Network epidemiology of cattle and cattle farms in Great Britain

Submitted by

Helen Fielding

to the University of Exeter as a thesis for the degree of

Doctor of Philosophy in Biological Sciences in September 2019

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: …………………………………………………………..

‘I have yet to see any problem, however complicated, which, when you look at it in the right way, did not become still more complicated.’

writer Poul Anderson,

quoted in ‘Infectious diseases of humans’ (Anderson and May, 1991)

Abstract

Infectious diseases of livestock can cause substantial production losses and have detrimental impacts upon human health, and animal health and welfare. To limit the impact of diseases, understanding more about the dynamics of transmission can assist in the control and prevention of infectious disease. In particular, understanding infection transmission on networks, ‘network epidemiology’, offers a flexible approach, incorporating between-host heterogeneity in potentially infectious contacts drawn from empirical study of interactions among individual animals, or among farms. Trading animals and optimising productivity are vital to the commercial viability of farms, however they necessarily involve compromises in biosecurity, animal health, and welfare. Better understanding of the relationships among these multiple factors might facilitate the development of sustainable livestock industries that are more resilient to disease outbreaks.

In this thesis I examine cattle interactions at two spatial scales, first at a national-level by studying the trading connections among farms, and then at a finer scale by analysing the social interactions among cattle. First, I introduce the concept of superspreaders, hosts that generate many more secondary infections than the rest of the population, and evaluate evidence for the notion that some farms might act as superspreaders of infection. I utilise the example of bovine tuberculosis (bTB) to illustrate this concept and find that farms might act as superspreaders in three main ways; first, via exceptional trading between farms, second, by factors that facilitate high within-herd transmission and trading of high-risk animals, and third, by harbouring undetected infection for long periods. I find mechanisms that align with all three processes in the cattle industry in Great Britain that might allow superspreader farms to contribute to the current bTB epidemic.

At a national level, I describe cattle movements among farms over time, finding that some farms consistently act as ‘hubs’ in trading networks, functioning in a similar way to markets, in that they are highly connected to other farms by many direct trades. Utilising the temporal network measure of ‘contact chains’, I quantify the farms that represent

potential sources of infection (ingoing contact chains) and the potential farms that a farm might infect over 1 year periods. Farms divide into two groups: those with very few connections (less than 10 farms) that are relatively isolated from the network, and those with very many connections (more than 1000 farms) that are highly connected within the network. I find that a substantial number of farms have over 10,000 farms in both their ingoing and outgoing contact chains, such that, if infected, they might potentially act as superspreaders by being more at risk of both acquiring and spreading infection.

Building on my previous analysis, I then characterise the 'source farms' in the ingoing contact chains, in terms of their location and bTB history. I find that after controlling for previously-established risk factors for bTB, having more source farms in areas of higher bTB risk in the ingoing contact chain increases the odds of a bTB incident on the root farm, whilst having more source farms in lower risk areas is associated with lower odds of a bTB incident on the root farm.

At a finer scale of contacts among animals, I explore interactions among dairy cattle in multiple herds using automated proximity sensors and GPS devices. When aggregated over long periods, cattle interactions appear dense and unstructured, however finer time spatial and temporal perspectives revealed structure and variation in contacts. Herds in our study had variable grazing and housing access, allowing us to determine that cattle interact with more other cows, for longer time periods when they are in buildings compared to contacts at pasture. Cattle exhibited heterogeneity in their number and duration of contacts, and although the majority of cattle interacted more equally with other cows, a small proportion of cows in each group showed evidence of stronger social ties.

Next, I consider associations between social interactions, production, and health. I review the existing literature on social parameters such as dominance rank and re-grouping of cattle, and find inconclusive outcomes regarding their impact on milk yield and somatic cell count, an indicator of udder health. I perform my own analysis to examine the relationship between the time cows spend with other cows, milk yield and somatic cell count, and do not find a statistically significant relationship. In considering social preference, cows that had experienced the same number of lactations were more likely

to interact, but cows spending more time with cows in the same lactation did not appreciably affect their milk yield or somatic cell count.

Finally, I draw together the findings of this thesis and reflect on how the identification of higher-risk farms might be useful in the control of livestock infections, and specifically bTB in Great Britain. I conclude that network analysis is a valuable tool to study the interactions of cattle and cattle farms, identifying unique opportunities for targeted approaches to disease control.

Acknowledgements

It has been an absolute privilege to do this PhD, I have learnt so much from so many people (mainly ecologists, who'd of thought?!) and will not be able to name them all here. If you care enough to read this, you probably helped me along the way, please consider yourselves all truly thanked.

First, I must thank Robbie McDonald for giving me the opportunity to work on this project and taking a chance on a farm vet, thank you for your support and guidance along the way. Thanks to Dez Delahay for sage guidance when it was needed, and to TJ McKinley for bearing with me when I asked what a model was, for introducing me the wondrous ways of R, and for teaching me the importance of the Oxford comma. Along with my supervisors, one person to whom I am hugely indebted is Matt Silk, who was always available to help with networks advice, general academic know-how, and a calming influence.

The majority of this thesis would not have been possible without a massive (some might say too massive) dataset from the APHA. I am grateful to the epidemiology and data teams for providing me with the data and helping me to understand it. Thank you very much to Laura Ozella and the team at the Institute for Scientific Interchange, and Jared Wilson-Aggarwal for their assistance and expertise in using and processing data from the proximity sensors.

Without the kind help of nine fantastic Cornish dairy herds, I would have had a lot of collars and no cows, so thank you to the farmers and thank you to all but one (you know who you are) of the 871 dairy cows that participated. I am hugely grateful to Ed Firth, Kelly Astley, Charlotte Epps, and Maddy Green for helping me both on farm, and assisting in the construction, destruction, and cleaning of hundreds of collars at all times of the day and night. Thank you to Becky-Clare Akers, Will Jones, Rich Woods, Cath Ratcliffe, and Amy James for lending me your time and cows for trial collar work.

To all members past and present of the Clan McDonald research/PhD support group, to name but a few; Olivia Bell, Jared Wilson-Aggarwal, George Swan, Rich Woods, Martina Ceccheti (for always being more depressed about the mizzle than me), and other ESI

lovelies, Owen Exeter and Tom Horton, you are just the best, keep being wonderful. Thank you to Nell Williams Foley, who has kept me in tea and chocolate over the past few months, with constantly encouraging words. Particular thanks go to Katie Sainsbury, Nigel Sainsbury, Cecily Goodwin, Sarah Crowley, and Cat McNicol for their friendship, support, coffee, and dog walks that kept me going. I am grateful to 'the Cave', which, for the last four years has unburdened me from such distractions as the outside world and daylight.

In much-needed breaks for the PhD, I have enjoyed the Cornish countryside with many of you, thank you to 'the Richs' (Inger and Sherley) for making me jump in the Cornish sea, and convincing me I always had time for a swim, run, circuits, or all three, when I thought I didn't - you were right! Thanks to James Duffy and Dan Padfield for being on hand for GIS, cluster, or R related problems, and the odd bike ride in the Cornish hills.

Thank you to my siblings Sarah and Matt, and my parents Heather and Mike, for understanding when I needed to work and for always being there if and when I needed you, thank you for your support and encouragement. My family also includes a small Labrador who, despite being neglected by me over the last few months, has brightened every single day since her arrival, thanks Ivy.

Finally, to Ed, I am so grateful to you for agreeing to come on this adventure with me, and for being a constant source of support and encouragement. Over the last few months, you have taken on everything so I could focus on this thesis, it is a truly joint effort.

So long and thanks for all the fish.

Contents

List of figures and tables

Author’s declaration for co-authored manuscripts

Chapters 2, 3, 4, 5, and 6 have been published or written for publication as co-authored academic papers. I conceived of the papers with the assistance of Robbie McDonald (RM), TJ McKinley (TM), Richard Delahey (RD), and Matthew Silk (MS).

For Chapters 3 and 4, all cattle movement and bovine tuberculosis testing data was provided by the Animal and Plant Health Agency. I designed the fieldwork for Chapters 4 and 5 and implemented it with the assistance of Edward Firth, Kelly Astley, Madeleine Green and Charlotte Epps and with the kind permission of eight Cornish dairy farms.

The proximity data was first processed by Laura Ozella, using code written by Ciro Cattuto at the Institute for Scientific Interchange Foundation, Torino. I performed all analyses in each chapter, with the assistance of MS and TJM.

I wrote all 5 manuscripts and amended them in response to comments, advice and suggestions from RM, TJM, RD and MS, and from anonymous reviewers for Chapters 2 and 3.

Chapter 1

Introduction

Chapter 1: Introduction

Infectious diseases

Advances in scientific and medical knowledge over the 20th century and, in particular, the development and large-scale implementation of vaccination, has led to the global eradication of smallpox in humans (Henderson, 2011) and Rinderpest in cattle (Roeder, 2011), and more general and dramatic reductions in human and animal mortality due to infectious disease (Anderson and May, 1992; Woods, 2011). Nevertheless, there are still many challenges in the control of infectious disease. Over 300 new communicable diseases of humans emerged between 1940 and 2004 (Jones et al., 2008) and in 2015, over 7 million people died of infectious diseases (Mathers et al., 2017). Globalisation means populations are exposed to a much wider range of pathogens and pandemics are more likely (Fauci, 2001). Animal and human health are closely linked; over 60% of newly emerging infectious diseases of people are caused by zoonotic pathogens (Jones et al., 2008). Although most of these infections have their origins in infections of wild animal populations, (e.g. human immunodeficiency virus (HIV), Ebola), domestic species can often act as a maintenance reservoir for transmission between wild animals and humans, for example avian influenza, rabies, brucellosis (Kilpatrick et al., 2009), thus making the study of livestock infections integral to safeguarding public health of human and non-human animals alike (Cleaveland et al., 2001).

Rising human populations and concerns over food security have stimulated the intensification of agriculture (The Government Office for Science, 2011) and substantial increases in the growth rates of pigs and poultry and milk yields of dairy cattle have been achieved through genetic profiling and selective breeding (Pryce and Veerkamp, 2001). Herd sizes have increased in order to increase efficiency and take advantages of economies of scale, while producers are becoming more specialised towards specific production stages, e.g. breeders, fatteners, finishers (Morgan and Prakash, 2006). However, there are potential conflicts between animal health and welfare, and commercial efficiency (Dawkins, 2017). Larger herds are frequently identified as being at higher risk of a disease incursion (Brooks Pollock and Keeling, 2009; Gardner et al., 2002) and increased trading of animals between more-specialised producers increases the risk of between-farm spread of infection (Gates, 2014; Robinson and Christley, 2007). Maximising productivity brings compromises in terms of reduced longevity, lameness, poor fertility, increased susceptibility to disease, and compromised welfare (Mench, 2002; Oltenacu and Broom, 2010) leading to increasing concern from consumers about animal welfare in intensive systems (Simm et al., 2001). Keeping animals at higher stocking densities and maximising their production has been facilitated by the non-therapeutic and therapeutic use of antimicrobials in livestock, and has most likely facilitated and even promoted the development of antimicrobial resistance (Gilchrist et al., 2007).

Understanding the intertwined relationships of a changing livestock industry and the health and welfare of animals is crucial in order to maintain a sustainable, commercially-viable livestock industry.

Quantifying infections

Since the emergence of germ theory, mathematicians and epidemiologists have sought to document and analyse the interactions between pathogens and host populations, starting with Daniel Bernoulli, who in 1760, successfully used mathematical methods to prove the efficacy of early forms of vaccination against smallpox (Bernoulli and Blower, 2004). Mathematical models, as applied to infectious disease, are simple frameworks that provide basic rules governing the spread of infections. A founding principle of epidemiology, based on the mass-action theory, is that the rate of spread of an infection is proportional to the product of the density of susceptible individuals and the density of infected individuals (Anderson and May, 1992, 1979), and this is outlined in the differential equations used in compartmental models of disease transmission (Kermack and McKendrick, 1927; Fig. 1.1). Each compartment contains individuals in different states of infection, which classically includes susceptible individuals (S), infected individuals (I) and recovered individuals assumed to have some degree of immunity (R).

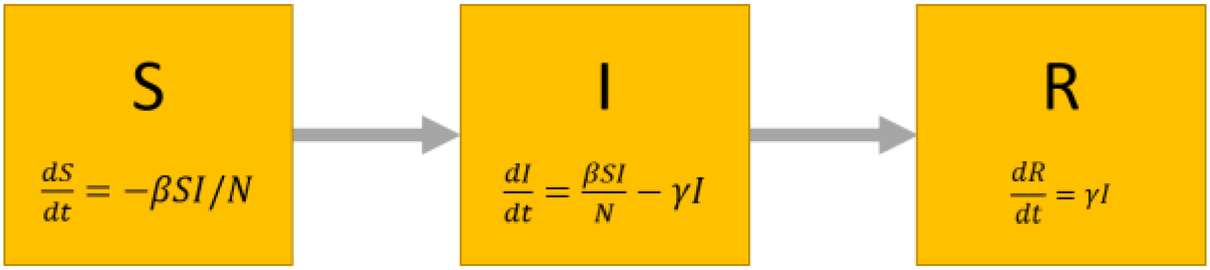


Figure 1.1. *Schematic of a frequency dependent compartmental deterministic SIR model. The rate of change in each compartment is governed by their respective equations; where N is the number of individuals in the population, S is the number of susceptible individuals, I is the number of infected individuals, R is the number of recovered individuals, β is the transmission rate and γ is the removal rate. In this model* $\frac{1}{\gamma}$ *is the mean length of the infectious period (Anderson and May, 1992).*

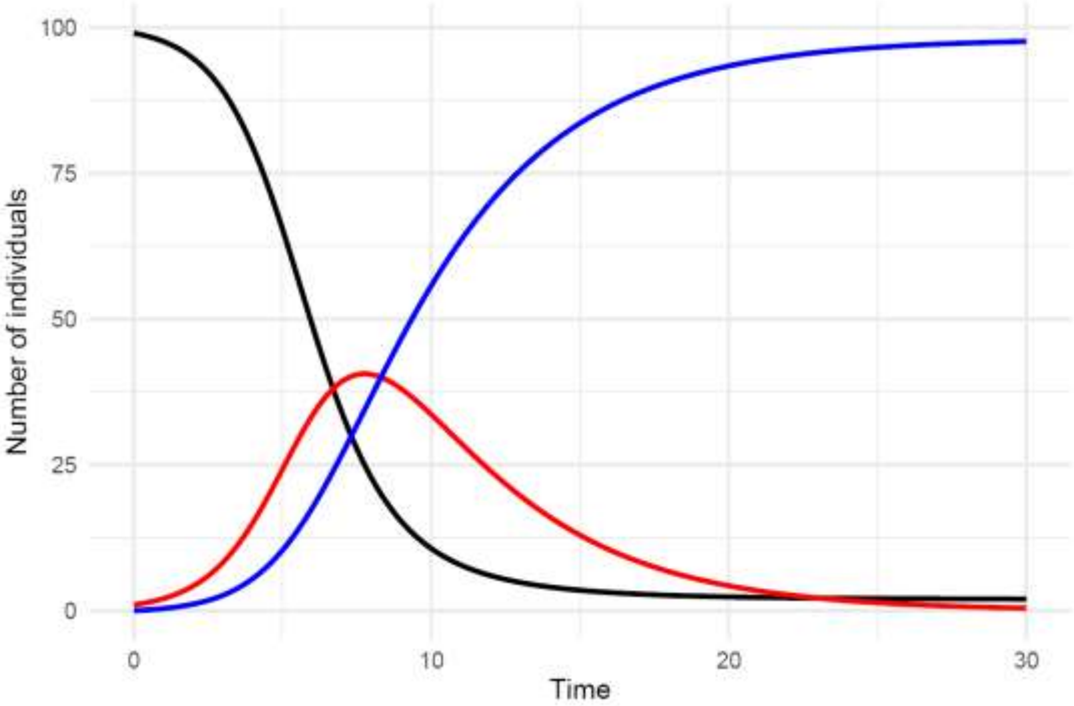


Figure 1.2. *Frequency of susceptible (black), infected (red) and recovered (blue) individuals over time from a simple SIR model simulation. Model was initiated with 1 infected and 99 susceptible individuals in closed population and parameterised with β = 1, γ = 0.25. The shape created by the number of infected individuals generated also represents the epidemic curve.*

When the density of susceptible individuals is above a critical threshold, an epidemic will occur, and when the density goes below that threshold, the epidemic will start to reduce and typically will die out, thus forming the classic ‘epidemic curve’ (Kermack and McKendrick, 1927; Fig. 1.2). From these models, we can estimate the basic reproduction ratio ‘R_0’, the average number of secondary infections that would arise from a single infected host in a population of susceptible individuals (Anderson and May, 1992; Diekmann et al., 1990)

$$R_0 = \frac{\beta}{\gamma}$$

If R_0 is less than or equal to one, infection is unlikely to proliferate, however if each host generates more than one secondary infection ($R_0 > 1$), the infection is likely to persist in the population, at least in the short-term (Keeling, 1997). The calculation of R_0 allows the testing of control strategies, potentially by calculating the proportion of the population that would need to be vaccinated (Keeling, 1997), or the number of animals that might need to be culled to contain an epidemic (te Beest et al., 2011).

Heterogeneity in R_0

R_0 is a population average and whilst it provides good estimates of disease spread where infection parameters follow normal distributions, its utility is reduced when heterogeneity arises in contact rates, infectiousness or the length of the infectious period (VanderWaal and Ezenwa, 2016). The impact of contact heterogeneity is well-illustrated in the case-studies of the first people infected with Severe Acute Respiratory Syndrome (SARS) in Canada in 2003, where the spread of infection differed markedly between Vancouver and Toronto due to the contact behaviour of the initial hosts in each region (Poutanen et al., 2003). Initial estimates of R_0 for SARS were overestimated as they were based on data collected in a hospital and a crowded apartment building, where contact rates were particularly high (Shen et al., 2004). Thankfully, the predictions of a large-scale pandemic were not realised and estimates for R_0 in the general population were much lower. However, an important factor in the spread of SARS was the high number of infections generated by some individuals, when models allowed R_0 to vary, epidemic predictions were much more accurate (Meyers et al., 2005). Individuals that generate disproportionately large numbers of secondary cases than the majority of other infected individuals are termed ‘superspreaders’, and have been identified in many infectious diseases, including SARS (Lloyd-Smith et al., 2005; Stein, 2011). Infected individuals may act in this way via high contact rates, being infected for a long duration or by being highly infectious (Shen et al., 2004; Small et al., 2006). If the majority of individuals do not transmit infection to many others but some act as superspreaders, the likelihood of an epidemic occurring is decreased, as infection is more likely to die out if it encounters one of the more numerous poorly-connected individuals, however if a superspreader is infected in the initial stages, many hosts will be infected quickly and the resulting epidemic is likely to be larger overall (Lloyd-Smith et al., 2005). When superspreader dynamics exist, modelling has suggested that targeting control efforts towards those individuals with particularly high rates can be much more effective than population-wide control measures; and thus identifying them can be a key step in the control of some infections (Lloyd-Smith et al., 2005).

Network epidemiology

Using networks as a framework for understanding disease spread requires and allows incorporation of heterogeneity in contact structure (Bansal et al., 2007) and spatial aggregation of infection (Keeling, 1999). The study of networks is cross-disciplinary, with strong conceptual origins in the social sciences and underpinnings in mathematical graph theory. In the context of epidemiology, nodes in a network typically represent individual hosts, or groups of hosts (e.g. a farm or school) and edges represent a link between those nodes. In a contact network, edges represent a connection between nodes that has the potential to transmit infection, which should be distinguished from a transmission network, where edges represent the actual transmission of infection (Craft, 2015). Edges may be directed or undirected; directed edges are used to describe the movement of animals or people between locations, whereas undirected edges describe a mutual connection such as a contact or relationship (Fig. 1.3). Edges can be weighted, typically by the duration or frequency of the contact between nodes.

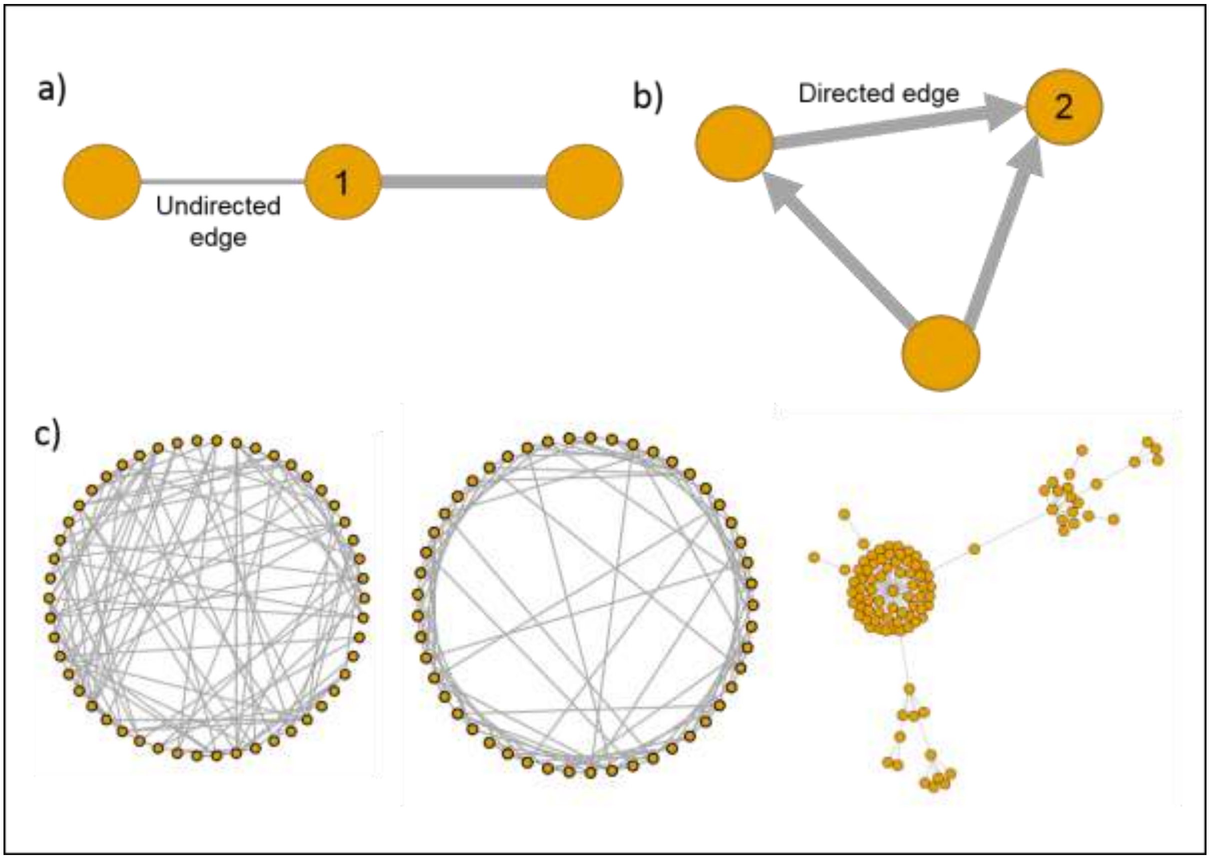


Figure 1.3. *Elements of a network and examples of different network structure. Nodes are depicted as orange circles and edges as grey lines. a) shows an undirected network comprising three nodes connected by two edges of different weights, illustrated by the width of the edge. The degree of node 1 is 2. b) shows a directed network comprising three nodes with three directed edges, the in-degree of node 2 is 2 and the out-degree is 0. The three nodes form a triplet as they all have a mutual node. c) shows examples of simulated networks based on random (Erdös-Rényi), small-world (Watts-Strogatz), and scale-free (Barabasi-Albert) algorithms implemented in the R package 'igraph' (Csardi and Nepusz, 2006) from left to right.*

The utility of integrating social relationships into the study of disease transmission was first demonstrated in 1985 where a study of sexual contacts and the occurrence of AIDS (acquired immune deficiency syndrome) demonstrated a likely infectious cause for the newly-emerged disease (Klovdahl, 1985). Networks are particularly useful in analysing sexually-transmitted infections due to the direct method of transmission and the heterogeneity between individuals in terms of contact rate (May and Anderson, 1987). It was quickly apparent that in addition to the characteristics of the host and pathogen, described by R_0, network structure could also influence the critical threshold over which an epidemic would occur (Keeling, 1999).

Network analysis characterises the local and global structure of the network in many different aspects. ‘Clustering’ or ‘transitivity’ describes the number of completed triangles in the networks relative to the number of possible triangles (Silk et al., 2017a), more simply put, the friends of your friends are also your friends (Appendix B Table B1). Networks displaying a high level of clustering are considered to have ‘small world’ properties (Keeling and Eames, 2005) and the differences between transmission in a homogeneously mixing population and on a network are most keenly felt when these local connections are more numerous. In this type of network, for any given node, the pool of susceptible contacts is limited (due to more shared connections) and quickly becomes locally saturated with infected individuals. This limits the initial spread of the epidemic and ultimately, the final epidemic size (Keeling, 1999; Keeling and Eames, 2005). Where networks are less clustered and have many connections distributed among all nodes, each node has a more wide-ranging pool of susceptible neighbours. In this scenario, the effect of the network is reduced and simulated epidemics quickly spread through the population, thus behaving more similarly to those on random networks (Keeling and Eames, 2005).

Degree and strength describe the number and combined weight of edges connected to a node respectively and are used to assess heterogeneity in contact rate between individuals. Networks with extreme heterogeneity in their degree distribution such as the internet and livestock movement networks, are termed ‘scale-free’ networks (Pastor-Satorras and Vespignani, 2001). Nodes with very many more connections than the majority of other nodes are called ‘hubs’ and are able to function as superspreaders if they become infected, creating a steep epidemic curve (Keeling and Eames, 2005).

Dynamic networks

A classical assumption in network analysis, and an important distinction between networks and homogenous mixing theory, is that the edges are long standing. This is a valid assumption for many contact networks that describe social relationships. However, in network epidemiology, the permanence of the edge is most important relative to the timescale of pathogen transmission (Cross et al., 2005). In a static representation, connections should be aggregated over a defined time period suitable for the pathogen being studied, and this is useful if edges change at a similar or slower rate than the spread of the pathogen (Cross et al., 2005; Kao et al., 2007). However if the network changes faster than the pathogen can spread, we may find that the resulting network is so dense that the transmission routes of the pathogen are effectively infinite. In this case the transmission of the pathogen might actually be estimated just as well by an assumption of homogenous mixing or a random network structure (Enright and Kao, 2018). In these cases, the use of dynamic networks is crucial, and the analysis of time-respecting paths can help us better understand the subtleties of networks at different temporal aggregations.

Infectious diseases of livestock

The study of animal populations has revealed heterogeneity in social contacts and trade networks (Craft, 2015), and the use of network analysis as a tool to study disease has become very popular in veterinary epidemiology (Martínez-López et al., 2009). Infections of livestock have been modelled at individual and herd level (Marcé et al., 2011; Volkova et al., 2010b).

Most within-herd models assume homogenous mixing of animals, however the groupings that occur on most commercial cattle farms, e.g. separating calves from dams and separating male and female youngstock, break down such mixing assumptions. Some studies have included more-relevant divisions, models of bovine viral diarrhoea (BVD) and *Mycobacterium avium* subsp. *paratuberculosis* (MAP) transmission have incorporated the age and sex structure of cattle herds into compartmental frameworks due to the heterogeneity of transmission in both infections (Courcoul and Ezanno, 2010; Marcé et al., 2011). Turner et al. (2008) explicitly created a contact network parameterised with estimates from previous literature. They found that infections mainly transmitted by direct contact might more easily die out in sub-groups where there were fewer average contacts. Duncan et al. (2012) compared infection transmission between empirically-derived contact networks of two small beef herds (made up of dams and calves) and random networks of equal size and found that epidemics on the empirical network occurred less often and when they did occur, fewer cattle were infected. Further studies showed that the differences encountered could be adjusted for by increasing the probability of infection in the observed networks (Duncan et al., 2014). However, it is not clear if the findings from extensive small beef herds can be extrapolated to other herd types and sizes.

At a larger scale, networks have provided a useful framework for modelling between-farm transmission. Livestock movements have been characterized for sheep, pigs, cattle, poultry and fish in many countries with a view to inform control and surveillance policies (Green et al., 2009; Kiss et al., 2006; Lentz et al., 2016; Martin et al., 2011; VanderWaal et al., 2015). Modelling of livestock infections and control programs at this scale allows policymakers to compare and evaluate national strategies without great expense or compromises to animal welfare, and thereby to explore where limited resources might best be focussed.

As a consequence of the bovine spongiform encephalopathy epidemic in the 1990s, it is mandatory for all bovines in the European Union to have individual information (date of birth, dam and sire, breed, etc.) and movements recorded on a unique passport. Cattle keepers are required to report movements between animal holdings, markets, slaughterhouses and shows to the relevant country authority. These data represent a valuable spatio-temporal, long-term, individual-level dataset that been used to simulate the transmission of foot and mouth disease (FMD), BVD, and bovine tuberculosis between farms (Ferguson et al., 2001a; Gates et al., 2014; Gilbert et al., 2005; Kao et al., 2007; Keeling et al., 2001).

In February 2001, FMD was confirmed in pigs in Great Britain, thereafter followed an epidemic which resulted in a national animal movement ban, the culling of over 6.9 million animals, and catastrophic impacts on the farming industry and the national economy (Thompson et al., 2002). Several groups of modellers developed complex simulation models in order to account for heterogeneity in spatial clustering of disease, and in the susceptibility and infectiousness of individual farms (Ferguson et al., 2001b, 2001a; Keeling et al., 2003, 2001; Morris et al., 2001). Among other control strategies, the models suggested rapid culling on premises adjoining infected premises would be influential in halting the epidemic. Despite controversy on this policy (Mansley et al., 2011), further modelling after the event has still shown this policy to be successful in limiting epidemic size (Tildesley et al., 2009) and that this policy implemented earlier would likely have reduced the overall epidemic size (Keeling, 2005). Vaccination was not employed in the 2001 epidemic due to concerns over the time to onset of immunity and the loss of national disease-free status (Keeling et al., 2003). However, subsequent modelling has suggested that even limited vaccination resources would be potentially useful in the event of future outbreaks, if more information on vaccine deployment logistics and efficacy can be collected (Bradbury et al., 2017; Tildesley et al., 2006). This catastrophic epidemic demonstrated the benefits and caveats of different model structures and the element of compromise which is inevitable and necessary in constructing models that describe complex infection dynamics (Kao, 2002; Keeling, 2005).

The British cattle industry

Trading of cattle between farms is seen by some as vital to the sustainability and profitability of the industry (Robinson and Christley, 2007). Nevertheless, trade of cattle between farms has been identified as a factor in the transmission of multiple endemic and non-endemic infections in cattle (Broughan et al., 2016a; Chase-Topping et al., 2007; Gates et al., 2014). Markets are a key feature of cattle trading in Great Britain, they facilitate the longer-range movements of animals and the dissemination of cattle to multiple premises (Robinson and Christley, 2007). Network analyses of cattle movements in Great Britain have shown that most farms have few direct connections, but that a few farms and most markets have very many connections (Christley et al., 2005b; Woolhouse et al., 2005). This reduces the local connectedness of the network and may increase the risk of larger epidemics (Keeling and Eames, 2005).

Bovine tuberculosis in Great Britain

Bovine tuberculosis (bTB) is a zoonotic, chronic respiratory disease of cattle caused by the bacterium *Mycobacterium bovis.* Although many wildlife and domestic species can be infected with *M. bovis*, badgers have been identified as a significant reservoir host of *M. bovis* (Krebs et al., 1997) and a contributor to infection in cattle (Donnelly et al., 2003; Woodroffe et al., 2006). The socio-economic impacts of the disease are large, with disease management causing disruption and distress to the farming community and costing the UK taxpayer around £100 million annually (Department for Environment Food and Rural Affairs, 2013). bTB has been intimately related to the British cattle industry since the early 1900s. In 1934, approximately 40% of cattle were reported to have been infected with *M. bovis* (Krebs et al., 1997). A test and slaughter policy introduced in 1960 dramatically reduced prevalence in cattle. Cattle testing, combined with the rise in pasteurisation of milk between 1920s to the 1950s (Atkins, 2000), meant zoonotic cases in humans decreased (Reynolds, 2006). The annual number of reactor cattle culled in Great Britain dropped from 23,000 in 1935 to less than 1000 in 1980 (Goodchild et al., 2008). Between 1980 and 2010 though, incidence in cattle had risen in endemic areas and spread to areas in the North and West of England and through Wales (Animal and Plant Health Agency, 2014; Brunton et al., 2015; Fig. 1.4 and 1.5). Since 2010, up to data collected for 2017, incidence has become more stable overall, but with increases in bTB incidents on the edge of the endemic area (Fig. 1.5; APHA, 2018).

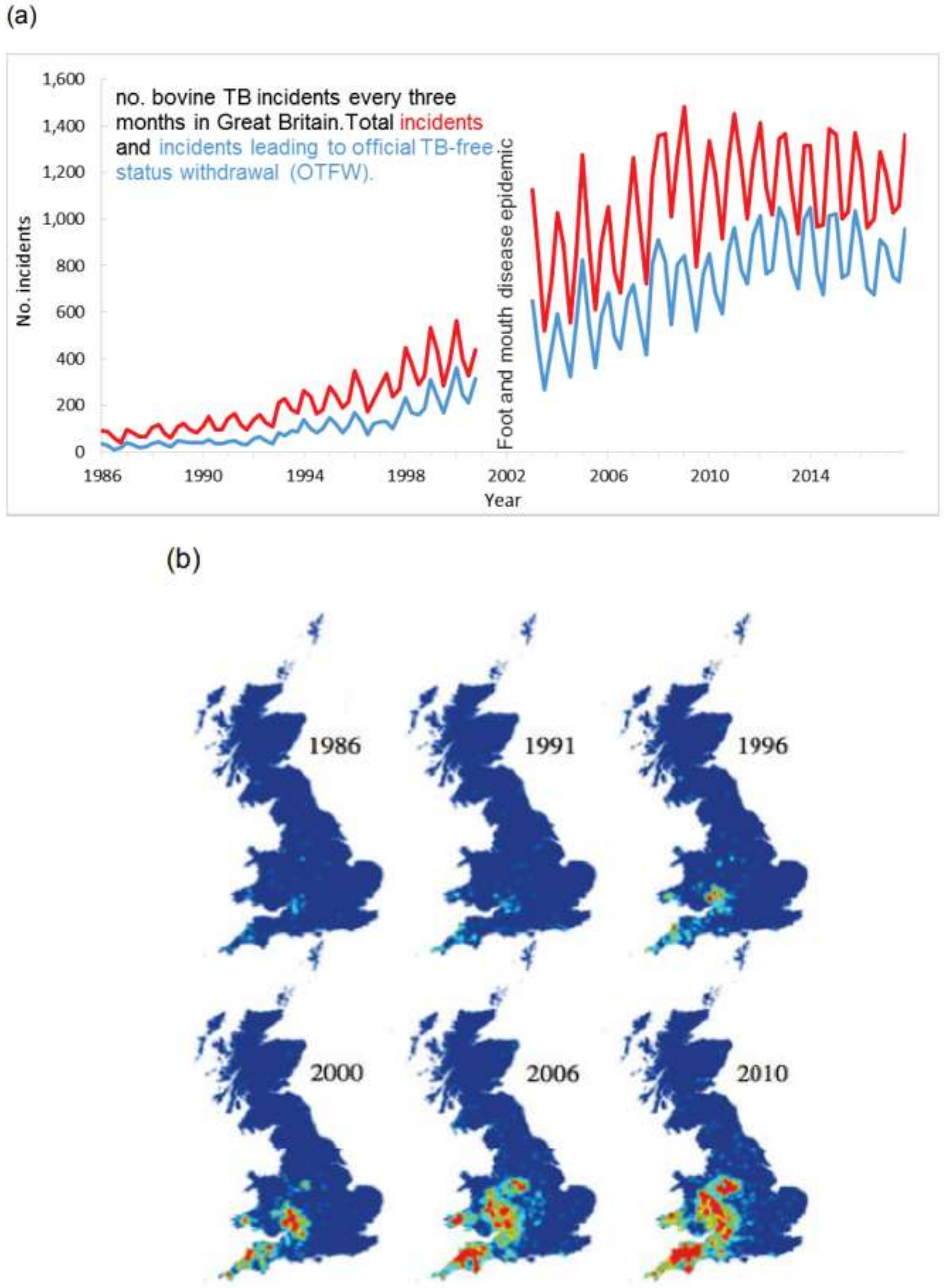


Figure 1.4. *Changes in incidence and distribution of bovine tuberculosis (bTB) in Great Britain from 1986 to 2012. (a) shows changes in incidence, which varies seasonally and demonstrates the interruption of bTB testing during the 2001 foot and mouth epidemic. (b) shows the increase in the geographical area affected by bTB from 1986 to 2010. Density of herds with officially tuberculosis free status withdrawn (OTF-W) status are demonstrated on a heat scale from blue (low incidence) to red (high incidence). Reproduced from the Bovine TB strategy review (Godfray et al., 2018).*

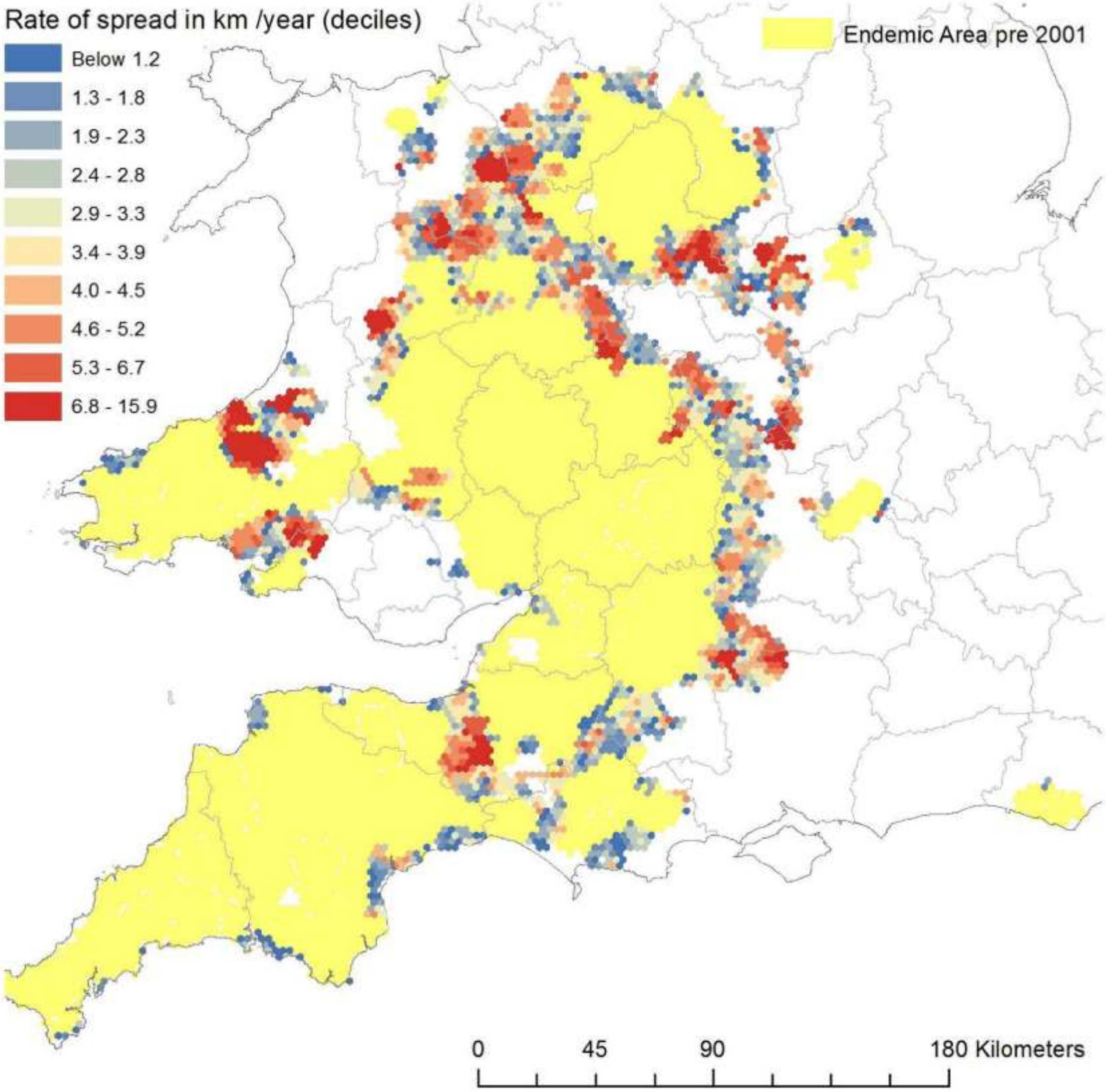


Figure 1.5. *Map illustrating the hexagonal areas to which endemic bTB (defined by the presence of three OTF-W incidents within 7 km) spread between 2001 and 2012. Rate of spread in km per year is calculated for the 2 year time period during which endemic bTB first passed through the hexagon. Reversion to non- endemic status may have occurred but was not calculated. Hexagons which were classed as endemic prior to 2001 are coloured yellow and are considered to be the 'core' endemic area.' Reproduced from Brunton et al. (2015).*

Routine surveillance of *M. bovis* infection is required under European Union (EU) law and the main form of disease control in the United Kingdom is cattle testing using the Single Intradermal Cervical Comparative Tuberculin (SICCT) test and culling of individuals that return positive test results. Frequency of SICCT testing differs by region and country, based on local incidence and epidemiology. Scotland is classified as Officially Tuberculosis Free due to very low incidence of infection. Scottish farms are tested every four years, but some farms are exempt, based on the likelihood of detecting cases in slaughterhouse surveillance (Kao, 2011). Herds in Wales have been tested annually since 2011. England is divided into three risk areas, herds in the Low Risk Area (LRA) are tested every 4 years and herds in the High Risk Area and the Edge area are tested at least annually (see Appendix C, Fig. C1 for map). The specificity of the SICCT test is high, yet sensitivity is estimated to be low (Conlan et al., 2012; Nuñez-Garcia et al., 2018) and in some cases, e.g. chronic incidents, the more-sensitive gamma interferon (γ-IFN) test is used (de la Rua-Domenech et al., 2006; Nuñez-Garcia et al., 2018). Since 2006, cattle in England and Wales over 42 days old must test negative to a SICCT test before movement from herds in an area that is tested annually, or more frequently (Animal and Plant Health Agency, 2018). In order to increase the test sensitivity associated with the movement of animals, post-movement testing of animals moved from herds in annual (or more frequent) testing areas to the Low Risk Area of England was introduced in England in 2016 and in Wales in 2017. Pre- and post-movement testing had been in place in Scotland since 2005 (Animal and Plant Health Agency, 2018b). OTF status is suspended (OTF-S) if at least one animal tests positive, a ‘reactor’, or if an ‘inconclusive reactor’ returns a subsequent inconclusive test result. OTF status is withdrawn (OTF-W) if *post-mortem* pathology characteristic of *M. bovis* is found, or if *M. bovis* is isolated by culture. A series of whole-herd tests are then scheduled (depending on the type of incident, risk area or region and herd history), which the herd must pass in order to have their OTF status reinstated. In general, no cattle movements on or off the farm are allowed during either type of incident (OTF-S and W), except movements off the farm directly to slaughter and some other exemptions under licences granted by the APHA. Post-mortem surveillance is performed at all GB slaughterhouses. If suspicious lesions are detected, in common with a SICCT test on farm, the OTF status of the source farm is suspended

(OTF-S) and a whole-herd test is triggered. OTF status is withdrawn (OTF-W) on isolation of *M. bovis* from the sample.

Modelling bovine tuberculosis

Since their implication as a reservoir host of *M. bovis* in the 1970s, the control of badger populations has been the focus of much attention. Working with limited knowledge, the modelling of bTB in badgers started with simple compartmental age-structured models of infection, which suggested that *M. bovis* was likely to persist in badger populations at a moderate prevalence (Anderson and Trewhella, 1985). As more information became available on badger ecology and infection, more complex spatial, stochastic models were developed, which showed no clear relationship between population reductions and prevalence of *M. bovis* infection (White and Harris, 1995). Comparisons between fertility control and lethal control considered that if culling caused the disruption of social groups, the success of lethal methods might be reduced (Swinton et al., 1997). It was clear that information to parameterise models was lacking (Smith, 2001), and in 1998, the Randomised Badger Culling Trial was set up to empirically assess the relative impacts of different approaches to badger culling on the local incidence of bTB in cattle; reactive culling, proactive culling (reduction of overall badger density in a specified area) or no culling of badgers (Donnelly et al., 2003). Reactive culling was stopped before the end of the trial due to a large increase in cattle bTB in treatment areas compared to no cull areas. By the end of the study, proactive badger culling reduced the risk of bTB in cattle in cull areas, but increased in surrounding areas. This finding was thought to be due to the perturbation of badger social groups creating increases in contact rate and subsequent rise in *M. bovis* infected badgers (Carter et al., 2007; McDonald et al., 2008). Overall, proactive culling areas were associated with a net increase in cattle bTB compared to areas where no culling was performed (Donnelly et al., 2003).

Within-herd models of *M. bovis* infection in cattle have aimed to estimate the transmission rate between cattle, the duration of time between infection and infectiousness, the performance of the SICCT test (as no gold standard is available for comparison), the relative force of infection from wildlife, and the effectiveness of differing control measures (Álvarez et al., 2014). Typically models of within-herd transmission of bTB insert 'Occult' and ‘Reactive’ states into the classic SI compartmental model framework to create a SORI models (Álvarez et al., 2014). There is no ‘Recovered’ compartment for *M. bovis* infection models as cattle testing positive to the infection are culled and thereby removed from the study system. These additional stages account for the variable incubation period of bTB where cattle are considered unreactive to currently available diagnostic tests (Occult) and the period of time for which cattle are likely to test positive on current diagnostic tests (Reactive). SORI models assume that during the O and R states the animal is not infectious to other cattle, however there are likely to be exceptions to this rule (Neill et al., 1992), which are described in SOR models, where states O and R are considered infectious.

Some models have acknowledged that differences in cattle herd types and sizes are likely to result in altered transmission dynamics. Conlan et al. (2012) ran SORI and SOR models with different herd sizes and found both showed substantial differences in basic reproduction ratios (SORI model: $n = 30$ $R_0 = 1.5$, $n = 400$ $R_0 = 4.9$) suggesting that there may be a trade-off between efficiencies associated with larger herds, and an increased risk of within-herd disease transmission. Higher transmission rates have been suggested for dairy herds due to more intensive management (Barlow et al., 1997). Studies distinguishing a difference between transmission in different herd types, or management, i.e. cattle at grazing or housed, found that transmission was five times lower between cattle outdoors than indoors in French herds (Bekara et al., 2014), and the cattle-to-cattle transmission rate was higher for dairy herds than beef herds in Spain (Álvarez et al., 2012). As milk yields are driven higher (AHDB Dairy, 2019a), more cattle are being housed indoors to meet rising nutritional demands (Charlton et al., 2011), and this could have implications for within-herd incidence of bTB. In their review of within-herd transmission of bTB, Álvarez et al. (2014) note that more empirical contact data on cattle farms would be useful to better inform transmission models such as these.

Single-host cattle and badger models have been developed into two-host models, which aim to understand transmission dynamics between the two species. A simple model by Cox et al. (2005) estimated the net R_0 of the bTB epidemic at approximately 1.1. More complex models have suggested a similar value for within-species R_0 as close to 1, and suggested that cattle measures were essential in the elimination of this infection (Brooks Pollock and Wood, 2015). They suggest the risks posed from badgers may vary with differences in other risk factors on farms and highlight that reducing the transmission from badgers to cattle would most likely be more effective than reducing the prevalence of *M. bovis* infection in badgers (Brooks Pollock and Wood, 2015).

It is clear from risk factor studies that bTB is a multifactorial disease; herd size, proximity to farm with a previous history of bTB, cattle movements were the most consistently identified risk factors in a review of herd-level risk factor studies (Skuce et al., 2012). The transmission of *M. bovis* in Great Britain occurs at multiple spatial scales; transmitted among cattle within a farm, between cattle and badgers and between contiguous farms at a local level, and among farms via cattle movements at a national level. Incorporating the interaction of infection at each of these scales makes the holistic modelling of the bTB epidemic very complex and computationally demanding. Brooks Pollock et al. (2014) constructed a dynamic stochastic model using approximate Bayesian computation that included within- and between-herd transmission of bTB based on national cattle movement and bTB testing data. Including the real animal movement network allowed farms to vary with respect to their contact rate with other farms and the mechanistic framework allowed the relative transmission pathways to be estimated. The model estimated that via their movements, some farms were responsible for many more infections than the majority of farms, acting effectively as ‘superspreader farms’, suggesting that if these farms could be identified, they might represent an opportunity for targeted surveillance and control strategies.

Thesis outline

The existence of farms clearly relies on their commercial viability, yet maximal productivity and efficiency can compromise biosecurity of farms and the health and welfare of animals. In this thesis, I examine features of the cattle industry at two spatial scales; between-farm trading networks and cattle-to-cattle interactions within farms and explore their relationship with animal disease, health, and welfare.

