysis is a one-mode network. A one-mode network is composed of a set of similar nodes (Figure 1); contacts between nodes (individuals) are considered equivalent and each node can have links to any other node. However, farm animal movements (especially swine movements) often occur by animal batches, moved from one holding to another one in a given truck. The truck may then play an epidemiological role and support the transmission of an infectious agent. The movements of animal batches performed by a single truck during a single day will be termed 'round'. But, rounds must be considered as a different class of nodes than holdings. Such a network is called an affiliation or a two-mode network in which nodes are partitioned in two mutually exclusive sets such that there are no links between nodes within a given set (Figure 1). Two-mode networks are built with one set of nodes and one set of events such that the adjacency matrix is given by nodes $\times$ events, rather than nodes $\times$ nodes as for one-mode networks (Borgatti and Everett, 1997).

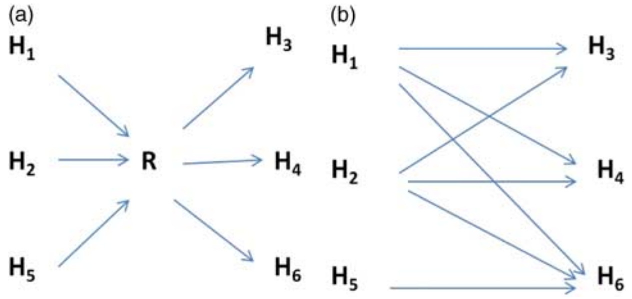


Figure 1 Network types: (a) two-mode network: two sets of nodes (H = holdings and *R* = round) and (b) one-mode network: one set of nodes (H = holdings). Two mode-networks were transformed into one-mode networks by replacing movements involved in a round by direct movements between holdings. The figure indicates the chronology of truck collection.

The objective of the study was to evaluate the vulnerability of the French swine trade networks to infectious diseases using network analysis indicators. The methodology described above was applied to swine movements between French holdings by studying different networks and different time scales: 6-month, monthly and weekly movements. Two types of networks (two-mode and one-mode networks) were built and examined to compare them and ascertain which type provided the best information. Besides, the similarity to cattle trade networks was discussed.

Material and methods

Data source

French swine movement data were obtained from the National Swine Database of Identification (BDporc) recognized by the French Ministry of Agriculture and managed by swine industry professionals. Systematic recording of swine movements started in France in January 2010, and the data set corresponding to the first 6 months of 2010 was used. Movements of pigs, as well as information related to livestock holdings, are registered. The database contains the production type and approximate number of animals kept on each farm based on self-reporting by the farmer. One agricultural enterprise can have several units operating under a unique identification number.

Movements of pigs are reported at group level by animal batches and not at individual level.

Reporting for a specified shipment (incoming or outgoing) includes the date of the event, the number of the round and the holding of origin (or of destination). An animal shipment was defined as a batch of animals gathered or dispatched simultaneously by a given transport vehicle. Holdings with incoming shipments 'receive' pigs, whereas holdings with outgoing shipments 'dispatch' pigs. These operations of loading and unloading (outgoing and incoming shipments) by a given truck in a given day is termed 'round'. Each round is identified by a unique numbering system. Only records of complete rounds with at least one outgoing shipment and one incoming shipment were included in the study.

Three kinds of holdings were distinguished: farms, dealers (trade operators) and slaughterhouses. Pig farms were classified as belonging to one of the following six categories: (1) breeding, (2) farrow-to-grow, (3) farrow-to-finish (4) growing, (5) grow-to-finish and (6) finishing herd.

Using the information on reported movements by type of animal, number of rounds and number of animals moved between the different holding types (farms, dealers and slaughterhouses) were compiled. The Euclidean distance between the origin and destination holdings was calculated. Zero distances indicated swine transferring between two units in the same holding.

Network analysis

Networks representing the swine trade were built that included each holding involved in any movement during the period. Imports and exports were not taken into account.

Two-mode-oriented and one-mode-oriented networks were constructed for three different time scales: the global period (6-month network), the month (6-month networks) and the week (26-week networks). All of the movements between two specific holdings during the time period were then condensed into a single arc. Initial analysis involved calculating several parameters to characterize networks and then identifying components or communities of nodes for different time scales (Table 1). Once subgroups had been identified within networks, they were extracted and described.

Network constructions. Movements to slaughterhouses were not taken into account. Although these movements were numerous, it can be assumed that they represented little or no risk for the transmission of swine diseases to other swine. An affiliation network was built with one set of nodes (holdings) and one set of events (rounds; Borgatti and Everett, 1997).

To build one-mode networks with a set of equal nodes (holdings), network data were altered (Figure 1). Outgoing and incoming shipment movements were ranked by date/h. Movements involved in a round were replaced by direct movements between holdings. We assumed that animals unloaded at a holding could have had contacts with all of the animals loaded previously during the round. Therefore, holdings where animals had been unloaded during a given round were linked to each of the holdings where animals had been loaded before, during the same round.

Network description. The networks were analysed separately and then compared. We first compiled descriptive parameters for all networks and then analysed them to detect significant connected components (Table 1).

First, several descriptive parameters (Wasserman and Faust, 1994) were calculated for all networks: the size (number of nodes and links), the average degree (average number of nodes directly linked to a node), the average path length (the average number of links along the shortest paths – or geodesics – between all pairs of nodes) and the diameter (the longest geodesic).

