JOURNAL OF THE ROYAL SOCIETY Interface

Recent network evolution increases the potential for large epidemics in the British cattle population

S.E Robinson, M.G Everett and R.M Christley

References	**This article cites 14 articles, 1 of which can be accessed free** http://rsif.royalsocietypublishing.org/content/4/15/669.full.html#ref-list-1 Article cited in: **http://rsif.royalsocietypublishing.org/content/4/15/669.full.html#related-urls**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *J. R. Soc. Interface* go to: **http://rsif.royalsocietypublishing.org/subscriptions**

Recent network evolution increases the potential for large epidemics in the British cattle population

S. E. Robinson[1],[*], M. G. Everett[2] and R. M. Christley[1]

[1]*Epidemiology Group, Department of Veterinary Clinical Sciences, University of Liverpool, Leahurst, Neston, Wirral CH64 7TE, UK*
[2]*Marylebone Provost, University of Westminster, 35 Marylebone Road, London NW1 5LS, UK*

Following the foot and mouth disease epidemic in Great Britain (GB) in 2001, livestock movement bans were replaced with mandatory periods of standstill for livestock moving between premises. It was anticipated that these movement restrictions would limit each individual's contact networks, the extent of livestock movements and thus the spread of future disease outbreaks. However, the effect of behaviour changes on the global network in adapting to these restrictions is currently unknown. Here, we take a novel approach using GB cattle movement data to construct week-by-week contact networks between animal holdings (AH) to explore the evolution of the network since this policy was introduced, the first time network theory has been used for this purpose. We show that the number of AH moving cattle as part of the giant strong component (GSC), representing the region of maximal connectivity, has been increasing linearly over time. This is of epidemiological significance as the size of the GSC indicates the number of holdings potentially exposed to disease, thus giving a lower bound of maximum epidemic size. Therefore, despite restriction of cattle movements, emergent behaviour in this self-organizing system has potentially increased the size of infectious disease epidemics within the cattle industry.

Keywords: network evolution; cattle industry; disease; epidemics; component

1. INTRODUCTION

More than half of all new or emerging infectious disease agents in humans are zoonotic in origin (Taylor *et al.* 2001). Farmed animal populations present the opportunity for targeted disease surveillance, paramount for the protection of human and animal health. Emerging zoonoses in Great Britain (GB) in recent decades prompted the demand for the creation of a central database, which would contain comprehensive details of the origin, identity and location of cattle. To address the demand, at European Union (EU) level, Council Regulation 820/97 (later replaced by Regulation 1760/2000) established a common EU framework of rules which requires all member states to hold computerized databases containing full information on animal identity and location.

Through the launch of the Rapid Analysis and Detection of Animal Risk (RADAR) information management system by the Department for the Environment, Food and Affairs (DEFRA), it is now possible for researchers to include the GB cattle movement data in disease modelling. Similar monitoring is, or soon will be, undertaken in other EU countries.

Although data handling, quality issues and biases have been reported (Bourn 2003; Mitchell *et al.* 2005; Robinson & Christley 2006), this unique database includes several full years of data detailing the movements of the entire GB cattle population (approx. 8 million) between animal holdings (AH) including farms and auction markets (MA). These data have previously been used in models of disease spread for bovine tuberculosis (Gilbert *et al.* 2005) and foot and mouth disease (FMD; Shirley & Rushton 2005; Ortiz-Pelaez *et al.* 2006).