First, in **Chapter 2**, I explore how the concept of superspreaders might be applied to livestock premises and the spread of infection between farms. I look at the components of superspreading broken down into contact rate, infectiousness and duration of infection and discuss how these might contribute to a farm becoming a superspreader. I use bovine tuberculosis as the example to illustrate this concept.

In **Chapter 3**, I describe and quantify the network of cattle movements within Great Britain over time from 2001 to 2015. Using network analysis techniques, I construct networks of farm to farm trades and identify that, similar to markets, some farms can also act as hubs in the static movement network by virtue of the number of cattle movements on and off their premises. I utilise the concept of contact chains to explore how connected farms are within the network and reveal that farms are quickly connected to thousands of others within short time periods.

In **Chapter 4**, I test the hypothesis that the connections identified in contact chains are related to their risk of infection. Using the same approach as in **Chapter 3**, I calculate ingoing contact chains for all GB farms from 2012 to 2014, and additionally characterise the ‘source’ farms in the chain by individual farm bTB history and their bTB risk area for farms in England, or by country for farms in Scotland and Wales. I use a multivariable logistic regression model to evaluate the effect of the number and characteristics of source farms in ingoing contact chains on the odds of a bTB incident occurring on the root farm.

In **Chapter 5**, I change to considering social networks of animals as opposed to trading networks of farms. I use network analysis as a tool to describe the social interactions of cattle on seven commercial dairy farms and combine the use of high-resolution proximity and GPS data of individual cattle. I compare the interactions of cattle in buildings and at pasture and examine individual variation in terms of the number and duration of their contacts. I also investigate how the network characteristics change when different temporal sections of the data are analysed.

In **Chapter 6**, I investigate the relationship between cattle social interactions in terms of the time they spend with other cattle and milk yield and somatic cell counts, an indicator of udder health. I apply network analysis techniques to assess the extent of social preference in relation to characteristics of individual cattle.

Finally, in **Chapter 7**, my general discussion explores the themes of scale and temporal and spatial variation running through the thesis. I evaluate the usefulness of network analysis as a tool to describe infection transmission among cattle and among farms, and I show how my network analyses might be applied specifically to the challenge of bTB control in Great Britain, and more broadly to the management and control of other livestock diseases.

Chapter 2

Characterisation of potential superspreader farms for bovine tuberculosis

Chapter 2: Characterisation of potential superspreader farms for bovine tuberculosis

This chapter has been submitted to Veterinary Research as:

Fielding HR, McKinley TJ, Silk MJ, Delahay RJ, McDonald RA. Characterisation of potential superspreader farms for bovine tuberculosis.

Abstract

Variation in host attributes that influence their infectiousness and their contact behaviours can lead some individuals, known as ‘superspreaders’, to make disproportionate contributions to the spread of infections. Where heterogeneity exists in infection transmission, understanding the influence of individuals in an epidemic can be crucial in deciding where to direct surveillance and controls for greatest effect. In the epidemiology of bovine tuberculosis (bTB) in Great Britain, it has been suggested that a minority of farms, or herds, might act as superspreaders of *Mycobacterium bovis*, showing significant heterogeneity in their farm reproductive number - R_f. We identify characteristics of farms infected with *M. bovis* that potentially may lead to exceptional values in the three main components of R_f; contact rate, infectiousness, and duration of infectiousness, and thereby might characterise potential superspreader farms. Marked variation among farms in the scale of animal movements has been suggested as a driver of superspreading farms. We discuss this and consider the risk of those movements actually transmitting infection, governed by pathogen attributes, within-herd prevalence and characteristics of the animals being traded, which, in the case of bTB especially, may introduce further disparity in R_f. We emphasise the contributions of farm management and imperfect diagnostic testing to increased and persistent infectiousness, which may exacerbate the effects of other potential superspreader characteristics. Deployment of selective control methods on putative superspreader farms in Great Britain could yield the disproportionate disease control gains much needed in this costly epidemic.

Introduction

In disease systems, superspreading individuals are defined by their tendency to generate many more secondary infections than other hosts (Lloyd-Smith et al., 2005) and thereby to exert a disproportionately strong influence on disease dynamics. Epidemics in populations with superspreaders tend to be larger and more 'explosive' when outbreaks occur, yet exhibit greater chances of infection dying out when the more numerous individuals with relatively low transmission rates are infected (Garske and Rhodes, 2008; Lloyd-Smith et al., 2005). As the heterogeneity of individual reproductive number (R_i – the number of secondary infections created from a single infected individual in a completely susceptible population) increases, there is wider variation in potential epidemic size, but the usefulness of the population mean reproductive number (R_0) decreases (Garske and Rhodes, 2008). Hence, epidemics have, for some diseases, been more effectively modelled by incorporating variation in R_i rather than assuming that the host population is homogeneous with regard to R_0 (Lloyd-Smith et al., 2005; Stein, 2011).

In the epidemiology of diseases of livestock, epidemics are often modelled using the farm as the epidemiological unit that acquires and spreads infection. The individual farm reproductive number, R_f, is thus defined as the number of secondary farms infected by a primary infected farm (Mardones et al., 2011) and seems to show the same variation as R_i, with a minority of farms making a disproportionate contribution to secondary cases (VanderWaal et al., 2015; Woolhouse et al., 2005), apparently driven primarily by their trading behaviour (Woolhouse et al., 2005). R_f has been calculated in epidemics of foot and mouth disease (FMD; Tildesley and Keeling, 2009), highly pathogenic avian influenza (HPAI; te Beest et al., 2011) and salmon infectious anaemia (Mardones et al., 2011). In FMD and HPAI models, a reduction in epidemic size was achieved by targeting control measures on farms with higher R_f, showing the importance and benefit of identifying and targeting superspreader farms.

Whilst the impact of transmission heterogeneity has been evaluated for these highly-transmissible diseases, its role in endemic infections has been less well developed, despite literature suggesting that superspreader farms may be important in their dynamics (Brooks Pollock et al., 2014) and that long incubation or periods of latency may allow undetected infection to spread further between farms (Dubé et al., 2011). Brooks-Pollock et al. (2014) constructed a dynamic, stochastic, spatial model of bovine tuberculosis (bTB) in Britain, using farm movements from the Cattle Tracing System (CTS) and bTB testing results to fit the model. They suggested that just 10% of farms may be responsible for the majority of onward transmission to newly infected farms, implying that a disproportionate contribution from some superspreader farms may play an important role in this epidemic. BTB is caused by infection with *Mycobacterium bovis* and is an ongoing problem for the British cattle industry (Allen et al., 2018). Test and slaughter policies have, in the past, reduced incidence within cattle herds (Department for Environment Food and Rural Affairs, 2014), though since the 1980s increases in herd incidence have been accompanied by geographical spread from South-West England to areas in the North-West and the midlands of England and South and West Wales (Brunton et al., 2015). Control of bTB costs UK taxpayers about £100 million annually and the financial and emotional impact on farmers is substantial (Department for Environment Food and Rural Affairs, 2013). In 2017, 43,564 bovids (Department for Environment Food and Rural Affairs, 2018) and 19,274 European badgers *Meles meles* (Department for Environment Food and Rural Affairs, 2017a), which can constitute a wildlife reservoir of the infection (Godfray et al., 2013), were culled as part of bTB control measures in England (Animal and Plant Health Agency, 2017). Cattle testing, predominantly using the Single Intradermal Cervical Comparative Tuberculin (SICCT) test, is currently mandatory on an annual basis in Wales and in the bTB High Risk Area (HRA) and Edge Area of England with some counties within the Edge area and potentially soon the HRA, subject to six-monthly testing. Testing is required on a four-yearly basis in the Low Risk Area (LRA) of England and in Scotland, apart from some exempt farms in Scotland, where the country is classed as Officially Tuberculosis Free (OTF). A positive reaction to the SICCT test, or lesions consistent with bTB found at slaughter, triggers movement restrictions on the affected farm. Initially, their individual farm OTF status is suspended (OTF-S) and then withdrawn (OTF-W), upon detection of *M. bovis* in a culture sample. Brooks Pollock et al. (2014) suggest that the majority of new infections are caused firstly by movements, and secondly through the local environment. They report a skewed distribution such that a small number of farms create a disproportionate number of cases, however, they do not elaborate further on what might characterise these particular farms.

As bTB is a chronic, notifiable disease with mandatory control measures designed to find and eliminate disease, it might seem unusual to consider the existence of superspreader farms in this context. BTB superspreader farms are likely to present differently to superspreaders of more highly transmissible diseases and could transmit infections over long periods of time, in contrast to those which cause a steep rise in the epidemic curve. We should consider that we are not discussing commonalities but those circumstances that occur rarely, concern a small number of farms and that might manage to evade current control policies in the ways we will outline in this review. Much bTB research concerns herds that have had a bTB incident (and are therefore under movement restrictions), however we focus on those herds that may be infected with bTB but where infection has not been disclosed by testing and so are able to sell potentially infected cattle. The mechanisms by which farms might function as bTB superspreaders in terms of their infectiousness, are precisely the attributes which will make them difficult to identify.

The transmission rate of an infection is governed by three components: contact rate, infectiousness, and duration of infectiousness (Fig. 2.1; VanderWaal and Ezenwa, 2016), offering a useful framework for discussion of how, by extreme traits in any or all of these three components, a superspreader might act. We first consider variation in contact rates among farms that arises from heterogeneity in the scale of cattle movements, both directly and as part of large and complex trading networks. Second, we study how the characteristics of *M. bovis* infection and of individual farms affect the risk of those movements transmitting infection, effectively governing the 'infectiousness' of farms. Third, we look at factors specific to bTB, such as imperfect diagnostic testing (de la Rua-Domenech et al., 2006), the common absence of clinical symptoms in infected animals in Great Britain (Neill et al., 2001), and environmental sources of infection, including wildlife, and how these might contribute disproportionately to the prolonged duration of bTB infection on some farms. We explore how their characteristics may differ from those of a 'classic' superspreader and suggest how this makes them problematic to identify.

Finally, we discuss what control options might be particularly appropriate for putative superspreader farms, within current bTB control policies in Great Britain.

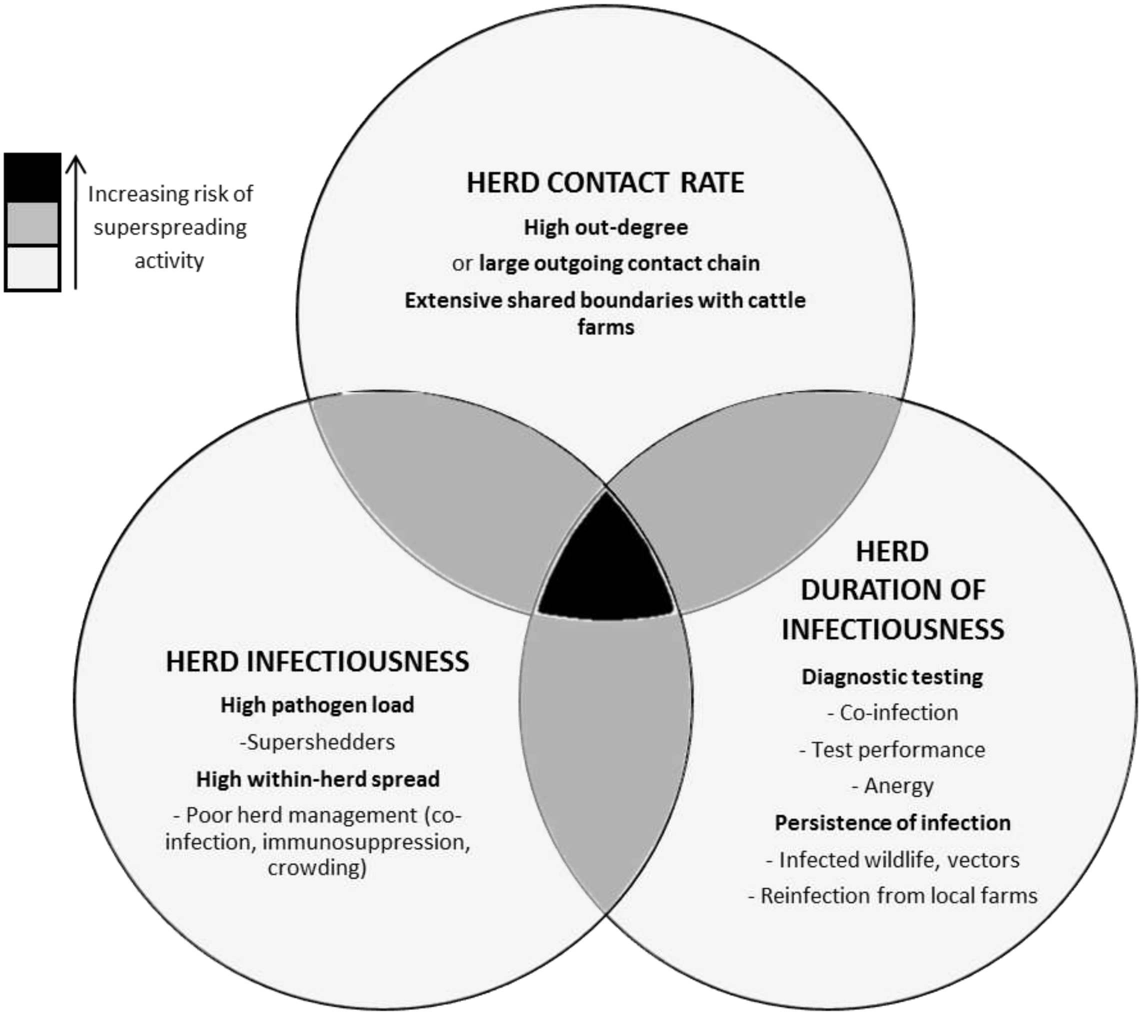


Figure 2.1. *What makes a superspreader farm? Venn diagram showing factors involved in extreme components of R_f and the increasing risk of superspreading activity when these factors are combined. These factors are not mutually exclusive and interactions occur between these components, for example the product of infectiousness and contact rate are typically combined to describe the 'transmission rate' and factors affecting both infectiousness and the duration of infectiousness such as immunity and co-infection are common.*

Herd contact rate

The buying and selling of cattle and their movements among farms are the most obvious and comprehensively-recorded type of interaction among cattle herds, and constitute a major potential mechanism for pathogen transmission among farms (Gates, 2014). Heterogeneity in these movements (VanderWaal et al., 2015) is likely to be a driving force behind variation in R_f (Woolhouse et al., 2005) and quantification of this trait has been effective in identifying potential superspreaders. Whilst the definition of a superspreader is specifically related to a tendency to seed a disproportionate number of secondary infections, achieved in this scenario via selling livestock to an comparatively high number of farms; here we also discuss trading as having an impact on the potential for a superspreader farm to become infected, e.g. increased exposure to infection through purchasing livestock from multiple sources. This could enable a farm to either exert its superspreader effect via disproportionate outward connections or seed disproportionate numbers of secondary infections cumulatively by virtue of being infected or continually re-infected for a long duration of time. Multiple analyses of farm trading networks have found a power-law distribution (Clauset et al., 2009) for in-degree (the number of farms animals are purchased from), out-degree (the number of farms animals are sold to), and overall degree (the sum of them both) in livestock movement networks, characterised by many premises with few contacts and a few premises with a disproportionately high number of contacts (Dutta et al., 2014; Mweu et al., 2013; Rautureau et al., 2011). These directed centrality measures can be used as proxies for a farm's ability to acquire (in-degree) and transmit (out-degree) disease (Dubé et al., 2009). For example, the in-degree of a farm was found to relate to herd seroprevalence of bovine coronavirus in Sweden (Frössling et al., 2012), and bTB in East Africa (Sintayehu et al., 2017) and many studies have shown that individuals in highly connected positions in the contact network are more likely to be infected (Corner et al., 2003; Godfrey et al., 2009). However there have also been studies involving *M. bovis* which suggest the opposite effect (Drewe, 2010; Weber et al., 2013b), revealing the complex nature of this particular infection. Second order connections (the contacts of contacts) should also be examined as they can influence the role an individual might play in disease spread; for instance in human sexual contact networks, the risk of *acquiring* human immunodeficiency virus infection was better predicted by the behaviours of partners of partners than by the individual's first order partners alone, and consequently was also preferable for estimating *onward* transmission (Ghani and Garnett, 2000). Eigenvector centrality, a measure that considers both direct and second order contacts, was helpful in determining how influential a farm might be in the spread or maintenance of a theoretical, highly infectious epidemic in Italian cattle movement networks (Natale et al., 2009). In a study of FMD outbreaks the reproduction number of second order contacts provided good predictions of epidemic size and, when combined with R_f, provided good estimation of heterogeneities in dynamics of FMD outbreaks (Tildesley and Keeling, 2009). Furthermore, Fu et al. (2015) found that combining global clustering and centrality metrics of simulated epidemics, with node centrality measures performed well in detecting superspreading nodes. The measures we have mentioned thus far however, only analyse static networks, which do not consider the temporal sequence of events. The analysis of dynamic networks is less well developed as it is methodologically and computationally more complex, however in some cases it is crucial to our understanding of how a pathogen might be transmitted through a network (Enright and Kao, 2018), allowing the possibility of transmission between nodes only when the links between them are present. The calculation of time-ordered paths or 'infection chains' is one such technique applied in movement networks which respects temporal order and gives an indication of the influence of individual nodes. The 'ingoing infection chain' is the network of farms connected to a farm as a result of movements onto that farm (Nöremark et al., 2011), and represents the possible sources that may have contributed to acquiring infection during a specified period. Typically these chains are positively skewed with many farms having small contact chains but some having very extensive chains (Fielding et al., 2019; Nöremark et al., 2011), similar to the pattern found for direct contacts on static networks. These very extensive chains of farms aggregated over a period of 5 years have been associated with increased risk of *M. bovis* infection in French cattle herds (Palisson et al., 2016), showing that chain magnitude may be useful in the predicting which farms might be more at risk of infection (by their ingoing infection chain) and those that might be more able to spread chronic infections (via their outgoing contact chain). In choosing the time period over which to study the network, the independent timescales of the movement network and the pathogen should be considered (Kao et al., 2007). If bTB spreads very slowly in comparison to a quickly evolving network, it may be that the contact networks we need to select become so saturated that they actually approximate towards homogenous mixing (Enright and Kao, 2018), making it more difficult to identify potential superspreader farms. However, as advances in computational power and dynamic network theory are made, it may become more feasible to interrogate these large networks. For bTB, there is a significant amount of data available on testing results and cattle movements, however we are constrained by unknown pathogen characteristics such as incubation, infectious and latent periods. The use of approximate Bayesian computation techniques (Kosmala et al., 2016; McKinley et al., 2009) and dynamic network models (Silk et al., 2017b) are likely to be useful in tackling these modelling challenges.

Herd infectiousness

While the number of movements and trading partners will undoubtedly be a principal driver to increase the influence of certain premises, farm and pathogen factors govern whether those movements actually transmit infection, i.e. the infectiousness of the farm. For highly transmissible infections with high within-herd prevalence, it is more likely that any movement would transmit infection regardless of the farm characteristics, however, where the disease spreads slowly within a herd, as is typically the case with bTB, the risk of selling an infected animal is more variable and farm factors become more influential. We now discuss farm factors, such as disease prevalence, herd immunity, presence of multiple infections, and the type of animals being sold, that can all influence the chances of selling infected animals.

Supershedders, highly infectious individuals releasing more infectious agents than others in their group (Chase-Topping et al., 2008), can increase herd prevalence. Their presence in the herd also reduces the efficacy of whole-herd control measures, due to heterogeneity generated in transmission rates (Lanzas et al., 2008). Supershedding can be driven by genotype, behaviour, signalment (age, sex, and breed; Craft, 2015), co-infection (Lass et al., 2013), immunosuppression (Stein, 2011), enhanced susceptibility, or strain pathogenicity of the infecting agent (Matthews et al., 2009). Heterogeneity in bacterial shedding has been found in cattle infected with *Escherichia coli* O157 (Chase-

Topping et al., 2008), *Salmonella enterica* (Lanzas et al., 2008), and *Mycobacterium avium* subspecies *paratuberculosis* (MAP; Pradhan et al., 2011). Supershedders of MAP, the causative agent of Johne's disease in cattle, are suggested to cause passive infections in other cattle within the herd, which are then able to shed and spread the bacteria, but remain tissue culture negative (Pradhan et al., 2011). Stress from movements and from weaning has been implicated as a risk factor for supershedding of *E. coli* O157 (Chase-Topping et al., 2007). Supershedders of bTB have been identified in red deer and badgers (Santos et al., 2015; Wilkinson et al., 2000). In cattle, the most likely route to being a supershedder of bTB, i.e. an animal excreting more pathogen than others, would be one with late-stage undetected infection (Houlihan et al., 2008), which we discuss below when considering diagnostic tests.

The risk of movements transmitting infection depends on the the type of animal being traded. Breeding cattle seem to be associated with a higher risk of various infections than other types of cattle. Gates et al. (2014) found that the purchase of cows that had calved made a disproportionate contribution to Bovine Viral Diarrhoea (BVD) infection risk on that farm. A risk factor analysis showed that the presence of supershedders of *E. coli* was more likely if the farm bought in female breeding cattle (Chase-Topping et al., 2007). In contrast, incidents of bTB were more likely if a herd purchased a bull (Griffin et al., 1993). The trading of male or female breeding cattle might present a relatively greater risk of infection transmission.

At times, herd make-up and farm practices may interact to drive variation in herd infectiousness. Some fattening herds may be more at risk of higher infectiousness, and therefore be more likely act as superspreaders, by the nature of their practices. Those herds rearing animals to sell directly to slaughter do not pose a disease transmission risk, however it is commonplace in Great Britain for animals to be sold as 'stores' (animals reared for beef but not ready for slaughter) and passed between multiple farms (Robinson and Christley, 2007). This type of herd tend to house cattle from different farms (usually purchased via markets (Robinson and Christley, 2007). Mixing of cattle from many source farms can have physiological effects that may increase susceptibility to infections (Proudfoot et al., 2012). Combined with exposure to a wide range of pathogens, this can often facilitate within-herd transmission of multiple infections (Griffin et al., 2010). Subsequently, co-infection can alter host immune responses and increase pathogen shedding (Lanzas et al., 2008), additionally increasing herd infectiousness. Vaccination, diagnostic testing, good management, and sourcing of animals from fewer, disease-free farms may all reduce this risk, however, for diseases such as bTB where vaccination is not available and tests have low sensitivity, these risks are more difficult to manage.

A final farm-level characteristic that will influence both herd infectiousness and its duration is husbandry and hygiene. Farm conditions can potentially lower resistance to infection and hence increase herd infectiousness. Proudfoot et al. (2012) describe the direct impact of the physical environment on the individual (poor hygiene, exposure to multiple pathogens, injury) which increases the risk of disease. They also note the indirect impact of social stressors (overcrowding, mixing of groups, isolation) on host physiology (immunomodulation, low resilience, chronic inflammation), which can then increase risk of infection and disease progression. Winter housing of cattle can cause social stress due to mixing of groups and crowding and, where there is also poor ventilation, transmission of airborne pathogens can increase (Gorden and Plummer, 2010). Increased seroconversion of dairy cattle to bovine Herpesvirus 1 has been associated with winter housing (Woodbine et al., 2009) and there is some evidence for housing as a risk factor for bTB transmission (Vial et al., 2015). Although most bTB incidents now have very few reactors (Animal and Plant Health Agency, 2017), suggesting low within-herd transmission, the sharing of a confined, poorly ventilated environment by supershedders or many high-risk cattle and susceptibles may be sufficient to cause a superspreading event within a farm (Lloyd-Smith et al., 2005), and thus increase herd infectiousness.

Duration of infectiousness

Prolonged infectiousness of a herd, through misdiagnosis (of novel infections, or rare infections where more common infections are assumed at first), undetected infections (if asymptomatic; Drosten et al., 2003), or poor test sensitivity, can facilitate the spread of disease. In the case of bTB infection in Great Britain, most infections do not present with clinical signs and the variable sensitivity of routine SICCT testing can miss infected animals. Combined, these factors can lead to farms being infectious for long periods of time, giving these farms more opportunity to sell on infected animals.

The infectious period of livestock can be managed by treatment, vaccination, or culling (Thurmond, 2003). Decisions on whether and how to apply these control methods are generally based on results of diagnostic testing. However, tests for *M. bovis* often have low sensitivity, which can be markedly lower in certain circumstances, allowing some herds to have an extended infectious period and thereby be able to act as superspreaders.

Co-infection with pathogens such as bacteria, helminths and viruses may alter the host response to infection and host infectiousness (Lass et al., 2013) but can also alter the non-specific immune responses measured by diagnostic tests. In cattle co-infected with liver fluke *Fasciola hepatica* and *M. bovis*, shifts in immunity from T-helper cell 1 to T-helper cell 2 responses have been implicated in reducing the immune response to the tuberculins used in SICCT testing in England and Wales (Claridge et al., 2012). However, another study in Northern Ireland found no significant relationship between liver fluke and response to the SICCT test (Byrne et al., 2018), and therefore further research is warranted. MAP is a chronic enteric infection affecting many cattle in Great Britain (Animal and Plant Health Agency, 2015a). There have been several experimental studies indicating that co-infection with bacteria in the *M. avium* complex reduces the sensitivity of SICCT and gamma interferon (IFN-γ) testing, through cross reaction of antigens and an increase in response to the comparative avian tuberculin injection (Álvarez et al., 2009). Evidence of further progression of *M. bovis* infection was found in cattle positive to both the *M. bovis* and *M. avium* purified protein derivative injected at SICCT testing than those with only an *M. bovis* reaction (Byrne et al., 2018). Questionnaire data suggested that dairy farms that had experienced MAP infection in the last 12 months were 4.7 times more likely to have a bTB incident (Broughan et al., 2016b). Bovine Viral Diarrhoea Virus (BVD) is widespread in England and Wales (Charleston et al., 2001) and immunosuppression in acute viral infection leaves animals susceptible to concurrent infections. Animals infected with *M. bovis* and acute BVD infection showed suppression of IFN-γ production when stimulated with tuberculin (Charleston et al., 2001) that has been associated with a particularly severe outbreak of bTB in a group of calves (Monies, 2000). However, recent studies in Northern Ireland have found no positive association between BVD infection and bTB infection at an individual or herd level (Byrne et al., 2018, 2017). In summary, co-infection with certain pathogens can lead to changes in the performance of diagnostic tests which reduce their chance of detecting or confirming infection, and subsequently may leave a reservoir of infection in the herd. Herds with high prevalences of these diseases have a higher chance of prolonged infectiousness increasing the risk of them becoming superspreader farms.

Immune responses change during the course of disease progression and the host's lifetime and diagnostic tests can therefore vary in their performance, depending on when the host is tested (Schukken et al., 2015). For some infections, longitudinal testing is required to increase test performance where single test results are not sufficiently robust (Schukken et al., 2015). As *M. bovis* infection in cattle progresses, the initial cell-mediated immune response wanes and some infected animals can become unresponsive or 'anergic' to SICCT testing (Neill et al., 2001). Undetected by routine tests, these animals remain in the herd, and over time may develop severe lesions and the capacity to disseminate infection in the herd, acting as supershedders (Houlihan et al., 2008). Annual SICCT testing in areas of higher bTB risk means that cattle should typically be detected before the natural cell-mediated response wanes, so the number of naturally-anergic cattle in Great Britain is likely to be low. However, their potential to persist undetected in an infectious state may be epidemiologically significant. In addition to the natural progression of infection, the sensitivity of the test may decrease over time due to repeated exposure to tuberculin (Coad et al., 2010), for instance in prolonged bTB incidents where herds are SICCT tested at 60 day intervals. Temporary anergy to the SICCT test can also develop in periods of stress, around parturition (Li, 2016), and when corticosteroids are administered (Phillips et al., 2002), allowing evasion of diagnosis if testing is performed at this time. Some animals not fully anergic may exhibit a partial cell-mediated response to the SICCT test and appear as inconclusive reactors (IRs). These animals are tested 60 days later and if they test negative (as one would expect of the temporarily anergic animals) can remain in the herd. However, an Irish study found that IRs which retest negative after 60 days and remain in the herd have 12 times greater risk of testing positive at the next routine SICCT test or at slaughter (Clegg et al., 2011). This suggests that they are false negatives at retesting, perhaps due to co-infection, anergy, or test sensitivity, and therefore pose a prolonged risk of infection spread within the herd. In England, as of 2017, these animals must not be moved off the farm to reduce the chance of further disease transmission between-herds (Department for Environment Food and Rural Affairs, 2017b), however the risk of within-herd transmission remains. Prior to 2018, the SICCT test and the IFN-γ test were the only ante-mortem tests approved to diagnose bTB in British cattle, both detecting cell-mediated immunity, however changes to GB legislation now allows the exceptional use of a non-validated test, which detects antibodies to *M. bovis*, if they are present in these anergic animals (Animal and Plant Health Agency, 2018c).

Although local transmission is not necessarily a superspreading trait in isolation, its combination with a high contact rate or high infectiousness may enable a superspreader farm to exist and focusing on single species or single transmission paths is unlikely to fully capture infection dynamics for diseases with more than one route of transmission or multiple hosts (Brooks Pollock et al., 2015). Farm-to-farm contact at boundaries, transmission from wildlife, fomites, and vectors are important sources of infection that can contribute to the persistence of infection on a farm (Sibley, 2010), in addition to livestock movements and undetected infections. Local farm density, proximity to an infected farm, and presence of biosecure boundaries have been used to estimate local disease spread in FMD epidemics (Keeling et al., 2001). Brennan et al. (2008) studied the contacts of cattle farms in North-West England in respect to contractors and companies, shared equipment, and employees. They found that the frequency of contacts in this network exhibited the same heterogeneity as found in animal movements. For bTB specifically, local spread between neighbouring cattle or wildlife is considered important (Brooks Pollock et al., 2014). A study of *M. bovis* transmission in France, where the infection is rare, combined the cattle movement network with a ‘spatial neighbourhood’ based on geographic proximity of farms (Palisson et al., 2016). They estimated that 73% of infection (the population attributable fraction) could be removed if local transmission was eliminated. Spatial clustering analysis of bTB data from England in 2005 showed only weak evidence for clustering of disease on a county level (Green and Cornell, 2005).

However, herd-level risk factor studies have found that risks of bTB are greater for farms whose neighbours have a history of bTB (Gardy et al., 2011). Badgers infected with *M. bovis* represent a potential source of transmission to cattle, most likely via indirect contact at latrines and via contamination of pasture and feed (Woodroffe et al., 2016). The persistence of infection in local badgers may therefore facilitate persistence in cattle, particularly where transmission occurs in both directions, thus creating a cycle of reinfection. In addition, the longer a pathogen is able to survive in the environment, the more opportunity there is for maintaining infection within the herd. If present, *M. bovis* can survive in infected cattle faeces in slurry for up to 6 months (Scanlon and Quinn, 2000) and on pasture for 1-6 months (depending on season; Williams and Hoy, 1930). The application of slurry from infected farms may present a risk of bTB persistence on infected farms, and could be a source of infection on uninfected farms through aerosolisation of bacteria when slurry is applied to nearby fields or use of slurry contractors where equipment is not disinfected between farms.

Identification of superspreader farms and targeted control - where should we direct effort for most efficient and most effective controls?

In addition to one of these extreme components of R_f, a superspreader must also be infected and have some contact with other hosts. For example, if infection is removed from a farm with a high contact rate, or if a highly infectious farm no longer contacts other farms (e.g. movement restrictions have been applied), they can no longer function as superspreader farms, however they might still be considered to hold superspreading potential. It might still be prudent to target these potential superspreader farms for additional surveillance if the uninfected farm is likely to become infected (i.e. large herd, higher-risk area or previous bTB history) or be harbouring undetected infection, or if the isolated farm is likely to make connections with other farms. Control measures should suit the specific characteristics of each farm and address the potentially transient nature of this phenomenon. The three components of R_f can combine to increase the risk of superspreading occuring (Fig. 2.1), and therefore increase the impact of an individual farm on disease dynamics. In this section we look at which existing control measures and which novel approaches might be used to target superspreader farms, should they be deemed influential in the dynamics of the bTB epidemic.

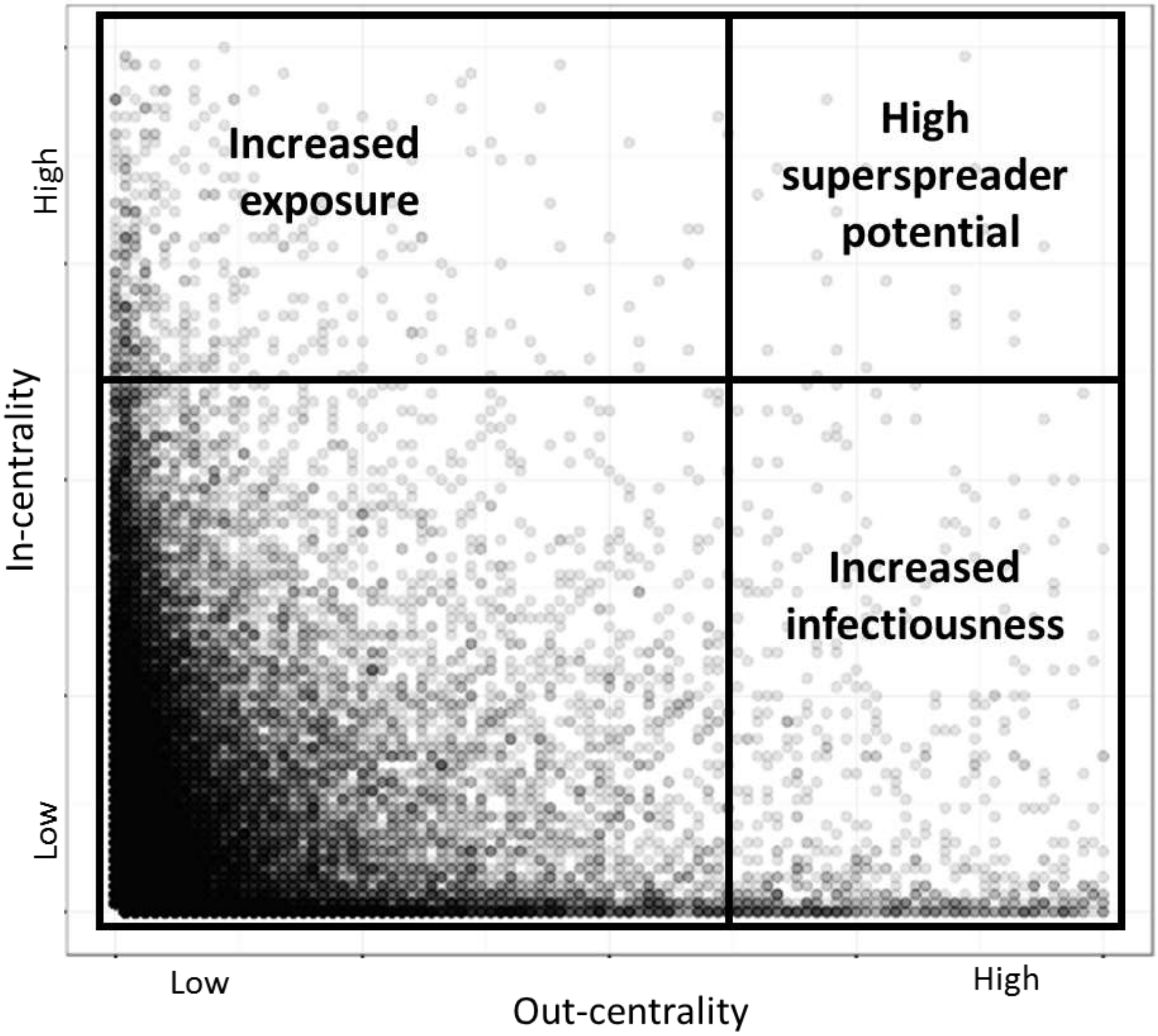


Figure 2.2. *Schematic to highlight the potential for superspreader behaviour associated with farms in animal movement networks having high in-centrality and/or high out-centrality and how they might combine to result in high potential for a farm to act as a superspreader. Points represent farms and the shading density increases to represent more farms at an individual point.*

The identification of superspreaders can provide an opportunity to focus or intensify control measures such as treatments, vaccination, isolation, restrictions, to gain disproportionate benefits. Lloyd-Smith et al. (2005) calculated that where half of all control effort is focused on the 20% of population responsible for the majority of disease transmission, it is up to three times more effective than random control. Livestock markets and some farm premises have a high throughput of animals, termed ‘hubs’ in networks analysis (Robinson and Christley, 2007), and are especially important in governing the size of epidemics of highly transmissible infections. Emergency disease control measures targeted at these hubs have been effective in limiting epidemic size in FMD and HPAI outbreaks (Green et al., 2006; Molia et al., 2016). Büttner et al. (2013) found in pig movement networks that a 75% reduction in giant component size in the network, which can be used as a proxy to measure the extent of epidemics, was achieved by removing the smallest proportion of farms (1.4%) when based on the number of farms to which they sold animals (out-degree). The second smallest proportion of farms (1.5%) to be removed was achieved by using the outgoing contact chain. Models based on disease transmission through cattle movement networks have shown that removal of the 20% of farms contributing most to R_0 resulted in a 97-99% reduction of R_0 (Volkova et al., 2010a). These studies have targeted the upper distribution of individual reproductive numbers or out-degree to effectively reduce estimates of final epidemic size, showing that focusing resources on these relatively small but highly influential groups can be more effective than random population-wide control.

Ideally livestock movement restrictions should facilitate a reduction in high-risk trading for an appropriate time period, whilst maintaining the ability to move low-risk stock onto and off the farm where necessary. More bespoke restrictions between trading partners are discussed by Gates and Woolhouse (2015), where farms with high in-degree (mixing of animals from multiple sources) are prevented from selling to farms with a high out-degree, thereby reducing the inherent risk associated with this behaviour. These trading restrictions decrease the likelihood of similar high-risk trade paths reconfiguring and allow the farm other outlets for trade to sustain their business. If farms are completely removed from a network, the farms connected to them tend to find new partners with which to trade (Brouwer et al., 2012). However, these may have a similar risk of disease transmission as the original partners, creating a new structure as risky as the old one (Brouwer et al., 2012). For bTB control, movement restrictions are placed on the individual farm when a reactor is found, and are later lifted conditional on two consecutive clear SICCT tests. However, undetected infection on the farm (Conlan et al., 2012) may mean these restrictions are lifted prematurely. Extended movement restrictions on those farms with recurrent bTB incidents, high out-degree or an exceptionally large outgoing contact chain, might provide an effective, risk-based addition to the current reactive approach. Limiting sales from high-risk farms to only approved finishing units (and to slaughter thereafter) or direct to slaughterhouses may be effective in limiting the spread of infection, as is currently allowed for some farms under bTB restrictions (Animal and Plant Health Agency, 2014b).

To discourage trading from higher to lower-risk farms, herds can be given a risk score based on various predictive measures. Farms in New Zealand have a designated bTB status score from 1 to 10, denoted by the number of years since the last bTB incident, that defaults to 10 if they have never had a bTB incident. Their score is, however, superceded by the lowest score of a farm with which they trade (Enticott, 2014), thereby encouraging farmers to trade with ‘less risky’ farms. In Britain, Adkin et al. (2015) developed a scoring system based on the previous bTB history of the farm, movements from higher risk areas, local bTB prevalence and herd size, to inform risk-based purchasing and give farmers the ability to make informed decisions. Although this specific risk-score has not yet been used in practice to any significant extent, the Cattle Health Certification Standards (CHeCS) scheme have launched a fee-driven, voluntary bTB accreditation-scheme based similarly on years free of bTB incidents, but not influenced by trading history. These schemes attempt to lower the risk of purchasing cattle, yet allow the farm as a business to continue. However, the success of such voluntary schemes is clearly dependent on industry uptake, which up to 2018 in Britain had been very low.

Surveillance is crucial to controlling infection within and between-herds. Targeting existing surveillance efforts at farms with the potential to become superspreaders, by means of their high connectivity, could therefore potentially reduce their impact. Frössling et al. (2014) found that using in-degree and ingoing infection chains to target surveillance detected more positives than random approaches. For bTB, it is crucial improve detection of infected animals on potential superspreader farms. Increasing the frequency of routine testing on high-risk superspreader herds, reduces the time to detection and removal or treatment of animals, and so can reduce the duration and intensity of infectiousness at the farm scale. The sensitivity of routine testing can also be increased by using the severe interpretation of the SICCT test (decreasing the cut-off criteria which defines reactors), a non-comparative cervical Single Intradermal Test (see (de la Rua-Domenech et al., 2006), or the IFN-γ test. However, increased sensitivity is typically associated with a loss of specificity, which if applied across a very large population would lead to unacceptably high numbers of false positive reactors being culled and the imposition of unnecessary restrictions upon farm business. The IFN-γ test is used in addition to the SICCT when specific criteria are met and the test can be particularly useful in detecting bTB-positive animals that have become desensitised to the SICCT test as a result of repeated testing (Coad et al., 2010), those co-infected with other *Mycobacterium* spp., and animals in early stages of infection (de la Rua-Domenech et al., 2006). Use of more specific antigens such as ESAT-6/CFP10 in the IFN-γ test may offer additional diagnostic power in herds with animals vaccinated with Bacillus Calmette-Guerin (BCG; van Pinxteren et al., 2000) and/or known co-infection with other mycobacteria (Aagaard et al., 2010), although sensitivity is unknown in this latter group. Targeted use of more sensitive and less specific measures or combined testing protocols, on superspreader farms, or farms which hold potential to be superspreaders would minimise the impact on the wider industry whilst potentially maximising disease control benefits.

Vaccination can be effective in limiting spread within a homogenous population. However, as heterogeneity of R_i increases, vaccine efficacy must be higher to achieve the same level of control (Lanzas et al., 2008). Therefore, preferential targeting of potential superspreader farms for vaccination might be better directed towards those farms which 'superspread' via higher contact rates rather than higher infectiousness. In models where 80% of individuals with high contact rates were vaccinated against influenza, there was a 91% disease reduction predicted for the whole population (Weycker et al., 2005). The use of BCG vaccination for control of bTB in cattle is currently prohibited under EU law (European Economic Community, 1977). Variable efficacies have been reported for the use of BCG vaccine in cattle (Waters et al., 2012) and its use in combination with the existing test and slaughter policy requires a test that differentiates infected from vaccinated animals (DIVA; Conlan et al., 2015). Although DIVA testing is being developed (Swift et al., 2016), there are concerns regarding low specificity (Conlan et al., 2015), economic viability, and the practicalities and regulation of field trials to prove reduction of transmission (Conlan et al., 2018). The use of vaccination against *M. bovis* in UK cattle remains speculative, yet if it became available some superspreader farms may represent effective targets for vaccination to reduce disease spread.

Conclusion

There are mechanisms, within current farm practices and the current bTB controls, which could feasibly allow the existence of bTB superspreader farms. Farms with influential roles in cattle movement networks are able to sell animals to many different premises. The risk of these animals being infected might be increased by the trade of certain animals and where farm management favours within-herd spread. The opportunity to sell infected animals increases with the duration of infection on the farm and this can be increased by factors which impair the sensitivity of routine testing and continual reinfection from within the herd, or local sources. In a minority of cases, we expect that a high contact rate might be combined with high infectiousness to create a superspreader farm. The challenge will be to identify these disproportionately important farms in ‘real time’. Where the three aspects of superspreading vary over time, superspreading may be a transient phenomenon. It is important that the additional restrictions placed on these farms occur only while they are at a high risk of transmitting disease, and ideally in a way that minimises impact on individual farm businesses. We have given examples where some network characteristics might be used to identify potential superspreaders. However, determining the most salient measures appropriate to a chronic endemic disease, such as bTB, evaluation of their ability to predict future behaviour will require further analysis.

It is probable that the operators of farms that may be exhibiting superspreader characteristics are unaware of their potential wider impact on bTB disease dynamics. Further research may reveal whether identification of these farms can encourage such farmers to adapt their practices and mitigate potential risks for the benefit of the majority. The categorization of some farms into this higher-risk bracket is intended to create an opportunity for greater resources to be directed at these farms. It is not intended to ‘remove’ or ‘eliminate’ these farms (or their practices) as to do so could represent a gross misunderstanding of both disease dynamics and of the cattle industry, and may not deliver the desired results in the long term. The aim of identifying superspreaders is to better understand the mechanisms by which they might operate and to adapt control methods to exploit their influential role in disease transmission, thereby enhancing control of the epidemic.

Chapter 3

Contact chains of cattle farms in Great Britain

Chapter 3: Contact chains of cattle farms in Great Britain

This chapter has been published in full as:

Fielding HR, McKinley TJ, Silk MJ, et al. (2019) Contact chains of cattle farms in Great Britain. Royal Society Open Science 6:180719.

Abstract

Network analyses can assist in predicting the course of epidemics. Time-directed paths or ‘contact chains’ provide a measure of host-connectedness across specified time-frames, and so represent potential pathways for spread of infections with different epidemiological characteristics. We analysed networks and contact chains of cattle farms in Great Britain using Cattle Tracing System data from 2001–2015. We focused on the potential for between-farm transmission of bovine tuberculosis, a chronic infection with potential for hidden spread through the network. Networks were characterised by scale-free type properties, where individual farms were found to be influential ‘hubs’ in the network. We found a markedly bimodal distribution of farms with either small or very large ingoing and outgoing contact chains (ICCs and OCCs). As a result of their cattle purchases within 12 month periods, 47% of British farms were connected by ICCs to >1000 other farms and 16% were connected to >10,000 other farms. As a result of their cattle sales within 12 month periods, 66% of farms had OCCs that reached >1000 other farms and 15% reached >10,000 other farms. Over 19,000 farms had both ICCs and OCCs reaching >10,000 farms for two or more years. While farms with more contacts in their ICCs *or* OCCs might play an important role in disease spread, farms with extensive ICCs *and* OCCs might be particularly important by being at higher risk of both acquiring and disseminating infections.

Introduction

Pathogen transmission among hosts may occur by a variety of routes, from different types of direct contact, to indirect contact via vectors, fomites and the environment (Warren et al., 1982). For livestock, animal movements between farms can be considered to form a directional link from the source to the destination farm, which may therefore indicate potential pathways for direct and indirect transmission of pathogens (Carrique-Mas et al., 2008; Gates et al., 2014, 2013a; Gilbert et al., 2005). Conceiving of farms as nodes and animal movements as edges in network analyses has been well developed in theory (Keeling and Eames, 2005), and has been applied to networks of holdings in multiple livestock species (Dutta et al., 2014; Green et al., 2011; Lentz et al., 2016; Natale et al., 2009; Nöremark et al., 2011; VanderWaal et al., 2015). Centrality measures can indicate the importance of a given farm within a trading network (Bell et al., 1999), and preferential protection, treatment or isolation of more central, or more influential, farms might enhance disease control measures (Natale et al., 2009; Rautureau et al., 2011; Volkova et al., 2010a). Network measures such as a farm's *degree* and *strength* (the number of other farms with which they trade and the number of animals traded, respectively) have been associated with infection risks (Christley et al., 2005a).

In a structured population, the transmission of infection depends on the frequency and nature of interactions among individuals and groups. Cross et al. (2005) noted that for a fixed frequency of movements, an acute disease with a short infectious period encountering a sparse network will be unable to spread extensively before extinction. However, a chronic disease in the same network might be able to persist and disseminate more widely, if it has a long infectious period, relative to the frequency of between-group interactions (Cross et al., 2005). Therefore, investigation of such networks requires consideration of the temporal aspects of infectiousness of the pathogen, relative to the frequency of movements (Vernon and Keeling, 2009). Static network analyses can be particularly useful in evaluating disease transmission where this period of risk is quantified on the same temporal scale as the network (Büttner et al., 2016a; Frössling et al., 2014; Kao et al., 2007). For example, in studying transmission of Foot and Mouth Disease (FMD), which has an incubation period of a few days (Orsel et al., 2009), networks encompassing one week of movements are appropriate. Whereas for a chronic infection such as bovine tuberculosis (bTB), infection can be asymptomatic or latent for months to years (Pollock and Neill, 2002; Probst et al., 2011), and a longer-term perspective is required. In reality, the edge of a static network permanently depicts what is truly a transient event, persisting as long as infection does, either in the incoming animal, a contaminated environment, or via secondary infections in other livestock or wildlife. It has been shown, in networks constructed from movements of adult dairy cows, that analysis of groups of statically connected nodes consistently overestimated the epidemic size of highly transmissible diseases, whereas measures that took into account the temporal order of movements provided a lower and more realistic estimate (Dubé et al., 2008; Vidondo and Voelkl, 2018). The concept of contact or infection chains has been reasonably well-developed in the field of veterinary epidemiology (Melmer et al., 2018; Nöremark, 2010; Nöremark et al., 2011) and the wider literature contains many examples of essentially the same approach, such as: *time-directed paths* (Holme and Saramäki, 2012), *source counts* (Riolo et al., 2001), *accessible worlds* (Webb, 2006), *output domains* (Dubé et al., 2008), and *reachability* (Dutta et al., 2014; Holme, 2005; Schärrer et al., 2015). All of these terms describe a temporally sequential network to identify the nodes that are accessible through edges to-or-from each index node within a selected time period. In this study, we have used the term *contact chain* for consistency with the current literature in our field. Ingoing contact chains (ICCs) identify the number of farms that could *potentially* transmit infection to the index farm over a defined period arising from the purchase and importation of animals. Outgoing contact chains (OCCs) quantify the number of 'downstream' farms that could *potentially* acquire infection from the index farm through its onward sale and export of animals. This structure of contacts may therefore help to predict the risk of the index farm acquiring and then passing on infection and to characterise patterns of risk across a national herd. Of course, not all movements result in the transmission of infection; at least one animal moved per edge must be infected and have the prospect of becoming infectious, to have a chance of infecting animals on other farms down the chain. Crucially, in our study we do not explicitly model transmission of infection and we use the term 'contact' chain, rather than 'infection' chain, representing only the *potential* for infection spread. For infections that can effectively be clinically hidden, such as bTB, contact chains can provide a scale, extent and map of potential transmission routes, which may improve our understanding of epidemiology, beyond that available through studying direct contacts. In a five-year study of the French cattle movement network, where bTB is rare (relative to the UK), farms in the highest quartile of ICCs traded indirectly with up to 84% of farms in the network (Palisson et al., 2016). It was shown that these farms in the highest quartile of ICCs were more likely to experience a bTB outbreak (Palisson et al., 2016), suggesting a link between the connectedness of farms through the purchase of animals and their risk of acquiring this chronic infection. Similarly, the magnitude of ICCs has been associated with the risk of acquiring an acute infection, bovine coronavirus, on Swedish cattle farms (Frössling et al., 2012).