For each global network, a power-law distribution using as $P(k) \sim k^{-\gamma}$ was fitted from distribution of node degree (denoted k). Estimating the parameters of a power-law distribution is complex as, most of the time, only the tail of the observed distribution follows a power law. As large fluctuations occur in the tail of power-law distributions, it is difficult to identify the range over which power-law behaviour holds. We used the approach proposed by Clauset *et al*. (2009) that combines a maximum-likelihood estimator of the scaling exponent (γ) with the Kolmogorov–Smirnov statistic for determining the threshold above which the power-law behaviour holds.

Some measurements have different meanings depending on network type and some have no meaning in two-mode networks (Table 1). Thus, the following indicators were calculated only for one-mode networks: the clustering coefficient (proportion of neighbours of a node that are linked to each other) and the assortativity (correlation between the degrees of linked nodes).

For the 6-month one-mode network, the distributions of the three main centrality measurements were computed for each holding type: degree, in-degree (number of different holdings from which a holding receives animals) and betweenness (the frequency at to which a node is on the shortest path between any pair of nodes; Freeman, 1978/1979). Moreover, a new measurement proposed by Nöremark *et al*. (2011) was calculated, the ingoing infection chain, which identifies the number of holdings that are connected to one holding taking into account the order of the movements (Nöremark *et al*., 2011). The order in which contacts occur is important, as animals leaving a holding before introduction of infection do not constitute a risk for the spread of disease.

Finally, weakly and strongly connected components were studied in the one-mode and two-mode networks. Weakly connected components (WC) are subnetworks for which a path exists between any pair of nodes, whatever the link direction. These WCs can be interpreted as independent

Table 1 *Network analysis glossary of terms interpreted in the context of livestock movement*

Name	Definition	Reference
Density	Proportion of contacts that could possibly occur in the network compared with those that are actually observed in the network	For one-mode networks (Wasserman and Faust, 1994) For two-mode networks (Borgatti and Everett, 1997)
Degree (in-degree)	In one-mode networks, the degree of a holding is equal to the number of its direct neighbours. The in-degree is defined as the number of holdings from which animals move onto a particular holding	For one-mode networks (Freeman, 1978/1979)
	In two-mode networks, the degree of a holding is the number of shipment rounds in which it takes part (rate of participation) and the degree of a shipment round reflects the number of holdings where animals were loaded	For two-mode networks (Borgatti and Everett, 1997)
Strong and weak components	Strong components are sections of the network where every holding can be reached from every other holding via directed paths, whereas weak components are sections of the network that are linked, but not every farm can be reached from every other farm.	(Christley *et al.*, 2005; Robinson and Christley, 2007)
Clustering coefficient	The amount of interrelationship that exist between all the nodes and one specific node indicated complicated partnership trade patterns because it implied that when two business partners sent animals to a third location they were also linked by an animal movement	(Watts and Strogatz, 1998)
Shortest path length diameter	The shortest path length or geodesic distance is the smallest number of links represented by animal movement sets required to travel from holding A to holding B	(Watts and Strogatz, 1998)
	The diameter is the longest geodesic distance Distance measurements of two-mode networks with links between rounds and holdings, must be divided by two to be compared to one-mode measurements Low values suggest that epidemic would spread quickly across the network	(Borgatti and Everett, 1997)
Assortativity	Preference of nodes to attach to others that are similar degree Negative assortativity is dissortativity; high-degree nodes tend to attach to low-degree nodes	(Newman, 2002)
Community	Subset of nodes in which there are significantly more links than expected by chance; group of preferentially linked holdings	(Newman, 2006)
Betweenness	Proportion of geodesic distances (represented by animal movement sets) between all pairs of holdings (excluding node A) that pass through node A	(Freeman, 1978/1979)
Ingoing infection chain	The ingoing infection chain measures all direct and indirect contacts through movements onto a holding. The sequence by which the movements occur is taken into account. It identifies holdings with many contacts through live animal movements and thus at potentially higher risk for introduction of contagious diseases	(Nöremark *et al.*, 2011)

subnetworks, as, in a given network, no connection exists between two WCs. Strongly connected components (SCs) are subnetworks for which every node can be reached from every other node via one (or several) directed path(s). They can be interpreted as areas of a network where connectedness is particularly high. A disease introduced into any holding of an SC can potentially reach any other holding in that SC (Rautureau *et al.*, 2011). Calculation of giant strong connected component (GSC) size in animal movement networks may then be used to estimate potential epidemic sizes (Robinson and Christley, 2007; Volkova *et al.*, 2010a and 2010b). The number of WCs and SCs and the size of the two largest WCs and SCs were determined. Both sizes were compared to qualify the largest WCs (or the largest SCs) as 'giant' weak (or strong) components (GWCs or GSCs).

Network communities. Many networks can be divided into communities. A community (or module) is defined as a subset of nodes in which there are significantly more edges than expected by chance. Community detection is characterized by the modularity function Q introduced by Newman (Newman, 2006), which scores the quality of a partition. The objective is to maximize the function Q, defined as $Q =$ (number of edges within communities) – (expected number of those edges).

We used the 'greedy algorithms' method proposed by Newman (Clauset *et al.*, 2004; Newman, 2004), which can only be applied on one-mode networks. We, therefore, identified communities in the global one-mode network with and without movements to a slaughterhouse. The spatial extent of the largest communities and the composition of these communities by type of holdings (proportion of farm productions and dealers) were determined.