Livestock farming in GB has undergone immense changes in a bid to remain viable amid the changing consumer trends and market forces of recent years. Emergent behaviour that came to light during the FMD epidemic in 2001 includes the frequent movement of cattle over long distances. Although these movements are intrinsic to the industry, they also form contact networks for disease propagation between farms which would not otherwise be in contact. In terms of disease transmission, the more connected a network, the more probable the disease will spread to a greater number of individuals. Therefore, in response to risks posed by moving diseased animals, movement bans were imposed during the FMD outbreak in 2001. Animal movements continue to be restricted up to this day in

*Author for correspondence (s.e.robinson@liv.ac.uk).

the form of standstill policies, which prohibits the removal of livestock from premises until a minimum 6-day period has elapsed following entry of an animal (although many exemptions exist; www.defra.gov.uk/animalh/movements/default.htm). Although this policy was originally introduced in order to reduce the number of farms potentially infected during periods when disease spread undetected in the population, it was also anticipated that this rule would reduce the number of animal movements, therefore cutting the risk of a disease outbreak (rules for livestock movements: standstill, http://www.defra.gov.uk/animalh/movements/standstill/default.htm). The rule is troublesome for the livestock industry and for the people who try to police it, not least because it is complicated by exemptions and regional variation (e.g. standstill period of 13 days exists in Scotland). DEFRA have acknowledged that some confusion exists and an independent review of livestock movement rules to examine the impact of government policies on farmer behaviour and risk of disease spread was commissioned (review of the livestock movement controls, http://www.defra.gov.uk/animalh/movements/pdf/livestock_movement_controls-review.pdf). Changes in farmer behaviour towards herd management and cattle movements since the FMD outbreak are difficult to quantify. Yet, they warrant examination as they affect the efficacy of the 6-day standstill policy.

Exploring cattle movement data over the last 3 years provides a unique opportunity to measure how animal keepers' responses to the restrictions have affected cattle movements. By forming a link between two premises in the direction of the cattle movement, a directed contact network can be created that enables exploration of the connectivity between different premises and the potential for disease spread. Applying network analysis methods to explore the effect of network structure, we can investigate the effect of restrictions on the potential for disease spread. Although network analyses have previously been conducted using cattle movement data (Shirley & Rushton 2005; Kao *et al.* 2006; Ortiz-Pelaez *et al.* 2006), the focus has often been on reporting individual node or subgroup-level properties.

Many networks exhibit community structure, the division of network nodes (AH in this case) into groups within which the network connections are dense, but between which they are sparser (Newman & Girvan 2004). Components, the simplest form of cohesive subgroup embedded within networks, are maximal connected subregions of a network in which all pairs of nodes are directly or indirectly linked; for directed networks (those in which the link exists in one direction only), components can be further classified as strong or weak. Strong components are regions of the network where every node can be reached from every other node via directed paths, whereas weak components are areas of the network in which nodes are linked but we ignore the direction of the tie. In many networks, a single large weak component exists, which contains a single large strong component. These form the simplest kind of core-periphery networks, the strong component forming the core with the remaining nodes in the periphery.

We shall call the dominant large weak component the giant weak component (GWC) and the strong component at the core the giant strong component (GSC). Additional weak and strong components may exist but these include substantially fewer nodes than the giant components. As each node in a strong component can be reached from every other node, the size of the giant strong and weak components can be important guides as to upper and lower bounds of the maximum size of an epidemic, i.e. the population at risk. The strong component represents the part of the network such that if any one of the AH are infected, then all other nodes in the same strong component are exposed together with some of the peripheral nodes.

Using social network analysis (Wasserman & Faust 1994), we investigated successive weekly networks of cattle movements in 2002–2004 in GB using the Cattle Tracing System (CTS) data, sourced from the RADAR information management system. The CTS captures all cattle movements, births and deaths in GB. From these data, we constructed directed contact networks in which a contact between two AH represents at least one animal moving between them. The direction of the contact assumes the direction of movement. Using this network, the potential for disease spread is investigated over time by examining the size of the strong component, the maximum population at risk should be an infectious agent introduced into an AH in the largest component.

2. METHODS

2.1. Rapid analysis and detection of animal risk database

In the UK, it is a legal requirement for owners of cattle to report movements, including births and deaths to the British Cattle Movement Service. Data on movements of cattle were derived from the UK Department for Environment, Food and Rural Affairs (DEFRA) RADAR database based on data download from the BCMS CTS database on 8 April 2005. The extract included data up to 8 April.