The cattle industry in Great Britain relies heavily on trade in animals among beef and dairy producers. Trading occurs privately, through a dealer or via livestock markets (Robinson and Christley, 2007) and each movement of a bovine animal is recorded by a national Cattle Tracing System. These records have been used to study both network structure and cattle demographics (Vernon, 2011); Green et al. (2006) analysed the initial spread of the FMD outbreak in 2001, before movement restrictions were implemented, revealing that livestock movements could result in widespread dissemination of the virus and that the timing of virus introduction affected epidemic spread through seasonal fluctuations in movements among farms. The susceptibility of a network to infection also depends on its overall connectivity (i.e. how many sections or components into which it is divided). Heterogeneity in British cattle movements is predicted to influence disease spread (Woolhouse et al., 2005) and so we have looked for known characteristics of farms that align with their trading behaviours. Production type and herd size have been found to be important in predicting movements among pig (Arruda et al., 2016; Lindstrom et al., 2012) and cattle farms (Nöremark et al., 2011; VanderWaal et al., 2015) and has been associated with persistence of bTB (Brooks Pollock and Keeling, 2009). We predicted variation in network measures and contact chains, based on herd size, production type and location, and thereby we expect that they might have varying influences on potential transmission of infection. We utilise both static and temporally relevant network analyses in the context of a chronic livestock disease in Great Britain in order to provide insight into the dynamics of cattle trading behaviour, investigate potential unobserved transmission routes and to characterise the important actors and practices within this network. We performed our network and contact chain analyses annually over an extended period to determine if changes within the British cattle industry have been reflected in the network structure or individual farm behaviour over time, and consequently if they might have the potential to affect disease transmission within the cattle population.

Methods

Population-level analysis

The Cattle Tracing System (CTS) records all movements, births and deaths of British bovines. For our study, the Animal and Plant Health Agency provided a cleaned, processed version of CTS data (see Green and Kao, 2007) on the recorded movements of cattle between locations in Britain from 1st January 2001 (when recording became mandatory; Green and Kao, 2007) to 31 December 2015. Data consisted of 158 million individual animal movements between premises. We removed births (41 million) and deaths (42 million) from the dataset and aggregated individual animal movements into batch movements of animals moved between the same farms on the same day. We included only animal holdings (farms) in this study, omitting 34 million movements to slaughterhouses as they represent sinks in the network where no epidemiologically significant transmission could occur. Twenty-six million movements (35%) took place via markets or showgrounds; we classed these as transitory and linked them as single edges from source to ultimate destination, removing the transitory node. Although we acknowledge the well-documented risks of livestock mixing at showing events and markets (Gibbens et al., 2001), and a market's role in concentrating and dispersing animals (Robinson and Christley, 2007), we considered premises that kept animals for longer than one day to be more relevant for transmission and persistence of slow-spreading infections, such as *M. bovis*. By directly linking the source and ultimate destination, the flow of animals through these premises remains in the analysis, whilst allowing us to focus on the farm premises, upon which opportunities for transmission of infection were most prolonged. We took 12-month periods from 1st January to 31st December for each year between 2001 and 2015 and grouped batch movements into single links between farms. These processes together left 9.5 million edges in the study network. Annual herd size was calculated as the mean of the daily number of animals on the farm over the same 12-month period. We used CTS data to define herd type for each year, defining it by the predominant classification (beef, dairy, or dual purpose) based on breed and then predominant sex within this classification. *Suckler* farms were defined by a majority of female beef animals, aiming to capture those herds where calves are reared by their dams before weaning (cow-calf systems). *Dairy* farms were defined by a female dairy majority, identifying herds producing milk commercially. *Fattening* units were defined by a male animal majority, identifying herds that did not breed cattle, but reared them for beef production. Any farms where the breed type or subsequent sex was not more than 50% were defined as *mixed herds*.

Network analysis

We constructed networks in which nodes were defined as unique animal holdings registered as keeping cattle, and directed edges were defined as a movement of one or more cattle between holdings. Directed edges were weighted by the number of animals moved to-or-from the same holdings during the network year, as we considered the number of animals to be proportional to the risk of a disease incursion, especially for a disease with low prevalence within herds (Volkova et al., 2010a). Only active holdings (those with a recorded movement, birth or death in the year of study) were included in each annual network. The network timeframe corresponded with previously defined 12-month periods between 2001 and 2015. Using a full year avoided bias from seasonal variation of movements (Robinson and Christley, 2006), yet was sufficient for transmission of a chronic infection (Kao et al., 2007). In and out-degree, in and out-strength, betweenness, edge density, degree assortativity, reciprocity, clustering coefficient, average path length, and the giant weakly and strongly connected components (GWCC and GSCC) were all calculated using the R (R Core Team Version 3.5.3, 2019) package igraph (Csardi and Nepusz, 2006). Definitions of all network measures and accompanying functions are provided in Appendix B Table B1. We compared measures from the observed networks in each year to values calculated from directed random networks of the same size and density, generated using the Erdös-Rènyi model (Erdös and Rényi, 1959; see Appendix A for further methods).

Contact Chains

We calculated ICCs and OCCs for one-year periods (starting and ending in January) from 2001–2015 using the R package 'EpiContactTrace' (Nöremark and Widgren, 2014). Overall, the effect of seasonal variations in movement patterns (Robinson and Christley, 2006) on our contact chains was likely to be minimal as we utilised a whole year of movements. However, should a farm have purchased a large group of animals in January every year and calculation of their chains began at the beginning of January, the chains would never have a chance to 'build', as the incoming movements to the farms from which they had purchased would not be included. Therefore this farm's 'true' chain would only be apparent in a chain that started in a different starting month, e.g. December. Any such effect could result in underestimation of the magnitude of the chain for some farms. Therefore, we calculated ICCs and OCCs starting at consecutive monthly-intervals from January 2012 to December 2013, a total of 24 one-year periods (see Appendix B Fig. B1 for schematic). We compared the results from the different starting months and then combined these 24 one-year periods to create a more robust summary of movements spanning 36 months (from the start of the earliest chain to the end of the latest), rather than a 12-month snapshot. We compared the mean, median and maximum number of farms from the combined 24 monthly-spaced chains and combined annually-spaced chains over the same time period (2012–2014) using Spearman's rank analysis. Summary values from both methods were very similar (Appendix A), and so our subsequent analysis utilised the mean of the 24 monthly-spaced chains. For comparison of chains 2001–2015, we used annually-spaced chains to reduce computational load. We set thresholds on a logarithmic scale for the number of farms in chains to aid their description; 0-10 = very small, 11-100 = small, 101-1000 = intermediate, 1001-10,000 = large, >10,000 = very large. The maximum number of farms in a contact chain in any one year represents the greatest extent of the potential impact any one farm may have on the network in that year. Correlations between network measures were randomised to account for non-independence of network data. Temporal stability of measures was assessed through ranking of nodes and calculating the standard deviation of mean rank over time (Wilson et al., 2013; Appendix A).

To characterise those farms at high risk of acquiring infection (defined as farms with very large ICCs) or of spreading infection (defined as farms with very large OCCs), or those that might be at high risk acquiring *and* spreading infection (farms with very large ICCs and very large OCCs), we used a logistic regression with a binomial error structure. We performed the analysis using a threshold of 100, 1000 and 10,000 farms (see Appendix B Fig. B8 for ROC values and Table B3 for predicted probabilities). The highest ROC values were achieved using a threshold of 10,000 farms and so this was used for the final models. Herd type, size, and region have previously been associated with contact chains of cattle farms in Sweden and Uruguay (Nöremark et al., 2011; VanderWaal et al., 2015) and were therefore used in our logistic regression. We grouped Great Britain into ten regions for this analysis (Appendix B Fig. B10). We tested the full model using backwards stepwise selection based on Akaike's Information Criterion (AIC; Burnham and Anderson, 2003) but found in every case that the full models had the lowest AIC. We calculated odds ratios and confidence intervals and performed ROC curve analysis to estimate the model goodness of fit (Hosmer and Lemeshow, 2000).

Results

National herd characteristics

The median number of cattle traded from a single farm to another over the 12-month period was two (interquartile range 1–4), and this remained similar across all years, apart from in 2001 when larger numbers of cattle were moved between farms (median = 3, interquartile range 1–8, max = 3990). Most cattle holdings (mean = 40,736, 56.8% of all farms) were characterised as suckler herds (Fig. 3.1). The total number of cattle holdings (number of nodes in the network for which we had sufficient information to classify the herd type) decreased by 20,840 during the study period, and this reduction was most evident amongst dairy herds (Fig. 3.1 and Table 3.1). Herd sizes increased over the study period; dairy herd size increased from a median of 166 in 2001 to 243 in 2015, while the numbers of animals in other herd types remained more stable (Appendix B Fig. B2).

Table 3.1. *Global network analysis metrics for observed directed, weighted networks of cattle movements between animal holdings (farms) in Britain from 2001 to 2015. Data are from the Cattle Tracing System.*

Year	Nodes (farms in network)	Edges (movements of at least one animal between two farms)	Active Farms (% of total farms)	Edge density	Reciprocity (0-1)	Clustering coefficient (0-1)	Average Path Length (no. of steps)
2001	85410	444514	95.2	0.000061	0.0319	0.0035	6.54
2002	82916	585262	93.4	0.000085	0.0351	0.0111	7.14
2003	80428	694775	92.7	0.000107	0.0368	0.0138	6.45
2004	79404	707915	92.7	0.000112	0.0397	0.0150	6.42
2005	77800	674776	93.2	0.000111	0.0422	0.0144	6.54
2006	76970	733241	93.7	0.000124	0.0382	0.0138	6.33
2007	73380	622011	92.2	0.000116	0.0379	0.0139	6.49
2008	72624	650563	93.0	0.000123	0.0353	0.0150	6.57
2009	71485	668151	93.3	0.000131	0.0339	0.0146	6.41
2010	70328	639541	93.1	0.000129	0.0342	0.0143	6.60
2011	69649	647205	93.3	0.000133	0.0329	0.0144	6.67
2012	67820	625378	92.9	0.000136	0.0315	0.0138	6.62
2013	67171	616215	93.0	0.000137	0.0305	0.0134	6.97
2014	66292	616627	93.0	0.000140	0.0303	0.0136	6.93
2015	64624	619632	92.0	0.000148	0.0306	0.0132	6.89
Mean	73753	636387	93.1	0.000120	0.0347	0.0132	6.64

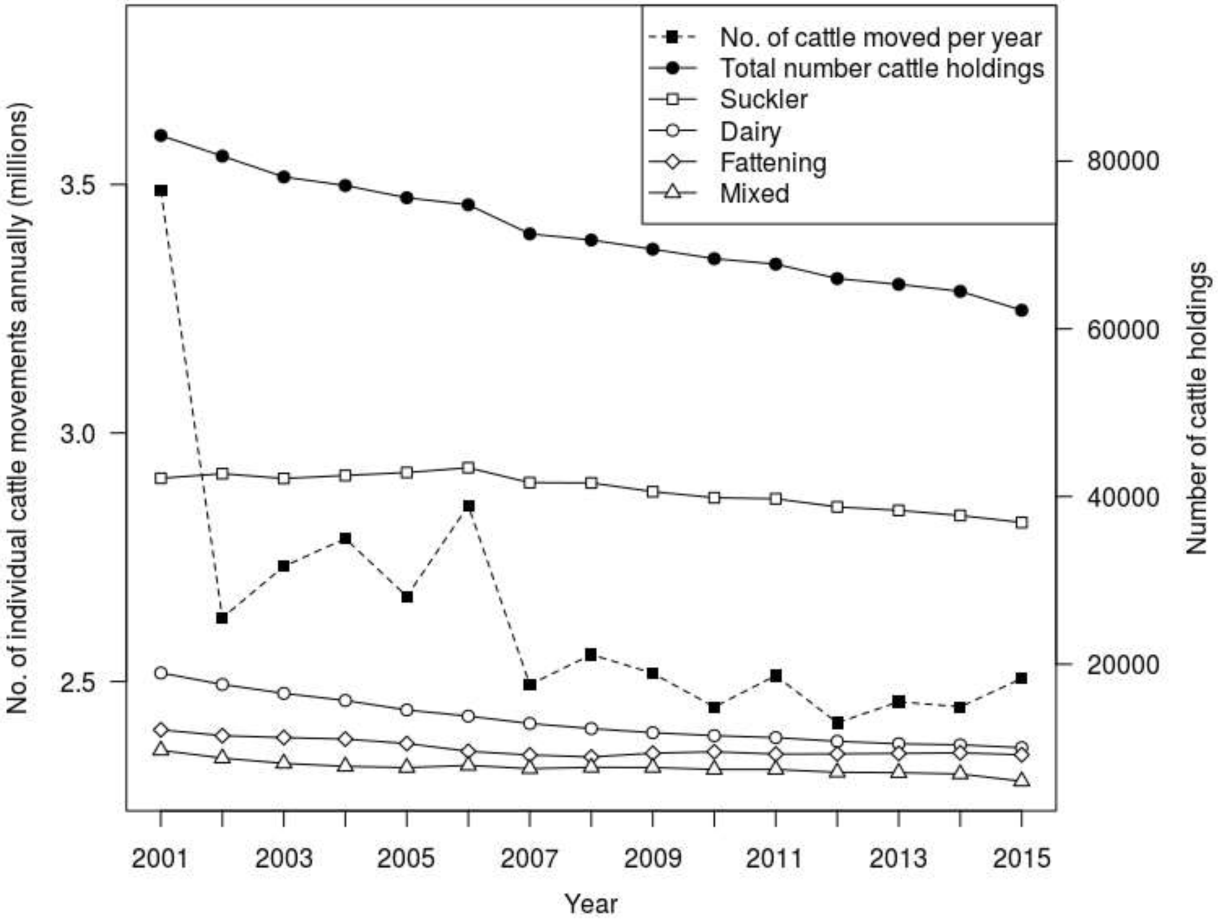


Figure 3.1. *Numbers of cattle moved between animal holdings (cattle farms) and the number of holdings, characterized by herd type, in Great Britain from 2001 to 2015. Data from the Cattle Tracing System. The numbers of cattle moved are shown with a dashed line and the numbers of holdings are shown with solid lines.*

The predominant between-herd flows of cattle were among suckler herds (12%), from suckler to fattening herds (9%), and from dairy to fattening herds (7%) (Fig. 3.2a). Suckler and dairy farms were the source for 23% and 18% of movements respectively between 2012 and 2014, despite the number of dairy farms being approximately a quarter that of suckler farms. Fattening farms received the most (21%) movements from other farms and dairy farms received the fewest (4%), of which most were from other dairy herds. 35% of movements were traded through markets, often from breeding herds (suckler and dairy) back to suckler herds or to fattening herds (Fig. 3.2b). The number of active farms (those with a birth, death or movement) stayed stable after 2001 (Table 3.1). ‘Isolated’, farms with a birth or death but *not* participating in the network make up 6.9% of farms (Table 3.1). On average, 34.4% of dairy holdings purchased no cattle in any one year (95% confidence interval = 31.8-37.1%), followed by suckler farms at 27.0% (23.9-30.1%), mixed holdings at 22.5% (19.1%-25.7%) and fattening units had the lowest percentage of closed farms at 9.64% (7.5-11.8%). Farms that did not report any inward movements in a five-year period from 2011-2015, and so could be considered ‘closed’ within this time, made up 9.6% of herds in the network. This is a similar value to dairy, suckler and mixed herds at 10.1%, 10.5%, and 10.6%. We found 5.4% of herds categorised as fattening enterprises were ‘closed’ for that period, suggesting that these were not typical ‘fattening’ units and that more than one form of enterprise was present.

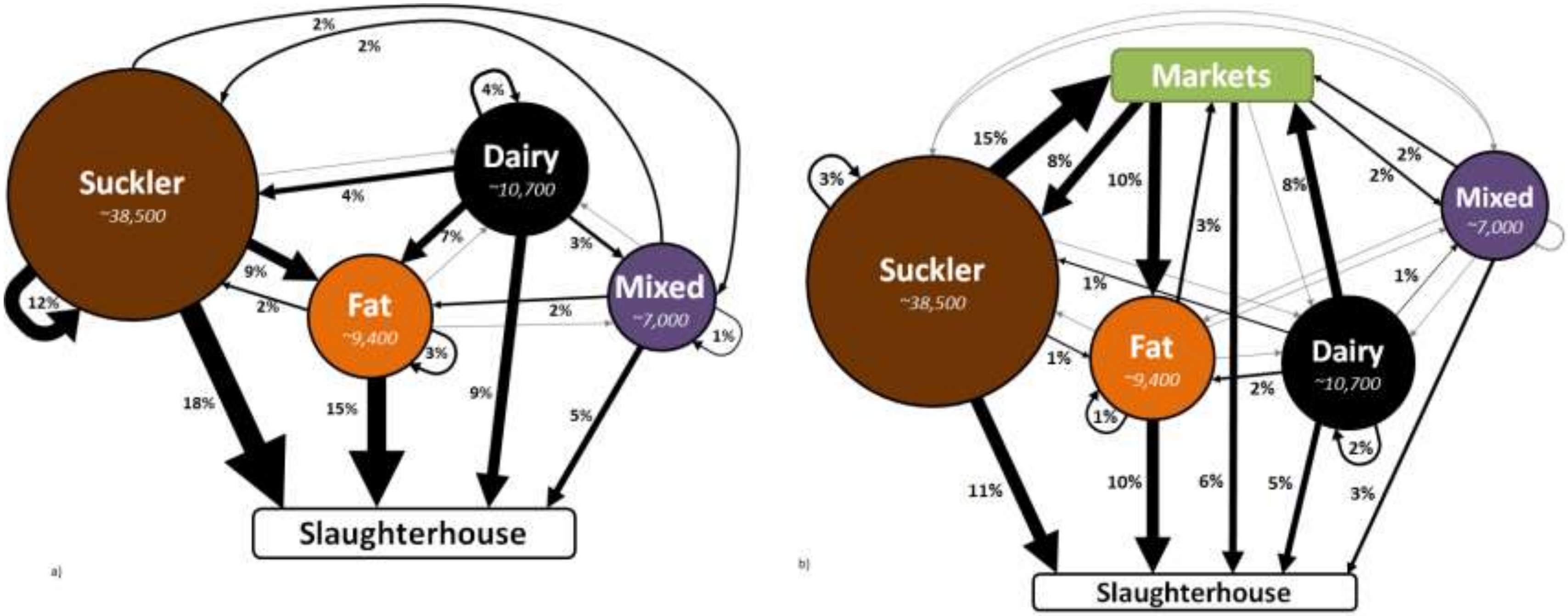


Figure 3.2. *Relative importance of pathways for cattle movements in Great Britain from 1 January 2012 to 31 December 2014. Movements are shown as a percentage of 13 910 851 movements over this period. Black arrows represent movements weighted as a percentage of total movements recorded. Flows of less than 1% are marked by grey arrows. The number of farms in each herd type is the mean number of unique, registered cattle holdings over the years 2012–2014, also represented by the area of the circle for each type. Our analyses are based on the characterization of movements shown in (a) which represents cattle movements among cattle holdings (farms) and slaughterhouses, and where movements made through markets are included as direct farm-to-farm/slaughterhouse movements. For comparison, in (b), markets are included explicitly as locations, to indicate the frequency of cattle movements via markets.*

Network analysis

The number of edges (created when at least one animal is traded between two farms) in the cattle movement network was lowest in 2001, after which it stabilised to between 60,000 and 70,000 between-herd links per annum (Table 3.1). In every year, all metrics of the observed network values lay outside the distribution of values from the random networks (Appendix B Table B2), showing that all of the observed networks differed from random in a range of key measures. The density of connections between farms in the network (edge density) increased by almost half over the study period (Table 3.1). Degree assortativity was more negative in all years in the observed network, when compared to random (Appendix B Table B2), meaning that farms with low degree were connected to those with high degree, and vice versa, indicating that some farms act as ‘hubs’ for movements. The reciprocity of edges was 3–4 orders of magnitude higher in the observed network than the random network (Appendix B Table B2), suggesting trading partners tend to reciprocate buying and selling cattle with one another. Clustering coefficients were higher, and average shortest paths were shorter, than in the random networks. GSCCs were smaller than in random networks, with the observed GSCCs containing fewer than half the number of farms of the random networks (Appendix B Table B2). Together, this suggests that the observed networks are modular, consisting of multiple smaller groups of well-connected farms, and therefore displaying small-world type properties.

Farm measures of movements remained consistent when they were ranked by the number of farms with which they traded (degree), and the number of cattle moved (strength), between years (Appendix B Fig. B4). This applied to trade both in-to and out-from the farm. Ranks of more global measures of connectivity for each farm (contact chains and betweenness) were also consistent between years but showed more variation than local measures (degree and strength; Appendix B Fig. B4). These results indicate that individual farm movements tended to stay consistent over time.

Observed degree was highly variable, compared to the random network (Appendix B Table B.2). In and out-degree of individual farms were positively skewed (skewness = 17.5 and 3.07 respectively; Fig. 3.3). Degree had power law exponents suggesting that the network might be characterised as scale-free, with many farms trading cattle with only a few direct partners and a small number of farms trading with many direct partners (Appendix B Table B4). The number of premises from which individual farms buy in cattle had much greater range (in-degree range = 0–4346) than the number of farms to which individual farms sell cattle (out-degree range = 0–305) (Fig. 3.4). The number of animals moved in-to and out-from individual farms also showed a positively-skewed distribution and range that was more marked for the number of animals bought in (in strength skewness = 16.3, range = 0–15,359) than the number sold (out strength skewness = 8.67, range = 0–6,472; Fig. 3.4). These patterns for degree and strength were consistent across all years. Larger farms tended to trade with more other farms consistently among years, demonstrated by a positive relationship between herd size and degree (mean r_s = 0.619, 95% confidence interval 0.603–0.634). Larger farms also traded more animals (mean r_s = 0.679, 95% CI = 0.662–0.696). Herd size and out-degree and out-strength (mean out-degree r_s =0.531, 95% CI = 0.501–0.562, SE = 0.014, mean out-strength r_s =0.587, 95% CI = 0.550–0.624) were more strongly correlated than herd size and in-degree and strength (mean in-degree r_s = 0.290, 95% CI = 0.280–0.300; mean in-strength r_s = 0.283, 95% CI = 0.274–0.292). Differences between some herd types in the number of farms that were traded with, and number of animals traded, were clear. Median in-degree and in-strength were higher for fattening farms, and median out-degree and out-strength were higher for dairy farms (Fig. 3.4).

Betweenness values were positively skewed (skewness = 29.2), with a large group of farms showing low betweenness scores and a smaller proportion of farms with very high betweenness scores, but no distinct differences between herd type (Appendix B Fig. B3). Mean betweenness was much larger in the observed network than in the randomised networks in all years but one, and median betweenness was much lower. This provides further evidence that some farms, not linked to a particular herd type, act as hubs within the network.

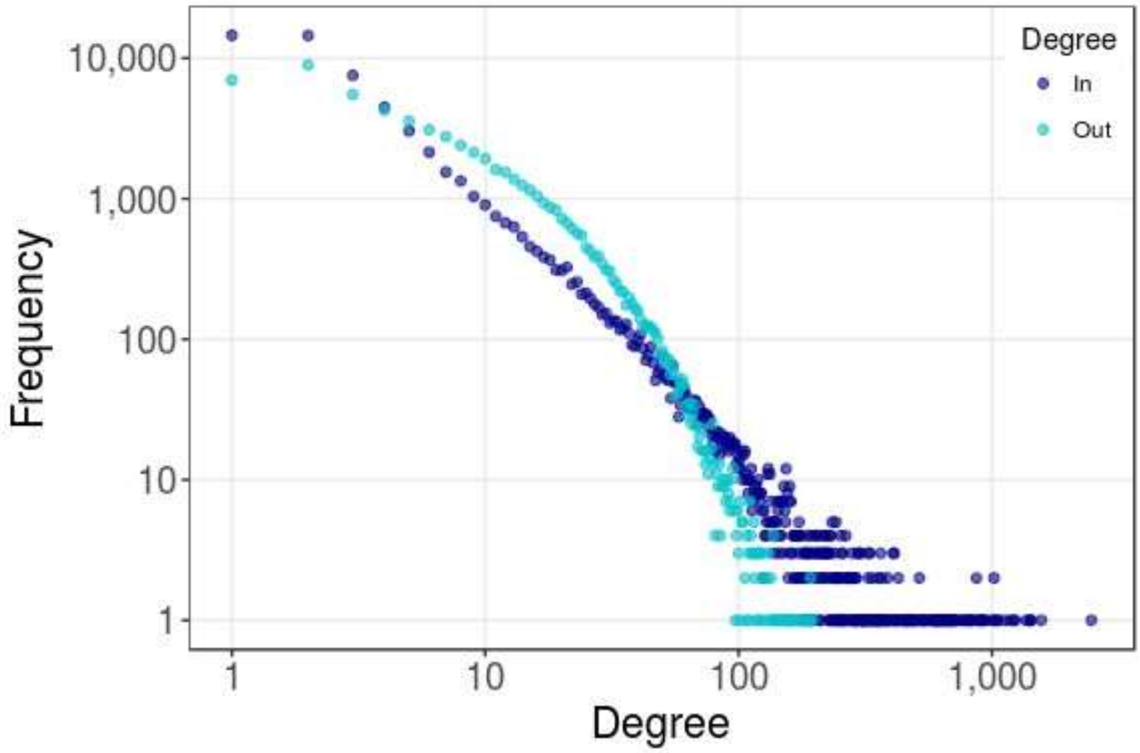


Figure 3.3. *Distribution of in- and out-degree for all farms in the annual network in 2015. Points represent binned data. Degree is n+1, for depiction on a log axis; therefore farms with no connections, i.e. zero degree, are depicted as degree = 1.*

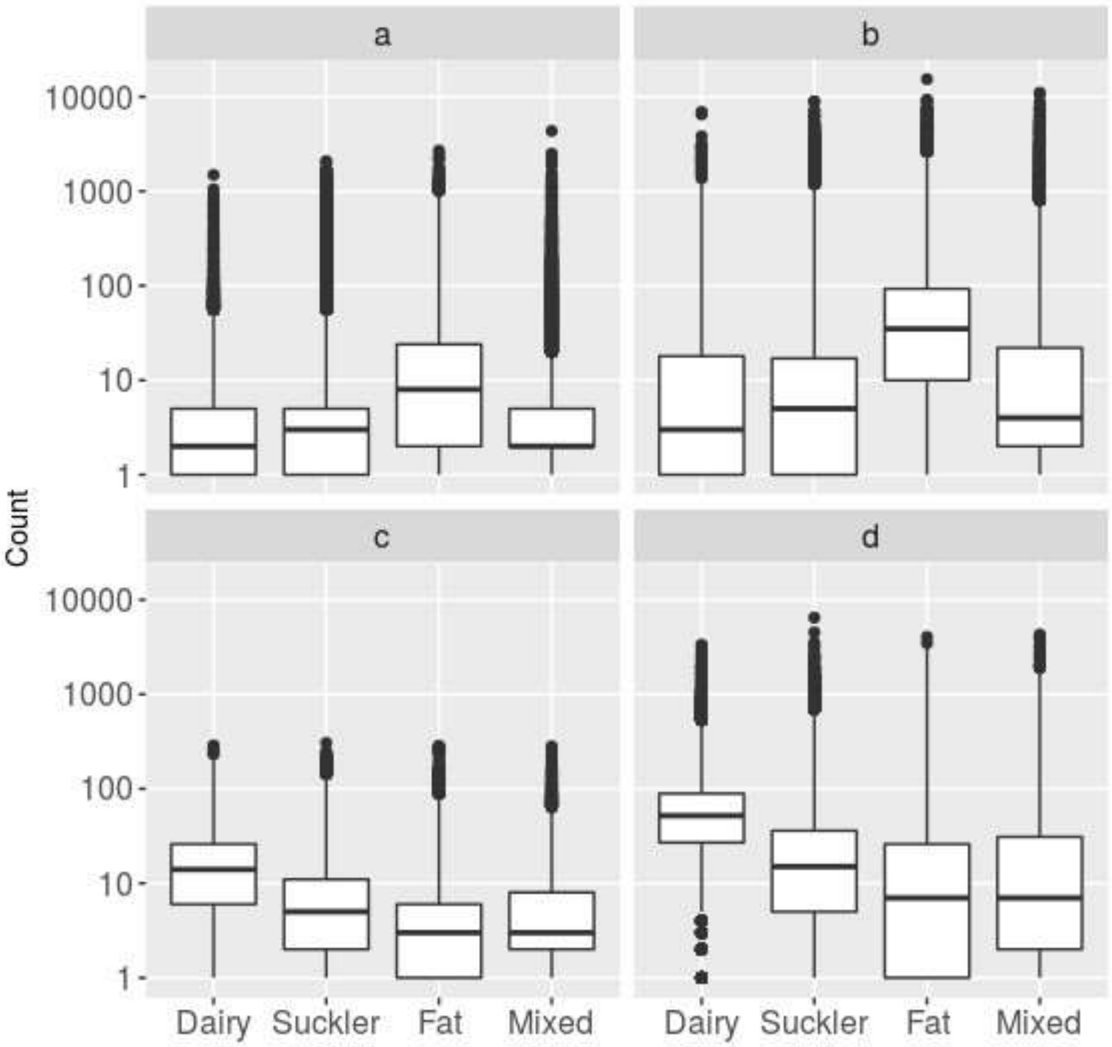


Figure 3.4. *The number of farms directly traded with (degree) and the number of animals traded (strength) in the annual network split by herd type for: (a) in-degree (number of farms cattle purchased from), (b) in-strength (number of animals purchased), (c) out-degree (number of farms cattle sold to), (d) and out-strength (number of animals sold). Combined data from network analyses from 2001 to 2015 are shown. Data are n+1, for depiction on a log axis, and include farms for which degree or strength were zero. Box plots show medians and interquartile ranges and the whiskers indicate the smallest or largest values no further than 1.5 times the interquartile ranges. Points outside this are outliers.*

Contact chains

Most farms had fewer than 100 farms in their ICC; however, up to 40% of holdings had large or very large ICCs over a one-year period, creating a strongly bimodal distribution (Fig. 3.5). The bi-modal distribution was also present for OCCs. The sizes of ICCs and OCCs for individual farms remained reasonably stable between all study years, though 2001 and 2002 have different characteristics (Fig. 3.6c and 3.6d). ICCs and OCCs were both positively skewed (skewness ingoing = 1.85, outgoing = 0.91). The maximum observed ICC of a single farm encompassed 86% of all British cattle holdings active in 2001 (n = 73,465 farms) and the maximum OCC occurred in 2004, encompassing 43% (n = 34,460 farms) of holdings. Across all years studied, approximately 50% of farms had very small ICCs, while 35-40% had large or very large ICCs (Fig. 3.5). More holdings in 2001 and 2002 have very small or small ICCs, than those from 2003 onwards (Fig. 3.6c). OCCs showed a different distribution; in most years, 50% of farms had very small to medium OCCs, and 50% had large or very large OCCs (Fig. 3.6d). More farms in 2001 and 2007 had large to very large OCCs, and in 2002, fewer farms had small or very small OCCs compared to other years (Fig. 3.6d). Clear differences were evident between herd types (Fig. 3.6a and 3.6b) with over 50% of fattening units but only 25% of breeding and mixed herds having large to very large ICCs (Fig. 3.6a). Over 80% of dairy farms, 55% of suckler farms and fewer than 40% of fattening farms had large to very large OCCs (Fig. 3.6b).

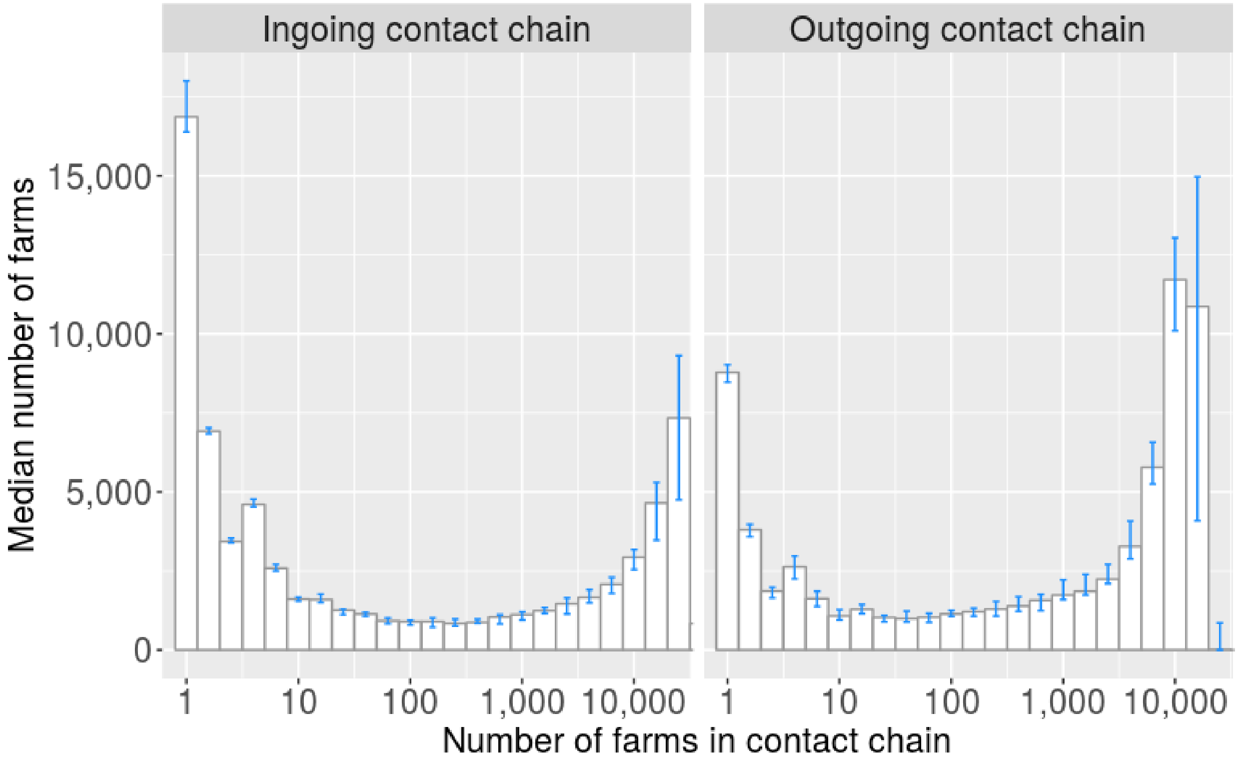


Figure 3.5. *Distribution of ingoing and outgoing contact chains with the number of farms in the chain for: ingoing contact chains (ICCs) and outgoing contact chains (OCCs), showing bimodal distribution for both types of contact chain. ICCs have more farms with fewer farms in their chains, whereas despite a lower mean and maximum number of farms, OCCs have more farms with very large chains. Data are n+1, for depiction on a log axis. Data include all active farms in Britain from 2001 to 2015. Error bars indicate the interquartile ranges.*

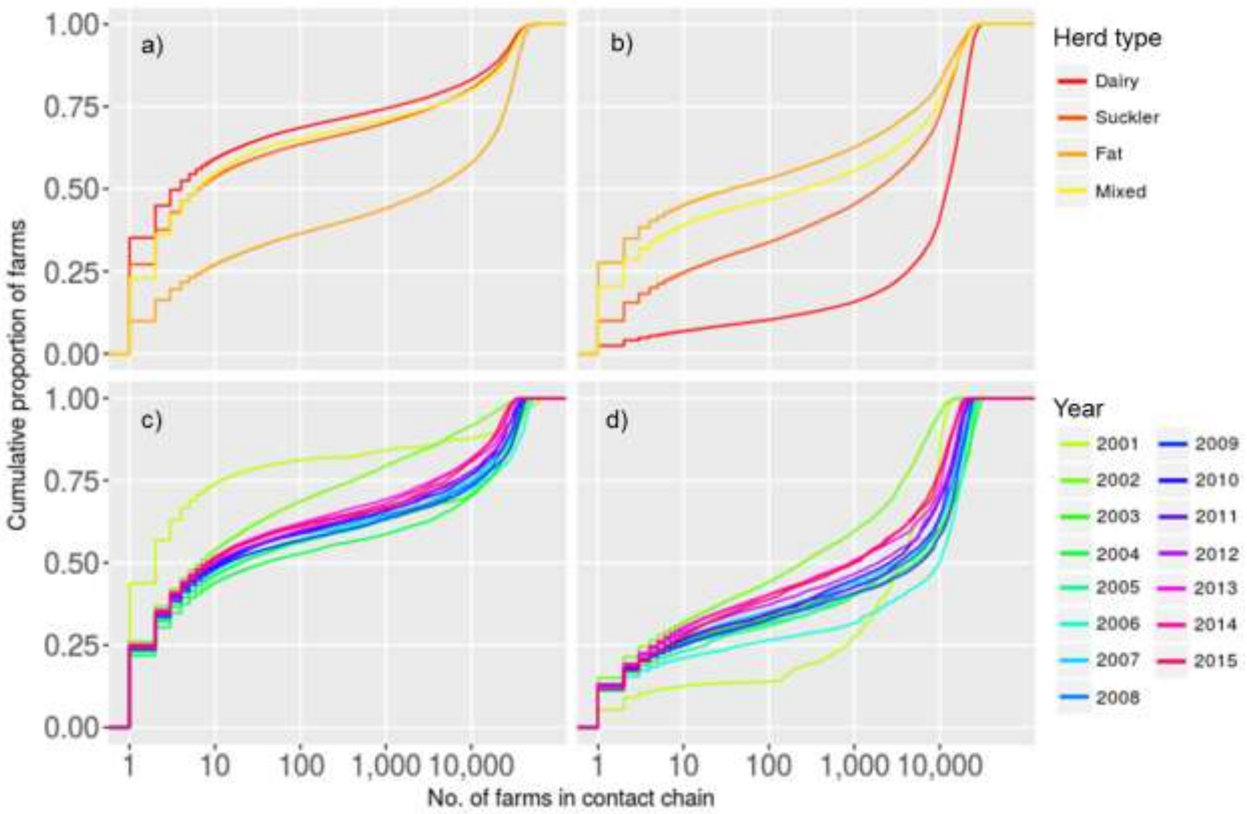


Figure 3.6. *Annual cumulative proportions of farms in Britain by the number of farms in their contact chains 2001–2015 and by herd type. Contact chains from 2001 to 2015 are included; therefore, some farms are represented multiple times for herd types. All years show a similar distribution shape; however, there appears to be some drift over time and marked deviations for 2001 and 2002, probably due to FMD movement restrictions, culling and restocking of herds in those years. Herd-type distributions have some similarities; however, fattening herds and dairy farms differ from other herd types for ICC and OCC, respectively. In (a,b), the colour of the line represents the herd type characterized by breed and sex CTS data, in (c,d), the colour of the line represents the year from 2001 to 2015. ICCs are shown in (a,c), OCCs are shown in (b,d). Data are n+1, for depiction on a log axis; therefore farms with chain length of zero, are depicted as chain length = 1.*

There was a weak correlation between ICC and OCC (r_s = 0.181, CI = 0.174–0.187, p <0.001, n = 76,031) for the mean values calculated from 24 monthly-spaced contact chains. This relationship was weaker when correlations of ICC and OCC were compared within all study years (mean r_s = 0.0398, CI = 0.020-0.594, p <0.001 in all years, n = 15). Farms tended to cluster at small ICC and OCC, large ICC and OCC and small ICC but large OCC (Fig. 3.7a). Regardless of ICC magnitude, dairy farms tended to have many farms in their OCC (Fig. 3.7b). Suckler herds tended to have more chains with large ICC and large OCC (Fig. 3.7c). Fattening herds had generally high ICCs with clustering at the low and high end of OCCs (Fig. 3.7d) and mixed herds tended to cluster more with low ICC and OCC (Fig. 3.7e). There was a strong positive and consistent correlation between in-degree and annual ICC over all years from 2001–2015 (mean r_s = 0.869, CI 0,858-0.879, p <0.001 in all years). This was true to a lesser extent between OCC and out-degree (mean r_s = 0.768, CI = 0.754-0.782, p <0.001 in all years). Over three-quarters of farms that sell animals to between 6 and 10 farms in the study period had an OCC linking over 1,000 farms, and farms that purchased animals from the same number of farms had a median ICC with 6486 farms. Some farms, despite low in- or out-degree, nevertheless had many farms in their contact chains (Fig. 3.8).

Logistic regression indicated that fattening farms were overall 9.5 (CI 8.7–10.5) times more likely to have very large ICCs (>10,000 farms in their ICC) than dairy farms (Table 3.2). There were also regional differences, where farms in the North of England were more likely to have very large ICCs (Table 3.2). Overall, dairy herds were more likely than any other herd type to have very large OCCs. However, with herd sizes above 500, mixed and suckler herds were more likely to have very large OCCs (Appendix B Fig. B7). Herds in the North of England and Scotland were much more likely to have very large OCCs than herds in the East of England (Table 3.2). Herds with over 500 cattle had an increased likelihood of very large ICCs *and* OCCs, especially in non-dairy holdings (Appendix B Fig. B7). Again, herds in the North of England and Scotland were at higher risk of having very large ICCs *and* OCCs (Table 3.2). ROC curve analysis showed all models had acceptable to excellent goodness of fit (Appendix B Fig. B8).

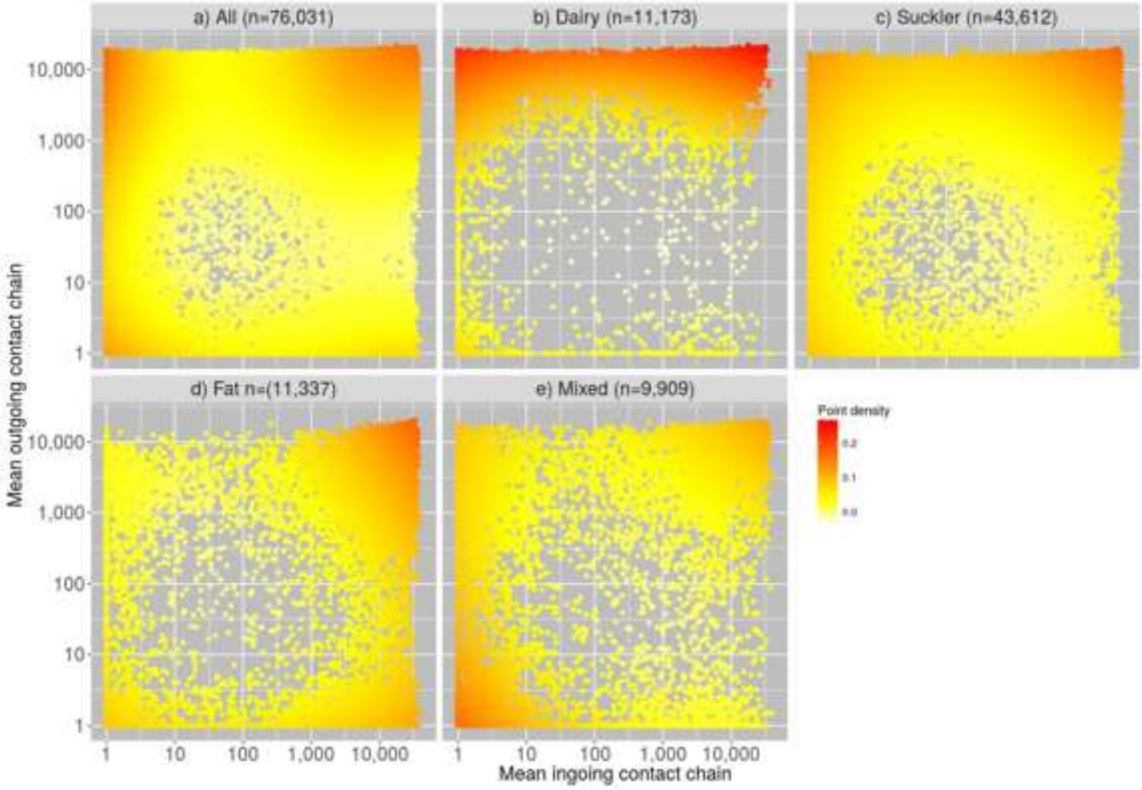


Figure 3.7. *Point density scatterplots to show relationship between mean ICC and mean OCC, where many farms lie in the top right of the plot. Data are split by herd type (a) all, (b) dairy, (c) suckler, (d) fat and (e) mixed farms with mean values of contact chains using combined data from 24 sequential chains 2012–2014, showing that the distribution of point-density changes with herd type. Point density is shown on a colour scale with lines smoothed by local polynomial regression fitting with a span of 0.6. Data are n+1, for depiction on log–log axes; therefore farms with chain length of zero are depicted as chain length = 1.*

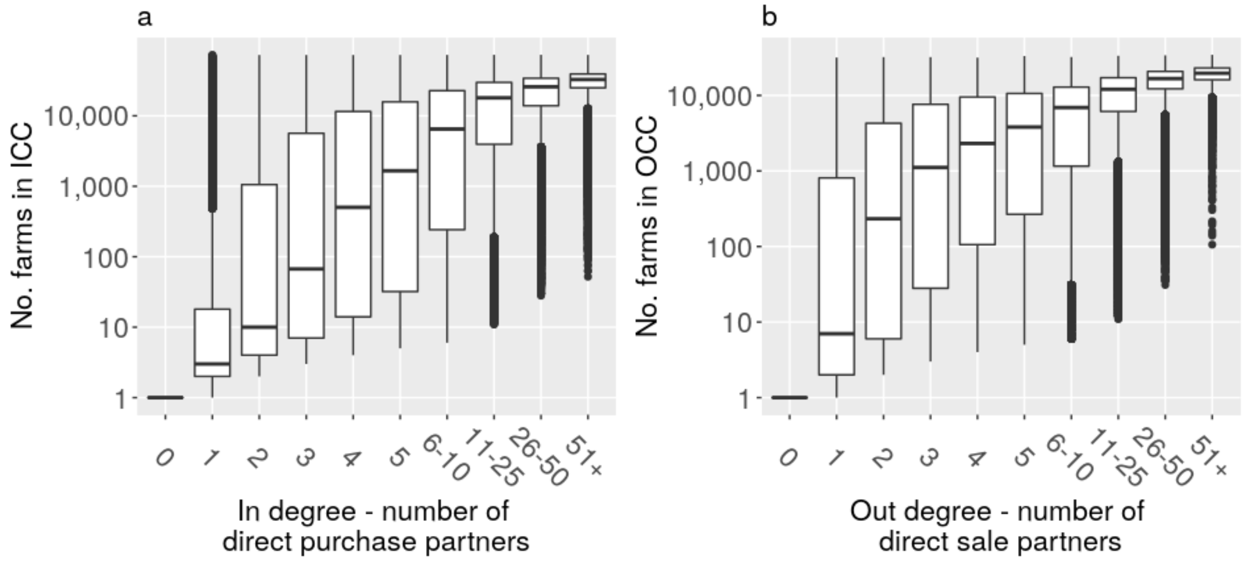


Figure 3.8. *The relationship between degree and contact chains showing the increasing number of direct trading partners and increase in number of farms in contact chain. (a) In-degree and ingoing contact chains and (b) out-degree and outgoing contact chains. Chain lengths are from 2001–2015 and are n+1, for depiction on a log axis. Box plots show medians and interquartile ranges and the whiskers indicate the smallest or largest values no further than 1.5 times the interquartile ranges. Points outside this are outliers.*

Table 3.2. *Odds ratios and 95% confidence intervals for three logistic regression analyses to identify characteristics of farms with very large contact chains. The response variable is whether or not farms have (a) ICCs, (b) OCCs, and (c) ICCs and OCCs containing over 10 000 farms (values from combined 24 monthly-spaced chains). Herd size has been mean centred and is displayed for increments of 10 cattle (see Appendix B, Fig. B5–7 for further herd size analysis).*

Response variable	Explanatory variable	Levels	Odds Ratio	2.5% confidence limit	97.5% confidence limit
a) Ingoing contact chain >10,000		(Intercept)	0.151	0.131	0.172
		Herd size (per 10 cattle)	1.004	1.002	1.006
	Herd type	Dairy	*Ref*		
		Fat	9.509	8.659	10.454
		Mixed	1.280	1.162	1.411
		Suckler	1.281	1.189	1.381
	Region	East of England	*Ref*		
		East Midlands	1.187	1.036	1.362
		North East	1.517	1.302	1.769
		North West	1.463	1.288	1.665
		Scotland	0.767	0.677	0.872
		South East	0.297	0.251	0.350
		South West	0.577	0.509	0.656
		Wales	0.484	0.425	0.553
		West Midlands	1.073	0.942	1.225
		Yorkshire	1.532	1.344	1.748
	Herd type: Fat	Herd size (per 10 cattle)	1.186	1.176	1.196
	Herd type: Mixed	Herd size (per 10 cattle)	1.052	1.047	1.057
	Herd type: Suckler	Herd size (per 10 cattle)	1.046	1.043	1.049
b) Outgoing contact chain >10,000		(Intercept)	0.211	0.161	0.272
		Herd size (per 10 cattle)	1.010	1.008	1.012
	Herd type	Dairy	*Ref*		
		Suckler	0.093	0.087	0.100
		Mixed	0.112	0.101	0.123
		Fat	0.051	0.046	0.057
	Region	East of England	*Ref*		
		East Midlands	1.526	1.153	2.047
		North East	11.584	8.817	15.442
		North West	18.075	14.009	23.716
		Scotland	15.212	11.815	19.923

Response variable	Explanatory variable	Levels	Odds Ratio	2.5% confidence limit	97.5% confidence limit
		South East	0.961	0.713	1.308
		South West	2.534	1.963	3.326
		Wales	2.305	1.780	3.034
		West Midlands	1.878	1.439	2.489
		Yorkshire	3.646	2.796	4.828
	Herd type: Fat	Herd size (per 10 cattle)	1.011	1.006	1.016
	Herd type: Mixed	Herd size (per 10 cattle)	1.040	1.035	1.045
	Herd type: Suckler	Herd size (per 10 cattle)	1.041	1.038	1.044
c) Ingoing and outgoing contact chains >10,000		(Intercept)	0.025	0.016	0.036
		Herd size (per 10 cattle)	1.003	1.001	1.006
	Herd type	Dairy	*Ref*		
		Fat	0.493	0.433	0.561
		Mixed	0.496	0.428	0.574
		Suckler	0.372	0.335	0.412
	Region	East of England	*Ref*		
		East Midlands	1.153	0.739	1.853
		North East	8.152	5.463	12.674
		North West	10.429	7.163	15.909
		Scotland	6.360	4.371	9.696
		South East	0.383	0.215	0.678
		South West	2.263	1.545	3.468
		Wales	1.197	0.801	1.863
		West Midlands	2.156	1.447	3.347
		Yorkshire	3.358	2.265	5.195
	Herd type: Fat	Herd size (per 10 cattle)	1.020	1.014	1.025
	Herd type: Mixed	Herd size (per 10 cattle)	1.031	1.025	1.036
	Herd type: Suckler	Herd size (per 10 cattle)	1.034	1.030	1.037

Discussion

Epidemics can be difficult to control if underlying transmission dynamics are not fully understood, especially in large networks where potential transmission pathways can be extensive and convoluted. The 9.5 million edges and ~70,000 nodes included in this study show that the British cattle network is complex, and potential transmission pathways can be extensive. Quantifying the extent of these chains is an important step in trying to understand the potential transmission routes for infections.