Network analyses were performed using the Igraph package (v 0.5.2, http://igraph.sourceforge.net) for *R* software (Anonymous, 2009).

Results

Swine trade data

Available data comprised all French swine trade information with ~23 084 pig holdings and their links in the pork production chain (breeders to slaughterhouses). Swine-holding density was clearly concentrated in the west of France: in Brittany, the primary region of French pork production; and in the southwest, near the Pyrenees (Figure 2).

During the 6-month study period, 155 154 shipment round connections between 13 968 holdings (13 809 farms, 19 dealer holdings and 140 slaughterhouses) were recorded. Rounds could involve animals of several production types, as shown in Table 2, where movements are presented in detail. The network was shown to have many intense flows as seen from the ratio of number of pigs moved over number of movements (78.1 animals). 78.3% of incoming pig shipments (75.4% of animals) were movements to slaughterhouses. In all, 84% of rounds made by a truck consisted solely or partly of these movements. Differences of animal numbers between outgoing and incoming shipments for the transfers by dealer holdings correspond to export movements. Thus, dealer holdings were the primary gateway for export animals in comparison with animals which would leave farms directly.

Average distance of an animal movement was 82.5 km (median distance 55 km and range 0.1 to 793 km); to slaughterhouses it was 86.4 km (median distance 58.4 km and range 0.1 to 725 km), which was significantly higher than to holdings, 75.2 km (median distance 51 km and range 0.1 to 793 km; Wilcoxon test: $P < 0.0001$).

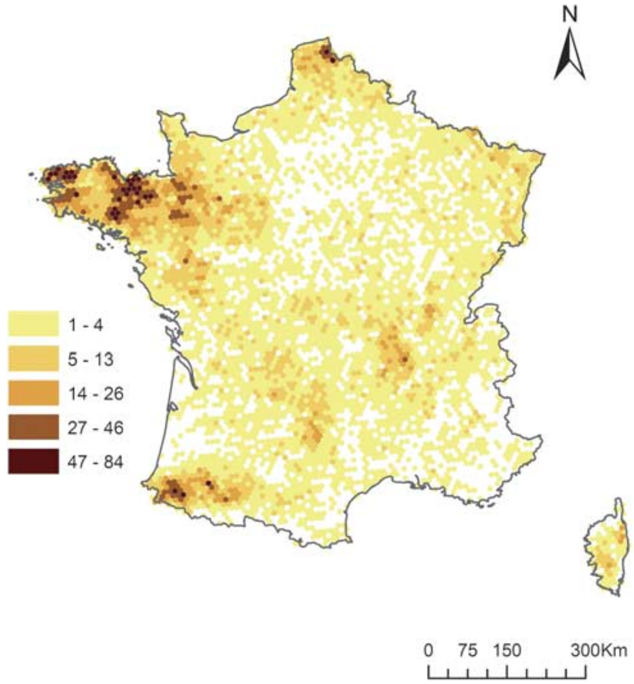


Figure 2 Spatial density of swine farms, France, 2010. Density: number of holdings/10-km diameter hexagon.

Network characteristics

Comparison between two-mode and one-mode networks. The two-mode global network contained 10 181 holdings linked by 24 835 rounds with 71 197 arcs. On average, 60% of holdings participated in the monthly (with 16.5% of rounds and arcs) and 21% in the weekly networks (with 3.8% of rounds and arcs; Table 3). A holding took part in approximately seven rounds (average degree of holdings) during the 6-month

Table 2 *Number of animal movements and of animals moving between French holdings from January to June 2010 according to outgoing and incoming shipments*

	Outgoing shipments			Incoming shipments		
	Rounds	Movements	Animals	Rounds	Movements	Animals
Farms	21 381	26 023	3 720 098	17 470	19 549	3 669 858
Piglets						
Fat pigs	116 104	157 810	11 845 976	225	236	9114
Breeding pigs	5065	7117	144 245	4985	15 587	144 139
Culled pigs	20 391	39 682	207 574	299	478	5245
Dealer Holdings	59	66	6956	164	177	14 842
Piglets						
Fat pigs	109	110	7974	479	482	28 020
Breeding pigs	9	9	35	66	66	401
Culled pigs	13	13	511	1547	1558	27 644
Slaughterhouses	–	–	–	3889	3897	47 068
Piglets						
Fat pigs	–	–	–	115 504	115 527	11 756 771
Culled pigs	–	–	–	18 534	18 534	173 441
Total (distinct)	155 154	230 830	15 933 369	155 154	176 091	15 876 543

Piglet = just weaned pig; fat pig = pig from fattening structure; breeding pig = sow and boar; culled pig = retired breeding pig.