2.2. Data extraction

Data from 1 March 2005 onwards were deemed to be incomplete at the time of data retrieval and were not included. The principal data files used for the network analysis included animal movement data. Each record contains the location of each animal throughout its life and provides a life history for each animal. This is in the form, animal i was at location j from date 1 to date 2. However, in order to create a movement network, movement data were needed in the form, animal i moved from location j to location k on date 2. Data with departures before 1 January 2002 were removed. Only movement records after this date were included as prior to 2001 movements, records were often incomplete because the FMD epidemic of 2001 greatly disrupted cattle movements in GB. When an animal departure date was missing, it was assumed that the animal was still alive on the last location at which it was recorded

as being present at the date at which the data ended. This provided the raw dataset which was used to rejoin movements.

A second field was added and set to equal the date of arrival minus the date of departure for each animal. This enabled the identification of multiple movements in a single day, with the 'transient locations' scoring zero. This correctly sorted the movement's records, except where more than one 'intervening' location was present in a single day (i.e. A→B→C→D). In such circumstances, the correct order of the intervening locations (B→C or C→B) could not be ascertained.

The data also contained movements to or from unknown locations. Hence, the movement data were reconstructed after removing movements to or from unknown locations (resulting, in total, in the loss of 16% of movements). This was done in such a way that movements such as A→unknown→B would appear as A→B. Where an animal moved to an unknown location, then back to the original location (i.e. A→unknown→A), loops were generated when the movements to unknown locations were ignored (i.e. A→A).

2.3. Network analysis

In order to examine the network evolving over time, we created networks of AH contacted by cattle movements during sequential weekly (Sunday to Saturday) periods. The selection of weekly intervals over which to construct the networks was partly informed by 6-day standstill for cattle in England and Wales, which prohibits the movement of any animal off the premises for a period of 6 days after an animal arrived onto the premises. Although the weekly interval and 6-day waiting period may not coincide, by restricting the network to weekly movements each AH receiving cattle will only do so once within any one week. Records of movements occurring between the location types, AH, landless keepers, common land and MA were selected as these were considered as important for propagating disease between cattle herds (as opposed to other location types in the dataset associated with death or slaughter). For the remainder of this study, these are referred to collectively as AH. Reports of births and death were also disregarded at this stage. Data regarding the movements of cattle between different AH for each week (Sunday to Saturday) during the time period between 1 January 2002 and 28 February 2005 were used to construct directed weekly trading networks in Pajek (De Nooy *et al.* 2005; http://vlado.fmf. uni-lj.si/pub/networks/pajek/). For each weekly network, we recorded the following network features: the number of ties existing between premises (i.e. the number of movements); the number of premises moving cattle; the number of cattle moved; the size of the GSC; the size of the GWC; the number of strong components; and the number of weak components.

2.4. Time-series analysis

The raw data of these network features were plotted to assess the temporal trends in the weekly cattle movement network. Data were smoothed and plotted with 3 and 53 week rolling average windows to highlight seasonal and long-term trends in the network. Smoothing and plotting routines were carried out in the statistical package R (www.r-project.org). Many forms of contact are transient and this temporal component of the network is an important feature of the data. Within the cattle movement database, there is an inherent temporal structure to the data, given the 6-day standstill period that cattle must remain on a farm before moving off. Within a 7-day period, there is no explicit temporal order of the contacts within that 7-day period. Movements between premises have been modelled as batch movements; in effect, we are assuming that if one farm is infected, then all cattle on that farm are infected.

3. RESULTS

As expected, during the study period, seasonal patterns existed for several features of the contact structure between AH. Within each year, spring and autumn peaks are evident in the number of cattle movements off premises, number of animals moved, size of the GSC, size of the GWC and the number of strong components (figure 1). For the majority of these network parameters, the autumn peak is larger than the spring peak.