In respect of their contact chain distributions, British cattle holdings form two groups; those with very few (fewer than 10) contacts in their chains, and those with very many (more than 1000). Variation in this measure within a relatively short period could reflect important differences in a farm’s risk of acquiring and spreading infection, and key opportunities for action at ‘critical control points’ in the network. Previous studies have reported similarly skewed data in annual cattle farm contact chains in Sweden and Switzerland (Nöremark et al., 2011; Schärrer et al., 2015), and ICCs of pig farms in Germany (Büttner et al., 2013). In our study, contact chains of individual farms were stable over time, however larger chains showed some variation.

Large herd sizes are commonly associated with increased risk of disease (Bessell et al., 2012) and this is often attributed to large numbers of animals being purchased (Brooks Pollock and Keeling, 2009), we found that although there is an overall positive relationship between herd size and the number of linked farms, numbers of animals traded and chain magnitude, these relationships varied among herd types. This suggests that there are other mechanisms, beyond more animals entering the herd, that contribute to the apparent increased risk of disease in large herds (Gardner et al., 2002). Dairy, suckler and fattening herds have distinct patterns of degree and contact chains, indicating that they play different roles in the network. Through the purchase of animals from many other farms, fattening herds may be more susceptible to acquiring infection, while being less likely to pass on infection via movements as, clearly, many fattening cattle move straight to slaughter. Through selling animals to many different farms, dairy farms that become infected may be disproportionately influential for disease spread in the cattle network, offering a potential target for control measures (Büttner et al., 2013). In addition to the established role of markets as a ‘mixing pot’ for highly-transmissible diseases while animals are on site, here we emphasise their part in facilitating the dispersal of animals to many premises from one source farm, thereby potentially amplifying the spread of fast and slow-spreading diseases alike. In characterising farms into only four groups, we inevitably simplify the diversity of cattle farming operations in the British cattle industry. Fattening farms represent animals most likely intended for beef production, but we do not distinguish between premises rearing animals from calves to slaughter weight and cattle ‘dealers’ purchasing store cattle to sell onto other dealers, markets or farms. These differing businesses may have varying impacts on disease dynamics, where dealers may exhibit properties similar to markets, in acting to disperse animals to many farms. Farms in the South-East and East of England, where cattle densities are lower (Agriculture and Horticutural Development Board, 2017), may be less well-connected due to fewer chances to trade. Pre-movement bTB tests are not required for movement within the North of England and Scotland (Animal and Plant Health Agency, 2015b) and this may be responsible for the higher connectivity of their farms.

The cattle movement network in Britain displays scale-free properties, typical of those seen in movement networks in other countries. The number of cattle farms in Britain has decreased, driven largely by a reduction in numbers of smaller dairy farms and the formation of larger dairy herds (Fig. 3.1. and Appendix B Fig. B2). However, similar numbers of cattle were traded between fewer farms over the study period, resulting in a substantial increase in network density. Movement restrictions for 9 months of 2001, due to the FMD outbreak (Vernon and Keeling, 2012), account for low numbers of separate movements (edges), smaller GSCC, and low edge density that year (Appendix B Table B2). However, the restocking of farms from larger batches of animals after resumption of movements gave rise to the larger number of cattle moved that year (Vernon, 2011). Although it is well documented that the cessation of bTB testing during 2001 (Abernethy et al., 2013) contributed to the spread of bTB (Carrique-Mas et al., 2008; Gopal et al., 2006), the increased volume of animals moved could have made a significant contribution to the subsequent increase and spread of *M. bovis* infection. The number of animals in a batch is likely to affect the risk of farms acquiring or transmitting infection (Lentz et al., 2016), especially for pathogens with relatively low transmissibility such as bTB. Some studies have combined the number of animals traded with contact chains (Frössling et al., 2014; Schärrer et al., 2015), and this could be incorporated in further analyses.

The density and clustering of the network are mid-range compared with other livestock movement networks (Lentz et al., 2016; Mweu et al., 2013; Nöremark et al., 2011; VanderWaal et al., 2015), apart from an Italian network which had much lower clustering (Natale et al., 2009), suggesting British farms trade in small communities, and exhibit small-world properties. In the British network after 2004, clustering coefficients stayed stable and, after 2005, reciprocity decreased, suggesting that although the network becomes denser, it also becomes more dispersed, perhaps due to an increased propensity to travel further to trade cattle. The pattern of GSCC size in our results reflects that seen in previous studies (Christley et al., 2005b; Vernon and Keeling, 2012), and the reduction in size seen from 2004–2009 is continued in our analysis, which extends to 2015. Reduction of the size of such key components has previously been associated with reduced risk of epidemics (Volkova et al., 2010a).

A small number of farms in our network act as hubs (nodes with many more direct trading partners than the majority). Whereas previously this network role was considered to be fulfilled predominantly by markets (Christley et al., 2005b), hub farms provide similar linkages in the network that might facilitate epidemic spread by creating potential transmission 'shortcuts' through the network (Shirley and Rushton, 2005). Negative degree assortativity, similar to Scandinavian networks, where direct sales between farms, rather than via markets, are the norm (Mweu et al., 2013; Nöremark et al., 2011), combined with highly skewed degree distribution makes this type of network highly receptive to control measures targeted at hubs, rather than random selection (Büttner et al., 2013). It may therefore be beneficial to apply control measures, which have previously been aimed solely at markets, to hub farms as well. Risk-based trading measures might employ a proxy value for 'superspreader potential' that used network measures, similar to an 'infection potential' value (Rossi et al., 2017a), a probability of disease ratio (Frössling et al., 2014), or selecting a threshold for the most highly connected farms. This value would provide additional information on which farmers might base their buying decisions. However, it may be commercially harsh to rate farms in this way, as they have little control over this value beyond their direct purchases, and some farms even with very few direct trading partners are connected to very large contact chains (Fig. 3.8).

We used contact chains as temporally relevant network characteristics by which to assess the potential for acquiring and /or transmitting a slow-spreading infection arising from a farm’s trading network. There are numerous methods by which to achieve similar proxy measures from analysis of movement networks, such as those used by Frössling et al. (2014) and Büttner et al. (2016b). The algorithms used by Rossi et al. (2017a) and Konschake et al. (2013), allow a very fine scale of analysis, such that each node has its own infectious period and time of infection. We adopted a broader measure of contact chains, applied over an extended period of 15 years, to detect how variable farms may be in their network positions over time and in the context of infections with long incubation and infectious periods.

Although moving an infected animal into the herd presents the clearest risk of disease transmission, farms have other “connections”, including neighbouring farms and contacts via fomites, service providers and wildlife, any of which might be important for transmission of infections (Brennan et al., 2008; Dommergues et al., 2012; Rossi et al., 2017b). These connections often occur at a local scale and have varying importance depending on the pathogen of interest. Animal movements, however, can be implicated at all spatial scales, from the 43% of movements which occur within 20km of the source farm to the substantial number of long-range movements documented in Britain (Mitchell et al., 2005). Animal movements between neighbouring farms are also likely due to be underestimated due to local practices (Skuce et al., 2012). Here we have focused on this movement of animals as a potential pathway for the transmission of chronic infections and suggest that due to the long timescale, infections transmitted via movements may be more extensive through the network than we expect and that investigations into direct contacts may not be sufficient to trace the source and reach of some infections. Thus more extensive contact chains may better guide us to some of the varied sources and transmission pathways of cattle infections.

By examining networks of cattle movements, we observed two distinct patterns of interaction; many farms quickly became connected to a large proportion of the national network, yet some remained relatively isolated. We have shown marked variation between farms, not only in degree and betweenness but also in the more complex contact chains among British cattle farms, sustained over a period of 15 years that has been characterised by change in the industry and recovery from catastrophic disease outbreaks. Farms that exhibit extremes in ICCs and OCCs may be more likely to be infected and then more likely to act as superspreaders, via their outward trading behaviour (Lloyd-Smith et al., 2005). The British networks' scale-free type properties suggests the industry may benefit from targeted control of these influential nodes (Büttner et al., 2013).

Risks associated with direct trading partners are relatively easy for farmers to consider. However, the chains to which they can become connected remain hidden, along with the potential risks of exposure to infection they bring. By increasing the number of their direct trading partners, farmers are likely to see large, and sometimes very large, increases in the number of farms in their contact chain. Chains can also quickly become large, even with very few direct contacts, perhaps leaving farmers who believe they trade 'carefully' with a false sense of security. Knowledge of contact chains, and the trading patterns and history of the farms from which they are buying, might better equip farmers to judge the exposure associated with their animal trading behaviour. Contact chains allow us to assess a farm's role within the network and further investigation should explore their application to target certain farms, herd types, or practices for improved control of diseases.

Chapter 4

Effects of trading networks on the odds of bovine tuberculosis incidents on cattle farms in Great Britain

Chapter 4: Effects of trading networks on the odds of bovine tuberculosis incidents on cattle farms in Great Britain

This chapter is to be submitted to Royal Society Open Science as:

Fielding HR, McKinley TJ, Silk MJ, Delahay RJ, McDonald RA. (in prep) Effects of trading networks on the odds of bovine tuberculosis incidents on cattle farms in Great Britain.

Abstract

The trading of animals between farms and via markets can provide a conduit for the spread of infections and so by studying trading networks we might better understand the dynamics of livestock diseases. We constructed ingoing contact chains of cattle farms in Great Britain that were temporally linked by their trading behaviour, to elucidate potential pathways for the transmission of infection and to evaluate their possible effect on the risk of a farm experiencing a bovine tuberculosis (bTB) incident. Our findings are consistent with variation in risk associated with region, herd size, disease risk area and history of previous bTB incidents on the root farm and nearby farms. However, we also identified effects of both direct and indirect trading patterns, such that connections to more farms in the England High Risk Area up to three movements away from the root farm increased the odds of a bTB incident, while connections with more farms in the England Low Risk Area up to eight movements away decreased the odds. Relative to many other risk factors for bTB, trading behaviours are arguably more amenable to change, and consideration of risks associated with indirect trading, as well as those of direct trading, might therefore represent an additional approach to bTB management in Great Britain.

Introduction

Understanding pathways for the transmission of infections is fundamental to devising efficient control strategies for livestock diseases. These pathways exist at many scales, describing transmission among individuals (Lloyd-Smith et al., 2005), farms (Lentz et al., 2016) and countries (Nigsch et al., 2013). Trading connections among farms can generate vast networks of animal movements (Fielding et al., 2019) that have been implicated in disease transmission (Kao et al., 2007; Mekonnen et al., 2019). Network epidemiology utilises these networks as a framework on which to model the spread of infection (Martínez-López et al., 2009), and measures describing the centrality of a farm in a trading network and the nature of its contact chains (assessing temporal connectedness) can be useful in assessing an individual farm's risk of acquiring infection (Palisson et al., 2016).

Mycobacterium bovis causes bovine tuberculosis (bTB) in cattle but can infect a wide range of mammalian species (Broughan et al., 2013). Despite intense study and significant resources invested in control measures, the infection remains endemic in large parts of Great Britain (Allen et al., 2018). Among other factors, wildlife reservoirs of infection and undetected infections and movements of cattle have been shown to contribute to disease persistence and spread. A dynamic transmission model using British cattle movement and bTB testing data from 1996–2011 predicted that movements alone accounted for 13% of bTB incidents, and played a further role in multifactorial bTB incidents (Brooks Pollock et al., 2014). Modelling by Green et al. (2008) suggested that in 2004, 16% of herd infections occurred as a result of cattle movements, and that local effects were more important. The effects of cattle movements on patterns of bTB incidence in Great Britain have varied across regions (Gilbert et al., 2005; Johnston et al., 2011) and over time as new policies have been implemented (Vernon and Keeling, 2012). However, in spite of the introduction of pre-movement and, in some locations, post-movement testing for animals leaving high-incidence areas (Animal and Plant Health Agency, 2018b; Gates et al., 2013b; Vernon and Keeling, 2012), studies using data from 2006 to 2013 still suggested that there was an increased risk of bTB incidents for Scottish herds that had purchased cattle from high-risk regions of England and Wales (Gates et al., 2013a, 2013b; Salvador et al., 2018).

When compared to other risk factors for bTB incidents, cattle movements represent an activity that might be more amenable to management, for example by the use of risk-based trading, regulation or legislation. While the role of direct cattle movements in the transmission of bTB in Great Britain has been well-documented, the risks associated with indirect movements arising from trade has not yet been quantified or explored in detail. Second order contacts (i.e. contacts of contacts) in livestock networks have been shown to improve estimates of simulated epidemic size in models of disease in the British cattle herd (Tildesley and Keeling, 2009) and, in Italy, have assessed the influence of individual farms in disease transmission (Natale et al., 2009). In order to interrogate the network beyond first-order, direct contacts, data on the sequential purchases of cattle can be used to construct temporally explicit contact chains over a specified period. The ingoing contact chain (ICC) represents the source farms that may contribute infection to the root farm, while the outgoing contact chain represents those farms to which the root farm might transmit infection. The magnitude of ingoing contact chains has been associated with the risk of bTB in French cattle (Palisson et al., 2016) and was used to inform risk-based targeting of farms for surveillance of bovine coronavirus and bovine respiratory syncytial virus in Sweden (Frössling et al., 2012).

We have previously constructed contact chains for all cattle herds in Great Britain and found that a large proportion of herds have remarkably extensive chains, extending to tens of thousands of other farms within 12-month periods (Fielding et al., 2019). In the present study we hypothesise that farms with larger numbers of farms in their ICCs are at greater risk of exposure and acquisition of infection. We predict that infection risks might therefore be greater for farms with ICCs that include more risky trading partners, such as those that have experienced a recent bTB incident or that are located in regions with high incidence rates. We also predict that the closer a farm is to the root farm, both in terms of geographical proximity and trading proximity, the more impact it will have on disease risk on the root farm. It has been established that risk factors are likely to vary across different geographical regions (Johnston et al., 2011; Skuce et al., 2012), therefore we performed multiple analyses using subsets of data from specific regions. Overall, our study aimed to assess the importance of new parameters from contact chains, based on trading networks, alongside established risk factors, on the risk of bTB incidents on cattle farms in Great Britain.

Methods

Study farms and bTB data

We obtained bTB surveillance and farm information data collated in the Cattle Tracing System (CTS) by the Animal and Plant Health Agency (APHA). Study farms were active premises (i.e. registered cattle holdings with a birth, death or movement) between 1st January 2012 and 31st December 2014 with full bTB testing and location data, resulting in a final dataset comprising 71,096 cattle farms. Annual herd size was derived from the mean daily number of cattle on the premises. We took the mean of these annual herd sizes across the three years of the study. Herd type was defined using CTS data, first by the predominant breed type (dairy, beef, or dual purpose) and then sex of animals on the farm (as outlined in Fielding et al., 2019). Farms were allocated to bTB risk areas for England based on the designated risk area of their county parish defined by APHA in 2015 (Animal and Plant Health Agency, 2017) (Appendix C methods and Appendix D Fig. D1). A farm was defined as having had a bTB incident if its Officially Tuberculosis Free (OTF) status was withdrawn (OTF-W) or suspended (OTF-S) on one or more days during the time period in question, i.e. for assessing previous bTB history this period was 2010–2014, and for our response variable it was 2015–2016. OTF status is suspended (OTF-S) if at least one animal tests positive (or returns two inconclusive test results) for *M. bovis* using either of the statutory live tests: the Single Intradermal Cervical Comparative Tuberculin (SICCT) test or the gamma interferon test, and is withdrawn (OTF-W) on finding *post-mortem* pathology characteristic of *M. bovis*, or isolation of the bacteria by culture.

Networks

A single directed network was calculated for each of the three study years, representing farms as ‘nodes’ and with farm-to-farm movements of at least one animal represented by ‘edges’, weighted by the number of animals traded between those farms in the study year (Appendix C methods). Betweenness centrality was calculated for each root farm in each network, providing a centrality measure of the importance of the farm in connecting together different parts of the network (Silk et al., 2017a). ICCs are comprised of farms linked by movements of animals *onto* farms in a chronological sequence, such than at least one animal from a movement could have potentially seeded an infection that could be passed on in a subsequent movement of animals from that farm. Any farm that may have contributed to infection on the root farm (via movements) within the study period is included in the chain as a ‘source farm’. The ‘level’ for each source farm corresponds to the minimum number of movements away from the root farm at which they appear in the ingoing contact chain. To avoid any seasonal effect of starting the chain in a specific month, we calculated twenty-four ICCs at one month intervals, each spanning 12 months of movements. We then grouped the source farms from each of the twenty-four ICCs. Where farms appeared more than once in the chain in any one year, we included them at the closest (minimum) level to the root farm and removed other instances. Each chain was curtailed at 8 levels away from the root farm due to computational limitations (Appendix C methods).

For each root farm, we quantified how many source farms from each region were at each level of the contact chain and how many source farms at each level experienced a bTB incident from 2010–2014. We used cumulative counts across the different levels, e.g. the number of source farms at level one, the sum of the number of source farms at levels one and two, etc., up to the eighth level. We recorded the mean distance between each root farm and all of their source farms at each level. To incorporate bTB risks from local farms, we calculated the proportion of farms that had experienced a bTB incident between 2010–2014, within a radius of 0–2 km, 0–4 km, 0–6 km, 0–8 km, and 0–10 km of the root farm. Location was defined by the (x, y) co-ordinates associated with a particular herd identity (County Parish Holding number) in the APHA data. We selected baseline variables for a regression model based on previously identified risk factors (Salvador et al., 2018; Skuce et al., 2012) listed in Table 4.1.

We specified a multivariable regression model using a baseline set of variables (Table 4.1), with the binary response of whether the root farm experienced a bTB incident in 2015 or 2016, i.e. immediately after the period over which the ICCs had been calculated. Various risk factors could be specified at different lags (e.g. distances or levels of the network). However, the inclusion of multiple variables corresponding to a single risk factor specified at different lags results in high multicollinearity, and thus model instability. Instead, we chose a single representative variable within each set of lagged variables by adding each lagged variable in turn to this baseline model and choosing the one with the lowest value of Akaike’s Information Criterion (AIC)

(Burnham and Anderson, 2003). This avoided the inclusion of a large number of highly correlated variables and problems with inference caused by singularity errors in the design matrix, while preventing excessive model selection and therefore reducing bias in the estimated coefficients of the model (Harrell, 2001). As part of preliminary analyses, a number of methods were trialled, each providing similar results to our final choice of model and method (see Appendix C methods for further details). As an example of our chosen method, to select the most suitable lag for the risk of bTB from the local area, we ran five multivariable regression models including each distance with the baseline risk factors described above (Appendix D Fig. D2). The model with the lowest AIC included the proportion of farms with bTB during 2010–2014 within a radius of 8 km from the farm, and was therefore included in the final model. This process was repeated to select representative variables from two sets of contact chain variables; those we expected to be associated with increased or decreased risk of a bTB incident on the root farm (Appendix D Fig. D3). The final model included all baseline explanatory variables, the proportion of farms in a radius of 0–8 km with a recent bTB incident, and the selected representative source farm variables: the number of farms located in the England High Risk Area at levels 1–3, and the number of farms located in the England Low Risk Area at levels 1–8 of the ICC (Table 4.1). To capture potential differences in aetiology between different regions, we grouped our root farms by their region (Appendix D Fig. D1) and performed the same multivariable analysis, only removing region as an explanatory variable.

To aid comparison of ORs between variables specified on different scales, we re-scaled the coefficients for continuous variables such that each estimate corresponds to the OR for a variable changing from the 10th percentile to the 90th percentile of the observed data (Table 4.1 and Appendix D Table D2). To make it clear that these do not relate to a per-unit increase OR, we henceforth refer to them as *standardised ORs* (note that these are not standardised in the usual statistical sense (using standard deviations), which would not be sensible here due to high asymmetry in the distributions of some of the explanatory variables). Model performance was assessed (Appendix D Fig. D4) and we tested for overfitting of the final model using bootstrapping (Harrell, 2001). Given the size of the data set the model showed negligible evidence of any overfitting, so we chose to present the unadjusted results here. All analysis was performed using R (R Core Team Version 3.5.3, 2019); network measures were calculated in ‘igraph’ (Csardi and Nepusz, 2006), contact chains were calculated using ‘EpiContactTrace’ (Nöremark and Widgren, 2014), and regression was performed using ‘lme4’ (Bates et al., 2015).

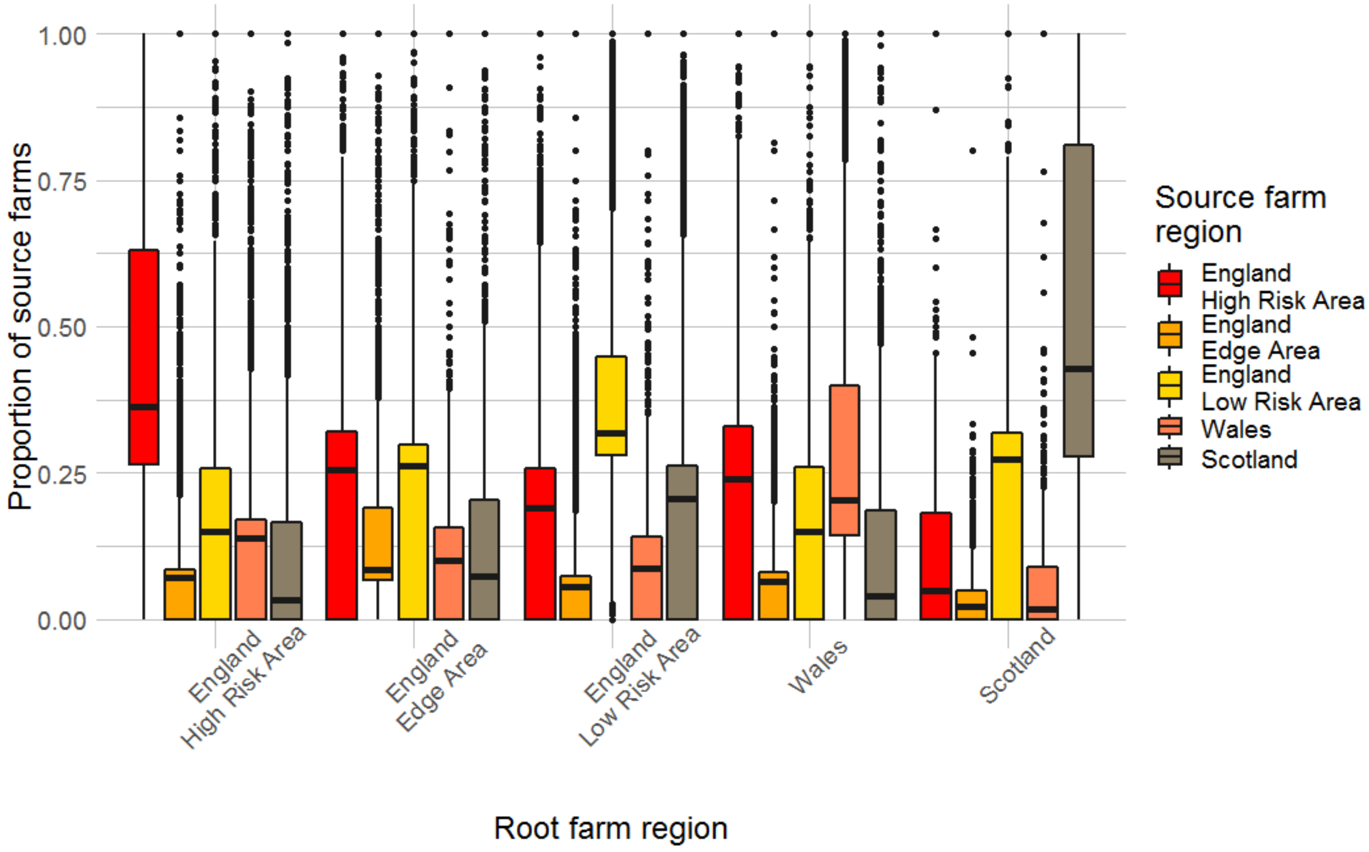
Results

The majority of source farms were located within the same bTB risk area or region as the root farm. However, the mean distance between root and source farms increased with the number of levels in the ICC (Appendix D Fig. D5) and an average root farm in the England Edge Area, England Low Risk Area or in Wales had about 20–25% of its source farms located in the England High Risk Area (Fig. 4.1). Also, 21% of root farms in Scotland had over 20% of source farms located in the England High Risk Area (Fig. 4.1). Apart from Scotland, all regions had some farms that had almost all of their source farms located in other regions (Fig. 4.1). The maximum number of farms at each level increased up to level 3 and declined thereafter, likely demonstrating a saturation effect in the most extensive chains (i.e. more than 10,000), where farms connecting at “higher” levels in the chain had already been counted at “lower” levels, and therefore did not increase the total number of farms in these extreme chains. In contrast, the median number of farms increased at each level up to level 8, likely demonstrating an overall amplification effect where the presence of more farms at each level, generally allowed more farms to connect at the next.

Table 4.1. *Effect sizes of explanatory variables on the odds of a bTB incident on the root farm in 2015–2016. Odds ratios (ORs) with 95% confidence intervals are from our multivariable logistic regression analysis using the full Great Britain dataset. ORs of continuous variables are standardised as the odds associated with the difference between the 10th and 90th percentiles of the raw data. Baseline variables are indicated by an asterisk (*).*

Region	Parameter		10th percentile (raw data)	90th percentile (raw data)	Odds ratio	2.5% confidence limit	97.5% confidence limit
Great Britain (n = 71096)	*Root farm risk area/country	Scotland			Baseline		
		Wales	-	-	6.67	5.19	8.68
		England Low Risk Area	-	-	2.88	2.23	3.77
		England Edge Area	-	-	10.58	8.24	13.76
		England High Risk Area	-	-	8.94	6.90	11.75
	*Root farm herd type	Mixed			Baseline		
		Dairy	-	-	1.33	1.19	1.49
		Fat	-	-	0.90	0.80	1.02
		Suckler	-	-	1.07	0.97	1.19
	*Root farm bTB incident 2010-2014 (binary)		-	-	2.79	2.62	2.98
	*Cattle purchased by root farm		-	-	0.98	0.90	1.07
	*Mean number of farms in ICC	1st quartile (0–1)			Baseline		
		2nd quartile (2–662)	-	-	1.04	0.94	1.15
		3rd quartile (663–6280)	-	-	1.12	1.01	1.24
		4th quartile (6281–39676)	-	-	1.17	0.99	1.38
	*Cattle purchased direct from England High Risk Area		-	-	1.23	1.12	1.34
	*Root farm herd size		4	280	19.41	17.09	22.06
	*Mean number of cattle purchased annually		0	198	1.01	1.00	1.03
	*Root farm betweenness		0	304239	1.00	0.99	1.00
	Proportion of farms within 8km with bTB 2010-2014		0	0.54	7.36	6.51	8.33
	No. farms in England High Risk Area at levels 1-3		0	627	1.11	1.07	1.15
	No. farms in England Low Risk Area at levels 1-8		0	10144	0.69	0.58	0.83

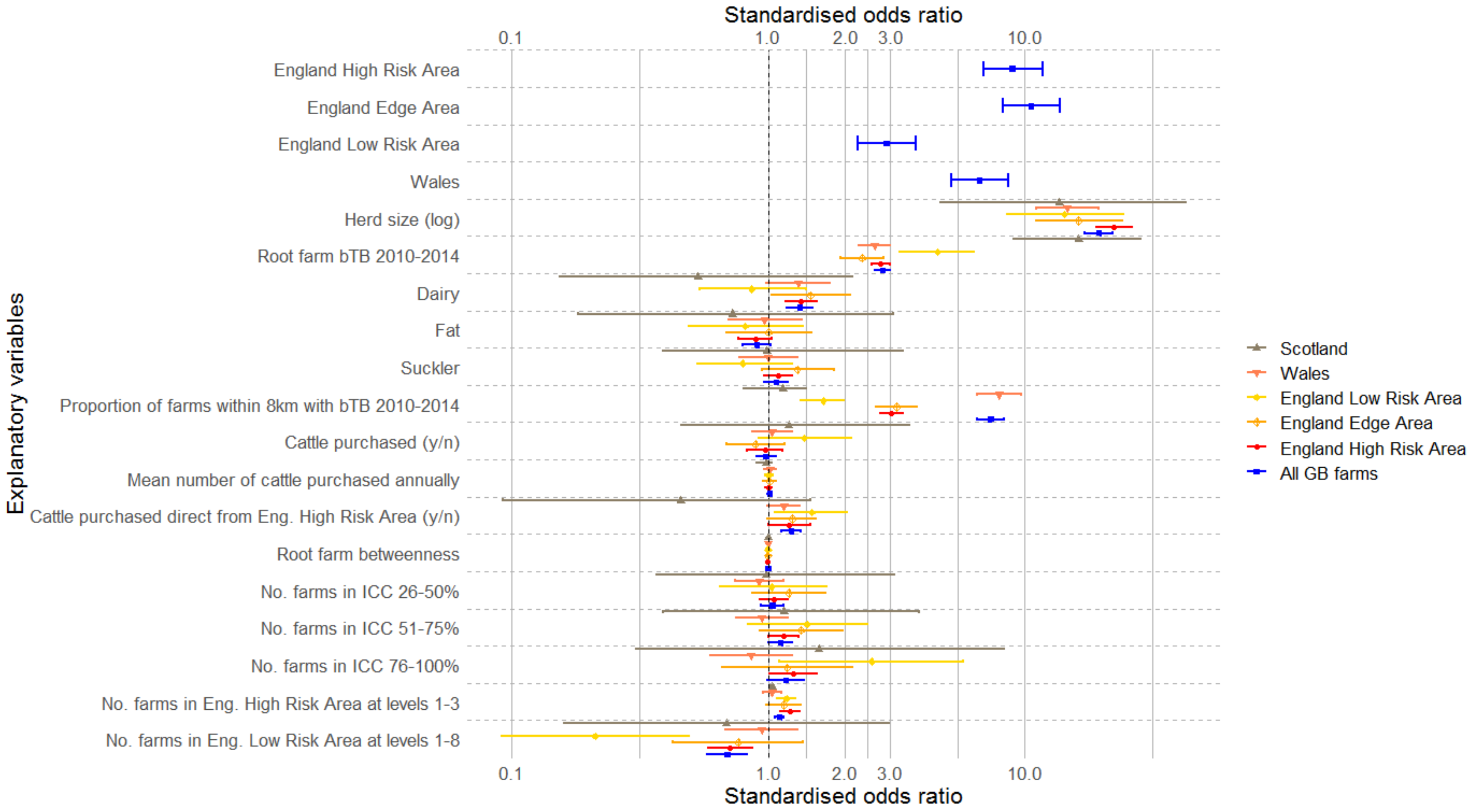
Figure 4.1. *Proportional contribution of each bovine tuberculosis risk region to the numbers of source farms comprising the ingoing contact chain of root farms in each of the disease risk regions in Great Britain. The majority of source farms are located within the root farm's region, however in the England Edge Area, England Low Risk Area and Wales over 25% of source farms are from the England High Risk Area. Boxplots show the median, 25th and 75th percentiles, and the upper and lower whiskers extend to the largest or smallest value no further than 1.5 times the interquartile range, data beyond this range are plotted as outlying points.*



In 2015–2016, 12.9% of root farms experienced a bTB incident, of which 78.5% were classified as OTF-W (Appendix D Table D1). Proportions of bTB incidents classified as OTF-W or OTF-S in the England High Risk Area and Wales mirrored those in GB as a whole, whereas in the England Edge Area the relative proportions of these classifications were more equal and in the England Low Risk Area and Scotland, there were more bTB incidents classed as OTF-S than OTF-W (Appendix D Table D1). Characteristics of the root farm and local bTB history were major risk factors associated with the likelihood of a bTB incident. Compared to farms in Scotland and the England Low Risk Area, root farms located in the England Edge Area or England High Risk Area, or in Wales, had greater odds of experiencing a bTB incident in 2015–2016 (Fig. 4.2). Root farm herd size was a strong predictor of bTB incidents with the odds of larger herds (the 90th percentile) having a bTB incident in 2015–2016 being 19.41 (95% CI = 17.09–22.06) times higher than for smaller herds (the 10th percentile—see Table 4.1). Root farms that had themselves experienced a bTB incident in the period 2010–2014 had 2.79 (2.62–2.98) times higher odds of experiencing another bTB incident in 2015–2016 (Table 4.1). Suckler, fat, and mixed herds had similar odds of a bTB incident, yet the odds of a dairy farm having a bTB incident in 2015–2016 were on average 1.33 (1.19–1.49) times higher than for a mixed farm (Table 4.1). Root farms with a higher proportion of farms within 8 km that had experienced a bTB incident (2010–2014) had 7.36 (6.51–8.33) times higher odds of experiencing an incident themselves (Table 4.1).

In selecting our variables, model performance was better (i.e. a lower AIC) when we included the number of source farms in the England High Risk Area, than those in the England Edge Area, Wales or the number of source farms with a bTB incident in the last 5 years (Appendix D Fig. D3). A root farm with more farms in the England High Risk Area at levels 1–3 of their ICC had 1.11 (1.07–1.15) times higher odds of having a bTB incident (Table 4.1). We found that having more farms in lower-incidence areas at any level of the ICC was associated with a decreased risk of bTB incident on the root farm, specifically root farms had lower odds of a bTB incident in 2015–2016 (OR = 0.69, 0.58–0.83; Table 4.1) if they had more source farms located in the England Low Risk Area at levels 1–8 of their ICC. The odds of having a bTB incident were 1.23 (1.12–1.34) times higher when purchasing cattle direct from the England High Risk Area compared to not purchasing cattle from the England High Risk Area (Fig. 4.2).

Figure 4.2. *(overleaf) The effect of root farm characteristics on the odds of a bTB incident on the root farm in 2015–2016 in Great Britain and in the disease risk regions. Odds ratios of continuous variables are standardised as the odds associated with the difference between the 10th and 90th percentiles of the raw data and are shown with 95% confidence intervals (whiskers). Variables include the number of farms from the England High Risk Area at levels 1–3 and the number of farms from the England Low Risk Area at levels 1–8 in the ingoing contact chain of the root farm (ICC), connectivity of the farm within the trading network (betweenness), the number of farms in the ICC, and previously established bTB risk factors.*


Standardised odds ratio
0.1
1.0
2.0
3.0
10.0
Explanatory variables
England High Risk Area
England Edge Area
England Low Risk Area
Wales
Herd size (log)
Root farm bTB 2010-2014
Dairy
Fat
Suckler
Proportion of farms within 8km with bTB 2010-2014
Cattle purchased (y/n)
Mean number of cattle purchased annually
Cattle purchased direct from Eng. High Risk Area (y/n)
Root farm betweenness
No. farms in ICC 26-50%
No. farms in ICC 51-75%
No. farms in ICC 76-100%
No. farms in Eng. High Risk Area at levels 1-3
No. farms in Eng. Low Risk Area at levels 1-8
Scotland
Wales
England Low Risk Area
England Edge Area
England High Risk Area
All GB farms

Our regional analyses showed broadly similar effect sizes to those obtained using the entire Great Britain dataset, however there were key differences in the effect sizes for bTB history on root farms and neighbouring farms (See Table 4.1, Appendix D Table D2, and Fig. 4.2 for analysis of all factors). In terms of network factors, having more farms in the ICC in the England Low Risk Area decreased the odds of a bTB incident more for root farms in the England Low Risk Area than for farms in the full Great Britain dataset. In terms of previously established risk factors, the odds of having a repeat bTB incident on root farms in the England Low Risk Area and in Scotland following an incident in the preceding 5-year period were over fifteen times greater than those for farms which experienced no bTB incident in that period. Similarly, the odds of bTB incidents on farms in the England Low Risk Area were almost five times greater if they had experienced a previous incident. For root farms in the England High Risk Area, England Edge Area and in Wales, a repeat incident had more than twice the odds of occurring than a first-time incident. Larger herds had at least thirteen times the odds of experiencing a bTB incident than smaller farms in all regions. Compared to other herd types, dairy farms still had increased odds of experiencing a bTB incident in 2015–2016 in the England High Risk Area and England Edge Area. Local effects were strongest in Wales where the proportion of farms within 8 km with previous bTB incidents was associated with almost 8 times higher odds (Appendix D Table D2). The effect was lower for farms in the England Low Risk Area and Scotland, where very few farms had neighbouring farms with recent bTB incidents.

Discussion

After accounting for previously identified risk factors, we have revealed additional associations between bTB risk and the characteristics and bTB history of indirect trading partners in the extensive trading networks of British cattle farms. The odds of a bTB incident were increased where there were more farms from the England High Risk Area at levels 1–3 in the ICC, while having more farms located in the England Low Risk Area at any level in the ICC was associated with a decreased risk.

Our study covers the period 2010–2016, four to ten years after the introduction of comprehensive pre-movement testing, and selective post-movement testing, and indicates that there remains an increased risk associated with direct *and* indirect trading from the England High Risk Area. While we found a weaker relationship associated with direct purchases from the England High Risk Area than seen in previous studies based on data from 2008–2013 (Salvador et al., 2018), our results suggest that there is scope for further reductions in the risks of acquiring bTB through trading links. For herds with non-negligible levels of infection, the SICCT test has reasonable sensitivity when considered as a herd-level test (Christensen and Gardner, 2000), however when used for pre- and post-movement testing on small batches of cattle (the median number of animals per trade between 2002 and 2015 was 2 with an interquartile range of 1–4; Fielding et al., 2019), its sensitivity is likely to be lower (Martin et al., 1992), resulting in frequent false negatives and the undesirable movement of truly-infected animals. Although farms tend to trade directly most often within their regions (Mitchell et al., 2005) and pre- and post-movement testing has been shown to deter movements between regions (Gates et al., 2013b), we show that extensive inter-regional connections are quickly formed in the contact chains of many farms. We observed that within a few movements, large areas of Great Britain might be traversed by cattle movements and that farms can be connected to many farms outside of their risk area or geographical region. While relatively few of these connections are likely to transmit infection, those that do may be effective in translocating disease into new areas (Gopal et al., 2006). The use of more sensitive diagnostics such as the gamma interferon assay (Nuñez-Garcia et al., 2018) in movement-associated testing might act both to further deter trading from regions of high to low risk, due to higher costs, and to reduce the occurrence of infected animals being traded. Risk-based trading protocols have been outlined elsewhere (Adkin et al., 2015), and quantify direct trading risk, but additionally including an estimate of the risk associated with indirect trading, using the locations and numbers of source farms might form a better understanding of the risk associated with each purchase.

Results of the present study show that having more farms in the ICC was initially associated with increased risk, as observed in a similar study of cattle contact chains in France (Palisson et al., 2016). However, unlike the French study, the number of source farms was not associated with a large increase in the risk of experiencing a bTB incident, suggesting that trade in cattle might play a different role in predicting bTB incidents in the two countries. BTB prevalence is much lower in France and there is a comparatively lower force of infection from the local environment. This likely results in overall trading strategies being more important relative to the specific identity of trading partners in the movement network. Our analysis in Britain showed that in terms of bTB incident risk, the location of source farms was more important than the total number of source farms. We predicted that trading with farms that had a history of bTB incidents would be a more accurate predictor of risk than trading with farms from the England High Risk Area, as reported by Green et al. (2008). However, we found that trading with farms in the England High Risk Area was more informative in predicting a subsequent bTB incident on the root farm. This may suggest that past bTB incidents do not fully describe bTB risk and that undisclosed infections in herds or environmental sources of infection present a risk that is captured by being in the England High Risk Area but not detected through routine testing (Conlan et al., 2012).

Spatial proximity to farms with recent bTB incidents was associated with an increased odds of bTB incident on the root farm in the full GB model and regional models except for Scotland. We expected that farms closest (i.e. 0–2 km) to the root farm would have the strongest association with the odds of a bTB incident as this would encompass potential contacts of cattle on shared boundaries and the risks from infected local wildlife (Broughan et al., 2016b). However, the variable which best explained this risk was the proportion of farms with a bTB incident within a radius of 8 km. This could potentially be related to the additional effects of unrecorded local cattle movements, the spreading of infected slurry (Green et al., 2008), and movement of cattle among fragmented land parcels (Broughan et al., 2016b). That this proxy measure was still strongly associated with the odds of a bTB incident after accounting for other major factors suggests that some or all of these factors may be important drivers of local risk.

Notwithstanding the low number of repeat bTB incidents in Scotland (n=21) and the England Low Risk Area (n=60) and broad confidence intervals of this variable, the association between previous incidents in the preceding 5 years (2010–2014) and the risk of a later incident in Scotland or in the England Low Risk Area was considerably greater than the association between prior history in the higher incidence areas of England and Wales and bTB incidents. As the environmental force of re-infection in non-endemic areas is likely to be low, this suggests that a small number of herds in Scotland and the England Low Risk Area may have had OTF status restored without having fully eliminated infection (Conlan et al., 2012), or alternatively, that there is a repeated external source of infection not accounted for in our study. For example, we have not included cattle imports from outside GB, and may therefore have underestimated the risk of bTB on farms that import animals from Northern Ireland and the Republic of Ireland (Salvador et al., 2018), where infection is also endemic in cattle.

An advantage of the regression framework employed in the present study is that it is more straightforward to fit these models to complex data than it would be to fit a fully mechanistic infectious disease model (Brooks Pollock et al., 2014; Conlan et al., 2012). However, although we included approximations for risk based on region and local bTB occurrence, we were unable to fully disentangle the relationships between movements, regions and finer-scale spatial risk. It may be that the impact of the network dynamics are more evident if explicit trading paths are modelled, for example through the use of a compartmental network-based epidemic model (Brooks Pollock et al., 2014).

The effect sizes of trading patterns in our study were relatively small compared to other risk factors such as herd size (Bessell et al., 2012; Brooks Pollock and Keeling, 2009) and prior bTB history which have been identified previously (Green et al., 2008; Orton et al., 2018; Palisson et al., 2016). Nevertheless, the major risk factors identified in this and previous studies tended to reflect the inherent characteristics of farms (i.e. location, bTB history of farm and neighbours), which are difficult, if not impossible to change. By contrast, trading behaviour might be more readily manipulated through regulation and incentivisation with support from industry and policy-makers. The present study reinforces the importance of direct trading links and highlights the additional risks associated with indirect trading links, revealed by a more thorough examination of the movement network, as necessary factors to consider in strategies to reduce bTB transmission in Great Britain. However, given the computational complexities of interrogating such a dense movement network in real-time, the next challenge will be in deciding how best these sources of risk might be integrated into current management policies.

Chapter 5

Spatio-temporal variation in social networks of commercial dairy cattle in Great Britain

Chapter 5: Spatio-temporal variation in social networks of commercial dairy cattle in Great Britain

Abstract

Acquiring infectious disease is a potential cost of group living and the nature of social relationships within groups can be important in facilitating or impeding the spread of pathogens. Networks constructed from contacts between hosts allow us to examine social relationships and how they might influence the spread of infections through groups. We collected high-resolution proximity and location data from nine groups of domestic cattle (mean group size = 85) in seven dairy herds employing a range of grazing and housing regimes. Networks were constructed from contacts aggregated by different temporal windows (2-hours, daily, entire study period) and grouped by farm area (Buildings, Pasture, and when cows were divided between the two). Networks of cattle contacts aggregated over the whole study were highly saturated, but dividing contacts by space and time revealed substantial variation in cattle interactions. Contact durations were longer in Buildings networks and cows housed in buildings contacted more other cows. Between-cow variation in terms of the numbers of cows they contacted and the frequency of their interactions was greater than random in Buildings and Pasture networks. A few cows in each group formed consistent strong relationships, though the majority of cattle did not. Counterintuitively, we found that milking time in some groups led to cows being in contact with fewer other cows. Cattle allowed free access to all farm areas spent most of the study period split between the buildings and pasture, and showed substantial variation in the number and duration of contacts at pasture. By analysing interactions in multiple commercial groups of cattle, we find that variations in management, e.g. grazing access, milking routine, can alter the social interactions of cattle, which might in turn enable better understanding of how infections are transmitted among dairy cattle and how transmission varies among herds.

Introduction

Group living offers many fitness benefits to a species, however, a theoretical cost of group living is the general increase in the likelihood of exposure to infections (Kappeler et al., 2015). Host contact rate is a crucial component of the reproductive rate of an infection and modelling contact rate as random process can provide an effective framework to simulate the transmission of some infections (Anderson and May, 1992). Where contact rates vary between individuals and infection is directly transmitted, infections might alternatively be better predicted using networks as a framework to model the transmission of infections (Bansal et al., 2007; Craft, 2015). Including empirically-determined contact networks of beef cattle and horses in transmission models has produced lower estimates of epidemic size and duration when compared to epidemics based on homogenous mixing of contacts (Duncan et al., 2012; Milwid et al., 2019a), and incorporating heterogeneous contact structure into within-farm disease transmission models has enabled the dynamics of *Mycobacterium avium* subsp. *paratuberculosis* (Marcé et al., 2011) and bovine viral diarrhoea infections (Courcoul and Ezanno, 2010) to be more accurately described. Modelling variation in individual contacts showed *Escherichia coli* infections were less likely to persist in smaller groups of cattle (Turner et al., 2008). In extreme cases, individual heterogeneity can alter the flow of an infection such that epidemics might be less likely to occur, but cause more explosive outbreaks when they do occur (Lloyd-Smith et al., 2005). Contact structures that create strong divisions between subgroups can reduce infection rates (Sumner et al., 2018) but can also lead to increased transmission within subgroups and ‘structural trapping’ of disease (Sah et al., 2017). In the interests of parsimony, many network studies consider contacts to be static through time, though in reality this assumption is often violated and the analysis of dynamic networks can be important predicting disease dynamics more accurately (Bansal et al., 2010; Silk et al., 2017b; Springer et al., 2017). For example, models incorporating temporal dynamics using hourly cattle-contact networks have suggested that including this fine-scale temporal variation in contact networks can alter transmission dynamics, especially for infections with smaller reproductive ratios (Chen et al., 2014).