Table 3 *Descriptive parameters for two-mode 6-month, monthly and weekly swine trade networks in France, 2010*

	6-month network	Monthly networks mean data (range)	Weekly networks mean data (range)
Size			
Holdings	10 181	6002 (5713 to 6390)	2153 (1824 to 2408)
Rounds	24 835	4103 (3821 to 4503)	955 (784 to 1052)
Links	71 197	11 755 (10 753 to 13 090)	2737 (2253 to 3068)
Average degree			
Holdings	6.99	1.96 (1.87 to 2.05)	1.27 (1.24 to 1.30)
Rounds	2.87	2.86 (2.81 to 2.91)	2.87 (2.74 to 2.96)
Density	1.408×10^{-4}	2.388×10^{-4} (2.275×10^{-4} to 2.463×10^{-4})	6.680×10^{-4} (6.056×10^{-4} to 7.888×10^{-4})
Av. path length	5.49	3.08 (2.89 to 3.47)	1.89 (1.71 to 2.13)
Diameter	26	13 (12 to 15)	8 (6 to 11)
Weak components			
Number	217	533 (478 to 598)	495 (433 to 542)
Largest size			
Holdings (%)	9523 (93.5%)	4479 (74.6%) (4076 (71.3%) to 5046 (79%))	314 (14.6%) (105 (4.2%) to 646 (27.5%))
Rounds (%)	23 498 (94.6%)	2932 (71.5%) (2566 (66.6%) to 3421 (76%))	120 (12.6%) (28 (2.9%) to 251 (24.2%))
2nd largest size			
Holdings	10	56 (26 to 100)	106 (46 to 197)
Rounds	81	38 (13 to 72)	36 (11 to 72)
Strong components			
Number	388	173 (164 to 184)	57 (44 to 72)
Largest size			
Holdings (%)	35 (0.34%)	5 (0.08%) (2 (0.03%) to 8 (0.14%))	2 (0.09%) (1 (0.05%) to 4 (0.19%))
Rounds (%)	211 (0.85%)	14 (0.34%) (7 (0.18%) to 21 (0.55%))	4 (0.42%) (2 (0.20%) to 6 (0.66%))
2nd largest Size			
Holdings	4	2 (2 to 2)	2 (1 to 3)
Rounds	54	10 (5 to 13)	3 (2 to 5)

period study and between one and two rounds for monthly and weekly networks. Whatever the time scale, a round concerned on average 2.8 holdings (average degree of rounds).

In the same way, the 10 181 holdings of the global one-mode network were linked by 22 231 arcs, whereas 60% of holdings participated in the monthly networks with 33% of the arcs and 21% also participated in the weekly networks with 8.8% of the arcs (Table 4). A holding had on average 4.4 different contact holdings (average degree) with which it exchanged animals during the 6-month period study and on average 2.47 and 1.82 contacts in monthly and weekly networks, respectively.

The degree distributions of holdings for global one-mode and two-mode networks had similar curves and the tail distribution seemed linear on a log–log scale with respectively a power-law exponent γ of 2.6 and 3.1 for the nodes with the degree above 7 and 8, respectively (Figure 3).

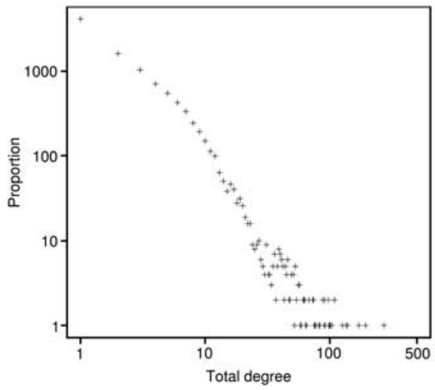
Irrespective of the network type, distance indicator measurements decreased with the time scale; a given pair of connected holdings was separated by approximately two animal movements (average path length), that is, there was one intermediary holding. In one-mode networks, there were from 2.55 arcs (in 6-month network) to 1.2 arcs (on average in weekly networks) between two holdings (Table 4) and in two-mode networks from 2.75 (i.e. 5.5/2) to 0.86 (i.e. 1.89/2) arcs (Table 3). The diameter (the longest geodesic between holdings) also decreased from 10 arcs (global network) to an average of 4 arcs (weekly networks) in one-mode networks (Table 4). In two-mode networks, the diameter varied from 13 (i.e. 26/2) arcs to an average of 4 (i.e. 8/2) arcs (Table 3).

Global and monthly networks were less fragmented, the largest WC including >70% of the nodes, whereas the weekly network fragmentation was higher, with the largest WC including only 14.6% of the holdings. On the other hand, no large strong components were observed, the size of strong components being in any case very low (~30 holdings for global networks and fewer than five holdings on average for monthly and weekly networks).

Specific one-mode network characteristics (Table 4). The global clustering coefficients in monthly and weekly one-mode networks were low and of the same order of magnitude as those for a random network (Newman, 2005). A higher value was obtained for the 6-month network. In this latter case, the local clustering coefficient appeared inversely linked

Table 4 *Descriptive parameters for one-mode 6-month, monthly and weekly swine trade networks in France, 2010*

	6-month network	Monthly networks mean data (range)	Weekly networks mean data (range)
Size	10 181	6002 (5712 to 6391)	2152 (1824 to 2408)
Links	22 231	7432 (6902 to 8143)	1957 (1563 to 2208)
Average degree	4.37	2.47 (2.41 to 2.57)	1.82 (1.71 to 1.93)
Density	2.145×10^{-4}	2.063×10^{-4} (1.994×10^{-4} to 2.116×10^{-4})	4.236×10^{-4} (3.809×10^{-4} to 4.630×10^{-4})
Average path length	2.55	1.72 (1.65 to 1.86)	1.20 (1.11 to 1.32)
Diameter	10	6 (6 to 7)	4 (3 to 5)
Clustering coefficient	0.01685	0.00910 (0.00835 to 0.0108)	0.00936 (0.00372 to 0.02368)
Assortativity	−0.15424	−0.14416 (−0.12806 to −0.16327)	−0.15520 (−0.12072 to -0.18529)
Weak components			
Number	217	532 (478 to 598)	495 (433 to 557)
Largest size (%)	9523 (93.5%)	4487 (74.8%) (4075 (71.3%) to 5046 (79%))	314 (14.6%) (105 (4.9%) to 646 (27.5%))
Second largest size	27	56 (33 to 100)	106 (46 to 197)
Strong components			
Number	80	29 (23 to 34)	8 (4 to 13)
Largest size (%)	30 (0.29%)	5 (0.08%) (3 (0.05%) to 7 (0.14%))	2 (0.09%) (2 (0.08%) to 4 (0.19%))
Second largest size	5	3 (2 to 3)	2 (2 to 3)