The size of the strong component exhibited a strong long-term trend, increasing linearly from approximately 275 AH (the 53 week rolling average for this time) in mid-2002 to approximately 800 AH in mid-2004 (figure 1a). Of the premises associated with the GSC, approximately 50 in each week were MA. The size of the GWC each week also increased although these changes were relatively small when compared with the changes in the GSC but, together with the increase in the size of the GSC, indicates decreased segmentation and increased cohesion of the overall network, thereby increasing opportunities for sustaining transmission. Although the number of strong components in the network was relatively constant (figure 1c), the number of weak components gradually declined due to the coalescence of smaller components into components of larger size (figure 1d). These changes in component structure have occurred with only moderate increases in the number of animal movements (figure 1b), the number of AH moving cattle and the number of strong components in the network.

Like most real-world networks, we found the cattle movement network to behave as a 'small world' with a greater degree of clustering than an equivalent random network, but with similar average path length (estimated on non-directed networks).

4. DISCUSSION

Our most significant finding is the linearly increasing trend in the number of AH moving cattle in the GSC based on weekly contact networks. These results have implications for the spread of disease within the cattle population within GB as both empirical and theoretical studies suggest that the size of the GSC influences the size, rate of spread and threshold of an epidemic (Moreno *et al.* 2002; Schwartz *et al.* 2002;
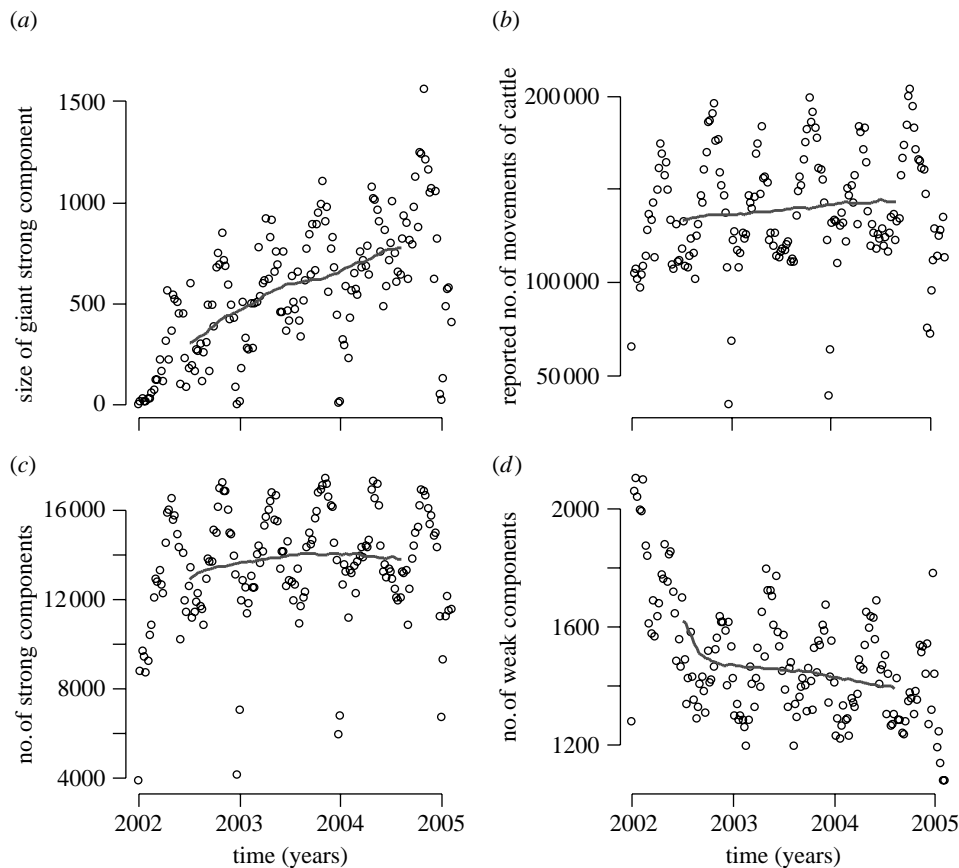
Figure 1. Temporal trends observed for cattle movement data between 2002 and 2005. (*a*) The size (number of premises) of the giant strong component arising from weekly contact networks of cattle movements between animal premises. Three week smoothing averages are plotted as circles, with the continuous grey line indicating 53 week smoothing average. (*b*) The number of movements of cattle. (*c*) The number strong components. (*d*) The number of weak components.