Behavioural studies of contacts among free-ranging and feral cattle show they live in matriarchal dominance hierarchies with evidence of long-lasting social preferences and low immigration and emigration (Bouissou et al., 2001; Reinhardt and Reinhardt, 1981). Studies of cow-calf relationships in small, extensively-kept, beef herds show strong preferences for dams to spend time with their own calf (Duncan et al., 2012; Handcock et al., 2009; Swain and Bishop-Hurley, 2007). However, group structure within modern dairy herds is quite different; dams and calves are separated early in life and dairy cows on farms that calve all year round are typically moved through groups in a ‘cascade’ system, according to their production stage (e.g. lactating or dry), meaning frequent additions to and removals from groups. Dairy herds in the UK are becoming larger (AHDB Dairy, 2019b), which typically requires further divisions into groups, usually according to milk yield. The dynamic nature of these groups can disturb social hierarchies and precipitate agonistic interactions (Bøe and Færevik, 2003), increase social stress (Proudfoot and Habing, 2015), and decrease milk production (Hasegawa et al., 1997). Typically, dairy herds in the UK graze outdoors from March to October and are housed in the winter, however, changes in the dairy industry such as the trend for higher milk yields have had impacts on herd management (Oltenacu and Broom, 2010). Metabolic demands of high-yielding cows are not always met on purely grass-based systems (Fike et al., 2010) and a subsequent rise in ‘zero-grazing’ herds means that more cattle are housed all year round so that they can consume a nutritionally complete ration (Haskell et al., 2006). This can impact upon individual behaviour; studies comparing the three main cow behaviours (feeding, ruminating and resting; Kilgour, 2012) on pasture and in housing show that cows tend to spend less time eating indoors, more time lying outdoors, and there is an increase in agonistic interactions indoors caused by competition for more focussed resources (Arnott et al., 2016). The ability of cattle to synchronise their behaviour, considered an indicator of positive welfare (Napolitano et al., 2009), has been shown to be more likely at pasture than in housing (Tuomisto et al., 2019). Social networks of dairy cattle have been described for cattle in housing (Foris et al., 2018; Gygax et al., 2010), in housing with limited pasture access (Boyland et al., 2016), and at pasture (de Freslon et al., 2019) using a range of methods including focal observations, video observation, proximity loggers and spatial positioning systems. Most studies on modern dairy herds show that they form a single, unstructured group, with some individual variation in sociability and some evidence for social preferences. Cows demonstrate very repeatable behaviour in the order that they enter the milking parlour (Beggs et al., 2018; Rathore, 1982), which is likely to be linked to social structure (Soffié et al., 1976; Wood, 1977), and changes to this order might also be an indicator of poor health (Polikarpus et al., 2015), reinforcing the evidence for links between social behaviour and animal health. Contacts between cattle have been shown to vary in different spaces *within* housing and over time (Chen et al., 2015; Gygax et al., 2010), however no studies have divided contact networks spatially between pasture and when cattle are in buildings.

To better understand patterns of social contacts among cattle on dairy farms, we recorded the interactions of nine groups of cows on seven commercial dairy farms in south-west England. We used automated proximity sensor technology to collect continuous high-resolution contact data and combined this with data from GPS devices to describe the cow's location on the farm. Our herds represented a range of management practices in terms of their milking regimes, housing, grazing access, and group sizes, all of which might be expected to affect the social interactions between herd members. Here we describe the social networks of these herds at different spatial and temporal scales; comparing networks of cattle in buildings to networks at pasture and analysing contact frequencies at weekly (the whole study period), daily and 2-hourly time aggregations. We predicted that interactions would differ among locations on the farm, with cattle more able to express social preferences at pasture, and more numerous contacts while in buildings. We focussed on network characteristics that might influence the transmission of pathogens: heterogeneity in contact rates and durations, the formation of sub-groups or 'communities', and the relative strength of connections within and between these communities.

Methods

Deployments and farms

Our herds were representative of small to average sized commercial dairy herds in the UK. We have anonymised farms and refer to them by a descriptive feature of the group (Table 5.1). On most farms (Night-housed, Strip-grazed, Free, Rotation 1a, Rotation 1b, Rotation 2), we recorded contacts among cows that were kept in a milking group separate to other animals in the herd. However, cows in the largest herd were split into several groups and we recorded interactions among the dry cows (Dry) and among the low-yielding milking cows (Housed). No changes were made from the normal routine of the farms, therefore cows were added to and removed from groups during most deployments as they calved in or finished their lactations. In contrast, the

Stable group included all milking and dry cows and thus no changes occurred to the group composition during the study period. Grazing management was also varied and ranged between allocation of additional strips of grazing after each milking (Strip-grazed), through grazing a different field or part of a field after each milking (Rotation 1a, Rotation 1b, Rotation 2, Stable, Night-housed), to set stocking in one field (Dry) and free range of all fields on the farm (Free). Supplementary forage (e.g. silage, hay) was often replenished after milking times.

In general, groups were kept out at pasture and brought into buildings only for milking, with four exceptions: 1) the Housed group were kept in cubicle housing during the study week due to insufficient grazing during a spell of hot weather, 2) the Night-housed group were in cubicle housing at night, 3) the Rotation 1b group were allowed access to buildings at all times and were kept in for two nights and days in the middle of the study period due to inclement weather and 4) the Free group were allowed free access to pasture, cubicle housing, and the automated milking system (AMS) at all times during the study. Group location at pasture or in buildings was always governed by the farms, except in the case of the Free group, where cows had free access to all areas of the farm, and the Rotation 1b group where they had access to the daily selected pasture and housing. ‘Buildings’ networks for the Night-housed, Rotation 1b, Housed, and Free groups represent milking times *and* being housed in buildings, whereas ‘Buildings’ networks for the Strip-grazed, Rotation 1a, Rotation 2, and the Stable groups represent just milking times (Table 5.1). Cows not present for the entire study period or for which we had incomplete data were removed from our analyses. All fieldwork was approved by the University of Exeter College of Life and Environmental Sciences (Penryn Campus) ethics committee (eCORN000087 v4.6).

Table 5.1. *Details of farm management and data collected from nine groups of cattle on seven Cornish dairy farms in Summer and Autumn 2018. Details include breed of cattle, grazing management, milking routine, group sizes, number of collars with GPS and proximity sensors that were deployed, and data quality. Details are omitted for two study groups removed from analyses due to data loss from more than 50% of cows.*

Deployment name	Start month of study	Breed of cattle	Source of replacement cows	Group monitored	Calving system	Cow access to farm locations		Grazing type	No. milkings per day	No. cows in group		Bulls in group	Study Period (days)	No. collars deployed	No. cows with complete proximity data	Percentage of the group with data
						Buildings	Pasture			Start	End					
Night-housed	Oct-18	Holstein Friesian	Multiple source farms	Milking	All year round	Milking and night	Day	Rotational grazing	2	98	97	2	6.77	98	87	88.8
Strip-grazed	Aug-18	Ayrshire	Home-bred	Milking	All year round	Milking	Day and night	Strip grazing	2	52	54	0	6.83	52	44	84.6
Free	Aug-18	Holstein Friesian	Home-bred	Milking	All year round	Access at all times	Access at all times	Free roam of pastures	Automated milking system	60	60	0	6.85	60	50	83.3
Rotation 1a	Aug-18	Holstein Friesian / Jersey cross	Home-bred	Milking	Autumn calving	Milking	Day and night	Rotational grazing	1	80	96	0	6.77	80	64	80.0
Rotation 1b	Oct-18			Milking		Milking and access at all times	Day and night	Rotational grazing	2	98	112	0	6.79	98	87	88.8
Dry	Sep-18	Holstein Friesian	Multiple source farms	Dry	All year round	No access	Day and night	Set stocking	0	33	32	0	6.76	37	22	59.5
Housed	Sep-18			Milking		Day and night	No access	NA	2	111	106	1	6.76	111	100	90.1
Rotation 2	Oct-18	Holstein Friesian	Home-bred	Milking	All year round	Milking	Day and night	Rotational grazing	2	175	181	0	7.04	177	95	53.7
Stable	Sep-18	Mixed	Multiple source farms	All cows	Spring calving	Milking	Day and night	Rotational grazing	1	59	59	1	6.84	57	37	64.9
									Mean:	85.1	88.6	0.4	6.8	85.6	65.1	77.1

Equipment

Nylon cattle collars with a plastic clasp (Suevia Haiges, Germany) were fitted with a proximity device and a GPS receiver such that one device lay at either side of the animal's neck. The GPS receivers (i-GotU GT-120 and GT-600 devices, Mobile Action Technology Inc., Taiwan) were configured to record fixes every ten minutes. The proximity device is based on a design by the OpenBeacon project (http://www.openbeacon.org/) and the SocioPatterns collaboration consortium (http://www.sociopatterns.org/) and has been used in contact studies of humans, horses, and dogs (Cattuto et al., 2010; Milwid et al., 2019a; Wilson-Aggarwal et al., 2019). The Radio Frequency Identification Devices (RFID) exchange low-power radio packets in a peer-to-peer fashion, using the signal strength to estimate the distance between devices (Cattuto et al., 2010). Using previous studies in humans and our own validation analyses (see Appendix E Validation), it was determined that attenuation of less than -70dBm was an appropriate threshold to approximate contacts of cattle between 1 and 1.5m. This distance was deemed appropriate for detecting affiliative behaviour or potential transmission of pathogens that may be spread oro-nasally or by direct contact (Cattuto et al., 2010). Contacts are best recorded when the devices are face-to-face, therefore when cows were positioned in close proximity but with tags facing opposite or the same direction, observations in our validation analysis (Appendix E Validation) suggested that contacts might have been underestimated. Observations during validation analyses (Appendix E Validation) also suggested that there may be a possibility of more distant contacts being recorded when metal structures are around due to signal propagation, possibly leading to over-estimation of contacts in this environment. However, in the validation analysis, post-data processing removed this false positive from the data and therefore the cleaned data are likely to still provide a good approximation of contact patterns between cattle, as they have done in previous studies in similar indoor barn environments (Milwid et al., 2019b). We selected a temporal threshold of 20 seconds to ensure contacts were not spurious, the contact was maintained as long as signal was sent or received every 20 seconds, if no signal was sent or received, the contact ended. Contact durations are therefore measured in 20 second blocks (Cattuto et al., 2010). Some tags (n = 16) recorded abnormally high contacts during short timeframes (Appendix E Fig. E1), therefore we removed any tags that recorded more contacts than 95% of the total contacts recorded by all tags within a 30 minute time frame, as we considered them highly unlikely to be biologically feasible. Some tags recorded very low numbers of interactions, however the timing of the contacts consistently through the study was considered biologically plausible, so they remained in the analysis.

Network and statistical analysis

1. Do contact patterns vary through time and space?

We calculated the mean frequency of contacts between all group members for each group for the whole study period, and at two-hourly and daily intervals. Contacts from the whole study period were aggregated to form the 'Full networks', where cows were represented as nodes and contacts were represented as undirected edges. Each combination of cow pairs is called a 'dyad'. Edges were weighted by the total duration of contact between the two cows for the relevant time period, at an appropriate scale, e.g. minutes-per-day on Full networks.

To detect differences in contacts when cows are in buildings or out at pasture, we approximated herd location based on GPS data. Areas of the farm were split into 'Buildings', comprising any part of the farm where cattle would be collected together under cover, including the collecting yard, milking parlour, housing, loafing area, etc., and 'Pasture' that was in use during the study period (using open-source QGIS software; QGIS Development team, 2019). For each 30-minute window of the study period, we defined a cow's location as where over 50% of GPS fixes were recorded. Herd location (Pasture or Buildings) was defined as the area containing over 75% of cows for the same 30 minute time window. Where less than 75% of the herd was in one location, we categorised this time as 'Split'. We aggregated these 30-minute windows by their location to create three 'Spatial networks', termed the 'Buildings network', 'Pasture network' and 'Split network'. We compared the relative proportion of time spent in each spatial location with the frequency of contacts that occurred in each location. We tested correlation between these networks using a quadratic assignment procedure from the R package 'sna' (Butts, 2016). Due to the high density of the Full and Spatial networks, with most cows in contact at some point during the study period, we assessed the networks based on stronger interactions by filtering out weaker edges. We removed edge weights below the 50th, 75th and 90th percentiles of edge weights in the unfiltered network, for example, F50 networks removed cow to cow relationships that were in the lowest 50th percentile of time spent together (James et al., 2009). We also tested correlations between Full and Spatial networks at each of these filtered levels (F50, F75, F90).

2. Do contact frequency and duration vary between cows?

We assessed the centrality of cows within each farm by calculating node degree (the total number of in-contact cows) and strength (the total amount of time each cow has spent with any other cow) for each cow. Heterogeneity in contact rate between cows is well-known for affecting disease transmission (Lloyd-Smith et al., 2005), and can be assessed by the coefficient of variation as per the following equations;

$$CV_{degree} = \frac{\sigma_{degree}}{\mu_{degree}} \qquad\qquad CV_{strength} = \frac{\sigma_{strength}}{\mu_{strength}} \qquad\qquad (1,2)$$

where CV is the coefficient of variation, σ is the standard deviation, and μ is the mean of the degree and strength distributions respectively (May, 2006). In general, higher CV values indicate increasing variation between values, therefore CV_{degree} that tends towards zero indicates a more homogenously mixing population (May, 2006). CVs were calculated for degree and strength on the Full network and Spatial networks.

As network parameters are inherently inter-connected and thus violate the assumptions of independence required of conventional statistical techniques, we constructed null, randomised networks to test if values from the random networks were statistically significantly different from values calculated from our observed networks (Croft et al., 2011). The static networks were randomised by creating new Erdös-Rényi graphs (Erdös and Rényi, 1959) with the same number of edges and nodes as the original network. We then randomly allocated edge weights from the observed network to the new edges in the Erdös-Rényi network. Metrics were considered statistically significantly different from random if their values lay outside the upper and lower 95% bounds of 4999 randomised values, i.e. where $P < 0.05$ (two-tailed).

3. Do cows exhibit social preferences and are they consistent over time?

To see if cows preferentially spent more time with particular cows, we measured the coefficient of variation in contact times that each cow spent with other cows (CV_{cddyad}) for each individual as per Equation 1 but using the distribution of contact durations for each dyad rather than the degree distribution. In order to randomise dyadic contact durations rather than just the overall strength distribution per cow, we took the raw list of contacts (the edgelist) and randomly re-allocated contact durations among individual contacts to create a random edgelist. The random edgelist was then aggregated by summing the weights from individual contacts for each dyad. We then created a new unweighted Erdös-Rényi network with the same number of nodes and edges as the random edgelist. The aggregated edge weights from the random edgelist were then randomly allocated to new edges in the Erdös-Rényi network, thereby keeping the overall amount of time cattle spent together the same as the observed network, but changing how this was distributed between animals. We tested the hypothesis that the CV_{cddyad} values of cows in observed networks were greater than the CV_{cddyad} values of cows in randomised networks and calculated an empirical *P* value (two-tailed) as per the following equation:

$$P = \frac{number\ of\ CV_{cddyad\ observed} \leq CV_{cddyad\ random}}{n} \qquad (3)$$

where n is the number of CV_{cddyad} from random networks (4999) plus the observed CV_{cddyad} as it might have feasibly come from a random distribution (Boyland et al., 2016). For example, *P* = 0.025 suggests that the observed value is less than or equal to than the random value in 2.5% of randomisations.

To assess the tendency of cows to repeatedly interact with the same cows over time, we measured variation in contact between each dyad over time. For each dyad, in each two-hour block of the study period, we attributed a 1 if there was contact and a 0 if there was no contact during that time. For a random comparison, at this stage we randomly distributed the 1s and 0s among all dyads, and thereafter performed the same analysis on the observed and random two-hour blocks. We calculated the proportion of two-hour blocks in which cows interacted per 24 hours, e.g. 6/12 = 0.5. We then calculated the coefficient of variation of the proportions from each study day ($CV_{proportion}$) for each dyad, where a low $CV_{proportion}$ would indicate little variation in the amount of time they spent together between each study day and a higher $CV_{proportion}$ would indicate cows spending increasingly varied amounts of time with each other. We removed dyads with no connections at all from the analysis to focus on the variation in contacts rather than non-contacts. We tested the hypothesis that the $CV_{proportion}$ of dyads in observed networks were more consistent, i.e. had less variation between days, than the $CV_{proportion}$ of dyads in random networks and calculated P values (two-tailed) as per the following equation

$$P = \frac{number\ of\ CV_{proportion\ observed} \geq CV_{proportion\ random}}{n} \tag{4}$$

where n is the number of $CV_{proportion}$ from random networks (4999) plus the observed $CV_{proportion}$ as it might have feasibly come from a random distribution.

4. Do cow's form discrete communities and are they consistent over time?

We performed community detection on unfiltered, unweighted Full and Spatial (Buildings, Pasture and Split) networks using the fast greedy algorithm (Clauset et al., 2004), as implemented in 'igraph' (Csardi and Nepusz, 2006). As the modularity value of a network is related to the number of nodes, and our networks are of different sizes, we used a relative measure, as outlined by Sah et. al (2017). We calculated the maximum modularity that could be achieved with a network of that size (Q_{max}; Sah et al., 2017) and the modularity of the observed network (Q), and then calculated the relative modularity (Q_{rel}) as per the equation below.

$$Q_{rel} = \frac{Q}{Q_{max}} \tag{5}$$

Due to the high density of the Full and Spatial networks, we also performed the community detection and modularity analysis on these networks filtered by edge weights below the 50th (F50), 75th (F75), and 90th (F90) percentiles of contacts.

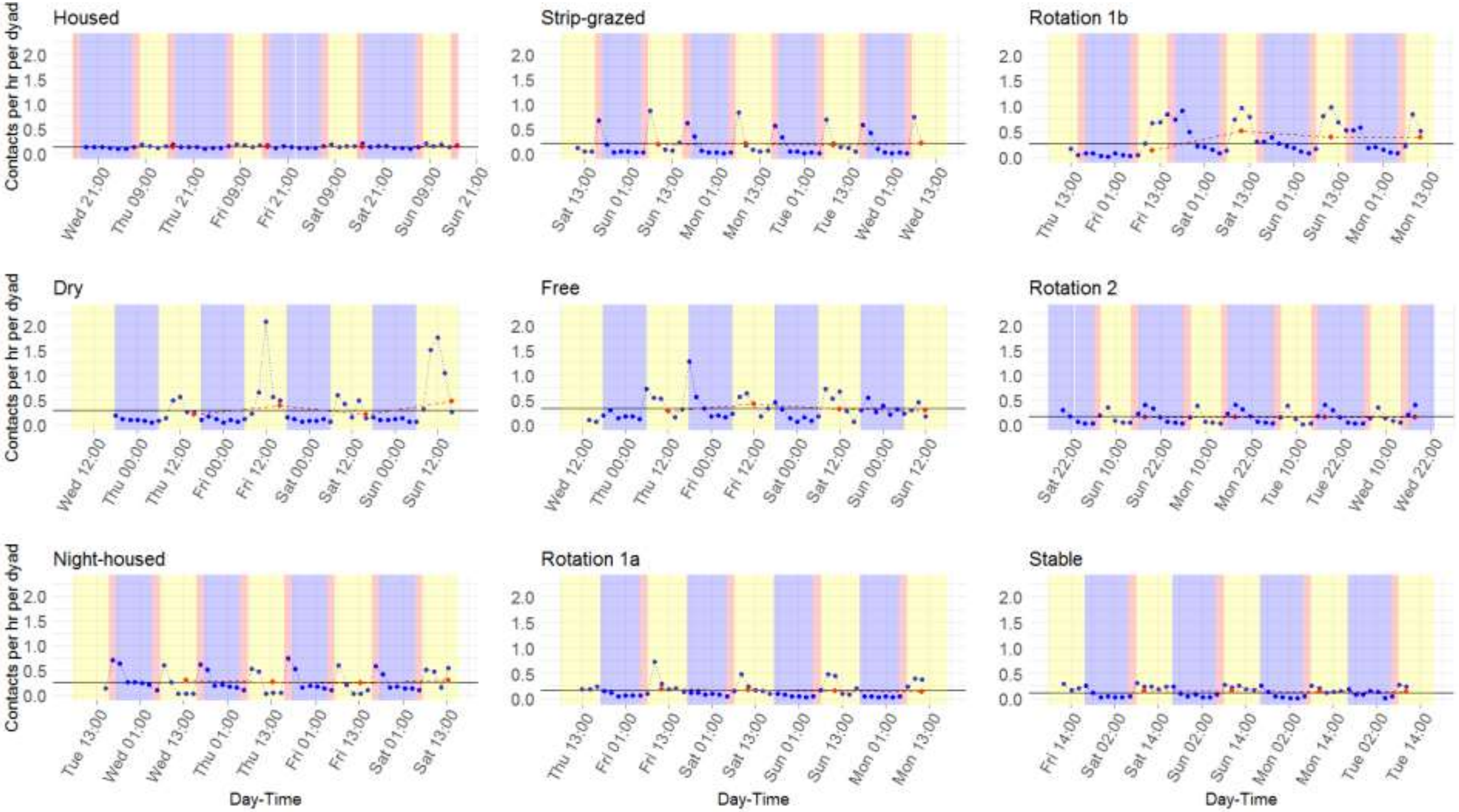
To assess if the composition of communities were repeatable over time and in different locations, we extracted the raw contacts that occurred in each location (Buildings, Pasture, and Split). We then grouped contacts that occurred during a similar time period, divided by a change in location or a change from day (defined as 07:00–19:00) to night, into several 'Temporal-Spatial' networks termed 'Temporal Pasture', 'Temporal Buildings' and 'Temporal Split' networks (Appendix E Fig. E2). If groups had less than three comparable networks for a certain category, they were removed from the analysis. For each time period we created a weighted network with aggregated contacts and calculated communities using the fast-greedy algorithm. For each dyad in each network, we assigned 1 if both cows were in the same network and 0 if they were in a different network. For a random comparison, at this point we randomly distributed the 1s and 0s among dyads (n = 4999) and then proceeded with the same analysis for randomised and observed data. We calculated the ‘repeatability’ (estimated from the variance component of a generalised linear mixed model using a binomial error structure in the package ‘rptR’ (Stoffel et al., 2017) to determine the stability of dyads being in the same community in multiple networks over time. Repeatability values can range from 0 indicating a lack of stability in community membership, to 1, indicating perfectly stable community membership.

All data analysis and manipulation was performed in R (R Core Team Version 3.5.3, 2019), unless otherwise stated. Networks were constructed and network measures (degree, strength, community detection and modularity) were calculated in the R package “igraph” (Csardi and Nepusz, 2006).

Results

We deployed collars on 11 groups of dairy cows over 8 farms. Group size ranged from 33 to 177, and a total of 770 cattle were collared over all deployments. We removed deployments completely if we had full proximity data for less than 50% of cows, which resulted in the loss of two groups, mainly due to water-damaged proximity loggers. We retained 9 groups of dairy cows on 7 farms for analysis. Complete proximity data were available for a total of 585 cows on these farms and the number of animals in our networks ranged from 22 to 100 (mean = 65 cows, standard deviation (SD) = 28 cows; Table 5.1). The mean duration of time cows spent in close proximity to other cows over all groups was 27 minutes per day (SD = 11 minutes, range among herds = 19–39 minutes). Average contact rate per dyad per hour over the study period among groups had a mean of 0.20 (SD = 0.07, range among herds = 0.13–0.32; Fig. 5.1) with the Free group and the Stable group having the highest and lowest mean contact rates respectively. Edge densities of all Full networks were very high, although networks were never fully connected (mean among herds = 0.95, SD = 0.03, range = 0.92–0.99).

Figure 5.1. *(overleaf) Temporal variation in numbers of contacts recorded between cows in nine groups of dairy cattle on seven commercial farms. Contacts were recorded by RFID proximity loggers and contact frequency is averaged per hour per dyad (cow-cow pair) and aggregated into varying time periods; the whole study period (mean = 7 days; black solid line), daily (red circles and dashed lines), and 2-hourly (blue circles and dotted lines). Daily and 2-hourly values are selected from the middle four days of the study period for brevity. Milking regimes (light red shading) varied by group and timings plotted here are those reported by farmers. There is little variation day-to-day within a farm but strong within-day patterns are evident in most farms, typically with higher contacts in the day (yellow shading) compared to night (blue shading) and higher contacts around dawn and dusk, which frequently align with milking times.*


Housed
Strip-grazed
Rotation 1b
Dry
Free
Rotation 2
Night-housed
Rotation 1a
Stable
Contacts per hr per dyad
2.0
1.5
1.0
0.5
0.0
Wed 21:00
Thu 09:00
Thu 21:00
Fri 09:00
Fri 21:00
Sat 09:00
Sat 21:00
Sun 09:00
Sun 21:00
Sat 13:00
Sun 01:00
Sun 13:00
Mon 01:00
Mon 13:00
Tue 01:00
Tue 13:00
Wed 01:00
Wed 13:00
Thu 13:00
Fri 01:00
Fri 13:00
Sat 01:00
Sat 13:00
Sun 01:00
Sun 13:00
Mon 01:00
Mon 13:00
Wed 12:00
Thu 00:00
Thu 12:00
Fri 00:00
Fri 12:00
Sat 00:00
Sat 12:00
Sun 00:00
Sun 12:00
Wed 12:00
Thu 00:00
Thu 12:00
Fri 00:00
Fri 12:00
Sat 00:00
Sat 12:00
Sun 00:00
Sun 12:00
Sat 22:00
Sun 10:00
Sun 22:00
Mon 10:00
Mon 22:00
Tue 10:00
Tue 22:00
Wed 10:00
Wed 22:00
Tue 13:00
Wed 01:00
Wed 13:00
Thu 01:00
Thu 13:00
Fri 01:00
Fri 13:00
Sat 01:00
Sat 13:00
Thu 13:00
Fri 01:00
Fri 13:00
Sat 01:00
Sat 13:00
Sun 01:00
Sun 13:00
Mon 01:00
Mon 13:00
Fri 14:00
Sat 02:00
Sat 14:00
Sun 02:00
Sun 14:00
Mon 02:00
Mon 14:00
Tue 02:00
Tue 14:00
Day-Time

1. Do contact patterns vary through time and space?

Contact frequency in most groups varied substantially when contacts were aggregated into 2-hourly windows (Fig. 5.1), with higher numbers of contacts coinciding with management practices, e.g. milking, which tended to occur early morning and early evening, or feeding, which tended to occur after milking. Although cows were not specifically gathered for milking or feeding in the Free group, there were still increases in contacts around dawn and dusk. Patterns of contact frequencies typically followed 24-hour (e.g. Fig. 5.1, Rotation 1a and 1b) or 12-hour cycles (e.g. Fig. 5.1, Strip-grazed, Rotation 2, Stable), thus the number of contacts aggregated over 24-hour periods were consistent between days. Mean daily contact frequencies were similar to mean values calculated from contacts during the entire study period (Fig. 5.1). Contact frequencies over 2-hourly periods in the Housed and Stable groups showed much less temporal variation relative to all other groups (Fig. 5.1). In groups with access to pasture, contact rates were close to zero at some points in the study, more often during the night, whereas contact rates did not fall this low when groups were housed (Fig. 5.1).

More contacts occurred in Buildings networks, compared to Pasture and Split networks, relative to the amount of time spent in these areas (Fig. 5.2). On average, cows had contacts of longer durations in Buildings than at Pasture (Fig. 5.3). For groups housed in buildings (Night-housed, Rotation 1b), and the Free and Rotation 2 groups, most cows contacted almost all other cows in the Buildings networks, and contacted fewer cows in Pasture networks (Fig. 5.3). However, where cattle were only brought into buildings for milking, most cows contacted fewer cows in Buildings networks (Stable, Rotation 1a), or showed no difference in proportion of cows contacted (Strip-grazed) compared to Pasture networks. When cows had access to all areas of the farm, the group was most often split between buildings and grazing, suggesting that cows prioritised preference for a particular location over staying together in one group.

Figure 5.2. *Relative amount of time and number of contacts recorded between cows in seven groups of dairy cattle divided by area of the farm. Pie charts show the relative proportion of time and number of contacts recorded by proximity loggers in Buildings (brown) and at Pasture (white). Herds are defined as 'Split' when less than 75% of cows are in either Buildings or Pasture (purple). Two study groups are not shown as they were at grazing or housed for the study duration. Most cattle only had access to buildings or pasture at specified times, however, the Free group was allowed access to all farm areas and elected to spend the majority of time split into two groups.*

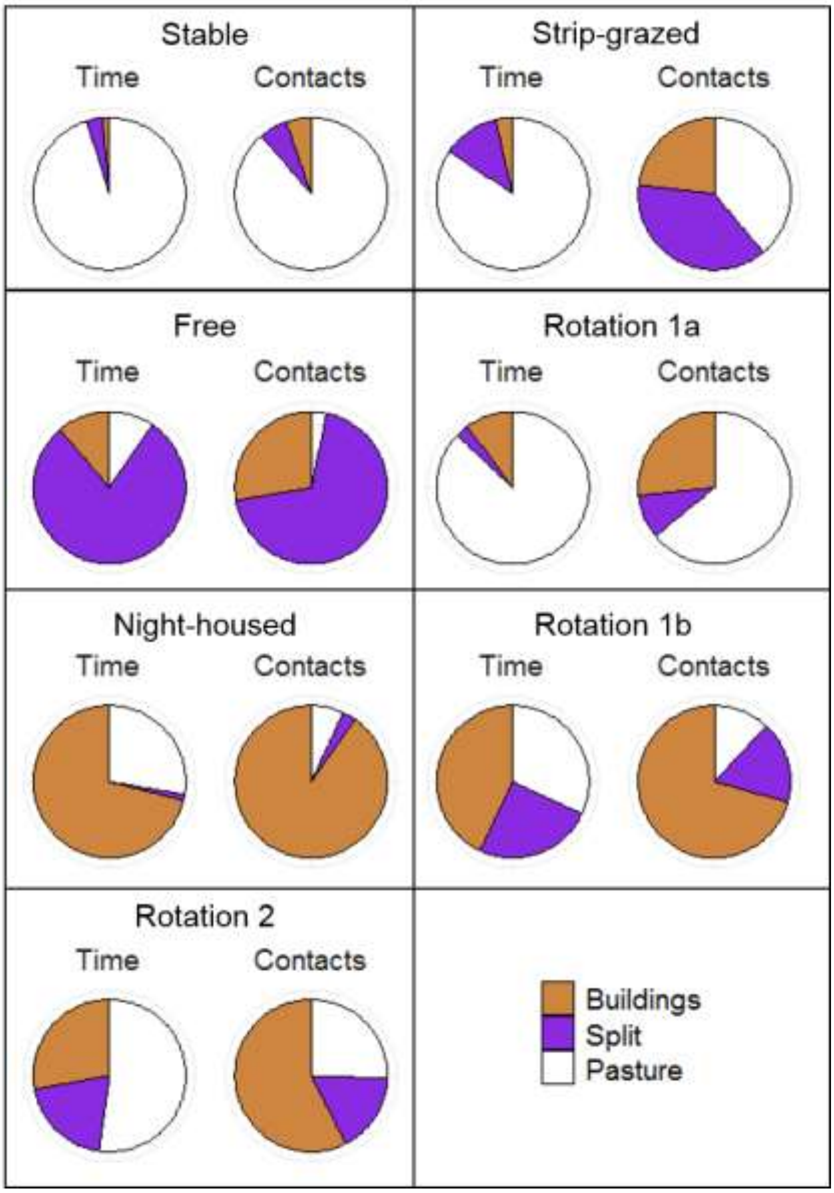
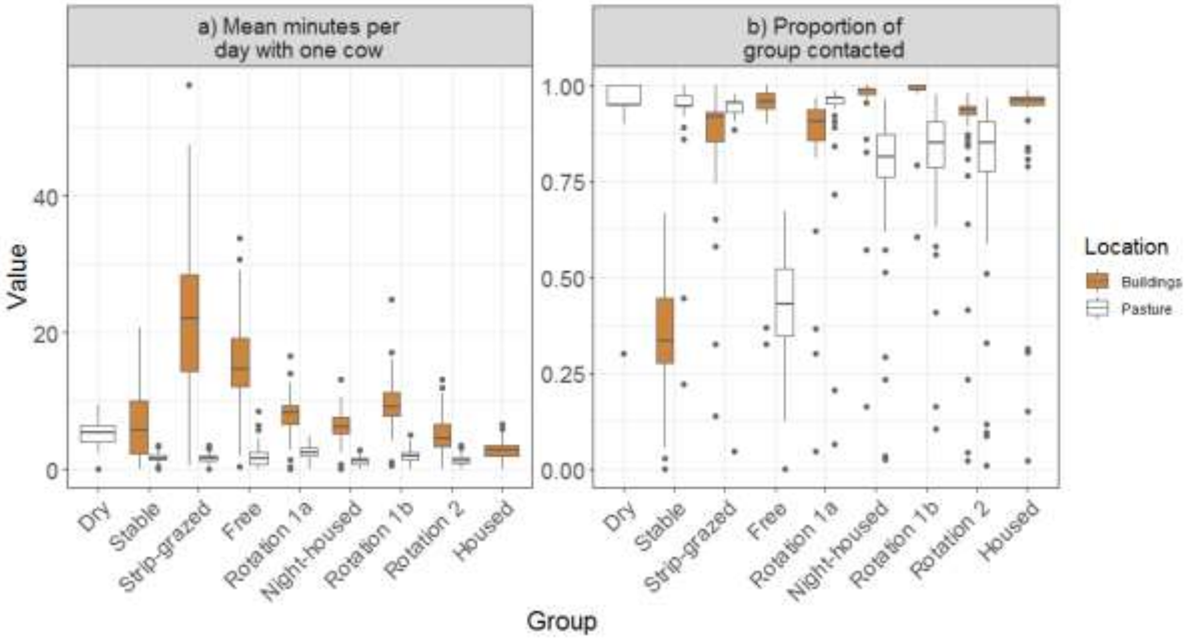


Figure 5.3. *Distributions of time cows spent with other cows and number of cows contacted in nine groups of dairy cattle at pasture and in buildings. Plot a) shows the mean amount of time cows spent with other cows. Plot b) shows the proportion of cows in the group with which they came into contact during the whole study period (mean = 7 days). Values are calculated for Pasture networks (white) and Buildings networks (brown). Groups are ordered by ascending group size from left to right. Boxplots show the median, 25th and 75th percentiles of values and and the upper and lower whiskers extend to the largest or smallest value no further than 1.5 times the interquartile range, data beyond this range are plotted as outlying points. On all comparable groups, less time is spent in social contact at pasture and on most farms fewer cows are contacted at pasture compared with buildings.*



Edge densities varied among Spatial networks, with Buildings networks (mean edge density among groups = 0.85, standard deviation = 0.20) and Pasture networks (mean = 0.82, SD = 0.17) more dense than the Split networks (mean = 0.73, SD = 0.20). In the quadratic assignment procedure analysis, Spatial networks per day were generally not correlated with each other (*r* = < 0.0000–0.0972, n = 14; Appendix E Table E1), except for the unfiltered Buildings and Split networks for the Strip-grazed group (*r* = 0.3, *P* < 0.001, Appendix E Table E1). Filtering networks to retain only the strongest contacts did not noticeably affect correlation scores.

2. Do the number of contacts and contact duration vary between cows?

Variation among cows in their number of contacts (CV_{degree}) and their total contact duration with other cows ($CV_{strength}$) was greater in the Full, Buildings, and Pasture networks than in randomised networks (Fig. 5.4). In all cases, there was more heterogeneity in strength than degree (Fig. 5.4), likely due to dense networks, limiting the extent of variation possible for degree. Greater between-cow variation was found in Pasture networks, and this was most marked in the Free group where CV values on Pasture networks were more than double those found in Buildings networks (Fig. 5.4). Groups which were in buildings for milking only had greater between-cow variation in Buildings networks compared to Pasture networks (CV_{degree}: Stable, Rotation 1a, $CV_{strength}$: Stable, Strip-grazed; Fig. 5.4). Differences in individual variation between farms did not seem to be associated with group size (Fig. 5.4).

3. Do cows exhibit social preferences and are they consistent over time?

Some cows in each group showed greater variation in the duration of time they spent with individual cows (CV_{cddyad}), i.e. evidence for social preference, compared to cows in randomised networks (Fig. 5.5). The mean proportion of cows that exhibited greater variation than random and spent a substantial amount of time (over 21 minutes per day) in contact with their closest contact was 7.5% among groups (SD = 5.1%, range = 2.0–16.0%). The majority of cows in each deployment interacted more equably with all other individuals in the herd, akin to what might be expected at random (Fig. 5.5).

A small proportion of dyads in each group (mean = 3.7%, SD = 1.5%, range = 2.1–5.2%) spent more-consistent ($P < 0.025$) amounts of time together from day to day than would be expected if contacts re-assorted at random (Appendix E Fig. E3). On average these ‘consistent dyads’ spent a mean time of 30 minutes (SD = 11 minutes, interquartile range (IQR) = 25–38 minutes) together per 120-minute window, compared to a mean of 18 minutes (SD = 10 minutes, IQR = 12–25 minutes) in the rest of the dyads. Over all groups, 14 dyads were recorded spending a mean of over 60 minutes with each other out of every 120 minutes during the study period.

Figure 5.4. *Variation between cows in number and duration of contacts between cows in nine groups of dairy cattle compared to variation in random networks. a) Variation in number of other cows each cow contacted (Coefficient of Variation of degree = CV_{degree}) b) Variation in duration of contact ($CV_{strength}$) between-cows. Upwards-facing triangles indicate observed values above the upper 95% bound of random network values calculated for the whole study period ('Full network'; red), and for contacts only on pasture ('Pasture network'; white) and in buildings ('Buildings network'; brown). Groups that stayed in a single area for the whole study period are coloured by that area. Groups are ordered by ascending group size from left to right. Boxplots represent the distribution of CV values calculated from randomised networks (n = 4999). Median, 25th and 75th percentiles of values are shown and the upper and lower whiskers extend to the largest or smallest value no further than 1.5 times the interquartile range, data beyond this range are not plotted. Variation in the strength of contacts was greater than the variation in number of contacts.*

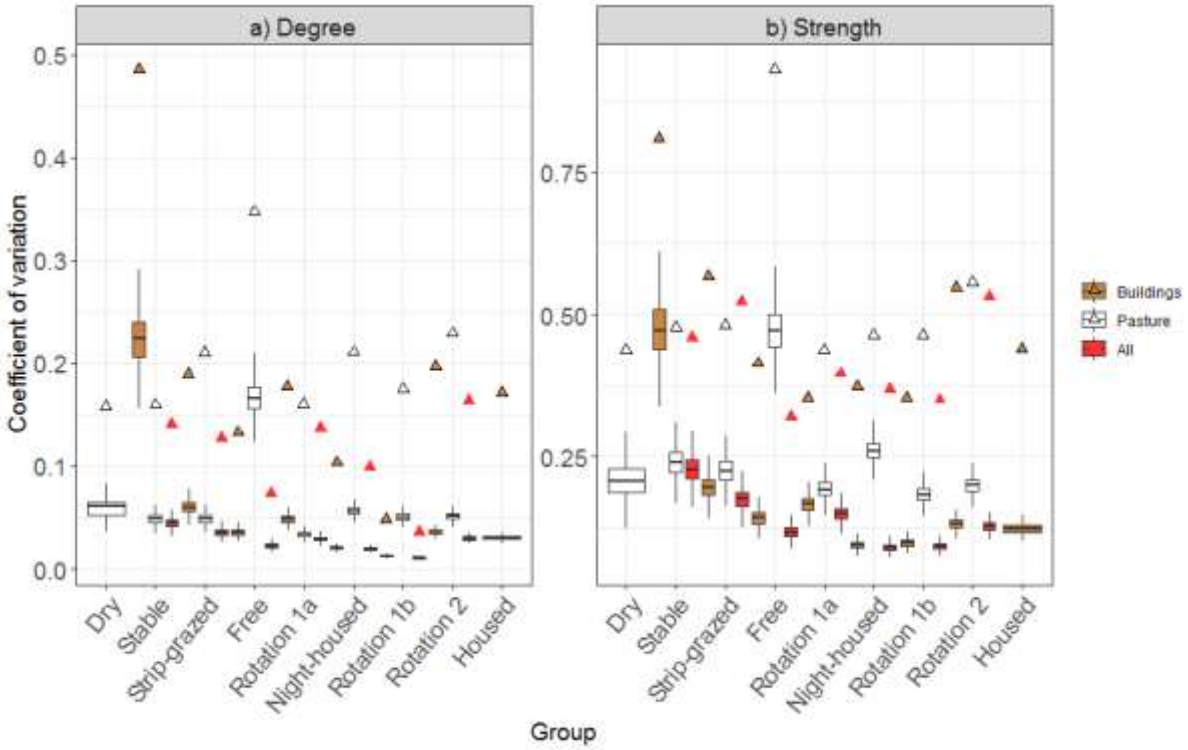
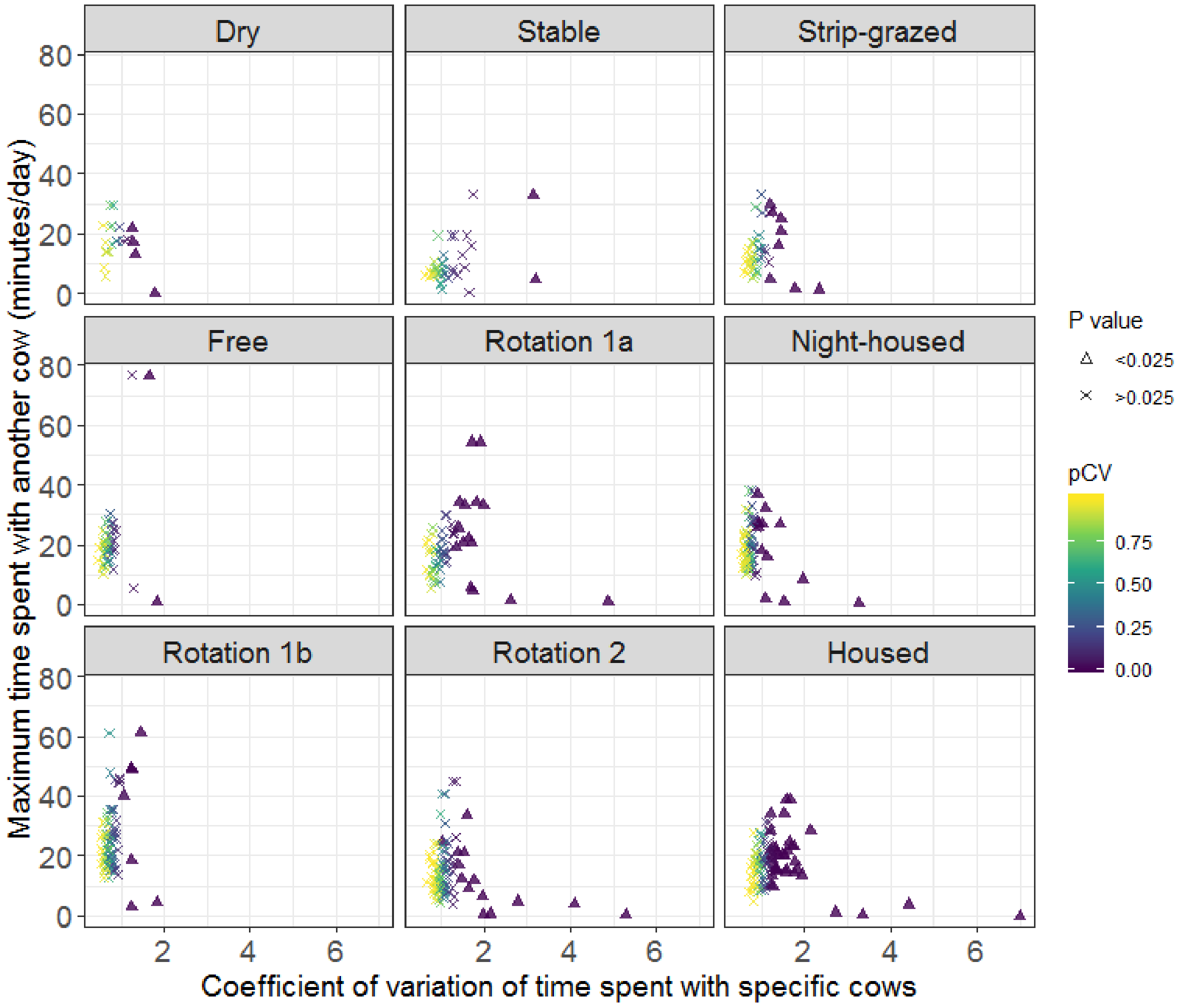


Figure 5.5. *Social preference based on cow-cow interactions in nine groups of dairy cattle. Plots show the variation in time spent with other cows (CV_{cddyad}), compared with the maximum time a cow spent with a single herd mate. Each cow is represented by a point, coloured by the P value for the coefficient of variation (CV) compared to CV values calculated on 4999 random networks (P values < 0.025 are shown as triangles and > 0.025 are shown as crosses). A small number of cows in each group show preference for spending more time with particular other cows, however, the variation in the interactions of most cows did not differ significantly from random.*



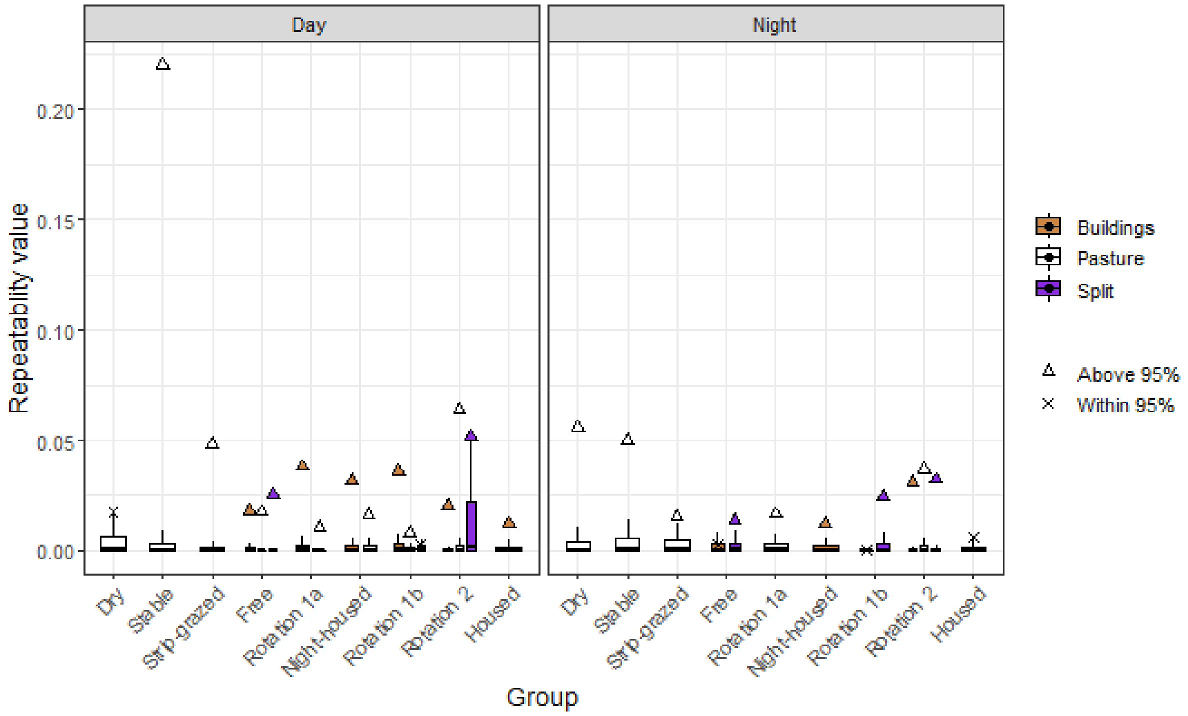
4. Do cow’s form discrete communities and are they consistent over time?

Community detection methods tend to work more effectively on sparse networks (Clauset et al., 2004) and thus only 2 or 3 communities were found in the denser, unfiltered, unweighted networks (Appendix E Fig. E4). These networks also tended to have very low relative modularity values (*Qrel* < 0.03, Appendix E Table E2 and Fig. E6). As networks were filtered and weaker edges were removed, more communities were detected. The largest numbers of communities (≥ 20) were found in F90 networks in the Strip-grazed Split network and the Free Pasture network (Appendix E Fig. E4), however the relative modularity between these communities did not differ from random (Appendix E Fig. E5).

Split networks in the Stable (NF), Strip-grazed (F50), and Free (NF and F50) groups had relative modularity that was significantly greater than random (Appendix E Table E2 and Fig. E5). Otherwise, observed modularity was rarely statistically significantly greater than that of random networks (Appendix E Table E2 and Fig. E5) and the lack of a consistent pattern of significant results suggests the few cases where observed values differ statistically significantly from random could potentially be a type I error associated with multiple testing.