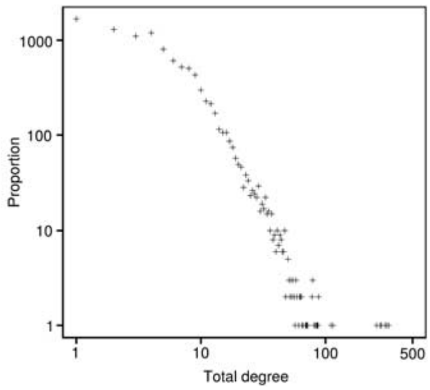


Figure 3 Distribution of holding degrees in the 6-month one-mode (left) and two-mode (right) networks of swine movements, France 2010 (log scale).

to the node degree (Figure 4). The assortativity of networks was negative indicating that networks were disassortative: nodes were more often linked to nodes with different degrees than to nodes with similar degrees.

Centrality values (degree, betweenness, in-degree and ingoing infection chain) appeared significantly higher for dealer holdings than for farms (Figure 5; Wilcoxon test: $P = 0.018$ for betweenness centrality and $P < 0.0001$ for the other centrality values). There were also significant differences between types of swine unit (Kruskal–Wallis test: $P < 0.0001$); considerable heterogeneity was observed within types of holdings (Figure 5). For example, degree centrality for farrow-to-finish herds ranged from 1 to 126 (median: 3) and ingoing infection chain centrality for finishing herds ranged from 1 to 573 (median: 4). However, herds with growing units had higher values for degree and betweenness centrality, representing central positions for both spreading and receiving disease, whereas herds with finishing units had higher values for in-degree and ingoing infection chain centrality values, indicators of vulnerability.

Modularity analysis

Networks including movements to slaughterhouses. Forty-five communities were identified ($Q_{max} = 0.58$). The number of holdings forming a community varied between 1 and 2629. Seven large communities contained about or more than 1000 holdings representing 98% of all holdings (Table 5). They were mostly spatially clustered. A number overlapped and shared their whole area as in Brittany (four out of seven), but communities could be located over several regions (Figure 6).

Network excluding movements to slaughterhouses. Three hundred and twenty communities were identified ($Q_{max} =$ 0.76). The number of holdings forming a community varied between 1 and 1857. Communities are described in detail in Table 5. Within these 320 communities, 57% of the holdings were part of the four largest communities (>1000 holdings). Although finishing herds and grow-to-finish herds mostly belonged to communities (on average 30% and 51% of holdings, respectively), holdings distribution by type was significantly different among communities (Pearson's χ^2

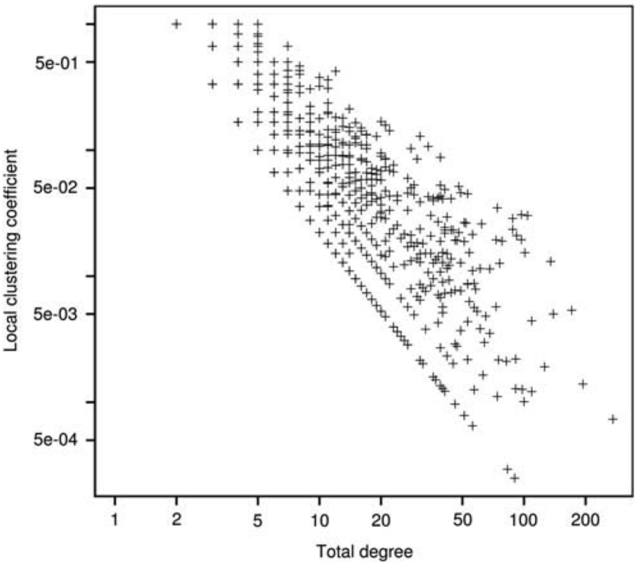


Figure 4 Distribution of local clustering coefficients according to node degree in the 6-month one-mode network of swine movements, France 2010 (log scale).