Newman 2003; Kao *et al.* 2006). The result of the network evolving into a less fragmented industry, with more holdings trading in the GSC each week (after allowing for seasonal variation), would be larger epidemics once disease entered this highly connected subregion of the network. Furthermore, it has been reported that highly connected populations, such as we have reported, could provide selective pressure that drives the more rapid evolution of pathogens in the population (Girvan *et al.* 2001). As with the overall numbers of cattle movements, the size of the GSC was found to vary with season, indicating that the risk and size of an epidemic varies across the year.

This analysis is not constrained to any one infectious disease agent; our findings are relevant to any disease agent that may be spread through the movement of infected cattle. However, they have less implication for epidemics where more stringent movement restrictions follow early (rapid) detection of disease. In contrast, our findings have direct application to the current epidemic of bovine tuberculosis in the GB which has been focus of recent concern (Gilbert *et al.* 2005; Donnelly *et al.* 2006). If farms within the GSC are found to be infected with bovine tuberculosis, then the current trend of increasing size of the GSC will result in more AH potentially becoming exposed. Our analyses have only considered the implications for disease spread in the direction of the

cattle movement. Specific features of different diseases that affect other farm animal species or bi-directional transmission could be explored using network analysis in which links between different AH assume different properties. However, further work may also need to explore the possibility that changes in on-farm management affecting a farms' susceptibility to disease may have accompanied the change in the industry's increased connectivity.

In this analysis (and most other recent network studies using these data), contact between AH is limited to the movement of cattle between them. In most disease contexts, this is unlikely to be a true reflection of reality. Other potential transmission routes contributing at the local level include environmental pathways, wildlife and personnel, evidenced during the FMD epidemic in 2001 (http://www.defra.gov.uk/footandmouth/pdf/fmdorigins1.pdf). Therefore, estimates of the size of an epidemic from network models that only include contacts between farms via cattle movements may underestimate the size, as well as the spatial spread of an epidemic. Therefore, more sophisticated disease-specific network models incorporating a range of contact types, and explicitly modelling local spatial transmission should aid in the development of more accurate predictions on which to base targeted interventions.

The chief aim of the 6-day standstill policy was to reduce the potential for disease to spread undetected within livestock populations in future outbreaks. It was also envisaged that this rule would limit the number of AH able to move cattle resulting in a less cohesive network. Our findings suggest that during the last few years, the cattle network is becoming more cohesive. The effect of the legislation was greatest immediately following its introduction but has decreased over time as cattle movements, essential for the viability of AH, have become increasingly organized in order to move cattle within current restrictions. The size of the GSC within weekly networks is surprising, considering the 6-day (England and Wales) and 13-day (Scotland) standstill rule that applies to AH moving cattle. This would suggest that some degree of organization must have been evolving within cattle industry to facilitate movements. We speculate that it may result from increased communication between animal keepers resulting in a more coordinated network to allow maximum movement within current movement restrictions or from increased reliance on markets for animal trade (increasing their role as movement hubs). However, these data are based on reported movements and hence it is possible that the observed changes in the network are the result of changes in reporting behaviour or collusion of farmers to circumvent the rules rather than actual cattle movements.

While this analysis has given an insight into the global network changes in animal movements, the drivers for these changes are harder to ascertain. There have been numerous changes in legislation affecting livestock farming in recent years, and therefore identifying which specific rules have resulted in the greatest impact on the cattle movement network cannot be determined from the data. However, the steady growth in the cattle population combined with a decline in the number of AH, MA and slaughterhouses in GB (a study of long-term trends affecting the farming industry, http://www.defra.gov.uk/farm/policy/regulation/charge/pdf/long-term-trends.pdf; Robinson & Christley 2006) is likely to be having an effect, such that those AH remaining are forced to trade as part of larger groups. Introducing measures to segment or fragment the network, important for interrupting transmission of disease (Potterat *et al.* 1999), may need to be explored. However, if the effect of introducing new legislation increases the rate at which AH, especially those in peripheral sites in the network, leave the industry, the overall effect may be to make the remaining AH more cohesive.