Although the consistency of our observed communities was statistically significantly higher than expected at random, the absolute observed repeatability values were still very low, indicating that the communities we detected in the Temporal networks did not consistently contain the same cows (Fig. 5.6). However, the Temporal Pasture Day networks of the Stable group were much more consistent than any other categories (day R=0.22 cf. night R = 0.05; Fig. 5.6)

Figure 5.6. *Consistency of community membership over time in contact networks of nine groups of dairy cattle. Several Temporal-Spatial networks were constructed for each group from the Temporal Pasture, Temporal Buildings and Temporal Split networks (between buildings and pasture) and divided into day (defined as 07:00–19:00) and night (see Fig. S2 for schematic). Categories (i.e. Split Night) were excluded if there were fewer than three comparable networks to define repeatability, therefore values for some networks are not present. Consistency of community membership was assessed by calculating the repeatability of cow-cow pairs being in the same community in sequential Temporal-Spatial networks. Repeatability values can lie between 0, indicating a lack of stability in community membership and 1, indicating perfectly stable community membership. Boxplots show the distribution of repeatability values from random networks and triangles and crosses represent observed repeatability values. Triangles indicate repeatability values that lie above 95% of random network values and crosses indicate repeatability values within 95% of random network values. Boxplots show the median, 25th and 75th percentiles of repeatability values calculated on randomised data (n = 4999) and the upper and lower whiskers extend to the largest or smallest value no further than 1.5 times the interquartile range, data beyond this range are not plotted. Groups are arranged by ascending group size.*



Discussion

Against a background of very dense, highly connected networks when viewed as a whole, our finer-scale spatio-temporal analyses show evidence of temporal variation in contact frequency, differing contact patterns between locations and evidence of stronger bonds between a small subset of animals.

Contact durations were longer in Buildings compared to those at Pasture, reflecting the evident space restrictions in bringing cattle closer together. For most deployments, the proportion of the group contacted was also higher in buildings, suggesting that opportunities for infection transmission via close contact to most animals are likely to be increased when animals are housed indoors, at least in the types of housing that were used on our study farms. In contrast to the groups housed indoors, in three groups that were only in buildings for milking, the average number of individuals contacted during these times was less than seen at pasture. We suggest this is due to the milking routine on these farms; cattle are herded into a collecting yard, small groups of cattle come into the parlour and then after milking are allowed to walk back to pasture in their own time. Contrary to typical assumptions that milking time tightly groups animals together (Barlow et al., 1997; Bekara et al., 2014), we have found that it might actually represent a time where cattle are more clearly divided into separate groups, depending on the routine of the farm. Despite cows in the Free group not being gathered specifically for milking (as they used an automated milking system), we still saw peaks in the frequency of contacts around dawn and dusk which may reflect a historical pattern of milking or feeding around these times (as the farm introduced the automated milking system in the previous year) or could suggest that there are times of day, regardless of management, where cows are naturally more interactive (Stoye et al., 2012). The networks formed by the Strip-grazed group during the transition to milking (Split networks) and during milking (Buildings networks) showed more similarity than seen in other groups, which could show a tendency for these cows to travel from pasture into milking and from milking in a similar order, as well as actually going into the parlour in a similar order, as demonstrated in other studies (Beggs et al., 2018; Bouissou et al., 2001).

We predicted that more communities might form and modularity might be higher at pasture, as cows would have the opportunity to space out. Indeed, when cattle were at Pasture there were some two-hour periods with hardly any contacts at all, particularly at night, perhaps showing a preference for cattle to be further apart at times when there is less competition for resources. However, there was no appreciable difference between the number of communities formed, or the modularity of the Pasture networks, compared to the Buildings networks, with one exception. On the farm where cattle were free to roam in any fields they chose, there were many more communities detected and there was greater variation in the number and duration of contacts among individuals. The limitations of one field, or part of a field, available to strip-grazing and rotational-grazing herds mean that cattle tend remain as one unit, whereas the additional options available when cattle can access multiple fields evidently facilitates formation of sub-groups.

Cattle are reported to form strong social bonds (Bouissou et al., 2001; Duncan et al., 2012) and we found evidence of consistent and extended durations of close contact, but between only a small proportion of individuals in each group, suggesting a manifestation of social preference in these individuals. Overall, we found that most cattle in the sampled herds interacted with each other in a manner more similar to random assortment. Studies have shown early experiences of unstable social groups can increase resilience to social stressors (Bøe and Færevik, 2003), suggesting that dairy cattle might adapt to dynamic social groups by creating fewer strong social bonds than might be expected in a less managed environment (Bouissou et al., 2001; Mclennan, 2013).

Contact structure among hosts can alter the transmission paths of infections (Bansal et al., 2007). Modelling studies have shown that modularity values of more than 0.45–0.6 start to alter the transmission rates of some infections within and between groups (Rozins et al., 2018; Sah et al., 2017). Only highly filtered networks of the Free and Stable groups in the present study had modularity values different from random that might be substantial enough to impede disease spread, and it is more likely that the weak and temporally unstable community structure in our networks would facilitate, rather than slow down, the spread of highly-transmissible infections to a larger number of herd members. Again, the exception is the Free herd, in which spatial divisions by cattle choice are enabled; in this herd, for large proportions of the study period, the cows were split between the pasture and the buildings, which might impede the spread of disease between animals. The differences seen in this herd suggest that, where choice is enabled by management, cattle may form social structures with divisions that hamper the spread of infections.

Repeated close contact at daily or twice-daily milking with specific individuals or cows in dyads that consistently spend long periods of time together might preferentially create opportunities between these specific cows for transmission of pathogens that require prolonged close contact, e.g. *Mycobacterium bovis* (Corner et al., 2003; Weber et al., 2013a). Heterogeneity in contact rate and strength can affect the speed and extent to which an infection is transmitted through a population (May, 2006). Despite cows being grouped closely together at many points during the study period, the number of conspecifics each cow connected with was more heterogeneous than was found on random networks in all groups, even when contacts were aggregated over the whole study period, suggesting that incorporating the contact structure of cattle into modelling studies might be prudent, even at coarse temporal aggregations.

When daily networks were calculated, contact frequency was similar between days for most groups; suggesting that when contacts are aggregated at this resolution, subtleties in cattle contacts might be missed, indeed in comparing simulation models on data aggregated into 12-hour time-steps, Duncan et al. (2014) found that infection transmission on empirical networks could easily be replicated by decreasing the transmission parameter in models based on corresponding random networks. When contacts from the present study were aggregated by shorter time periods (2 hours), variation in contact frequency was much more apparent. Modelling of disease transmission on high-temporal-resolution networks (Chen et al., 2014) has highlighted how aggregating contacts at the hourly or 2-hourly level might alter the predictions of epidemic spread for some pathogens compared to aggregating over longer time-periods (Chen et al., 2014). The time-scale over which contacts are aggregated should also be guided by the biology of the pathogen of interest (Dawson et al., 2019); data aggregated over short time windows are more likely to be informative for modelling the transmission of infections with short infectious periods (Perkins et al., 2009) and smaller R_0 (Chen et al., 2014). Yet for chronic infections with longer infectious periods, it is likely that incorporating such high resolution of contacts in transmission models will have less impact on transmission dynamics, while other networks parameters such as duration or strength of contacts may remain important (Read et al., 2008).

The proximity tags we used in this study most accurately record contacts when the devices fully face one another (Cattuto et al., 2010), therefore positioning of the sensors is challenging on large animals. As we were primarily interested in the transmission of infection via the oro-nasal route, it made practical sense to attach the tag to a neck collar, to which most cows in the study were already accustomed. We believe that due to the close proximity and mobility of the cattle, the number of contacts recorded provided a satisfactory representation of close contacts between cattle. However, contacts and, more likely contact durations, may have been underestimated when sensors were facing opposite directions. In indoor housing, where cows are frequently positioned with their heads side by side due to feeding along a barrier or lying in cubicles, this may have created a bias for contacts to be underestimated to a greater extent when groups were housed in buildings (Housed, Night-housed, Rotation 1b), rather than only in buildings for milking. Therefore, true differences in the contact frequency and duration in Buildings compared to Pasture may be greater than we observed. There were some individuals in most groups that were in contact with only a few other herd-members and had very low contact durations, which may have been due to altered positioning of the tags on these animals. However, voluntary isolation of sick individuals (Proudfoot et al., 2014) and changes in behaviour due to illness (Polikarpus et al., 2015) have been reported in dairy cattle, therefore sickness during the study period may have also contributed to low levels of interaction detected in some individuals.

The changing face of the dairy industry over recent years has resulted in many herds increasing in size and changes to how cattle are managed, with a trend towards more cattle being kept indoors for more of the year or, in some cases, entirely in indoor units. In these cases, the increased contact demonstrated in this study in buildings might impact upon the risk of transmission of infections (Arnott et al., 2016). In order to mitigate these effects, consideration might be given to housing design in order to allow enough space and spread out resources sufficiently that cattle are able to interact more freely whilst housed. Milking routine is often expected to increase contact between cattle, and thus increase the risk of disease transmission (Barlow et al., 1997; Bekara et al., 2014), however we find that this is not inevitable, and some milking protocols might reduce opportunities for disease transmission. Contrary to common thought regarding cow to cow interactions, and those demonstrated in beef herds

(Handcock et al., 2009), we found few social bonds in our sampled groups, perhaps exhibiting resilience to an unstable social structure. Whilst this response may reduce the impact of social stresses (Proudfoot and Habing, 2015), it is nevertheless likely that adapting to cope with an unstable social environment is less beneficial to cattle welfare than devising management strategies that might avoid social instability in the first place.

The structure of dairy cow social networks has typically been described as a single, unstructured group; indeed our study networks are densely connected during weekly and daily aggregations. However, when the network is divided in space and time, there are differences in cattle interactions that are not apparent at larger scales. From studying multiple groups of cattle, we have found that differences in management, even between dairy herds within the same region and with broadly similar management objectives, can influence social structure which might have implications for the transmission of disease, and should be considered when parameterising mathematical and statistical models of disease spread. Future research could focus on expanding contact studies to include whole herds, rather than single management groups to explore how movement between groups might affect disease spread. Recruiting larger groups might reveal stronger trends between social interactions and group size. So far, studies have focussed on contact networks as a means of mapping disease spread, however as diagnostic technologies, such as whole genome sequencing, develop and become more accessible, construction of transmission networks that combine sensor and molecular approaches during disease outbreaks is likely to substantially advance our understanding of pathogen and host relationships.

Chapter 6

Relationships between social interactions, production, and health in dairy cows

Chapter 6: Relationships between social interactions, production, and health in dairy cows

Abstract

The social environment experienced by livestock can have implications for their health, welfare and, subsequently, their productivity. Changes in the dairy industry, in order to maximise milk yields, have led to larger herd sizes and altered management of cows, which has impacted upon their social environment. Mixing of animals can lead to social instability of groups and expansion of herds can result in high stocking densities in housing. In this paper, we summarise existing studies assessing the relationship between social experiences (mixing of animals, dominance, grooming and different stocking densities) and production. We then directly investigated an association between social interactions and production indicators, using cattle proximity as a proxy for affiliative interactions between cows on three dairy farms over study periods of one week. In our review, studies of milk yield and dominance rank showed either no relationship or a positive correlation, to varying degrees of statistical significance. Mixing of animals had either no impact, or a negative association with milk yield, again to varying degrees of statistical significance. The mixed results we identified may be due to differences in study design, sample size, prior social experience of animals used in the studies, and the quality of analyses. Grooming was positively associated with milk production, and stocking density was not associated with milk yield, though fewer studies were identified. In our primary investigation, we did not detect biologically meaningful relationships between milk production or somatic cell count (an indicator of mastitis) and the total time a cow spent in social contact with other cows, or the mean time it spent with its four closest herd mates. Cows showed a preference for interacting more with cows that had experienced the same number of lactations, suggesting that grouping animals in terms of lactation number might encourage affiliative interactions. However, cows with more of their closest herd mates in the same lactation did not have demonstrably higher or lower milk yields. We suggest that further research on what constitutes a positive social environment for dairy cattle and how this could be promoted is warranted, thereby improving individual animal well-being and enabling a more robust assessment of the putative effects on production and health parameters.

Introduction

The social environment of animals impacts upon their health and welfare (Mellor, 2015). It has been shown that positive social experiences in livestock animals can reduce their stress (Laister et al., 2011; Takeda et al., 2003) and that calm animals have enhanced immunity (Dimitrov et al., 2005). Conversely, crowding, isolation, and social instability can precipitate stress and disease (Proudfoot and Habing, 2015; Proudfoot et al., 2012). Social stressors have been associated with reduced milk yield in cows, sheep and goats (Hasegawa et al., 1997; Miranda-de la Lama and Mattiello, 2010; Sevi et al., 2001) and reduced weight gain in pigs (Hyun et al., 1998). By contrast, cows involved in more affiliative interactions have been shown to produce more milk (Sato et al., 1991), and calm sheep have been shown to produce better quality milk (Sart et al., 2004). Improving the social environment to encourage positive interactions and experiences, may therefore be able to improve health and welfare, and increase production.

Grouping of cows by their stage of lactation allows more precise feeding to meet nutritional requirements (Sowerby and Polan, 1978), though moving animals between groups can be a source of social stress to dairy cows (Proudfoot and Habing, 2015). Industry-wide trends for larger herds (AHDB Dairy, 2019b) may exacerbate this problem as larger herds usually require more management groups and the subsequent increased movement of cows between groups can cause social instability. The trend towards higher milk yields (AHDB Dairy, 2019a; Pryce and Veerkamp, 2001) has resulted in the need for more intensive management of cattle, often indoors, to meet their nutritional demands (Charlton et al., 2011), which on some farms might lead to overstocking of animals (Krawczel et al., 2012) and concerns about a cow’s ability to perform natural behaviours (Arnott et al., 2016). By exploring the impacts these changes might have on the social behaviour of cows, we might be better able to understand how to ameliorate negative effects and enhance positive ones.

Two important indicators of productivity and cow health are milk yield, and the somatic cell count in milk, which can be an indicator of infection in the mammary gland, should counts exceed 200,000 cells/ml (Dohoo and Leslie, 1991). These measures have been used to assess the impact of various social experiences of cattle and we provide a summary of the literature in Table 6.1. We can split these studies into two broad categories; potentially negative experiences (e.g. mixing of cattle) that might be expected to reduce milk yield and increase somatic cell count, and potentially positive experiences (e.g. mutual grooming), the manifestation of which is thought to have positive effects on health and well-being, leading to animals being more able to reach their production potential. Regrouping of animals is a common management practice in modern dairy herds, however the introduction of new animals into a group has been associated with altered dominance hierarchies, exhibited by more aggressive interactions to re-establish the social order of the group (Raussi et al., 2005). The studies we identified that analysed the relationship between mixing animals between groups and milk production ranged from low quality (small sample size, no control group, no replicates and simple correlation analyses) to higher quality (larger sample size, multiple groups, control groups, replicates and analyses that control for confounding factors; Table 6.1). The negative association between regrouping and milk production has been identified in the short-term (1–5 days after regrouping; Brakel and Leis, 1976; Jezierski and Podluzny, 1984; von Keyserlingk et al., 2008) and longer-term (7–14 days after regrouping; Arave and Albright, 1976; Hasegawa et al., 1997). However, other studies have failed to detect a statistically or biologically significant association between mixing groups and milk yield (Clark et al., 1977; Silva et al., 2013; Sowerby and Polan, 1978; Zwald and Shaver, 2012), including one study where specifically dominant individuals were selected to change groups (Collis et al., 1979). Of the studies that included a control group (4 studies of 11), just one (Jezierski and Podluzny, 1984) detected a mean decrease in milk yield in the mixed group that was statistically different from the change seen in the control group. Larger studies (n > 500) did not detect a consistent, statistically or biologically significant relationship between moving cows among groups and milk yield. In replicating the mixing of animals between groups, Sowerby and Polan (1978) noticed that the response of cows to mixing changed over time and found that cows with a (pre-study) history of mixing exhibited a smaller change in milk yield than those that had no prior experience of group mixing. As most studies did not record this, differences in cows’ prior experience may have led to systematic bias in some studies, which could lead to the inconsistent study outcomes reported.

Early ethological studies of cattle concentrated on assessing dominance hierarchies of herds by observing agonistic pairwise interactions or recording the outcomes of ‘tests’ (e.g. food competition) determined by the researcher and allocating to each individual a relative position in the group hierarchy; their ‘dominance value’ or ‘rank’. A higher dominance rank is thought to confer priority over resources, e.g. water and food (Andersson et al., 1984), and thereby might be expected to increase milk yield; submissive cows may experience stress from bearing the brunt of agonistic interactions and therefore result in a lower milk yield. Studies examining the relationship between dominance and production provide mixed results (Table 6.1). We identified four studies that found no correlation between social rank and milk production (Beilharz et al., 1966; Collis, 1976; Dickson et al., 1969; Jezierski and Podluzny, 1984; Soffié et al., 1976), but many other studies have identified a positive correlation between the two (Andersson et al., 1984; Arave and Albright, 1976; Barton et al., 1973; Brakel and Leis, 1976; Hussein et al., 2016; Sambraus, 1969; Sato et al., 1991; Schein and Fohrman, 1955; Sottysiak and Nogalski, 2010). Of the five studies that demonstrated a statistically significant ($P < 0.05$) relationship between dominance rank and milk yield, there was moderate positive correlation with some variation in effect size (mean $r = 0.41$, standard deviation (SD) = 0.28). When cows were split into groups by their social status, e.g. dominant, subordinate, etc., groups of dominant cows and individual cows produced more milk than non-dominant cows (Andersson et al., 1984; Sottysiak and Nogalski, 2010). Many of these studies utilised correlation in analysing the relationship between milk yield and social metrics (Table 6.1), however they also demonstrated a covariance of both milk yield and dominance rank with age, body condition, and body weight. In the single analysis that used a multivariable framework, no relationship was detected between dominance rank and milk yield suggesting that confounding factors may have been influential in the mixed outcomes of these relatively simple analyses.

Studies assessing the relationship between social factors and somatic cell count have also demonstrated mixed results. Dominance value was not correlated with somatic cell count in individual cows (Arave and Albright, 1976). The mixing of animals was associated with an increased bulk milk somatic cell count in one group, but had no control group with which to compare (Kay et al., 1977). Two larger studies (n = 103 and 567) showed no difference in somatic cell count between control (stable groups) and treatment (new animals introduced) groups (Clark et al., 1977; Silva et al., 2013), suggesting there is more evidence for a lack of a relationship between these variables.

Historically, animal welfare principles were aimed at avoiding negative states, i.e. the freedom from fear, distress, hunger, etc., whereas more recently the focus has been drawn towards achieving positive welfare outcomes (Boissy et al., 2007; Mellor and Beausoleil, 2015), which is reflected in the relative higher numbers of early studies focusing on negative experiences or behaviours (e.g. mixing of animals, dominance) and the fewer studies assessing the impact of positive social experiences (e.g. mutual grooming) on production and health. Affiliative behaviours such as social licking in cattle, have been shown to calm the receiver (Laister et al., 2011) and be positively correlated with the milk yield of the cow receiving grooming ($r = 0.65$, $P < 0.01$, Sato et al., 1991; $r = 0.21$, $P < 0.025$, Wood, 1977). Cows involved in more grooming of other cows have been associated with higher milk yields in one study ($r = 0.55$, $P < 0.05$, n = 20; Sato et al., 1991), yet a negligible relationship was found in a larger study ($r = 0.098$, $P > 0.05$, n = 104; Wood, 1977). Cows with the same lactation number that are home-bred have often been reared in the same cohort as calves and will have spent similar amounts of time together in the herd. By recording contacts among cattle in a commercial dairy herd, Boyland et al. (2016) found that cows preferred to interact more with cows that had the same number of lactation periods, a measure which is strongly correlated with the age of the cow, and is hereafter defined as 'lactation number'. By recording grooming behaviour between a herd made up of twin cows grouped since they were 4–7 days old, Wood (1977) found that cows preferred grooming other cows in their own age group. These findings were thought most likely to be due to early bonding and increased familiarity between these cows. It has been shown that longer-term familiarity, such as cow's being reared together as calves, was more closely related to affiliative interactions than more recent periods of time together, e.g. cow's grouped together during their dry period (Gutmann et al., 2015). However, these studies have not explored the impact of being able to associate with more familiar conspecifics on their production.

Contacts recorded by proximity devices worn on neck collars have been shown to be a biologically and statistically significant predictor of affiliative rather than agonistic interactions (Boyland et al., 2016), perhaps due to the positioning of the sensor in a region where grooming interactions typically occur (Tresoldi et al., 2015). Although proximity sensors lack the ability to discriminate the giver and receiver of an interaction (Foris et al., 2018), pairs of animals in close proximity were more likely to engage in allogrooming than agonistic interactions (Tresoldi et al., 2015). We use cattle proximity as a proxy for a positive social interaction and continuously recorded interactions among dairy cows for one week in three herds in south-west England. Through network analyses, we also examine social preference in cows in terms of the number of lactations a cow has experienced, stage of lactation a cow is in, milk yield, and somatic cell count. We investigated the relationship between the social interactions of individuals and their productivity in terms of milk yield and somatic cell counts, whilst accounting for aspects of cow physiology and farm management factors, which we also expect to influence these responses. Management practices which promote a positive social environment might help to mitigate the negative effects of other practices required in modern dairy production, such as regrouping (Talebi et al., 2014), and the presence of familiar animals might reassure animals and allow freedom to exhibit natural behaviours (Patison et al., 2010). Better understanding of the social preferences of cows and possible associated benefits might encourage implementation of cattle management and grouping structures that create a more positive social environment and allow animals to achieve their production potential (Rault, 2012).

Table 6.1. *Summary of studies examining the relationship between social parameters and production parameters (milk yield and somatic cell count) in dairy cattle. Studies are split into categories of social parameters (grooming, dominance rank, regrouping and stocking density) and summarised by parameters, sample size, study design, analysis and outcomes. Studies are published in English between 1955 and July 2019. Grey shading indicates a statistically significant relationship (P < 0.05).*

	Social parameter	**Production indicator**	**Sample size**	**Study design**	**Outcome**	**Method of analysis**	**Author and year**
Grooming	Number of other herd members groomed	Milk yield	1 group of 104	Observational	No relationship (r = 0.117; P > 0.05)	Spearman's rank correlation	Wood, 1977
	Time spent grooming others				No relationship (r = 0.098, P > 0.05)		
	Number of herd members received grooming from				Positive relationship (r = 0.212; P < 0.025)		
	Time spent receiving grooming				Positive relationship (r = 0.214, P < 0.025)		
	Time spent receiving grooming	Milk yield	1 group of 20	Observational	Positive relationship (r = 0.65, P <0.01)	Correlation	Sato et al., 1991
	Time spent grooming others				Positive relationship (r = 0.55, P < 0.05)		
Dominance rank	Dominance value	Milk yield	1 group of 20	Observational	Positive relationship (r = 0.88, P < 0.01)	Correlation	Sato et al., 1991
	Dominance rank	Milk yield	1 group of 78	Observational	Small positive effect (r = 0.25, P < 0.05)	Correlation	Schein and Fohrman, 1955
	Dominance value	Average milk production	4 groups, total cows = 89	Observational	No relationship (t value = 1.85, P > 0.05)	Multivariate regression	Beilharz et al., 1966
	Dominance value	Daily and lactation milk yield	27 groups, total cows = 1017	Observational	No relationship (r = 0.02 (daily), -0.02 (lactation), no P value cited)	Correlation	Dickson et al., 1969
	Social rank	Milk yield	6 herds, cows in each group = 6–49	Observational	Significantly correlated in 2 out of 6 herds (P values and r values not available)	Correlation	Sambraus, 1970
	Dominance (Schein-Fohrman) rank	Milk production	1 group of 53	Observational	Positive correlation (r = 0.33, P > 0.05)	Correlation	Barton et al., 1973
	Dominance value	Individual and group milk yield	4 groups of 32, total cows = 128	Experimental (1 control group, 3 replicates)	No correlation between individual milk yield and dominance value.	Correlation	Jezierski and Podluzny, 1984
	Dominance value	Peak milk yield	1 group of 49	Observational	Negative correlation (r = - 0.21, P > 0.05)	Correlation	Collis, 1976

	Social parameter	Production indicator	Sample size	Study design	Outcome	Method of analysis	Author and year
	Dominance value (calculated by food competition test)	Milk production	1 group of 34	Experimental (no control group)	'No correlation between milk production and dominance value.'	Correlation	Soffié et al., 1976
	Dominance rank	Fat-corrected milk yield and water consumption	1 group of 12	Experimental (no control group)	Fat corrected milk yield and water consumption significantly higher in dominant cows (P<0.05), yet milk yield was not different between dominant and submissive cows.	Least squares means	Andersson et al., 1984
	Dominance rank	Milk yield	1 group of 126	Observational	The group of 3 dominant cows produced more milk	Split cows in to groups by social status and recorded milk yield	Sottysiak and Nogalski, 2010
	Dominance value and stocking rate	Milk production	1 group of 252	Observational	Positive relationship between dominance value and milk yield at low and medium stocking rates (r = 0.42, 0.32; P < 0.05)	Pearson's correlation	Hussein et al., 2016
	Dominance value of cows mixed between groups	Milk yield	2 groups of 24, total cows = 48	Experimental (no control group, 4 replicates)	Positive effect in moved cows (r = 0.49, P < 0.05) and non-moved cows (r = 0.27, P < 0.01)	Correlation	Brakel and Leis, 1976
	Dominance rank	Milk yield	2 groups of 17, total cows = 34	Experimental (no control group)	Small positive effect (r = 0.16, P < 0.05)	Correlation	Arave and Albright, 1976
		Leukocytes in milk			No relationship (r = 0.03, P > 0.05)		
Regrouping	Addition of new cows	Milk yield	2 groups of 17, total cows = 34	Experimental (no control group)	5% decline in milk yield during second week after additions in both groups	Method not stated	Arave and Albright, 1976
	Mixing of cows between groups	Milk yield	2 groups of 24, total cows = 48	Experimental (no control group, 4 replicates)	Milk yield depressed by 3% 1 day after cows mixed, then returned to normal	Least squares means	Brakel and Leis, 1976
	Addition of 2 dominant cows and 2 subordinate cows	Bulk milk SCC	1 group of 32	Experimental (no control group, 1 replication)	Increase in BMSCC after mixing and return to pre-mixing levels 10 days later	Change in response after intervention	Kay et al., 1977
	Mixing of cows between groups	Milk yield (5 day average)	4 groups, total cows = 103	Experimental (1 control group)	No change in pooled average milk yield before and after mixing	Analysis of variance (ANOVA)	Clark et al., 1977
	Moving increasing numbers of animals between groups	Milk yield	7 herds each divided into ≥ 2 groups, total cows = 6731	Experimental (no control group)	Some farms showed small changes in milk yield, others large changes. Evidence of cows becoming more resilient to moves over time.	Multivariate analysis	Sowerby and Polan, 1978

	Social parameter	Production indicator	Sample size	Study design	Outcome	Method of analysis	Author and year
	Dominant cows changing groups	Average group milk yield	2 groups of 15 and 2 groups of 12, total cows = 54	Experimental (no control group)	No change in average milk yield	ANOVA (Sheffé's method of multiple comparisons)	Collis et al., 1979
	Changing groups	Individual and group milk yield	4 groups of 32, total cows = 128	Experimental (1 control group)	Mean decrease in milk yield of 4% after group mixing significantly different from the control group.	Comparison of mean change in milk yield before and after intervention (method not stated)	Jezierski and Podluzny, 1984
	Individuals changing groups	3 day moving average of milk yield	2 groups of 51, total cows = 102	Experimental (no control group)	Moved animals had a 4.7% decrease in milk production 2 weeks after moving ($P < 0.01$). Middle ranked cows (not moved) showed a decrease of 3.8%	ANOVA	Hasegawa et al., 1997
	Individuals introduced to established groups	Average group milk yield	4 groups of 11 and 3 groups of 12, total cows = 80	Experimental (no control group)	Reduction in milk yield from 43.4 ± 1.5 kg/d to 39.7 ± 1.5 kg (~ 8.5% reduction, $P < 0.001$) on day 1	Change in mean milk yield after intervention (t-test)	von Keyserlingk et al., 2008
	Mixing of cows between groups	Milk yield	2 herds each divided into 2 groups, total cows = 585	Experimental (1 control group in each herd)	Mixing showed no relationship with milk production	Least squares means	Zwald and Shaver, 2012
	Traditional flow of cows through groups (TRD) compared to all in all out system (AIAO)	Monthly milk yield and SCC	2 groups, TRD (n = 308) and AIAO (n = 259), total cows = 567	Experimental (6 replicates each in control and treatment groups)	No difference in SCC or milk yield between stable (AIAO) groups and mixed (TRD) groups	ANOVA	Silva et al., 2013
Stocking density	Increased stocking densities	Average group milk yield	4 groups , total cows = 136	Experimental (1 control group, 3 stocking densities)	No difference between pens in milk yield	Mixed effects regression model	Krawczel et al., 2012
	Isolation and increased stocking density	Milk yield	1 group of 17	Experimental (no control group, 1 stocking density)	Milk yield showed no relationship with stocking density or isolation	Difference in unadjusted means	Arave et al., 1974

Methods

Farms and production data

The milking groups of three commercial dairy herds (two Holstein-Friesian herds – ‘HF1’ and ‘HF2’, and one Ayrshire herd – ‘Ayrshire’) were studied for one week between August and November 2018 (Table 6.2). All herds calved all year round, kept cows in separate milking and dry groups, and were milked twice daily. Animals in HF1 were the low-yielding cows of that herd and were housed during the study period, due to lack of grazing in a bout of hot weather. Cows in HF2 and Ayrshire were grazed on a rotational grazing system where the paddock was changed after every milking. We obtained details of the last calving date, lactation number, recent milk yield and somatic cell count of all cows in HF1 and a subset of cows in Ayrshire and HF2 by kind permission of the farmers.

Equipment and proximity data

Nylon cattle collars with a plastic clasp (Suevia Haiges, Germany) were fitted with proximity devices based on a design by the OpenBeacon project (http://www.openbeacon.org/ and the SocioPatterns collaboration consortium (http://www.sociopatterns.org/). The devices use radio frequency identification detection (RFID) technology to exchange low-power radio packets in a peer-to-peer fashion, using the difference in signal strength as a proxy for distance between devices (Cattuto et al., 2010). We defined contacts based on a spatial threshold of 1–1.5m to capture close contacts, assessed by validation analyses (see Appendix E Validation and Chapter 5 Methods for further details). We only included contacts in our analysis that lasted over 20 seconds in duration, aiming to avoid detecting ‘walk-by’ type interactions and to record more significant behaviours such as allogrooming events, which are reported in observation studies to last around 38 seconds (means of 37 and 39 seconds) (Sato, 1984; Tresoldi et al., 2015). Contacts are maintained as long as signals are transferred between devices at least once every 20 seconds, therefore contact time was measured in 20 second blocks. We removed contacts recorded during milking times reported by farmers, as these represent times when social grooming or affiliative behaviours were less likely to occur (Wood, 1977).

Social parameters

We constructed networks based on contact data described above, where cows were nodes and the duration of contact between them over the study period formed the edge between them: if there was no contact then no edge was present. Assortativity measured the tendency of cows to interact more with similar cows based on a particular characteristic, with a value of +1 indicating preferential association between cows with similar characteristics, and -1 indicating preferential association between cows with different characteristics. We calculated the assortativity of cows (for which we had data) by stage of lactation and lactation number (Ayrshire: n = 44, HF1: n = 71, HF2: n = 95), and somatic cell count and milk yield (Ayrshire: n = 39, HF1: n = 71, HF2: n = 87) in the R (R Core Team Version 3.5.3, 2019) package 'igraph' (Csardi and Nepusz, 2006). As the networks were densely connected, we calculated assortativity on networks filtered by edge weights to focus on the effect of increasingly 'strong' ties. We achieved this by removing edge weights (contact durations) below the 50th, 75th and 90th percentiles of edge weights to create F50, F75 and F90 networks respectively. We constructed null graphs with the same number of nodes and edges as the observed network where each edge has an equal, fixed probability of being present or absent based on the Erdös-Rényi model (Erdös and Rényi, 1959). We randomly allocated observed edge weights to the edges of the new random networks (n = 4999). We removed edges from random graphs below filtering thresholds (F50, F75, and F90) after constructing the randomised graphs from the original networks. Assortativity values of our observed networks were deemed to be statistically significantly different to those on random networks when the observed value lay outside 95% of randomised values.

We calculated the mean association strength, i.e. the time each cow spent with all other cows in the network divided by the number of cows the cow could have contacted (n-1). Individuals that spend longer amounts of time in contact with other cows may be considered to be more gregarious animals, however the definition of social bonds is complex and this metric does not reveal whether this time is distributed equally among many herd mates, or whether an individual has just a few partners with which they spend a lot of time (Shultz and Dunbar, 2010). It has been estimated that cows might be most comfortable with 2–4 individuals (Takeda et al., 2000); therefore, to estimate a cow's tendency to form stronger social bonds, we took the mean of the time each cow spent with its longest 4 contacts (mean top 4 contacts) and in order to explore the effect of a cow spending time with more familiar cows, we recorded the number of top 4 contacts with known lactation numbers that were in the same lactation as the selected cow.

Statistical modelling

We constructed two similar Bayesian mixed effects models with milk yield (measured by weight; Model 1) and somatic cell count (thousand cells per ml; Model 2) as response variables. Mean association strength, 'mean top 4 contacts', and the number of top 4 contacts in the same lactation were used as explanatory variables. In our models we accounted for individual physiology by including cow lactation number, stage of lactation, and somatic cell count (in Model 1 only) and milk yield (in Model 2 only), as fixed effects. Farm was included as a random effect in both models to account for farm differences in nutrition, husbandry, and breed.

Higher somatic cell counts are associated with a decreased milk yield (Hadrich et al., 2018). In order to account for this relationship in our analysis, we used data where both variables were recorded from the nearest recording to the study period (HF1 = 7 days after study period, HF2 = 1 day before study period, and Ayrshire = 11 days after study period; Table 6.1). To check if this single value for milk yield would approximate to milk yield over the study period, we additionally obtained daily milk yields during the study period for HF1, and milk yield and somatic cell count data from a second recording 16 days before the study period for the Ayrshire group. Mean daily milk yield in HF1 was highly correlated with the single milk recording value ($r = 0.9$, $P < 0.001$; Appendix F Fig. F1). For the Ayrshire group, the two values for milk yield and somatic cell count from milk recording days were moderately correlated ($r = 0.59$, $P < 0.001$ and $r = 0.67$, $P < 0.001$ respectively; Appendix F Fig. F1). We also ran a model similar to Model 1 using HF1 average daily milk yields as a response variable instead of the single milk yield value, and outcomes were similar to those identified in Model 1.

Lactation number represents the number of lactations a cow has had (and therefore correlates with age) and was classed as a categorical variable with the final level corresponding to lactation numbers of 5 or more. On our study farms, animals were home-bred, and therefore lactation number might also be a proxy for time spent together as calves and time spent in the milking herd. Lactation number has a non-linear relationship with milk yield, lowest in an heifer's first lactation, rising in the 2nd and 3rd lactation and then decreasing as cows mature (Vijayakumar et al., 2017). Lactation number also mediates the relationships between stage of lactation and milk yield (cows in their 2nd and 3rd lactations tend to reach peak yield more quickly (Vijayakumar et al., 2017), and somatic cell count and milk yield (where higher milk losses occur in older cows (Hand et al., 2012).

We quantified days in milk as the number of days between the most recent calving date and the milk recording date (when somatic cell count and milk yield were measured). Multiparous cows (2 or more lactations) with a similar number of days in milk are likely to have been re-introduced back into the milking group at a similar time and grouped together in the preceding dry period. The relationship between stage of lactation and milk yield is described by the lactation curve, the profile of which changes with lactation number. Therefore, because the lactation curve was not the same for all groups we could not describe this relationship with a single term and we grouped this variable into four stages of lactation; 1–70 days, 71–170 days, 171–270 days and >271 days in milk (Vijayakumar et al., 2017) to allow for variation in the relationship between groups.

To acknowledge the mediating effects of lactation number, we ran Model 1 with an interaction term between lactation number and lactation stage, and an interaction between lactation number and somatic cell count. We tested Model 2 with an interaction between lactation number and milk yield. Including the interaction terms in either model did not appreciably alter model fit (measured by the leave one out cross validation method 'LOOIC'; Vehtari et al., 2017) and did not alter the interpretation of our social variables, therefore we present the results of our models without interaction terms for ease of interpretation. Due to the predicted relationship between milk yield and somatic cell count, we ran Model 1 with and without somatic cell count, and Model 2 with and without milk yield; in both cases outcomes were similar to our reported results. To reduce bias, we did not perform variable selection except for the interaction terms (e.g. Harrell, 2001). Regression parameters of fixed effects were given improper flat priors and random effects were given non-standardised half Student-t priors based on the standard deviation of the random effect with 3 degrees of freedom and a scale parameter (Carpenter et al., 2017). Predictive posterior distributions showed good model fit (Appendix F Fig. F2 and F3) and good model convergence is demonstrated by Gelman-Rubin $\hat{R}$ values of 1.00 and large effective sample sizes (Appendix F Table F1). Effect sizes of continuous variables were scaled by the 10th and 90th percentiles of raw data so that they could be interpreted relative to the spread of raw data values (Fig. 6.2 and Appendix F Table F1). The effect size thus represents the change in the response variable associated in an increase in the explanatory variable from the 10th percentile of the data to the 90th percentile.

All data analysis was performed in R (R Core Team Version 3.5.3, 2019), with models constructed in ‘brms’ (Bürkner, 2017) and networks constructed in ‘igraph’ (Csardi and Nepusz, 2006). All fieldwork was approved by the University of Exeter College of Life and Environmental Sciences (Penryn Campus) animal ethics committee (Reference eCORN000087 v4.6).

***Table 6.2**. Details of three study farms and study periods providing information on farm management, group size, data gathered and mean production indicators obtained from the nearest milk recording date to the study period.*

Group name	Study dates	Breed	Group monitored	Housing/ Grazing	No. cows in group	Study Period (days)	Milk data recording date	Complete proximity data and cow lactation info	Complete milking and proximity data	Mean no. minutes spent in contact with others (standard deviation SD)	Mean no. minutes spent with 4 top contacts (SD)	Median no. cows in top 4 contacts in same lactation	Mean milk yield per day per cow kg (SD)	Median SCC (inter-quartile range)	Mean days in milk (SD)	Proportion (number) of cows SCC > 200,000 cells/ml
Ayrshire	31/08/18-07/09/18	Ayrshire	Milking	Strip grazing	52	6.8	18/09/18	44	35	9.2 (4.4)	37.6 (18.9)	1	21.3 (7.5)	49 (27–126)	143 (102)	21% (n = 11)
HF1	04/09/18-11/09/18	Holstein Friesian	Milking - low yield group	Indoor cubicles	111	6.8	18/09/18	71	71	14.2 (6.3)	74.6 (31.9)	1	21.0 (5.2)	128 (93–237)	244 (49)	31% (n = 34)
HF2	26/10/18-02/11/18	Holstein Friesian	Milking	Rotational grazing	177	7.0	25/10/18	95	87	12.8 (7.1)	62.8 (33.2)	1	31.0 (7.7)	43 (27–105)	175 (99)	14% (n = 25)

Results

After accounting for hardware and software performance, full proximity data (and lactation number and lactation stage data) were available for 85% (Ayrshire = 44/52), 64% (HF1 = 71/111), and 54% (HF2 = 95/177) of animals in the study herds (Table 6.2) and were used in calculating network parameters. Milk yield and somatic cell count data were available for all 71 cows in HF1 and for only a subset of the Ayrshire herd (n = 39/44) and the HF2 herd (n = 87/95) and were used in our statistical models (total n = 197; Table 6.2). The mean contact time per pair of cows among all cows was 12.6 minutes over the whole study period (standard deviation (SD) = 6.6 minutes). The mean time spent with the top 4 longest contacts among all cows was 62.0 minutes (SD = 33.0 minutes). Most cows (73.6%) had at least one cow in their top 4 contacts in the same lactation but only 2.0% of cows had all 4 top contacts in the same lactation (Table 6.2). There were similar numbers of cows in each lactation number (mean = 39 cows, SD = 5 cows) and the largest proportion of cows (40.1%) were between 171 and 270 days in milk. The mean milk yield of all cows was 25.6kg per cow per day over 2 milkings (SD = 8.4kg). Somatic cell counts were low for most cows, but 20% of cows had cell counts above 200,000 cells/ml, suggestive of mastitis (Dohoo and Leslie, 1991).

Cows in all groups showed a tendency to associate preferentially with cows of the same lactation number on all filtered networks (Fig. 6.1). For other measures, assortativity was no different to that expected on random networks in all cases apart from on the F75 network for stage of lactation (Fig. 6.1), though due to the lack of other statistically significant effects in other herd’s networks for this variable, we expect this could be a Type I error due to multiple testing.

We detected no statistically significant association between either mean association strength, spending longer times with certain cows and those cows being in the same lactation and either milk yield or somatic cell count (Fig. 6.2 and Appendix F Table F1). There was a statistically significant relationship between lactation number and milk yield in Model 1, showing the lowest milk yields in first lactation heifers and higher milk yield thereafter (Fig. 6.2 and Appendix F Table F1). Higher milk yields were associated with cows in their second lactation or higher (Fig. 6.2). There was a positive relationship between milk yield and lactation stage in the first 70 days of lactation in Model 1 and then this decreased through all subsequent stages of lactation, demonstrating the expected long tail of the lactation curve. In Model 1, there was no statistically significant association between somatic cell count and milk yield, or vice versa in Model 2 (Fig. 6.2 and Appendix F Table F1). Model 2 demonstrated increasing somatic cell count with the number of days in milk and a higher somatic cell count was associated with cows in their 5th or higher lactation (Fig. 6.2).

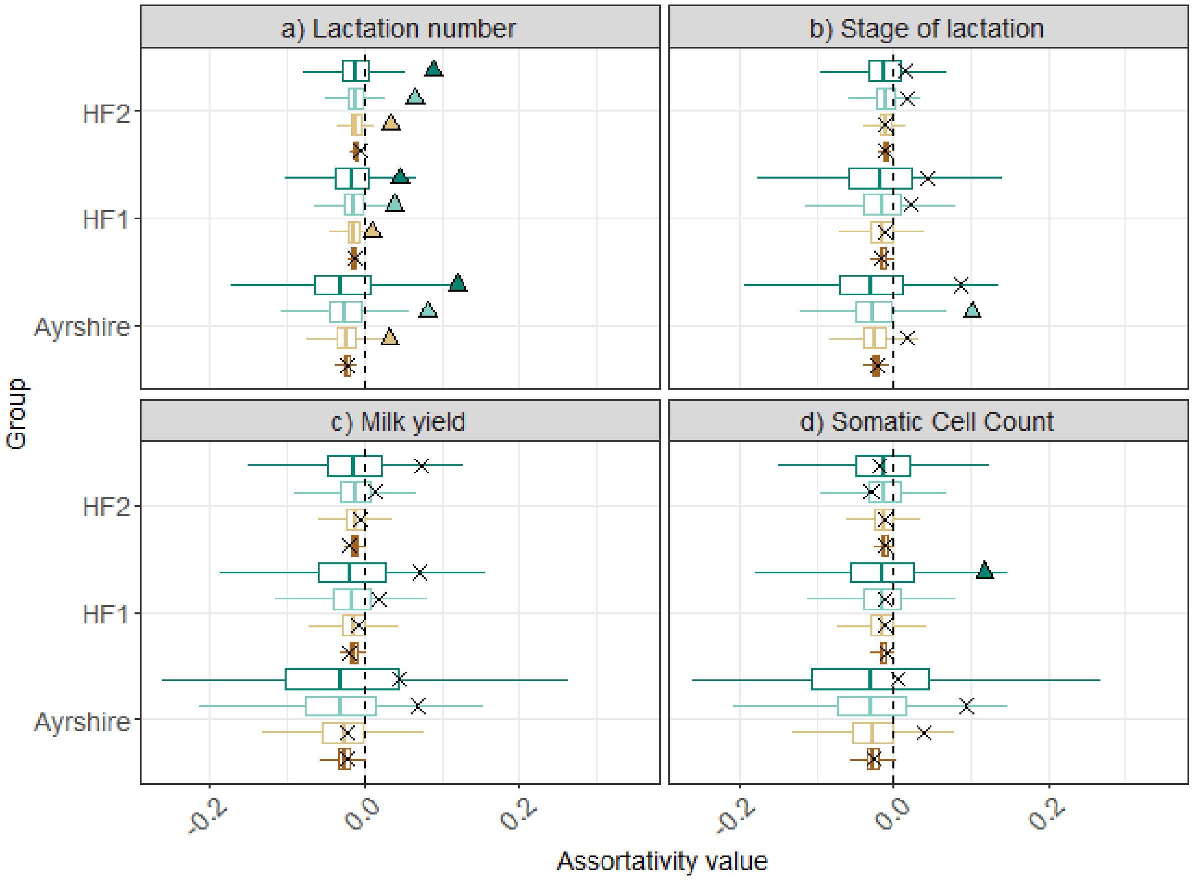


Figure 6.1. *Assortativity values from observed and random networks of cattle from three study groups. Triangles represent observed assortativity values outwith 95% of random network values, crosses represent observed values within 95% of random network values, and boxplots represent the distribution of random values. Assortativity indicates the tendency of cows to interact preferentially, based on similarities in a) lactation number and b) stage of lactation (Ayrshire n = 44, HF1 n = 71, HF2 n = 95), and c) milk yield and d) somatic cell count (Ayrshire n = 39, HF1 n = 71, HF2 n = 87). Assortativity was calculated on undirected, weighted networks that were not-filtered (NT – brown), and networks with 50th (F50 – gold), 75th (F75 – light blue), and 90th (F90 – dark green) percentiles of edge weights removed. The dashed line at zero represents the assortativity value where interactions are unrelated to characteristics of each cow. Cows tend to interact preferentially with cows of a similar lactation number (a), but they did not tend to interact preferentially based on their stage of lactation (b), milk yield (c) or somatic cell count (d).*

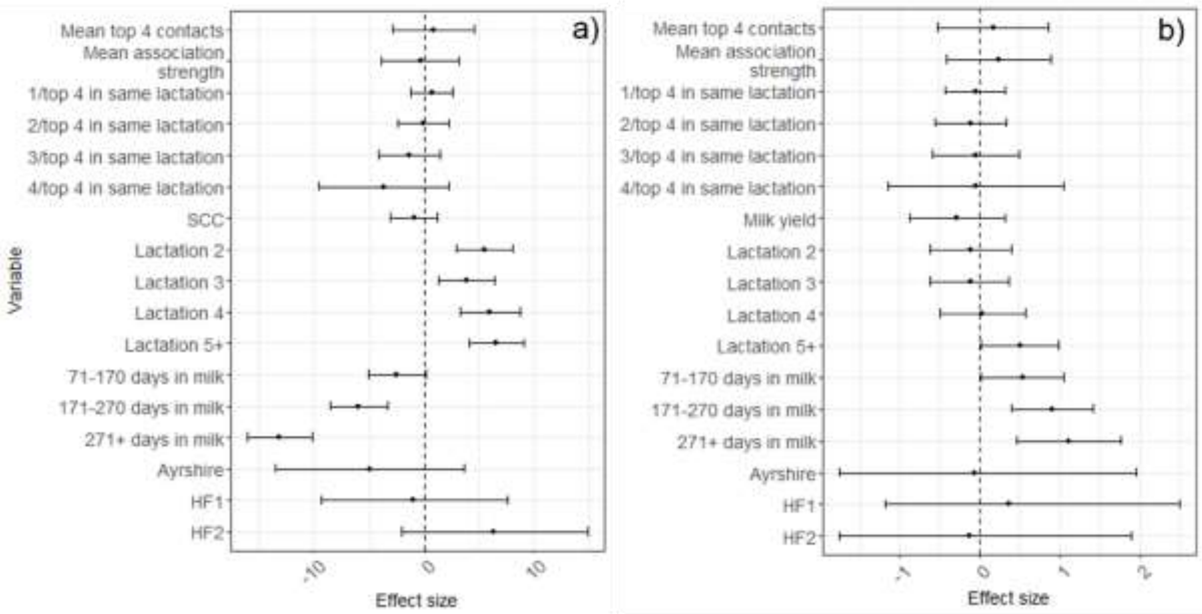


Figure 6.2. *Effect sizes and 95% credible intervals from two Bayesian mixed effects models with the response variables of milk yield (a) and somatic cell count (b). Studies included 197 cattle from three Cornish dairy herds. Fixed effects in the models were mean association strength, mean time spent with 4 closest herd mates, number of 4 closest herd mates in the same lactation, lactation number, stage of lactation, and somatic cell count (only in Model 1) and milk yield (only in Model 2). Farm is included as a random effect in both models. Continuous variables are scaled to represent the effect size associated with a change in the explanatory variable from the 10th to the 90th percentile of the raw data. Points show effect sizes and whiskers represent 95% credible intervals (see Appendix F Table F1 for details).*

Discussion

This study has found that cows prefer to associate with cows in the same lactation number. Similar lactation number in the herds in this study, where animals are home-bred, is likely to mean that cows may have been reared as calves in the same cohort and entered the milking herd at a similar time. In contrast, the lactation *stage* of cows, represented shorter-term familiarity, i.e. being more recently in a group of dry cows together, did not affect which cows they spent time with. Therefore, we show that in multiple herds, given the opportunity, cows preferentially associate with animals with which they share longer-standing rather than more recent ties, adding to evidence that long-standing familiarity is important in the formation of social bonds (Boyland et al., 2016; Gutmann et al., 2015; Reinhardt and Reinhardt, 1975).