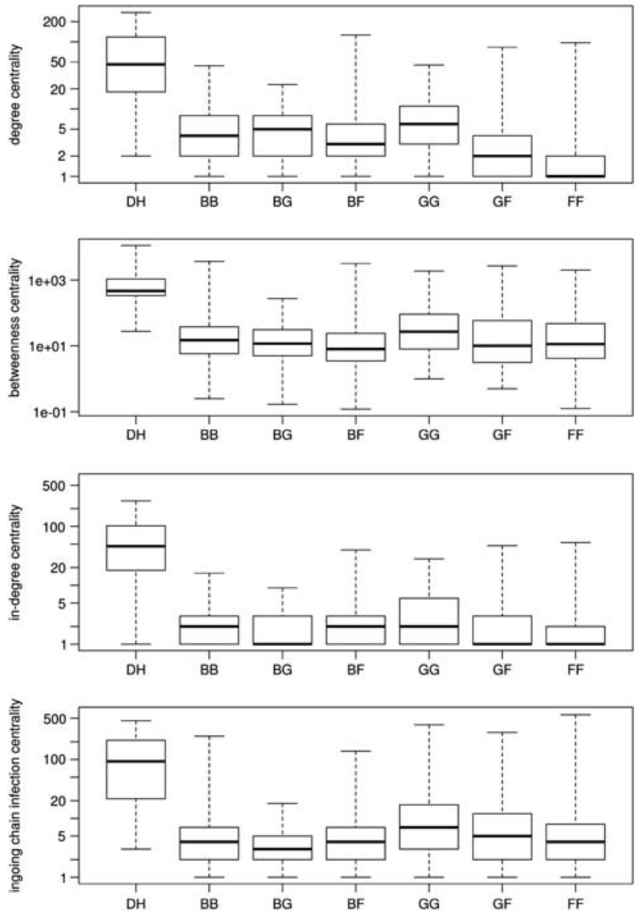


Figure 5 Distribution of degree, betweenness, in-degree and ingoing infection chain centrality in holdings involved in the 6-month one-mode network of swine movements, France, 2010. (DH = dealer holdings; BB = breeding; BG = farrow-to-grow; BF = farrow-to-finish; GG = growing; GF = grow-to-finish; FF = finishing herd).

test: $P < 0.0001$). However, three communities were clearly spatially clustered in Brittany, the fourth being in the north of France (Figure 6).

Discussion

The swine industry forms a complex network spread heterogeneously across the country, with holdings being dynamically linked by animal movements. This network was studied at three different time scales to explore structural vulnerability to the spread of any disease.

One- and two-mode networks

The French swine trade was described first by links between two sets of nodes at different levels of analysis: actors (holdings) and events (rounds). Data such as these involve two levels of analysis (i.e. two modes). Two-mode data offer some very interesting analytic possibilities for gaining greater understanding. Thus, holdings involved in swine trade were concerned by approximately seven rounds during the 6-month period study. And a round connected each time on average from two to three holdings. Although rich in their data content, the two-mode networks are, however, difficult to interpret and little work has been done concerning animal trade movements with this approach (Robinson and Christley, 2007). A common strategy in such cases is to examine instead a one-mode 'projection' of the network onto either the holdings or the gathering rounds. For this purpose, we altered the two-mode network to a one-mode network keeping the same sequence of outgoing and incoming shipments. The results obtained were close to those from two-mode networks. We could thus continue the network analysis with a wider range of network methods and indicators.

Moreover, the data set used for this study reflects trade activity for only 6 months, as registration has only recently been made mandatory. The observation period and the lack of precision or error-rate due to self-reporting could introduce some bias in the analysis. The 6-month period was, however, compatible with a complete cycle of swine production.

Scale-free structure

The node degree distributions were evaluated and appeared heavy-tailed suggesting a scale-free structure. Moreover, the correlation between local clustering coefficients and node degrees confirmed as well as the disassortativity (assortativity <0), the scale-free structure and the presence of hubs. Theoretical studies have shown that the scale-free structure of networks can influence the speed and the extent of a disease that would spread via network links (Pastor-Satorras and Vespignani, 2001; Kiss *et al.*, 2006a). Heterogeneity in the distribution of links within a network is a key factor and reveals the presence of central individuals (hubs) that are the most likely to spread a disease and that could be targeted by control measures.

Table 5 *Number of holdings and trade connections within trade communities in 6-month networks, one including slaughterhouses and one excluding slaughterhouses (France, 2010)*

Communities ranked by decreasing size	Number of holdings				Number of links
	Slaughterhouse	Dealer	Farm	Total	
6-month network including slaughterhouses					
1	36	0	2593	2629	6183
2	10	0	2389	2399	6956
3	38	15	2197	2250	5352
4	8	1	2155	2164	7189
5	6	1	1946	1953	5729
6	19	2	1312	1333	3421
7	5	0	930	935	2441
Others (nodes<80)				295	13 925
Total = 45				13 958	51 200
6-month network excluding slaughterhouses					
1	–	12	1845	1857	3890
2	–	1	1699	1700	3720
3	–	0	1148	1148	2828
4	–	0	1100	1100	2650
Others (nodes<480)				4376	8735
Total = 320				10 181	21 823

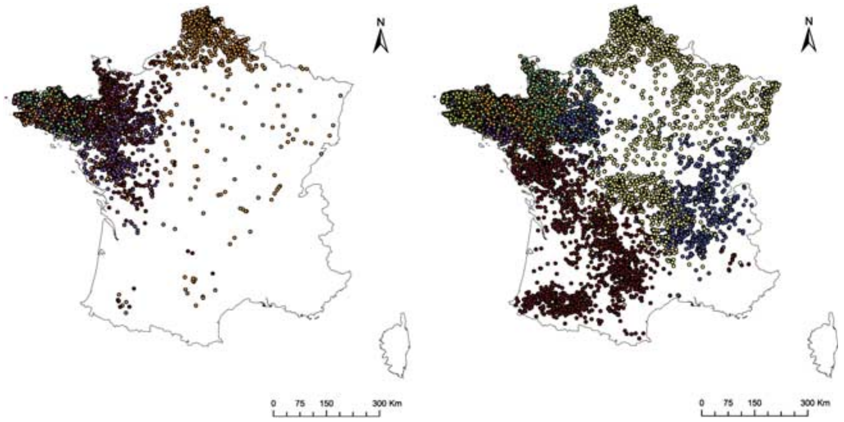


Figure 6 Spatial distribution of the largest communities (>1000 holdings) of the 6-month swine trade network including slaughterhouses, France, 2010; seven large communities in networks including slaughterhouses (right) and four large communities in networks excluding slaughterhouses (left).