We have presented evidence of behavioural adaptation to regulation. Regardless of the intended effects of the 6-day standstill, self-organization, a natural phenomenon within networks, is creating a more cohesive group in which cattle are traded. Current policy regarding cattle movement restrictions was informed by studies which suggested that this policy would interrupt spread of disease and reduce the risk of FMD spreading undetected (rules for livestock movements: standstill, http://www.defra.gov.uk/animalh/movements/standstill/default.htm). Owing to evolving behaviour within the network, its efficacy appears to have been reducing over time since its introduction.

The use of network theory for these sorts of issues will be essential to assess the real performance of future legislation.

5. CONCLUSION

To date, network analyses of livestock movement databases have been used to identify individual or subgroup-level network properties. Our study of network-level properties shows that current legislation may need to be reviewed if the risk of large epidemics is to be reduced, and that the ongoing collection of cattle movement data presents governments with a unique opportunity to monitor changes in global network topology in order to influence, create or validate regulation to enhance disease surveillance policy.

REFERENCES

Bourn, J. 2003 *Identifying and tracking livestock in England.* London, UK: Stationery Office, on behalf of National Audit Office.

De Nooy, W., Mrvar, A. & Batagelj, V. 2005 *Explanatory social network analysis with Pajek.* Cambridge, UK: Cambridge University Press.

Donnelly, C. A. *et al.* 2006 Positive and negative effects of widespread badger culling on tuberculosis in cattle. *Nature* **439**, 843–846. (doi:10.1038/nature04454)

Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R. & Wint, W. 2005 Cattle movements and bovine tuberculosis in Great Britain. *Nature* **435**, 491–496. (doi:10.1038/nature03548)

Girvan, M., Callaway, D. S., Newman, M. E. J. & Strogatz, S. H. 2001 A simple model of epidemics with pathogen mutation. *SIAM Rev.* **99**, 1–9.

Kao, R. R., Danon, L., Green, D. M. & Kiss, I. Z. 2006 Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. *Proc. R. Soc. B* **273**, 1999–2007. (doi:10.1098/rspb.2006.3505)

Mitchell, A., Bourn, D., Mawdsley, J., Wint, W., Clifton-Hadley, R. & Gilbert, M. 2005 Characteristics of cattle movements in Britain—an analysis of records from the cattle tracing system. *Anim. Sci.* **80**, 265–273. (doi:10.1079/ASC50020265)

Moreno, Y., Pastor-Satorras, R. & Vespignani, A. 2002 Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B* **26**, 521–529. (doi:10.1007/s10051-002-8996-y)

Newman, M. E. J. 2003 Properties of highly clustered networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **68**, 026 121.

Newman, M. E. J. & Girvan, M. 2004 Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **69**, 026 113.

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.

Potterat, J. J., Rothenberg, R. B. & Muth, S. Q. 1999 Network structural dynamics and infectious disease propagation. *Int. J. STD AIDS* **10**, 182–185. (doi:10.1258/0956462991913853)

Robinson, S. E. & Christley, R. M. 2006 Identifying temporal variation in the reported births, deaths and movement of cattle in Great Britain. *BMC Vet. Res.* **2**, 2.11. (doi:10.1186/1746-6148-2-11)

Schwartz, N., Cohen, R., ben-Avraham, D., Barabasi, A. L. & Havlin, S. 2002 Percolation in directed scale-free networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **66**, 015 104.

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. (doi:10.1017/S0950268880500453X)

Taylor, L. H., Latham, S. M. & Woolhouse, M. E. 2001 Risk factors for human disease emergence. *Phil. Trans. R. Soc. B* **356**, 983–989. (doi:10.1098/rstb.2001.0975)

Wasserman, S. & Faust, K. 1994 *Social network analysis.* Cambridge, UK: Cambridge University Press.