As we identified that cows preferred to interact with cows of the same lactation number, we assessed the relationship between having more cows of the same lactation number in their top four closest contacts and outcome variables, however there was no biologically or statistically significant association with either milk yield or somatic cell count of cows. Indeed, we found no biologically or statistically significant relationship between the social parameters we tested and milk yield or somatic cell count. There are several possible explanations for this outcome. First, there may not be a detectable effect of social experiences on the health and production of dairy cattle generally, or specifically in the herds sampled in this study; despite differences in their management, broadly speaking our study farms were similar in terms of well-managed, small to average-sized dairy herds within Cornwall. Second, the parameters we tested may have not captured social experiences that effectively represent positive interactions between cattle, or may not be those experiences that relate to productivity. If our proximity events, defined as contacts actually represented a mixture of positive and negative interactions, the resulting relationship with our response variables may have been balanced out. As the baseline milk yield of each cow is likely to influence the magnitude of the relationship with social parameters (Jezierski and Podluzny, 1984); cows already achieving close to their physiological potential may have less opportunity for further increases in production as a result of positive social experiences, yet more susceptible to decreases in yield due to negative experiences. In contrast, a cow yielding less than their potential might benefit more from positive social interactions but be less vulnerable to the impacts of negative experiences.

Previous literature shows differing relationships between moving of cows between groups and production, however, few studies have considered the characteristics of the newly introduced cattle. Two studies specifically selected dominant cows to change groups but this did not seem to affect study outcomes (Collis et al., 1979; Kay et al., 1977). However, it has been suggested that the removal of cows thought to contribute positively to the group, e.g. ‘social groomers’ might be detrimental to overall herd milk yield (Wood, 1977) and that the magnitude of the impact of removing animals from the group may depend on the number of social interactions that cow has had (Jezierski and Podluzny, 1984). Further studies might consider characterising the social nature of individuals e.g. as highly gregarious or less social, and experimentally removing both types of individual from the group, to see if this leads to a different magnitude of relationship with production.

We have described the contrasting results of studies examining social behaviour in livestock and identified that prior social experiences of the animals involved in the studies, the method of analysis, and the production level of animals may have contributed to a lack of consistency in outcomes. Studies with large sample sizes, multivariable analyses and those that control for previous social experience of the animals involved are likely to be able to identify a more accurate estimation of the relationship between the social environment of cows and production parameters. We found that rather more studies focussed on negative social experiences, perhaps as the impacts of negative social environments may be more easily detected, e.g. feather pecking in hens, tail-biting in pigs, than impacts of a positive social environment (Ellen et al., 2014). However, we identified a preference for cows to interact with cows of a similar lactation number across all herds, providing further evidence for social preference in cattle and suggest that future research to optimise grouping strategies on commercial farms would be beneficial to cow welfare.

Chapter 7

Discussion

Chapter 7: Discussion

Background

The health of animal and human populations is closely linked. Epidemics in livestock can compromise human public health (zoonotic infections) and can be costly to national economies in terms of direct losses from livestock, livestock-associated, and tourism industries (Thompson et al., 2002). However, aspects of commercial activity and maximising productivity can lead to compromises in terms of animal health and welfare (Dawkins, 2017) and, potentially, resilience to disease outbreaks. In this thesis, I have furthered understanding of how some factors that are driven by industry processes can affect the interactions among cattle farms and among cattle, and in turn, potentially affect the epidemiology of an endemic pathogen. At these different spatial scales, I have used network analysis as a tool to better understand potential pathways of disease transmission through cattle populations. In this discussion, I will review the key findings of this thesis, then discuss them in the context of the control of bTB in Great Britain. I then evaluate the utility of networks in the analysis of cattle epidemiology and make final conclusions.

Key findings

Heterogeneity in host contact rates can increase the extent and rate of spread of some epidemics (Lloyd-Smith et al., 2005). In **Chapter 2**, I applied this concept to the context of livestock infections and considered farms as the ‘host’, able to acquire and spread infection. Despite implications that superspreaders might be influential in the epidemiology of bovine tuberculosis in Great Britain (Brooks Pollock et al., 2014), so far, there had been no exploration of the characteristics these farms might exhibit or mechanisms by which they might operate, and thus I use bTB as an example throughout **Chapter 2**. I outlined three main components involved in superspreader characteristics, and identified ways in which a minority of farms might be able to spread infection to many other farms. First, I reviewed current literature studying heterogeneity in contacts between farms via animal trading and found evidence that it is common in livestock networks for a few farms or markets to account for very many trading connections (Dutta et al., 2014; Mweu et al., 2013; Rautureau et al., 2011; Woolhouse et al., 2005). Studies modelling the impact of control measures directed at these highly-connected premises found them to be influential in reducing the extent and spread of simulated epidemics. This highlighted the importance of trading connections and led to investigations performed in **Chapters 3** and **4**. Second, I suggested that the infectiousness of the farm might be driven by the within-herd prevalence of a pathogen. Factors which facilitate spread of pathogens within a herd, such as high stocking densities, poor ventilation, and animals being housed indoors might increase the risk of a farm spreading infection, and are further explored in **Chapter 5**. Third, farms that are infected for long periods of time have greater opportunity for seeding new infections. I highlighted that the poor sensitivity of current bTB tests might enable infected farms to go undetected for long periods of time and suggest that on-farm factors such as immunosuppression and co-infection might further reduce the performance of diagnostic tests and further exacerbate this issue.

After outlining the role of cattle movements in farms acting as superspreaders in **Chapter 2**, I performed a national-scale analysis of cattle movements in GB from 2001 to 2015 in **Chapter 3**. Networks constructed from these movements were dense and showed scale-free properties. Although most farms were not directly connected to many others, I consistently identified farms that acted in a similar way to that of markets, as hubs in the network, connecting to many other farms and thereby potentially acting as superspreader farms, should they become infected. I quantified temporal pathways through the network, where linked farms represented potential sources of infection in ingoing contact chains, and farms potentially exposed to infection in outgoing contact chains. I found that whilst many farms were indirectly connected to only a few farms, many farms were connected to substantial numbers of other cattle farms via ingoing and outgoing contact chains. Contact chains scaled with herd size, and we found differences between herd types. Dairy farms were more likely to have more farms in their outgoing contact chains and fattening farms were more likely to have more farms in their ingoing contact chains. We found evidence for potential ‘superspreader activity’ in a smaller subset of farms that had numerous chains in their ingoing contact chains, suggesting they might be more exposed to infection, and more farms in their outgoing contact chains, suggesting they might be more able to transmit infection and act as superspreaders if infected.

To explore the implications of these indirect trading connections on disease spread, in **Chapter 4**, I characterised the farms in the ingoing contact chains, i.e. potential ‘source farms’, in terms of their individual recent bTB history and their country (Scotland and Wales) or bTB risk area (within England). I used a logistic regression model in a frequentist framework to evaluate the effect of being indirectly connected to more farms in total, or more farms located in areas of lower or higher bTB risk, on the odds of a bTB incident on the root farm. The model showed that although the number of farms in the ingoing contact chain was not associated with a change in the odds of a bTB incident, there were increased odds associated with trading indirectly with more farms from higher-risk areas, and decreased odds associated with trading indirectly with more farms in lower-risk areas. These findings suggested that the large-scale connections detected in **Chapter 3** were relevant in the transmission of a chronic disease. In assessing the impact of direct and indirect connections, I also estimated the effectiveness of bTB pre-and post-movement testing protocols that were in place during the study period and concluded that limitations of the performance of diagnostic tests or strategies used, meant that during the study period there were still risks of disease transmission associated with trading of animals from higher-risk areas of England and Wales. This showed that the putative risks of trading infected animals that I identified in **Chapter 1** were well-founded. In **Chapter 2**, I discussed that in addition to movements, farms might be able to seed new infections, or repeatedly sustain infections in neighbouring farms via local connections i.e. shared boundaries, wildlife, etc. (Brooks Pollock et al., 2014). Modelling in **Chapter 4** demonstrated that local bTB incidents were a strong risk factor for subsequent bTB incidents nearby, highlighting the importance of this pathway in disease transmission.

My review in **Chapter 2** suggested that increased within-herd transmission might increase the risk of a farm transmitting infection to other farms. In order to better understand the interactions among cows within a herd, in **Chapter 5** I deployed proximity loggers and GPS devices and recorded contacts among nine groups of dairy cattle on seven commercial farms that exhibited a range of management types. I constructed multiple networks at differing temporal and spatial scales to assess how cattle interactions might be affected by their environment. We found that fully aggregated networks were dense, but on dividing into smaller time windows, that cattle had bursts of activity at certain times of day, typically associated with management practices. In looking at farms with cows outdoors and indoors, we could directly compare interactions in the two spaces; cows housed in buildings contacted more conspecifics, for longer durations, compared to cattle at pasture on the same farm. I found evidence for heterogeneity in contact rates between animals and identified that some individuals in each group showed evidence of social preference.

Cattle farming is a commercial enterprise, and as such, farm management decisions are typically focussed on improving productivity of animals. In light of the changes in the dairy industry that were suggested in **Chapter 3**, where we found a trend for fewer, larger dairy herds, in **Chapter 6**, I explored the relationship between productivity and social interactions of cattle. Using a subset of the farms previously studied, for which we obtained production data, I evaluated what factors might drive the individual heterogeneity and social preference identified in **Chapter 5**. I tested if interactions were based on cow characteristics and found that cows that have had a similar number of lactations are more likely to spend time together. In **Chapter 6**, I synthesised previous literature on cow sociality and milk yield, and cow sociality and somatic cell count, and found that most studies considered social aspects of dominance, mixing between groups and social grooming, but found inconsistent outcomes. Using the contact data from **Chapter 5**, I constructed a Bayesian mixed effects model and investigated the effect of three social parameters; the amount of time spent with other cows, with the four closest contacts, and the number of the four closest contacts that were in the same lactation (measured with proximity sensors), on milk yield and somatic cell count. Indeed, the social parameters I tested did not have a significant relationship with milk production or somatic cell count. The mixed outcomes from these studies perhaps reflect the myriad of factors influencing production parameters of dairy cattle.

High-risk farms

Network analyses offer an approach by which we can investigate disease spread and incorporate key processes such as heterogeneity in host contacts. Simulation studies allow us to explore how network structure can affect transmission of a pathogen, and thereby are also able to demonstrate the effectiveness of different control strategies. Modelling shows that the more effort that is targeted at individuals generating more infections, superspreaders, the lower the effective R_0, and the greater the efficiency of control measures (Lloyd-Smith et al., 2005). Between-farm transmission models have shown that removal of nodes from the network can have substantial impacts in reducing the final epidemic size (Volkova et al., 2010a). In this thesis, I have identified several farm characteristics which might make them particularly likely to spread onward infections, and thereby possible candidates for additional or novel control measures in order to limit the wider spread of infections. Here, I review the characteristics identified in this thesis and go on to suggest how control or prevention measures might be directed towards them for improved control of infection in the GB cattle herd.

Highly-connected farms

Analysis of networks in **Chapter 3** indicated that some farms were directly connected via trades to many more farms, and my analysis of contact chains identified a subset of farms with large numbers of other farms in their outgoing and ingoing contact chains, suggesting that, if infected, they had the potential to act as superspreaders. These highly-connected farms might therefore represent suitable targets for control in outbreaks of infection.

Targeting farms by the number of farms in their contact chain has been shown to be effective in surveillance strategies for moderately-contagious viruses (Frössling et al., 2012), and in predicting *M. bovis* infection in cattle (Palisson et al., 2016). However, analyses in **Chapter 4** show that, in the case of bTB in GB, the number of farms in the ingoing contact chain is considerably less important than the risk area in which the source farms are located. Consequently, risk-based control measures for bTB specifically might best be aimed at farms trading indirectly with more farms from high-risk areas.

Dairy herds

Throughout the studies presented in this thesis, I have outlined the changing structure of the GB cattle industry and changes to the composition of herds. In **Chapter 3**, I identified a reduction in the number of cattle premises and an increase in dairy herd size from 2001 to 2015. This effect was predominantly driven by a loss of smaller dairy herds and expansion of pre-existing herds. In **Chapter 3**, I found that although dairy herds were fewer in number than beef suckler herds, due to the onwards sale of most male dairy calves, they contributed a disproportionate number of animals to the beef industry. Via these onward movements we also found that dairy farms were more likely to have more farms in their outgoing contact chains (**Chapter 3**). In **Chapter 5**, I identified an increased odds of bTB incident on dairy farms compared to other herd types, as is established in the pre-existing literature (Conlan and Wood, 2016; Downs et al., 2016). Through their highly-connected position in the network and increased risk of infection, we suggest that dairy herds might be likely to be potential superspreaders, and thus in **Chapters 5** and **6** I performed a closer-analysis of within-herd dynamics on dairy farms with aims to better understand the impacts of these industry changes, at a herd and individual animal level.

Large herds

Larger herds are more likely to experience disease incursion (Brooks Pollock and Keeling, 2009), and in **Chapter 4**, my data supported previous evidence that herd size is strongly associated with bTB incidents in GB. Although it is clear that dairy herds are also more likely to be large herds, it can sometimes be difficult to disentangle these characteristics as risk factors for infection. The mechanisms that drive infections in larger herds are not yet fully explained (Gardner et al., 2002), nevertheless a number of explanations have been hypothesised: first, that larger herds partake in a larger volume of movements, indeed in **Chapter 3**, I showed that the likelihood of a farm being connected to many others in contact chains was higher, yet not inevitable, for larger herds. This might be partly attributed to the expansion of many farms; herd expansion occurs very slowly if replacement animals are home bred over time, therefore expansion is most likely achieved through purchasing more animals, inherently increasing the number of purchased animals and thereby the risk of acquiring new infections (Faust et al., 2001). Second, it has been suggested that within-herd transmission rates are much higher in larger herds (Conlan et al., 2012). Third, that there is not necessarily an increased risk of infection per se, just that there is an increased likelihood of detecting infection, if present, in infected large herds (Jordan and McEwen, 1998). As more dairy cattle in GB now exist in larger herds, a trend which seems likely to increase, it will be important to quantify the mechanisms that drive this susceptibility to infection, in order to provide effective guidance on how to mitigate the risks where possible (Villarroel et al., 2007). In summary, for the control of bTB infection in particular, the outcomes of this thesis suggest that we might consider targeting farms highly-connected to many farms in high-risk areas via their ingoing contact chains, and large, dairy herds for more optimised control of infection.

Targeted control strategies

The scales at which I have looked at cattle networks also form a framework for how one might implement control strategies targeted at these ‘high-risk’ farms; 1) by reducing the within-herd prevalence of infection on farm, and 2) by reducing the spread of infection between farms. Some control measures have advantages in their wide applicable to all pathogens, e.g. movement restrictions, whereas others are limited in their control of specific pathogens, e.g. vaccination.

Reducing within herd prevalence

In order to reduce within-herd spread, better understanding of how infections spread at a cow-cow level is required (Álvarez et al., 2014). Our findings in **Chapter 5** suggest that in future modelling, including a transmission or contact parameter based on herd type, or better still, based on housing and grazing regimes in within-herd transmission models, as has been done in some studies (Bekara et al., 2014) is appropriate and might result in more accurate predictions of infection transmission within cattle herds. The increased contacts in buildings compared to pasture suggest that this might contribute to within-herd transmission and might be particularly pertinent if dairy herds are increasingly being housed in indoor systems (Charlton and Rutter, 2017). With this in mind, whether cattle are kept indoors or outdoors, increasing available space might be beneficial in reducing disease transmission.

Brooks Pollock et al. (2014) found that targeting superspreader farms with more SICCT testing was largely ineffective, yet the use of a more sensitive, albeit less specific, test, e.g. γ-IFN testing, reduced the number of cattle reactors found over time. In using a test with lower specificity, a large number of uninfected animals would be culled and OTF status wrongly withdrawn from farms, yet fewer truly infected individuals remain. Targeting the more ‘risky’ farms that I have identified in this thesis with γ-IFN tests might reduce the prevalence of infection on farms, and also initiate movement restrictions on those farms wrongly missed by routine SICCT testing (**Chapter 2**; Conlan et al., 2012).

Due to costs and logistics in deployment, vaccinating a small number of farms for greater effect might be particularly useful. As I considered in **Chapter 2**, the most likely candidate for vaccination against *M. bovis* is the BCG vaccine, however efficacy is limited and for vaccination to be beneficial at a herd-level, tests that are able to differentiate between infected and vaccinated animals (DIVA) are required to be highly specific (Conlan et al., 2015).

Reducing high risk trading

The implementation of movement restrictions has a cost. At a farm-level, having OTF status suspended or withdrawn and being unable to sell stock can mean extra feed costs and increased stocking densities (Bennett and Cooke, 2006). An increased density of animals on the farms can also negatively impact upon animal welfare and facilitate the transmission of infections, as discussed in **Chapter 2** and **6**. Movement restrictions specifically targeted at specific ‘high-risk’ edges in the network are likely to be particularly effective in reducing spread along the network, but also limit the wider commercial impact on farms that have less influence on disease transmission (Enright and Meeks, 2015; Gates and Woolhouse, 2015).

Testing requirements for movements of cattle between farms, are based primarily on the bTB risk area, in which the immediate source farm is located. However, analyses in **Chapter 4** revealed that this direct trade does not fully account for the risk of a bTB incident, and that it might also be prudent to consider the risk regions of indirect trading connections, i.e. the locations of the source farms for the source farm. A recent study showed that the odds of an individual cow being found as a reactor in a slaughterhouse were best characterised by the region where the animal had spent most of its life (McKinley et al., 2018), thus it seems that the zoning of England into risk areas captures the risk of disease well. Using the ingoing contact chains of farms or an assessment of where the animal has lived for most time to assess the risk associated with a particular farm or animal, might be useful to include in risk-based trading protocols in the future (Adkin et al., 2015).

Ultimately, for farmers to assess indirect trading connections, there needs to be more communication and transparency around trading. The ibTB mapping tool, which shows the location and basic details of bTB incidents since 2015 (Animal and Plant Health Agency and Environmental Research Group Oxford, 2019) has substantially increased the ability of farmers to assess the risk of bTB from particular farms, however its full potential is likely to lie in integrating this information into auction markets. Information needs to be communicated clearly in real-time and be available at the point of sale.

In **Chapter 4**, I show that large portions of GB are traversed in relatively few movements in some cases, therefore increasing barriers to trade between risk areas is likely to assist in fragmenting the cattle movement network. Post movement testing in Scotland has been an effective deterrent to cross-border trade of cattle, most likely due to increased costs and extra time spent testing animals (Gates et al., 2013b). Since our study period, the introduction of post-movement testing for animals moved from high-risk areas might hopefully have had similar effects to reduce this inter-regional trade.

The future of bTB control

In studying bTB it is important to acknowledge that, although cattle-to-cattle transmission is likely to be responsible for the majority of new infections, *M. bovis* is a multi-host pathogen with a wildlife reservoir (Donnelly and Nouvellet, 2013). Modelling of control measures suggests that badger control will only have limited effects on overall bTB (Brooks Pollock et al., 2014). Industry-led culling of badgers is now widespread in the England High Risk and Edge areas, and tentative analysis from the first two years of culling suggests similar results to that found in the RBCT (Brunton et al., 2017). Further research is ongoing into how badger vaccination might provide an opportunity for non-lethal control (Buddle et al., 2018), and in so doing guarding against detrimental effects of social perturbation (McDonald et al., 2008).

The current political climate provides a period of uncertainty for the future of bTB control in Britain, as many laws governing bTB surveillance, diagnostic testing, and vaccination of cattle are mandated by the European Union. In theory, the separation of the UK from the EU would allow the UK and the devolved governments to change current policies, perhaps with the more widespread use of diagnostic phage (Swift et al., 2016), antibody testing (Waters et al., 2011) and cattle vaccination and DIVA testing.

The utility of networks

Simple mathematical models of transmission can use homogenous mixing assumptions to estimate the contact rate of a host population, or, as this thesis has shown, networks can provide an additional framework by which to understand disease spread (Craft, 2015). Networks are a versatile framework with many possible applications in a wide range of fields, yet, networks are not universally applicable. Network models compared to models of homogenous mixing have the most diverse outcomes where individual host contacts are heterogeneous and where infections are spatially aggregated (Keeling and Eames, 2005), thus making networks a particularly useful tool in these cases.

In studying cattle networks in this thesis, I found that networks at the scales of farms and of animals, were, on initial inspection, densely connected. National-level cattle movements formed a large network structure that was different to random networks in all aspects that were evaluated. It showed scale-free structure in that farms acted as hubs, and small-world clustering. Aiming to better comprehend this large structure I considered temporal pathways in the form of contact chains, which revealed that farms formed two broad groups in terms of being highly connected, or relatively isolated. As the chains built, and connections were added further from the root farm, connections to some source farms tended to be repeated, indicating that at these points the network neighbourhoods of these farms might be locally saturated, suggesting that the magnitude of the chain was comparable to a ‘local neighbourhood’ of each farm (Keeling, 1999). On a smaller scale, within-farm cow-cow interactions were also dense if contacts were aggregated over the entire study period, yet temporally and spatially dividing the network revealed more subtle interactions between cattle.

Many different methods have been used to analyse networks and I have demonstrated in this thesis that the methods chosen can be integral in shaping study outcomes (Dawson et al., 2019). At both scales, the measures by which I described the networks were, on the whole, different from random networks of the same size, thereby suggesting that using networks to study the interactions of cattle and cattle farms is worthwhile.

Conclusion

Maintaining a sustainable livestock industry in Britain is important for the 350,000 people it employs (Office for National Statistics, 2018) and for consumers concerned with high standards of animal welfare and the regulated use of antimicrobials and hormones (Clark et al., 2016). Livestock epidemics can have impacts upon the wider agricultural industry and, in large-scale epidemics, on the national economy (Thompson et al., 2002). It is therefore in the common interest to find a balance between commercial gain and farm biosecurity. Constructing networks from farms and individual cattle gives us an invaluable tool with which to study the transmission and potential control of infections. In this thesis, I have shown, through static, temporal, and spatial network analysis, and statistical modelling, that networks can add to our understanding of the epidemiology of cattle in Great Britain and can help to inform risk-based strategies for control.

Appendices

Appendix A (Chapter 3)

Methods

Network analysis

We generated random networks using the Erdös-Renyi model (Erdös and Rényi, 1959) to create 10 000 random networks per year (2001–2015). Edge weights from the observed network were randomly allocated to edges in the random networks for each respective year. We calculated the stability of the network position of individual farms across years using the methods of Wilson et al. (2013). For this analysis, we used nodes present in all years from 2001–2015, and within each year farms were ranked by the value of each network measure; in-degree, out-degree, in-strength, out-strength, and betweenness. We compared the standard deviation (SD) of ranks to the SD of ranks following node-based permutations (the swapping of attribute values between nodes of the network) of each yearly network, for each network measure. If the observed SD was lower than, and lay outside of the 95% confidence interval of, the SDs calculated from randomised networks, it was deemed to be repeatable through time. We additionally calculated what aspects of network position (in-degree, out-degree, in-strength, out-strength, and betweenness) correlated with mean herd size using Spearman's rank correlation coefficients. To account for the non-independence of network data, we calculated equivalent Spearman's rank correlations following node-based permutations for each year.

Contact chains

Using the same method employed to assess the stability of network measures on individual farms, we ranked ICC and OCC over all study years. In addition, using the same method used for correlations of mean herd size and network measures, we calculated the correlation of ICC with OCC, mean herd size with ICC and OCC, and node measures (in-degree, out-degree, in-strength, out-strength) with their corresponding ICC or OCC. Spearman's rho (rs) and p-values were obtained for each individual study year and rs is reported as a mean of all study years, with 95% confidence intervals.

Results

Temporal variation in contact chains

To investigate the variation in the contact chains of individual farms over time we performed additional analysis to assess our different methods. We compared the mean, median and maximum of the combined 24 monthly-spaced chains and the three combined annually-spaced chains (Appendix B Fig. B1) for years 2012 to 2014 using Spearman's rank analysis. The majority of farms showed little variation in the number of farms in their chains between different starting months of the 24 monthly-spaced chains. Farms with around 10 000 farms in their chain seem to exhibit a much larger degree of variation than those with chain lengths both below and above them (Appendix B Fig. B9). We were concerned that seasonal fluctuations in movements would cause variation in the number of farms in a contact chain depending on the starting month of the chain. However, no consistent pattern was apparent by visual examination of monthly differences in the number of farms in contact chains from 2012-2014. This suggests that the variation seen in farms with contact chains of around 10 000 is due to individual farm trading behaviour rather than population-level changes in movements. Overall, there was a strong correlation between the summary values (mean, median and maximum) of the 24 monthly-spaced chains and the 3 annually-spaced chains (mean ICC: rs = 0.932 (95% confidence intervals (CI) 0.931–0.0.933), mean OCC: rs = 0.909 (95% CI 0.908–0.910), median ICC: rs = 0.845 (95% CI 0.843–0.847), median OCC: 0.846 (95% CI 0.843–0.848), maximum ICC: rs = 0.923 (95% CI 0.922–0.924), maximum OCC: rs = 0.874 (95% CI 0.872–0.875); $p < 0.001$, n = 76,031).

Appendix B (Chapter 3)

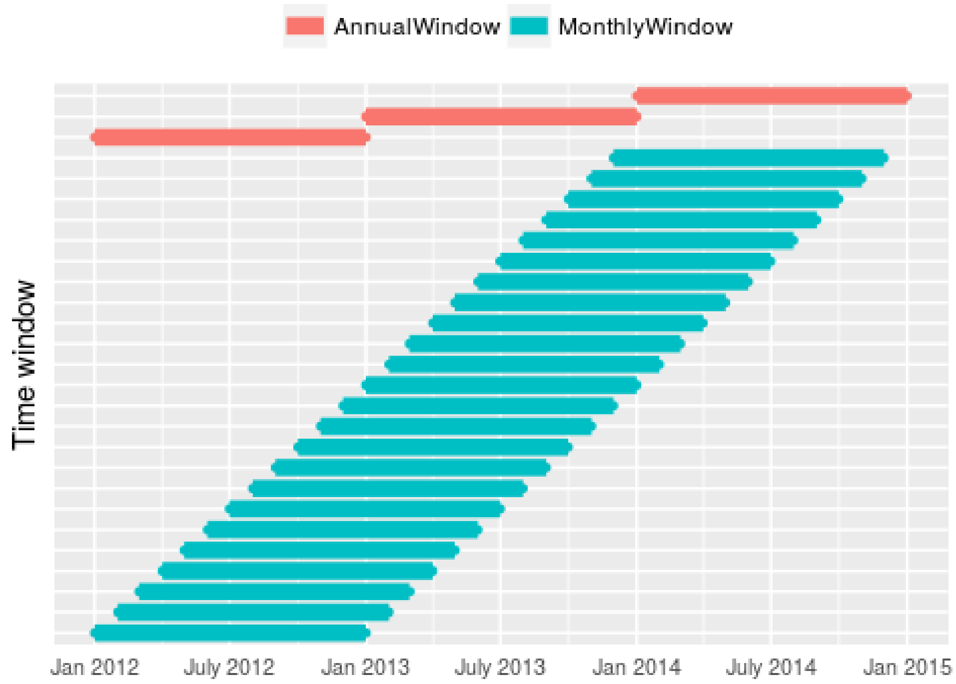


Figure B1. *Schematic showing the two different time windows over which contact chains have been analysed and compared. All periods span 12 months. Teal-coloured bars show the 24 x monthly-spaced periods and red-coloured bars span the 3 x annually-spaced periods from which movements are taken to calculate contact chain values.*

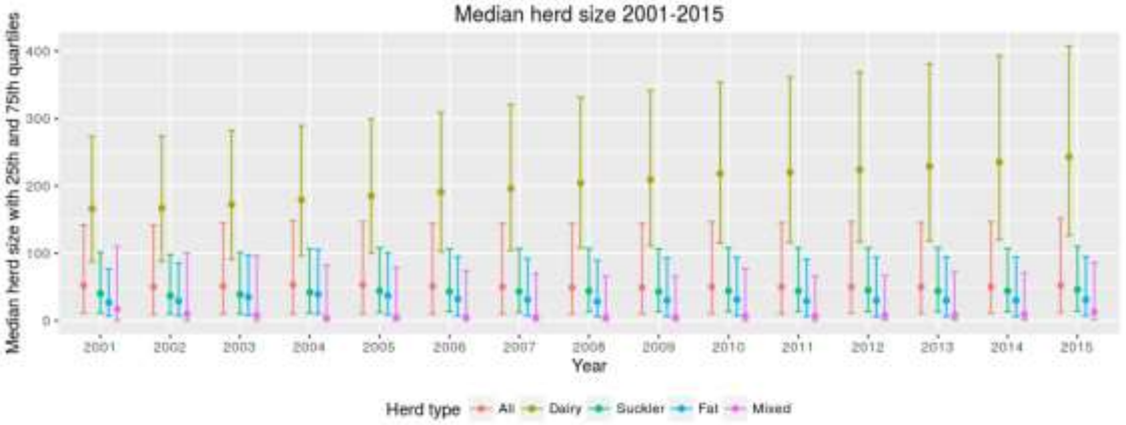


***Figure B2**. Summary measures for average herd sizes (median of the daily number of cattle on the premises for each year) from 2001 to 2015 and for each herd type calculated from CTS data. Data are shown as boxplots with dots representing median values and whiskers representing 25th and 75th percentiles of the data, showing an increase in the size and variation of dairy herds and stability in the herd size of other herd types.*

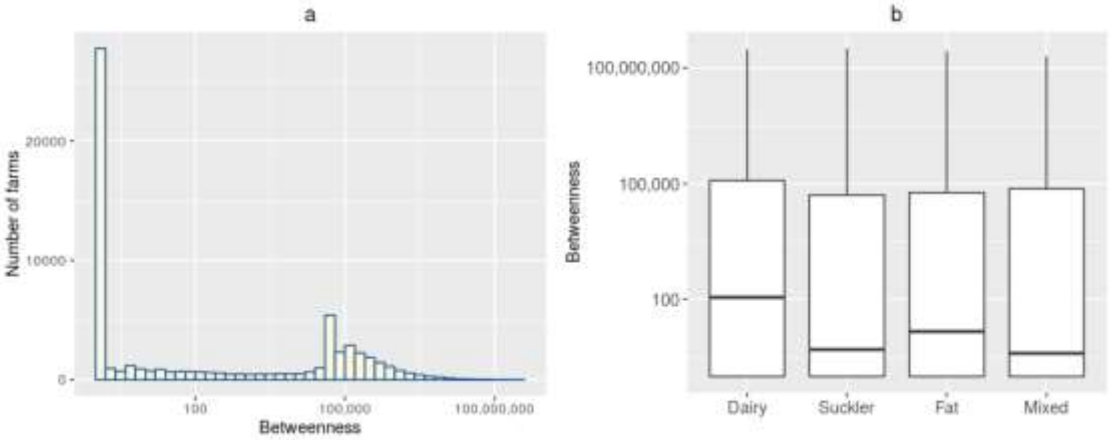


Figure B3. *Individual node betweenness centrality calculated using inverted edge weights: a) histogram with count log transformed (n+1) to show peaked distribution of betweenness and zero-inflated distribution, b) boxplot by herd type for the 2015 network of cattle movements showing little difference between herd types. Tops and bottoms of the boxes represent the 75th and 25th percentiles of the data, the black centre lines are the medians, and the whiskers extend from their respective hinge to the smallest or largest value no further than 1.5 * interquartile range.*

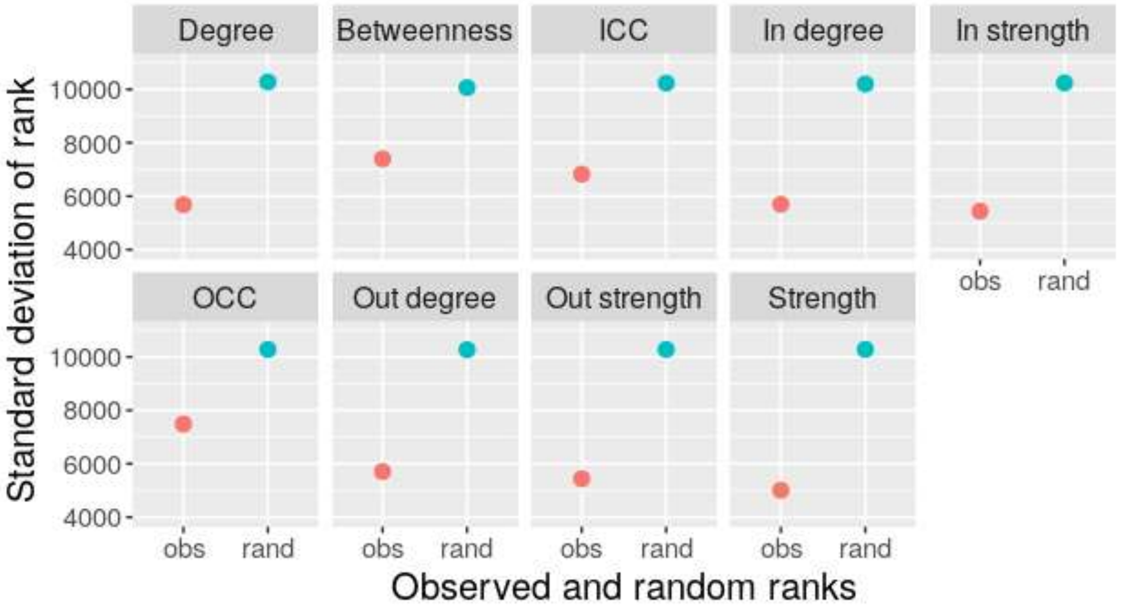


Figure B4. *Observed standard deviation in farm ranks, when ranked by network measure over multiple years. Randomised standard deviation is from 1000 randomly generated Erdös-Renyi networks. For each metric, observed standard deviation is less than random, indicating that farms were more consistent over time in their network measures than if the network were randomised.*

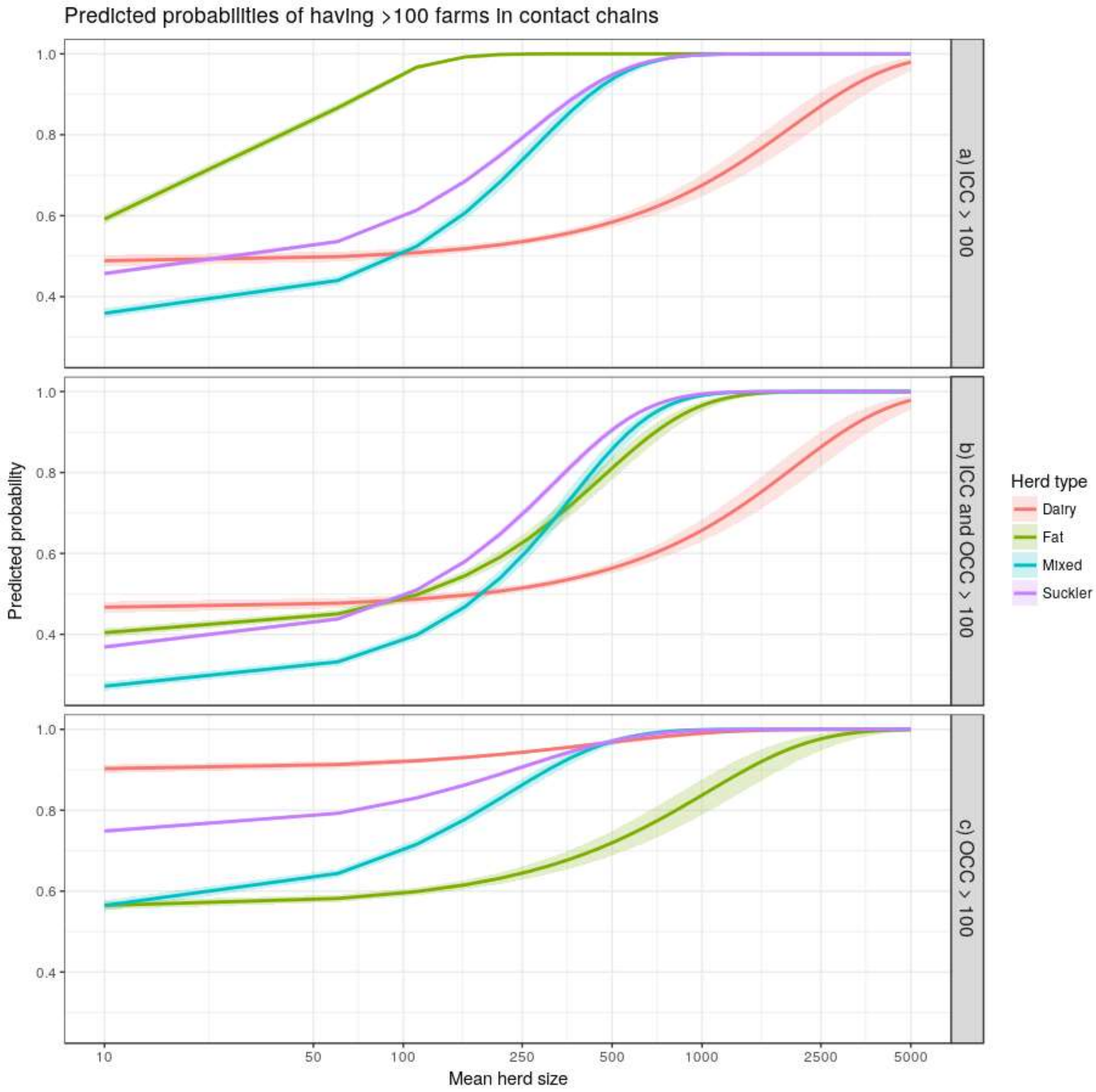


Figure B5. *Predicted probabilities from logistic regressions using mean herd size and herd type as explanatory variables of farms with over 100 farms in their a) ICC, b) both ICC and OCC and c) OCC as the response variable. Shaded areas represent the 95% confidence intervals. Uses the mean value of 24 sequential monthly-spaced contact chains from 2012-2014 as the value for contact chains.*

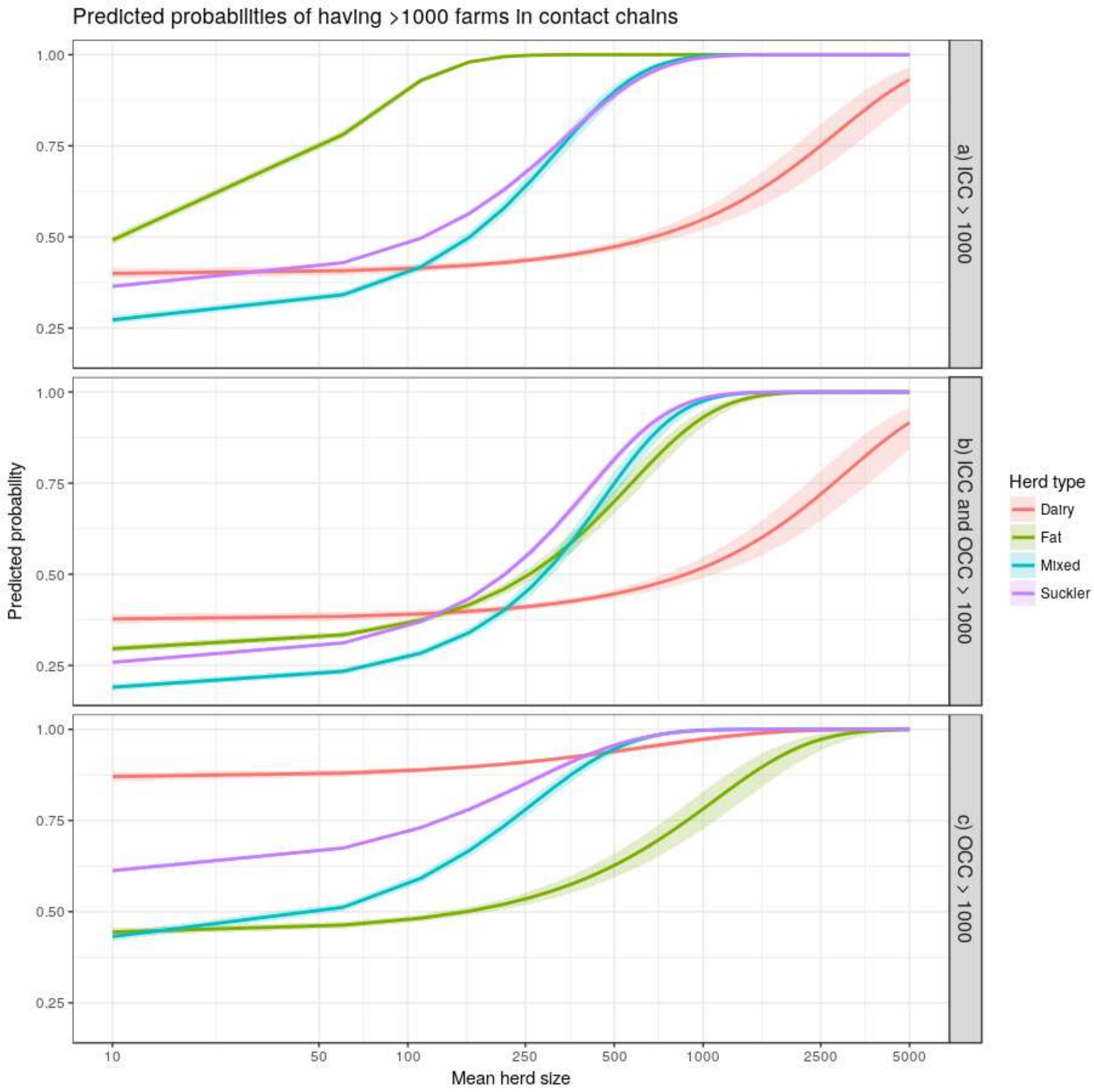


Figure B6. *Predicted probabilities from logistic regressions using mean herd size and herd type as explanatory variables of farms with over 1000 farms in their a) ICC, b) both ICC and OCC and c) OCC as the response variable. Shaded areas represent the 95% confidence intervals. Uses the mean value of 24 sequential monthly-spaced contact chains from 2012-2014 as the value for contact chains.*

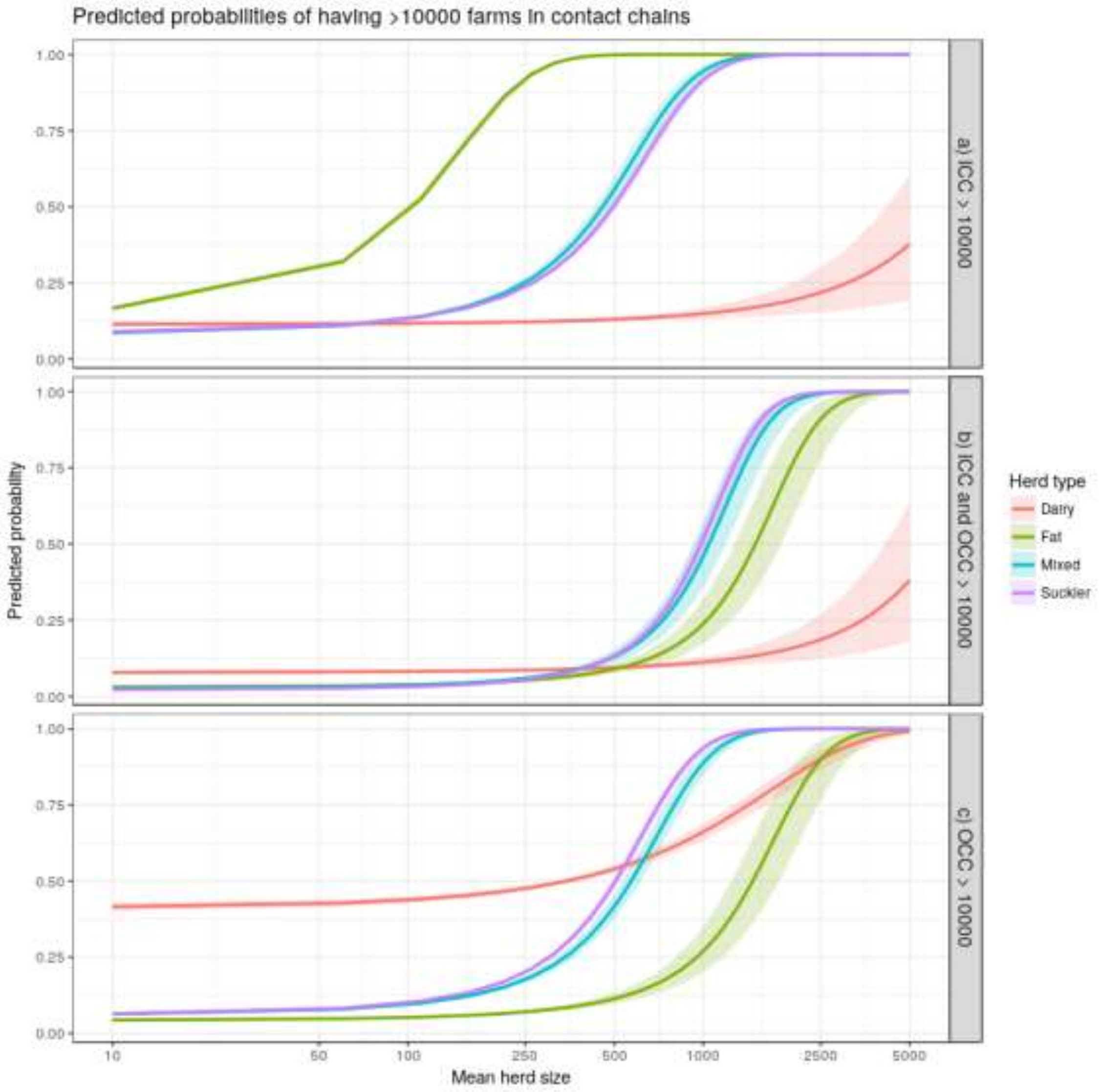


Figure B7. *Predicted probabilities from logistic regressions using mean herd size and herd type as explanatory variables of farms with over 10 000 farms in their a) ICC, b) both ICC and OCC and c) OCC as the response variable. Shaded areas represent the 95% confidence intervals. Uses the mean value of 24 sequential monthly-spaced contact chains from 2012-2014 as the value for contact chains.*

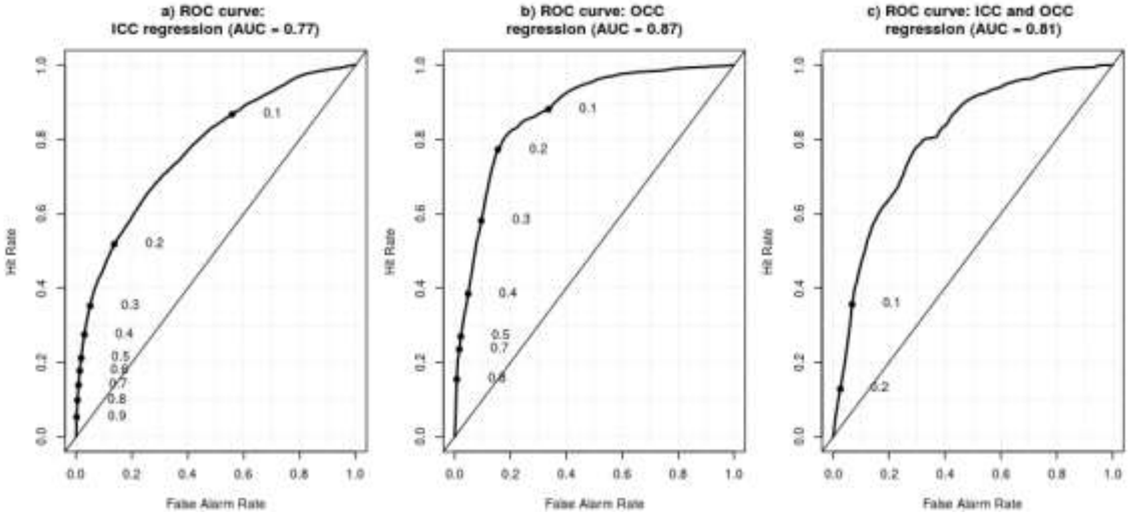


Figure B8. *Receiver Operating Characteristic (ROC) curves for logistic regression using farms with over and under 10 000 farms in their a) ICC, b) OCC and c) ICC and OCC as response variables. The area under the curve (AUC) indicates discriminatory power of each model (0.70-0.80 = adequate, 0.80-0.90 = excellent): 0.77, 0.87 and 0.81 respectively.*

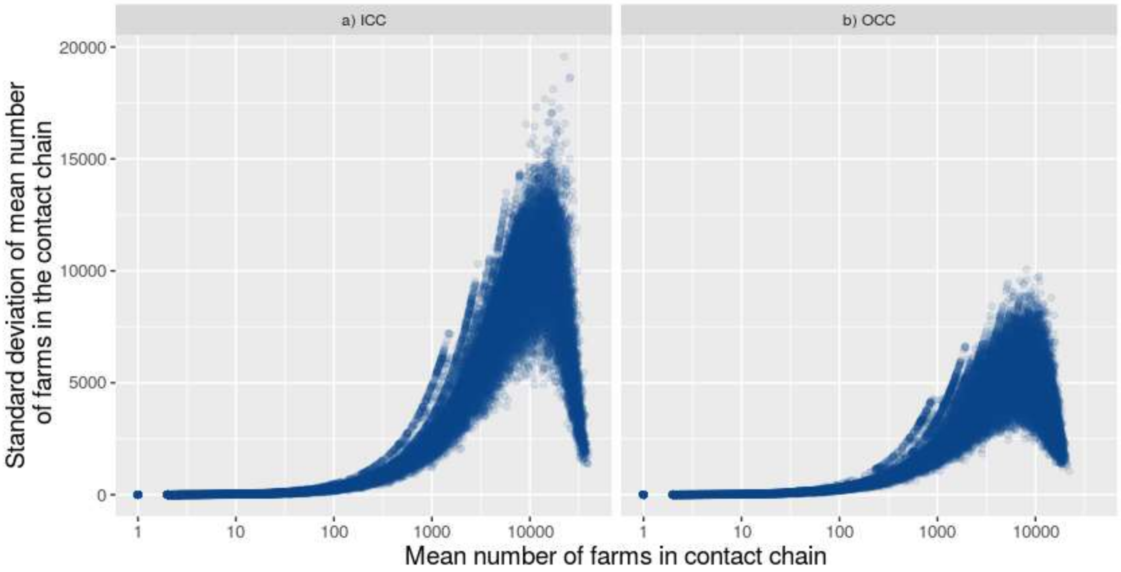


Figure B9. *Relationship between standard deviation and the mean of 24 monthly-spaced a) ICCs and b) OCCs of all active farms in GB over the period 1st January 2012 to 31st December 2013. Showing that variation between contact chains of individual farms increases as the mean contact chain increases, but then decreases for farms at the extreme right tail of the distribution with over 10 000 farms in their chain.*