The French swine trade network had the same topology as animal trade networks described elsewhere in Europe (Christley *et al.*, 2005; Bigras-Poulin *et al.*, 2006; Kiss *et al.*, 2006b; Bigras-Poulin *et al.*, 2007; Natale *et al.*, 2009; Volkova *et al.*, 2010a and 2010b).

Vulnerability to disease spread

No vulnerable structures emerged as GSCs. However, GWCs were identified and the swine network was well connected. Unlike SCs, WCs are subnetworks for which a link exists between any pair of nodes, whatever the link direction. These WCs are considered as independent subnetworks and show that the swine network was well connected but in a rather hierarchical fashion. We, therefore, focussed our attention on the community structure of the networks and associated values such as modularity. Community structures in networks are natural divisions of network nodes into densely connected subgroups. Identification of communities in the swine trade network showed that holdings were preferentially linked. In the analysis of network data, many methods have been proposed for finding communities, but few have been proposed for determining whether the distribution breakdown found is statistically significant or a result of chance. With the method used here, the significance of community structure can be effectively quantified by measuring its robustness to small perturbations in network structure (Karrer *et al.*, 2008), using the modularity *Q*. In principle, the maximum value of the modularity is $Q_{max} = 1$ and such a network could be considered as highly modular. Nevertheless, it is lower for most real networks (Good *et al.*, 2010) and modularity can be considered to be high from 0.5 (Clauset *et al.*, 2004). Here, the modularity tended to

0.6 for networks including slaughterhouses ($Q = 0.75$ for networks excluding slaughterhouses). Most large communities (>1000 holdings) overlapped and shared their whole area namely Brittany (four out of seven). For the German trade network over a period of 2.5 years, modularity was found to be 0.717 and the nine largest communities (>660 holdings) were also spatially clustered but covered the entire country (Lentz *et al.*, 2011). The groups of holdings represented by communities suggest groups in an integrated system or, for networks including slaughterhouses, the supply-pool of slaughterhouses. With another approach using a combination of SNA and cluster analysis, the holding pool representing pork industry activities has also been identified as made up of temporal–spatial clusters of holdings at risk for the introduction of disease by animal shipments in Spain (Martinez-Lopez *et al.*, 2009a).

Farms with growing units (unit in the middle of the production chain) appeared more central (betweenness value) and herds with finishing units appeared more vulnerable (in-degree and ingoing infection chain values). Holdings were very close with only on average one intermediary holding between a given pair of connected holdings (deducted from the average path length). The low values suggest first that epidemic would spread quickly across the network. Second, this reflects the dynamic nature of the swine industry with downward movements through the pyramidal structure where all holdings could play a role in disease spread along the production process. The scale-free structure and holdings with high centrality values allow us to assume that these holdings may control flows in part of the network, maintain connectivity and promote dynamic dissemination. Here, dealers stood out as having high centrality values. However, centrality values for given types of farms were close to one another and considerable heterogeneity was observed. Thus, some holdings should certainly be targeted, as, for example, sow pool herds where sows are inseminated in a central unit and farrow in satellite herds or, inversely, finishing herds, which collect piglets from numerous breeding herds. But, because of the trend of French herds to integrate several production units, no production types could be suggested for targeting.

Comparison with the French ruminant trade network

Some characteristics of the network were remarkably different compared with cattle or sheep movement networks. In the French cattle, trade network was more disassortative (-0.18 *v.* -0.14 in monthly networks) and centrality measurements of dealer and market appeared markedly higher than those in the French swine trade network (Rautureau *et al.*, 2011). Here, dealer holdings had higher centrality values than the other holdings but represented 0.19% of the population compared with 0.54% for the cattle population (Rautureau *et al.*, 2011). Moreover, dealer holdings in the swine trade seem to be dedicated to animal export. Thus, dealers definitely play a role as hubs at the level of international trade but they do not take part in national movements in the same way as described in cattle and sheep networks (Robinson and Christley, 2007). Besides, no GSC was found out in swine networks. In the French cattle trade network, GSC emerged whatever the time scale, a large subset of holdings existed and were continually linked and widely spread all over the country. GSCs included, 45% of network nodes in the yearly network (i.e. 108 904 holdings), and on average, 8% in the monthly networks (10 277 holdings) and 4% weekly networks (2115 holdings; Rautureau *et al.*, 2011). GSCs are considered as areas where connectedness is particularly high. Thus, a disease introduced into any holding in the subnetwork can potentially reach any other holding in the same subnetwork. Large detected communities in the swine trade network can also be interpreted as structural risk but they showed only that holdings were preferentially linked (holdings are not all linked).

For these reasons, the swine trade network thus appeared less structurally vulnerable than ruminant trade networks.

Comparison with other countries

In fact, the swine trade is a specific trade network and should be compared above all with other swine movement networks. The characteristics of the French monthly swine trade network are similar to those of the Swedish monthly swine trade network (Nöremark *et al.*, 2011; density, clustering coefficient and assortativity). The Danish swine industry network was described as being strongly heterogeneous within and between farms (Bigras-Poulin *et al.*, 2007). Nöremark *et al.* (2011) also found associations between in-degree or ingoing infection chain centrality and production type by classifying the herds according to production type 'highest in the pyramid'.

Our study applied an alternative method defining and using general SNA indicators (i.e. model and disease independent) to characterize animal trade network vulnerability to disease spread and tested this approach in two different network types. To a certain extent, the observed communities can be considered as risk structures. However, because of this fragmentation, the swine trade network appeared less structurally vulnerable than ruminant trade networks. This fragmentation is explained by the hierarchical structure, which thus limits the structural vulnerability of the global trade network. However, within communities, the hierarchical structure of the swine production system would favour the spread of an infectious agent (especially if introduced in breeding unit herds).

Acknowledgements

The authors thank members of BDporc for making holdings and animal movement data available. They are also grateful to Maria Nöremark (National Veterinary Institute, Sweden) and Nina Håkansson (University of Skövde, Sweden) for providing computing support for calculation of the ingoing infection chain centrality.

References

Ammendrup S and Barcos LO 2006. The implementation of traceability systems. Revue Scientifique et Technique 25, 763–773.

Anonymous 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Bell DC, Atkinson JS and Carlson JW 1999. Centrality measures for disease transmission networks. Social Networks 21, 1–21.

Bigras-Poulin M, Barfod K, Mortensen S and 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 RA, Chriel M, Mortensen S and 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.

Borgatti SP and Everett MG 1997. Network analysis of 2-mode data. Social Networks 19, 243–269.

Christley RM, Robinson SE, Lysons R and French N 2005. Network analysis of cattle movement in Great Britain. In Proceedings of Society for Veterinary Epidemiology and Preventive Medecine, Nairn, Scotland, pp. 234–243.

Clauset A, Newman ME and Moore C 2004. Finding community structure in very large networks. Physical Review E 70, 066111.

Clauset A, Shalizi CR and Newman MEJ 2009. Power-law distributions in empirical data. SIAM Review 51, 661–703.

Dube C, Ribble C, Kelton D and 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.

Fèvre EM, Bronsvoort BMdC, Hamilton KA and Cleaveland S 2006. Animal movements and the spread of infectious diseases. Trends in Microbiology 14, 125–131.

Freeman LC 1978/1979. Centrality in social networks, conceptual clarification. Social Networks 1, 215–239.

Good BH, de Montjoye Y-A and Clauset A 2010. Performance of modularity maximization in practical contexts. Physical Review E 81, 046106.

Karrer B, Levina E and Newman ME 2008. Robustness of community structure in networks. Physical Review E 77, 046119.

Kiss IZ, Green DM and Kao RR 2006a. Infectious disease control using contact tracing in random and scale-free networks. Journal of the Royal Society, Interface 3, 55–62.

Kiss IZ, Green DM and Kao RR 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.

Lentz HH, Konschake M, Teske K, Kasper M, Rother B, Carmanns R, Petersen B, Conraths FJ and Selhorst T 2011. Trade communities and their spatial patterns in the German pork production network. Preventive Veterinary Medicine 98, 176–181.

Martinez-Lopez B, Perez AM and Sanchez-Vizcaino JM 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 AM and Sanchez-Vizcaino JM 2009b. Social network analysis. Review of general concepts and use in preventive veterinary medicine. Transboundary and Emerging Diseases 56, 109–120.

Natale F, Giovannini A, Savini L, Palma D, Possenti L, Fiore G and Calistri P 2009. Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. Preventive Veterinary Medicine 92, 341–350.

Newman ME 2002. Assortative mixing in networks. Physical Review Letters 89, 208701.

Newman ME 2004. Fast algorithm for detecting community structure in networks. Physical Review E 69, 066133.

Newman ME 2005. Random graphs as models of networks. In Handbook of graphs and networks (ed. S Bornholdt and HG Schuster), pp. 35–68. Wiley-VCH Verlag GmbH & Co. KGaA, Germany.

Newman ME 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the USA, 103(23), pp. 8577–8582.

Nöremark M, Hakansson N, Lewerin SS, Lindberg A and 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.

Pastor-Satorras R and Vespignani A 2001. Epidemic spreading in scale-free networks. Physical Review Letters 86, 3200–3203.

Rautureau S, Dufour B and Durand B 2011. Vulnerability of animal trade networks to the spread of infectious diseases: a methodological approach applied to evaluation and emergency control strategies in cattle, France, 2005. *Transboundary and Emerging Diseases* 58, 110–120.

Robinson SE and Christley RM 2007. Exploring the role of auction markets in cattle movements within Great Britain. Preventive Veterinary Medicine 81, 21–37.

Volkova VV, Howey R, Savill NJ and Woolhouse ME 2010a. Sheep movement networks and the transmission of infectious diseases. PLoS One 5, e11185.

Volkova VV, Howey R, Savill NJ and Woolhouse ME 2010b. Potential for transmission of infections in networks of cattle farms. Epidemics 2, 116–122.

Wasserman S and Faust K 1994. Social network analysis: methods and applications. Cambridge University Press, Cambridge, UK.

Watts DJ and Strogatz SH 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.