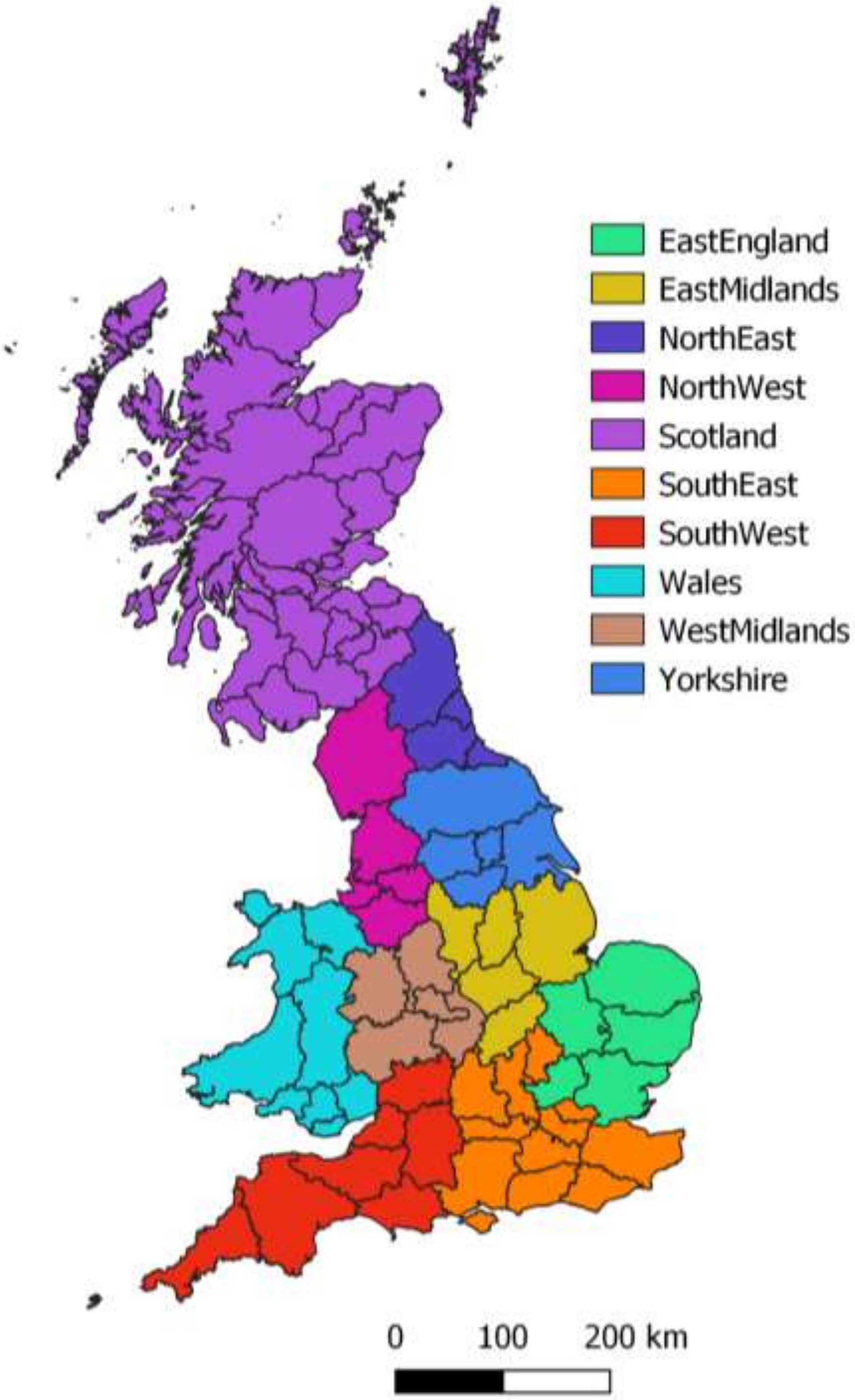


Figure B10. *Map of Great Britain showing counties grouped into regions and countries used in our logistic regression analyses.*

Table B1. *Glossary of network analytical terms and accompanying igraph functions.*

Network measure	igraph function (R language)	Description
Shortest path	-	The path between two nodes that traverses the least number of edges in the network.
Betweenness centrality	`estimate_betweenness()*`	A global measure of centrality indicating the number of times a node lies on the shortest-paths between all other nodes (Newman, 2015) in the network. Note* *In weighted networks, edge weights influence these shortest paths and betweenness values. For disease transmission, edges with high weight (more animals moved) represent an increased risk of pathogen transfer* (Natale et al., 2011)*. However, the algorithm used to calculate weighted betweenness by igraph treats high weight as a cost to edge traversal, therefore, we calculated betweenness with inverted weights (1/weight) to represent this higher risk of transmission.*
Edge density	`edge_density()`	Defined as the ratio between the number of edges in the network and all possible edges giving a value for density that is relative to the number of nodes. In general, a more dense network has more edges along which transmission can occur (Shirley and Rushton, 2005).
Degree assortativity	`assortativity_degree()`	The extent to which nodes connect with those with a similar degree to them and gives an indication of the presence of hubs in the network, i.e. those nodes that connect to many more other nodes than the majority. Hubs facilitate rapid spread of infection to their connected nodes, temporarily increasing the speed of epidemic spread (Kiss et al., 2006a).
Reciprocity	`reciprocity()`	Indicates the extent to which nodes reciprocally connect, by both buying and selling, to one another in a directed network.
Clustering coefficient	`transitivity()`	The *clustering co-efficient* detects network clustering and shows the tendency for nodes to be connected to other nodes that are themselves directly connected.
Average path length	`average.path.length()`	The average of all shortest paths in the connected network. It indicates in how many steps the GSCC can be traversed.
Giant strongly-connected component (GSCC)	`strong <-components(g, mode='strong')` `gscc <- max(strong$csize)`	The largest group of connected farms from which any node may reach any other node via directed links (Pastor-Satorras et al., 2015), and has been used to estimate the potential extent of an epidemic (Kiss et al., 2006b).
Small-world networks	-	Low density networks with high clustering exhibit *small-world* type properties, which can increase disease transmission but may reduce the spatial extent of the epidemic (Christley et al., 2005).
Scale-free networks	-	Small-world type networks with a power-law (Clauset et al., 2009) degree distribution are considered to be *scale-free*. Specifically in scale-free networks disease spread can occur regardless of epidemic thresholds, which often govern the spread of disease in other network structures (Pastor-Satorras and Vespignani, 2001). This can facilitate the transmission of infections with low reproductive rates, and therefore makes a population more susceptible to a greater range of pathogens.

Table B2. *Observed network values for cattle movement networks in Great Britain 2001–2015. The p-values indicate the probability of the values calculated in the 10 000 Erdös-Renyi random networks being greater than the equivalent value calculated from the observed network. We applied a Bonferroni correction to α=0.05 to account for the fact that we tested the same network values over 15 years. Therefore, p-values are considered statistically significant when p<0.003 or p>0.997.*

	Degree Assortativity		Reciprocity		Clustering coefficient		GSCC (% of network)		SD (in degree)		SD (out degree)		Median betweenness		SD betweenness (inverted weights)		Average Path Length	
	O	*p*	O	*p*	O	*p*	O	*p*	O	*p*	O	*p*	O	*p*	O	*p*	O	*p*
2001	-0.0916	>0.9999	0.0319	<0.0001	0.0035	<0.0001	*34.0*	>0.9999	71.7	<0.0001	5.83	<0.0001	*0.00*	>0.9999	9202867	<0.0001	6.54	>0.9999
2002	-0.0306	>0.9999	0.0351	<0.0001	0.0111	<0.0001	*47.6*	>0.9999	24.0	<0.0001	10.2	<0.0001	*2.00*	>0.9999	7228185	<0.0001	7.14	<0.0001
2003	-0.0572	>0.9999	0.0368	<0.0001	0.0138	<0.0001	*54.8*	>0.9999	29.6	<0.0001	12.5	<0.0001	*5.00*	>0.9999	6046683	<0.0001	6.45	<0.0001
2004	-0.0711	>0.9999	0.0397	<0.0001	0.0150	<0.0001	*57.5*	>0.9999	30.8	<0.0001	12.7	<0.0001	*9.00*	>0.9999	6384252	<0.0001	6.42	<0.0001
2005	-0.0877	>0.9999	0.0422	<0.0001	0.0144	<0.0001	*56.5*	>0.9999	32.6	<0.0001	12.1	<0.0001	*6.00*	>0.9999	6439959	<0.0001	6.54	<0.0001
2006	-0.0785	>0.9999	0.0382	<0.0001	0.0138	<0.0001	*56.0*	>0.9999	42.6	<0.0001	12.9	<0.0001	*4.00*	>0.9999	7726244	<0.0001	6.33	<0.0001
2007	-0.0853	>0.9999	0.0379	<0.0001	0.0139	<0.0001	*52.6*	>0.9999	35.8	<0.0001	11.7	<0.0001	*2.00*	>0.9999	7108390	<0.0001	6.49	<0.0001
2008	-0.0872	>0.9999	0.0353	<0.0001	0.0150	<0.0001	*52.5*	>0.9999	38.2	<0.0001	12.5	<0.0001	*2.00*	>0.9999	5940278	<0.0001	6.57	<0.0001
2009	-0.0854	>0.9999	0.0339	<0.0001	0.0146	<0.0001	*53.9*	>0.9999	38.9	<0.0001	13.2	<0.0001	*4.00*	>0.9999	7127118	<0.0001	6.41	<0.0001
2010	-0.0859	>0.9999	0.0342	<0.0001	0.0143	<0.0001	*52.6*	>0.9999	38.0	<0.0001	12.7	<0.0001	*4.00*	>0.9999	6082075	<0.0001	6.60	<0.0001
2011	-0.0883	>0.9999	0.0329	<0.0001	0.0144	<0.0001	*52.1*	>0.9999	39.2	<0.0001	12.9	<0.0001	*3.00*	>0.9999	5407043	<0.0001	6.67	<0.0001
2012	-0.0890	>0.9999	0.0315	<0.0001	0.0138	<0.0001	*50.4*	>0.9999	37.6	<0.0001	13.1	<0.0001	*3.00*	>0.9999	4746989	<0.0001	6.62	<0.0001
2013	-0.0916	>0.9999	0.0305	<0.0001	0.0134	<0.0001	*49.8*	>0.9999	38.7	<0.0001	12.9	<0.0001	*4.00*	>0.9999	4557005	<0.0001	6.97	<0.0001
2014	-0.0911	>0.9999	0.0303	<0.0001	0.0136	<0.0001	*49.3*	>0.9999	38.9	<0.0001	13.0	<0.0001	*5.75*	>0.9999	6587732	<0.0001	6.93	<0.0001
2015	-0.0846	>0.9999	0.0306	<0.0001	0.0132	<0.0001	*50.0*	>0.9999	40.0	<0.0001	13.2	<0.0001	*7.00*	>0.9999	4934042	<0.0001	6.89	<0.0001
	White = Observed lower than random network							Grey highlight = Observed higher than random network										
	O = Observed, *p* = *p*-value of comparison between observed and random networks.																	

Table B3. *Receiver operating characteristic values for logistic regression models using different thresholds as a response variable for the number of farms in a contact chain.*

ROC value	Threshold number of farms in response variable		
Chain	100	1000	10000
ICC	0.71	0.71	0.77
OCC	0.75	0.76	0.87
ICC and OCC	0.68	0.69	0.81

Table B4. *Power-law exponents from the degree distribution in the observed networks from 2001–2015. Test statistics are from a likelihood-ratio test performed to compare models that fitted the observed network to a power-law distribution with that from a log-normal distribution. In all years, apart from 2001 when trading patterns were grossly perturbed by movement restrictions during a foot and mouth disease epidemic, there was no evidence that either model fitted better.*

Year	Power-law exponent	R test statistic (log-likelihood ratio)	One-sided *P*-value
2001	3.14	4.9940	2.96E-07
2002	3.00	0.0321	0.487
2003	2.91	-0.7544	0.775
2004	2.89	-0.5700	0.716
2005	2.82	-0.2211	0.587
2006	2.71	-0.7489	0.773
2007	2.76	-0.1060	0.542
2008	2.77	0.0019	0.499
2009	2.77	0.4970	0.310
2010	2.78	0.4315	0.333
2011	2.75	-0.2503	0.599
2012	2.81	-0.0519	0.521
2013	2.78	-0.2134	0.584
2014	2.80	0.0744	0.470
2015	2.83	-0.2063	0.582

Appendix C (Chapter 4)

Methods

Regional definitions: England risk areas and Wales

During the study period (2010–2016) there were multiple changes to bovine tuberculosis (bTB) surveillance policies in Great Britain. Up until the end of 2012, herds in England were tested every 1, 2, 3, or 4 years according to the classification of their parish (smaller geographical areas within counties). On 1st January 2013, three risk areas were implemented in England; the England High Risk Area, where infection is considered endemic in badgers and cattle, the England Low Risk Area, where there is low incidence, and the England Edge area, which encompassed regions deemed to be at highest risk of endemicity in the near future (Lawes et al., 2016; see Fig. S1 for map). In most cases, whole counties lie within a single risk area, however five counties (East Sussex, Oxfordshire, Warwickshire, Derbyshire and Cheshire) were divided between the England High Risk Area and the England Edge Area along parish boundaries. For the purposes of this study, we defined these risk areas at the parish-level using reference data obtained from the APHA and as per the 2015 surveillance report by the APHA (Animal and Plant Health Agency, 2016). After the risk areas were implemented (2013 onwards), herds in the England High Risk Area and England Edge Area were tested at least annually and herds in the England Low Risk Area and Scotland were tested every four years, although in Scotland some farms were exempt from live animal testing according to risk-based analyses (Kao, 2011). Despite marked regional differences in herd incidence (and devolved bTB policy) within Wales, we elected to categorise the entire country as one risk area, since during the study period all Welsh regions were tested annually.

Herd type definition

Using Cattle Tracing System (CTS) data, *suckler* farms were defined by a majority of female beef animals, aiming to capture herds where calves are reared by their dams (cow-calf systems). *Dairy* farms were defined by a majority of female dairy cattle, identifying herds producing milk commercially. *Fattening* units were defined by a male animal majority, identifying herds that mainly reared cattle for beef production. *Mixed* farms were those where no one breed type or sex constituted more than 50% of the herd.

Constructing networks

In constructing our networks, only farm to farm movements were analysed, omitting premises where animals resided for less than one day. Movements between farms via transitory locations, such as markets, were classed as a direct link with the transitory location excluded, as we considered short-term locations of minimal importance to the transmission of bTB (Skuce et al., 2011), in contrast to their importance in the spread of highly infectious diseases. Multiple animals moving to and from the same farms on the same day were aggregated to form a single directed connection.

Constructing contact chains

The ‘root farm’ is at the start of the chain. Any farms that have directly sold animals to the root farm in the previous year are considered to be at level one in the contact chain. Farms initiating *previous* movements *onto* level one farms (i.e. those made before the move to the root farm) are considered to be at the second level away from the root farm (i.e. two movements in a temporal sequence connect the root farm and a farm at the second level of the chain), farms initiating movements of cattle onto second level farms are considered to be at the third level, and so on. All movements creating connections up the chain precede the movement between the lower levels and thus maintain a possible infection pathway. The resulting chain is fully-connected by animals that have been on the same farm for at least one day and therefore have the potential to transmit infection between farms. Any farm in the ICC of a root farm is considered a ‘source farm’. We generated twenty-four ICCs for each active root farm using movements recorded during the study period. The twenty-four chains end sequentially one month apart and each include the previous 12 months of movements. We chose 12 month periods as a conservative estimate of how long a bTB infected farm may trade without being detected. We calculated the *number* of farms in each ICC (source farms) but encountered computational limitations in *identifying* farms at more than eight levels away in extremely large chains, and so each chain was curtailed at this point. We estimate approximately 28% of root farms had source farms above level 8, and these source farms were not included in the analysis. However, given that the likely effect of farms diminishes at greater distances in the ICC and that farmers have less and less control over connections as the chain builds, we consider this is sufficient to capture the more important risk factors related to bTB transmission.

Trialling of different analytical approaches

Due to the complexity of including network parameters in statistical analysis and the number of variables that were included in the original model, I trialled a number of analytical approaches in terms of selecting variables for the multivariable analysis. Backwards and forwards AIC selection produced similar results, as did stepwise selection based on the Bayesian Information Criterion. We assessed models additionally with Bayesian model averaging and random forest analyses which also suggested similar variables were important in the full model. As each method showed very similar selection of major variables, this provided supportive evidence that the chosen method did not have a strong influence on our outcomes. Additionally, all methods struggled with multiple co-linear lagged variables for levels of the contact chains and prompted us to select the chosen method of selecting lagged variables as a more robust approach.

Selecting contact chain variables

Using the method described in the multivariable analysis section in Methods, we split the contact chain variables into two groups based on if we expected them to increase or decrease the odds of bTB on the root farm in the univariate analysis (Fig. S3). First, we ran the baseline model with each cumulative variable from both groups. We selected the variable from each group that was included the model with the lowest AIC. For the increased risk group this was the number of source farms in the England High Risk Area at levels 1–3 and for the protective group this was the number of source farms in Scotland at levels 1–8 (Fig. S3). Second, to determine if the selected variables had an impact on the AICs of one another, we ran a sensitivity analysis in which we repeated the previous steps by including each variable from each group (e.g. the protective group: number of farms in the England Low Risk Area at levels 1, 1–2, 1–3, etc. up to levels 1–8 and the number of farms in Scotland at levels 1, 1–2, 1–3, etc. up to levels 1–8) with the baseline model, but also added the variable from the other group that was selected in the first stage (e.g. selected variable from the increased risk group: the number of farms at levels 1–3 in the England high risk area). Initially, the variable in the model with the lowest AIC was the number of Scottish farms at levels 1–8 (Fig. S3, however, after adding in the ‘best’ variable from the increased risk group, and testing all variables from the protective group, the model with the lowest AIC became the number of farms at levels 1–8 in the Low Risk Area (Fig. S3). The best performing representative variable from the increased risk group (number of farms at levels 1–3 in the England high risk area) consistently had the lowest AICs with the either the number of farms in the England Low Risk Area or Scottish farms at levels 1–8 (Fig. S3). As the model containing both the number of farms in the England High Risk Area at levels 1–3 and the number of farms in the England Low Risk Area at levels 1–8, gave the lowest AIC, we selected the England Low Risk Area variable from the protective group. We also performed the above sensitivity analysis with and without the local risk factor (the proportion of farms with a bTB incident in 2010–2014 at a radius of 0–8km from the root farm). AIC values were consistently higher, by a mean of 3.22% (standard deviation = 0.0465), when the local risk factor was not included, therefore it was included in the final model (Fig. S3).

Appendix D (Chapter 4)

Figure D1. *Map of Great Britain illustrating the regional areas used in this study: Scotland, Wales, England High Risk Area, England Low Risk Area and England Edge Area. Bovine tuberculosis risk areas of England were defined at parish level for our analyses, as defined by the Animal and Plant Health Agency bovine tuberculosis surveillance report 2015 (Animal and Plant Health Agency, 2016). Counties that were partly in the England High Risk Area and partly in the England Edge Area from 2013 to the study end (Cheshire, Derbyshire, Warwickshire, Oxfordshire, East Sussex) are coloured on the map as the England High Risk Area and denoted with a white asterisk.*

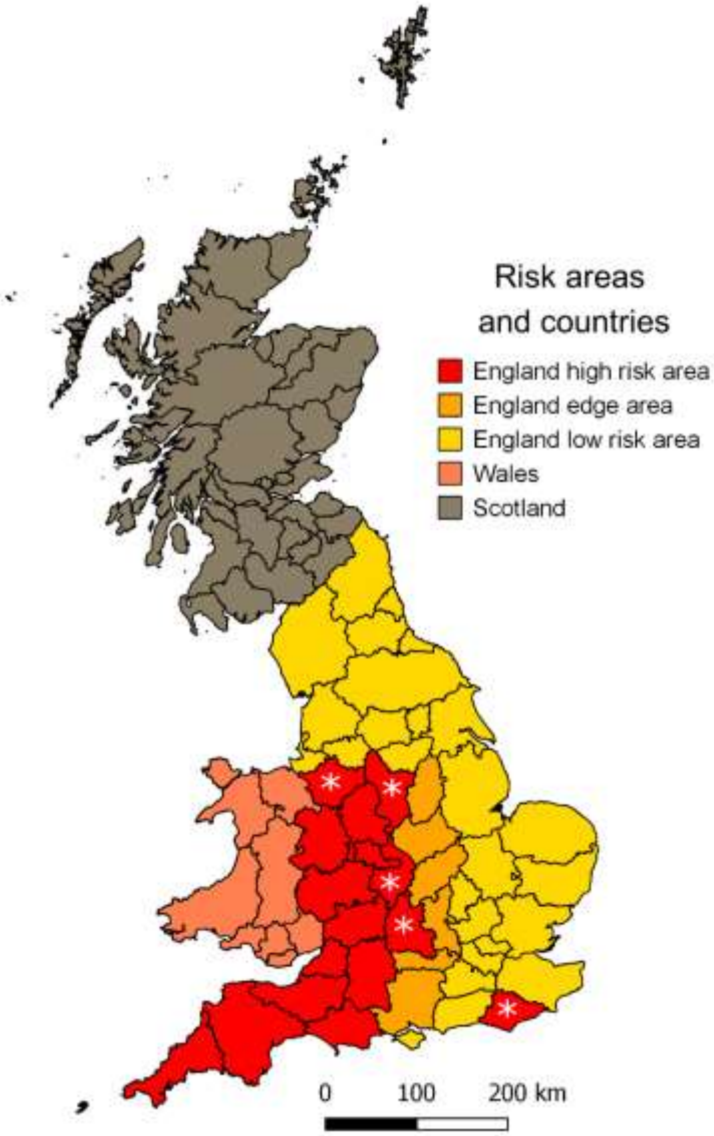


Figure D2. *Relative performance of test models to inform variable selection for inclusion in the final GB multivariable model. Variables tested are for the proportion of farms with a bTB incident 2010–2014 at increasing distance from the root farm. AIC values are shown for the baseline model, without any local bTB variable, and the five competing multivariable logistic regression models which included the proportion of farms within 0–2 km, 0–4 km, 0–6 km, 0–8 km, and 0–10 km of the root farm respectively. All models used the response variable of whether the root farm had a bTB incident 2015-2016 and all included the baseline explanatory variables: root farm risk area, root herd size, root herd type, whether or not root farm had bTB incident 2010-2014, root farm betweenness, the mean number of source farms in the ingoing contact chain (ICC), whether or not the farm purchased cattle, the mean number of cattle purchased annually, whether or not the farm purchased any animals from the England High Risk Area. The test model that included the proportion of farms within 8 km of the root farm with a bTB incident 2010–2014 had the lowest AIC value and was therefore included in our final multivariable model.*

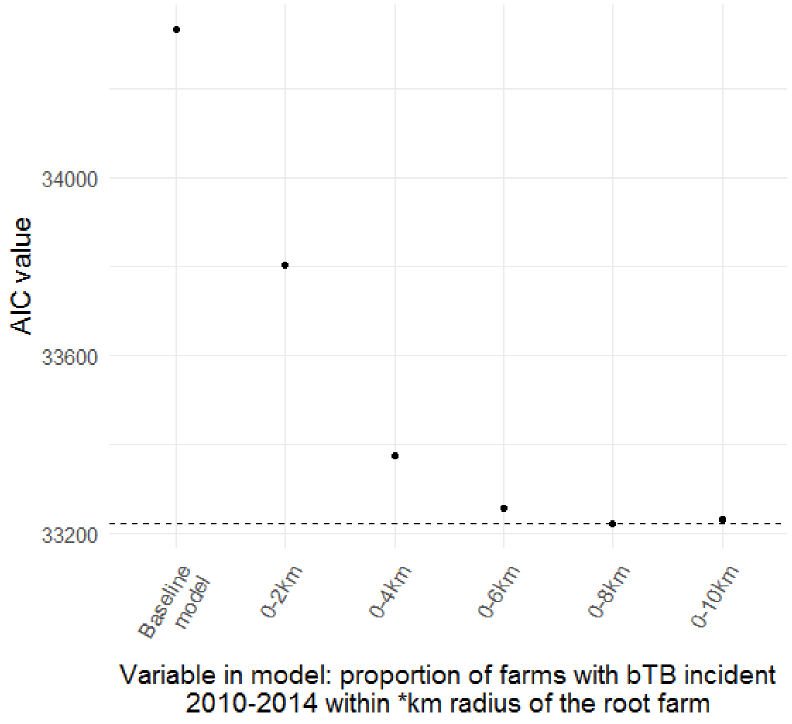


Figure D3. *Relative performance of test models to inform variable selection for inclusion in the final GB multivariable model. Variable selection was based on the combination of parameters included in the model with the lowest AIC value. Each variable within a category (e.g. increased risk) was tested with the selected variable from the other category (e.g. decreased risk) and the baseline model. Additionally, all variables (from left to right on the x-axis) were run with (bottom row) and without (top row) the local bTB incident variable (0–8 km see Fig. S2) and models always performed better with this variable included. The optimum combination of variables defined by lowest AIC values included the number of farms in the England High Risk Area at levels 1-3 in the ICC (grey line with circles, bottom left panel) and the number of farms in the England Low Risk Area at levels 1-8 in the ICC (orange line with triangles, bottom right panel). All models include the baseline explanatory variables root farm risk area, herd size, herd type, bTB incident 2010-2014 (yes/no), betweenness, the mean number of source farms in the ingoing contact chain (ICC), purchase of cattle (yes/no), the mean number of cattle purchased annually, and purchase of any animals from the England High Risk Area (yes/no).*

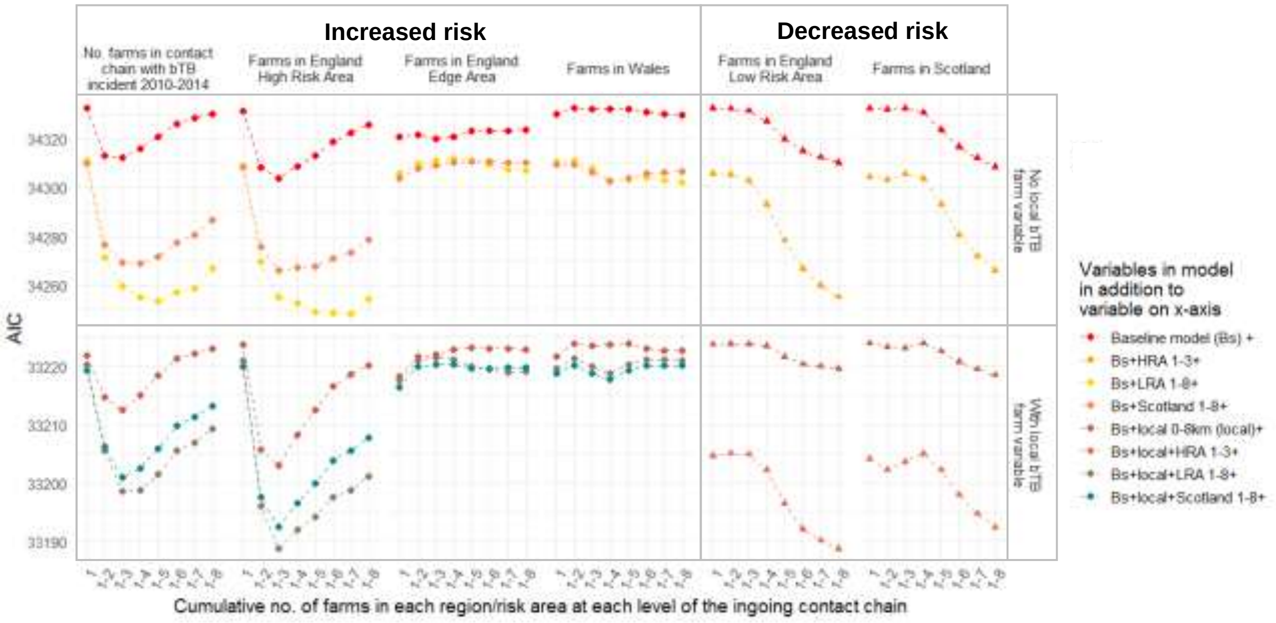


Figure D4. *Performance of final GB and regional multivariable models in predicting bTB incidents on root farms 2015–2016. Plots show model a) sensitivity, b) specificity, c) area under receiver operating characteristic curve, d) positive predictive value, e) negative predictive value, and f) accuracy. Models include the explanatory variables; root herd size, bTB incident 2010–2014, herd type, betweenness, cattle purchased (yes/no), the number of source farms in the ICC, the proportion of farms within 8km of the root farm with bTB incident 2010–2014, cattle purchased from England High Risk Area (yes/no), number of farms in England High Risk Area at levels 1–3 of the contact chain, number of farms in the England Low Risk Area at levels 1–8 of the contact chain, and root farm region (GB model only). We were unable to calculate a positive predictive value for the Scottish model as no bTB incidents were predicted to occur.*

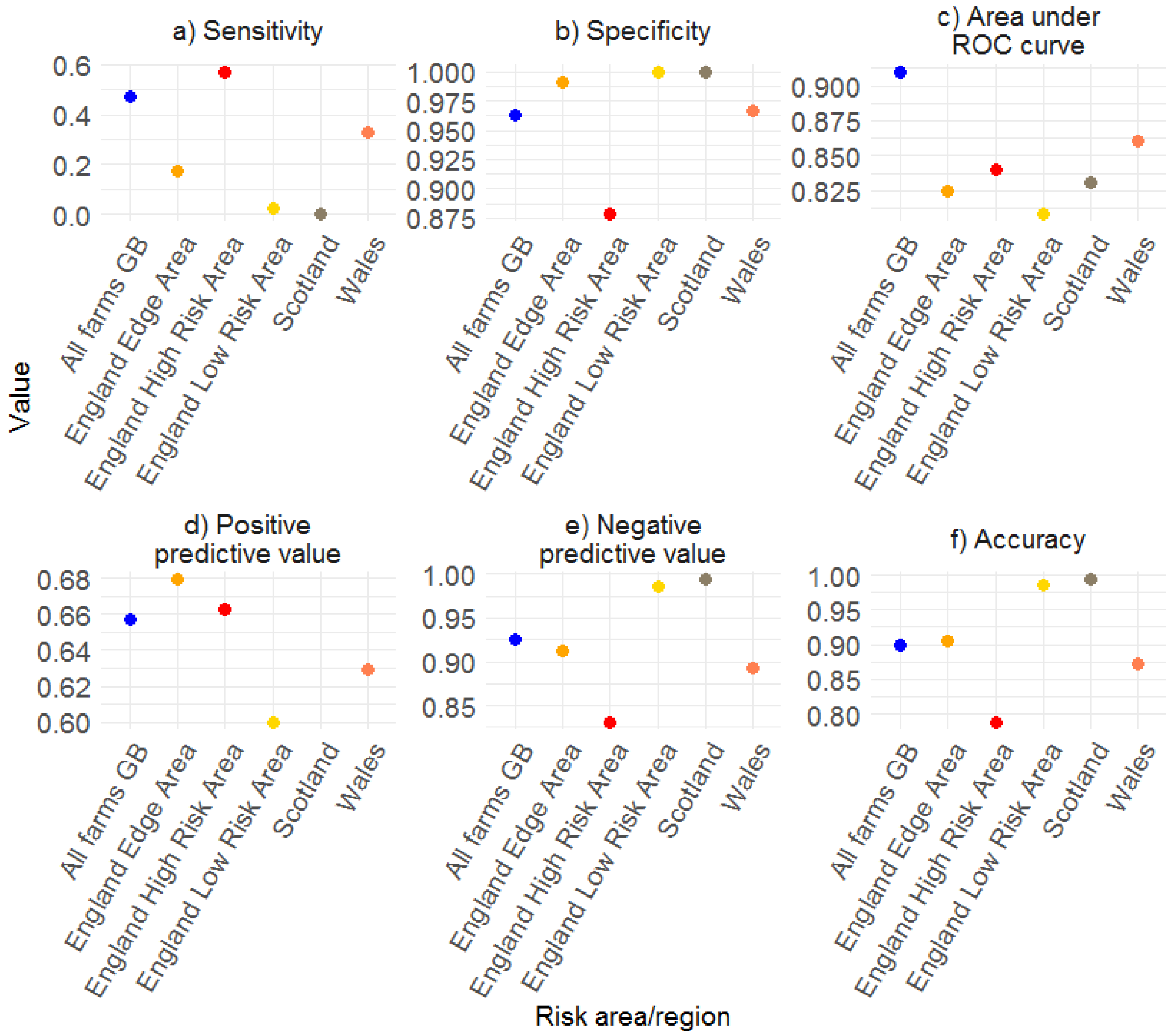


Figure D5. *Mean distance (km) to the root farm from source farms at each level of the ingoing contact chain in a trading network of cattle farms in Great Britain 2012 to 2014. The box plots indicate the median and 25th and 75th percentiles, the upper and lower whiskers extend to the largest or smallest value no further than 1.5 times the interquartile range and data beyond this range are plotted as outlying points.*

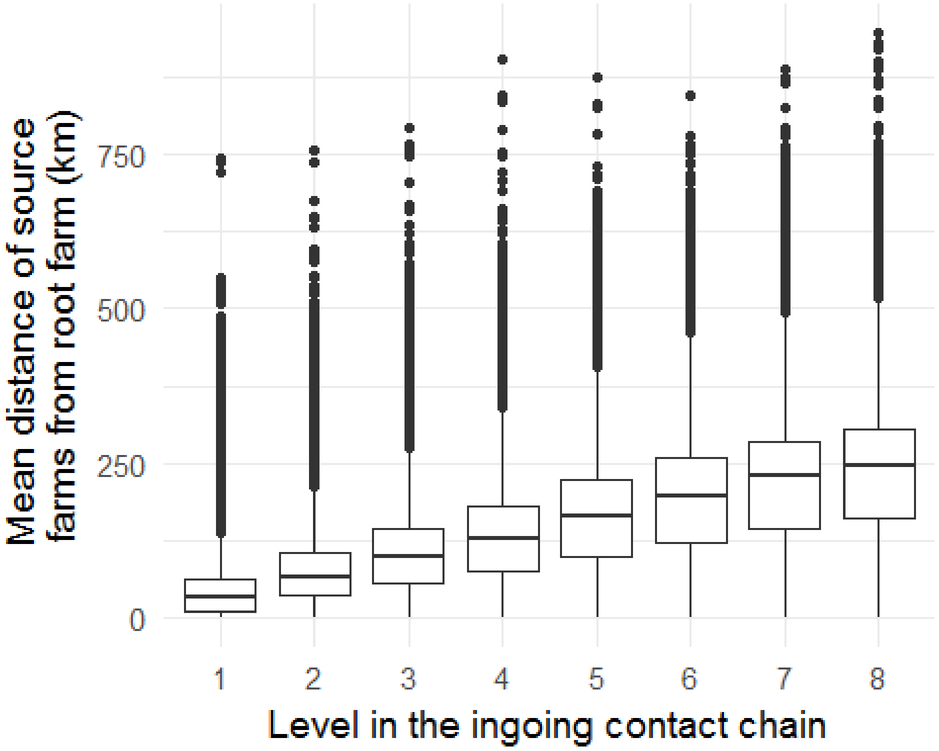


Table D1. *Relative numbers of study farms with bovine tuberculosis (bTB) incidents during 2015–2016. Incidents are classified as Officially Tuberculosis Free suspended (OTF-S) or withdrawn (OTF-W) and numbers of study farms with bTB incidents 2015–2016 that also had a bTB incident 2010–2014 in different regions and Great Britain as a whole are shown. Percentages shown are of those farms that experienced a bTB incident, not all study farms. *Five farms had unclassified bTB incidents, 4 in the England High Risk Area and 1 in the England Edge Area.*

	Region					Great Britain
	England High Risk Area	England Edge Area	England Low Risk Area	Wales	Scotland	
Number of root farms with OTF-S bTB incident 2015–2016	1234 (19.1%)	304 (42.3%)	181 (63.5%)	210 (12.4%)	53 (72.6%)	1982 (21.5%)
Number of root farms with OTF-W bTB incident 2015–2016	5213 (80.8%)	413 (57.5%)	104 (36.5%)	1486 (87.6%)	20 (27.4%)	7236 (78.5%)
Number of root farms with bTB incident (OTF-S or W) during 2015–2016 *and* during 2010–2014	5092 (78.9%)	322 (44.8%)	60 (21.1%)	1229 (72.5%)	21 (28.8%)	6724 (72.9%)
Total number of root farms with bTB incident (OTF-S or W) 2015–2016	6451*	718*	285	1696	73	9223

Table D2. *Effect sizes of explanatory variables on the odds of a bTB incident on the root farm in 2015–2016. Odds ratios (ORs) with 95% confidence intervals are from our multivariable logistic regression analysis using regional data from the England High Risk Area, the England Edge Area, the England Low Risk Area, Wales, and Scotland. ORs of continuous variables are standardised as the odds associated with the difference between the 10th and 90th percentiles of the raw data.*

Region	Parameter		10th percentile (raw data)	90th percentile (raw data)	Odds ratio	2.5% confidence limit	97.5% confidence limit
England High Risk Area (n = 21892)	Root farm herd type	Mixed			Baseline		
		Dairy	-	-	1.34	1.17	1.55
		Fat	-	-	0.89	0.77	1.04
		Suckler	-	-	1.09	0.97	1.24
	Root farm bTB 2010-2014		-	-	2.75	2.54	2.98
	Cattle purchased by root farm		-	-	0.97	0.83	1.13
	Mean number of farms in ICC	1st quartile (0–1)			Baseline		
		2nd quartile (2–421)	-	-	1.06	0.93	1.20
		3rd quartile (422–5601)	-	-	1.15	1.01	1.32
		4th quartile (5602–39676)	-	-	1.25	1.02	1.54
	Cattle purchased direct from England High Risk Area		-	-	1.21	1.00	1.47
	Root farm herd size		4	291	22.34	19.00	26.31
	Mean number of purchased cattle		0	201	1.00	0.97	1.03
	Root farm betweenness		0	277523	0.99	0.99	1.00
	Proportion of farms within 8km with bTB 2010-2014		0.26	0.59	3.03	2.74	3.35
	No. farms in England High Risk Area at levels 1-3		0	1181	1.21	1.11	1.33
	No. farms in England Low Risk Area at levels 1-8		0	8581	0.71	0.58	0.87
England Edge Area (n = 6960)	Root farm herd type	Mixed			Baseline		
		Dairy	-	-	1.46	1.03	2.10
		Fat	-	-	1.00	0.69	1.47
		Suckler	-	-	1.30	0.95	1.81
	Root farm bTB 2010-2014		-	-	2.33	1.91	2.82
	Cattle purchased by root farm		-	-	0.89	0.69	1.16
	Mean number of farms in ICC	1st quartile (0–1)			Baseline		
		2nd quartile (2–344)	-	-	1.20	0.87	1.67
		3rd quartile (345–6204)	-	-	1.35	0.93	1.95
		4th quartile (6205–37934)	-	-	1.19	0.66	2.12
	Cattle purchased direct from England High Risk Area		-	-	1.24	1.00	1.54
	Root farm herd size		3	257	16.23	11.11	23.96
	Mean number of purchased cattle		0	241	1.01	0.96	1.07
	Root farm betweenness		0	349309	1.00	0.99	1.00
	Proportion of farms within 8km with bTB 2010-2014		0.04	0.25	3.16	3	3.81

Region	Parameter		10th percentile (raw data)	90th percentile (raw data)	Odds ratio	2.5% confidence limit	97.5% confidence limit
	No. farms in England High Risk Area at levels 1-3		0	900	1.15	0.99	1.35
	No. farms in England Low Risk Area at levels 1-8		0	10715	0.76	0.43	1.36
England Low Risk Area (n = 18942)	Root farm herd type	Mixed			Baseline		
		Dairy	-	-	0.86	0.54	1.41
		Fat	-	-	0.81	0.49	1.36
		Suckler	-	-	0.80	0.53	1.24
	Root farm bTB 2010-2014		-	-	4.57	3.25	6.35
	Cattle purchased by root farm		-	-	1.38	0.92	2.11
	Mean number of farms in ICC	1st quartile (0–2)			Baseline		
		2nd quartile (3–1522)	-	-	1.03	0.65	1.68
		3rd quartile (1523–8464)	-	-	1.42	0.84	2.44
		4th quartile (8465–39185)	-	-	2.54	1.10	5.79
	Cattle purchased direct from England High Risk Area		-	-	1.48	1.06	2.04
	Root farm herd size		3	255	14.29	8.52	24.38
	Mean number of purchased cattle		0	225	1.01	0.97	1.04
	Root farm betweenness		0	395549	1.00	0.99	1.00
	Proportion of farms within 8km with bTB 2010-2014		0	0.06	1.64	1.34	1.98
	No. farms in England High Risk Area at levels 1-3		0	494	1.18	1.08	1.28
	No. farms in England Low Risk Area at levels 1-8		0	11932	0.21	0.09	0.49
Wales (n = 11525)	Root farm herd type	Mixed			Baseline		
		Dairy	-	-	1.31	0.99	1.75
		Fat	-	-	0.97	0.70	1.35
		Suckler	-	-	1.00	0.77	1.31
	Root farm bTB 2010-2014		-	-	2.60	2.27	2.99
	Cattle purchased by root farm		-	-	1.04	0.86	1.25
	Mean number of farms in ICC	1st quartile (0–1)			Baseline		
		2nd quartile (2–459)	-	-	0.92	0.74	1.15
		3rd quartile (460–4404)	-	-	0.95	0.75	1.20
		4th quartile (4405–33893)	-	-	0.86	0.60	1.24
	Cattle purchased direct from England High Risk Area		-	-	1.15	0.99	1.33
	Root farm herd size		5	219	14.65	11.12	19.39
	Mean number of purchased cattle		0	127	1.02	0.97	1.07
	Root farm betweenness		0	225091	1.00	0.99	1.01
	Proportion of farms within 8km with bTB 2010-2014		0.06	0.54	7.95	6.52	9.73
	No. farms in England High Risk Area at levels 1-3		0	317	1.04	0.95	1.13
	No. farms in England Low Risk Area at levels 1-8		0	7065	0.95	0.68	1.31
Scotland (n = 11777)	Root farm herd type	Mixed			Baseline		
		Dairy	-	-	0.53	0.16	2.13
		Fat	-	-	0.72	0.18	3.09
		Suckler	-	-	0.99	0.39	3.35
	Root farm bTB 2010-2014		-	-	16.26	9.06	28.38

Region	Parameter		10th percentile (raw data)	90th percentile (raw data)	Odds ratio	2.5% confidence limit	97.5% confidence limit
	Cattle purchased by root farm		-	-	1.20	0.46	3.55
	Mean number of farms in ICC	1st quartile (0–1)			Baseline		
		2nd quartile (2–632)	-	-	0.98	0.37	3.10
		3rd quartile (633–6682)	-	-	1.15	0.39	3.90
		4th quartile (6683–36712)	-	-	1.58	0.31	8.33
	Cattle purchased direct from England High Risk Area		-	-	0.46	0.09	1.46
	Root farm herd size		4	364	13.64	4.69	42.68
	Mean number of purchased cattle		0	198	0.98	0.90	1.03
	Root farm betweenness		0	286021	1.00	0.99	1.00
	Proportion of farms within 8km with bTB 2010-2014		0	0.05	1.14	0.80	1.41
	No. farms in England High Risk Area at levels 1-3		0	21	1.04	1.02	1.05
	No. farms in England Low Risk Area at levels 1-8		0	10386	0.69	0.16	2.96

Appendix E (Chapter 5)

Validation of proximity tags

The proximity loggers used in Chapters 5 and 6 of this thesis have been previously validated on humans, and used in studies on humans, dogs, and horses (Cattuto et al., 2010; Milwid et al., 2019b; Wilson-Aggarwal et al., 2019). Their validation for our studies consisted of three approaches; two video analyses with cattle and one static analysis with posts.

The static validation analysis was carried out by colleagues at the University of Exeter (Professor Darren Croft) and analysed by collaborators at the Institute for Scientific Interchange (ISI; Laura Ozella). Multiple tags were fixated to four wooden posts at different heights in a square pattern, posts were sequentially placed at increasing distances away from each other, with and without an obstruction between the tags. This allowed assessment of how tags interacted at different heights and orientations. The relative attenuations between all tags were measured and findings suggested that -70dbm was an appropriate threshold to measure contacts between 1 and 1.5m.

The first video analysis recorded cattle feeding along a trough in a barn on one of our study farms, cattle were parallel to one another and this analysis suggested that in these circumstances, where tags were oriented in the same or opposite directions, interactions may not be recorded if the tags were not fully facing each other, as per the outlined function of the tags in Chapter 5 Methods.

The second video analysis recorded three cows in differing positions in a yard, however unfortunately data were only successfully downloaded for the interactions of two cows. Interactions were recorded and compared to tag data. The attenuation of these recorded contacts was consistent with previous validation of attenuation in human validation studies, suggesting that the attenuation chosen (contacts less than -70dBm) was appropriate to use as a proxy for 1–1.5m contact between cattle. In this study, one contact was recorded when cattle were further away than we would expect a contact to be recorded between tags. Propagation of this signal was most likely due to the environment in which the analysis was carried out, with metal barriers surrounding all sides of the yard (see Photo E1). Although metal is likely to be commonly found inside farm buildings, this contact had an attenuation higher than the threshold and would have been removed in data cleaning, demonstrating that the post-data collection processing was effective in selecting appropriate contacts between cattle.

Photo E1.

Figure E1. *Example of the contacts recorded from two proximity tags (green and purple lines) recording very high numbers of contacts over a short time period compared to other tags in the same deployment (all other coloured lines). All data recorded by proximity tags (including the two shown) that recorded more contacts than 95% of the total contacts recorded by all tags within a 30 minute time frame were removed from the study.*

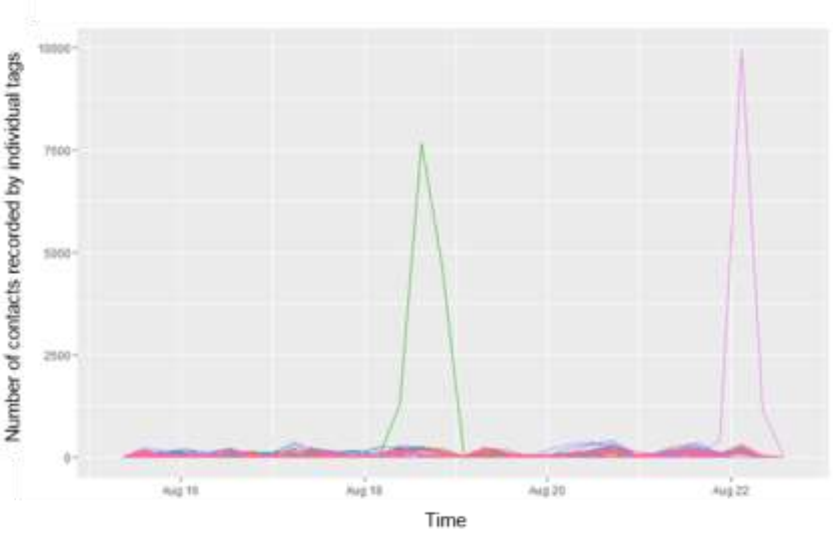


Figure E2. *Schematic showing grouping of contacts to construct Temporo-spatial networks (Temporal Pasture, Temporal Buildings, and Temporal Split networks) by different times of day (day or night) to analyse the repeatability of dyads being in the same community over time (examples from Strip-grazed and Dry groups). Each number represents a separate network formed from the contacts below it (represented by multiple overlapping points), the location is denoted by the y-axis and the colour represents the time of day (yellow = day, blue = night), where day is defined as 07:00–19:00. For example, the yellow '1' in the top left of the plot represents the first Pasture, day network for the Strip-grazed group. We calculated communities in each of these networks, noted for each dyad if they were in the same community in each network and then tested the repeatability of dyads being in the same community across the sequential networks for each network category, (e.g. the pasture day category compared yellow 1, 2, 3, 4, 5, 6, 7 in the Pasture group – circled in yellow on schematic). As the Dry group stayed in one location for the majority of the study, the networks were only divided by time of day, not location.*

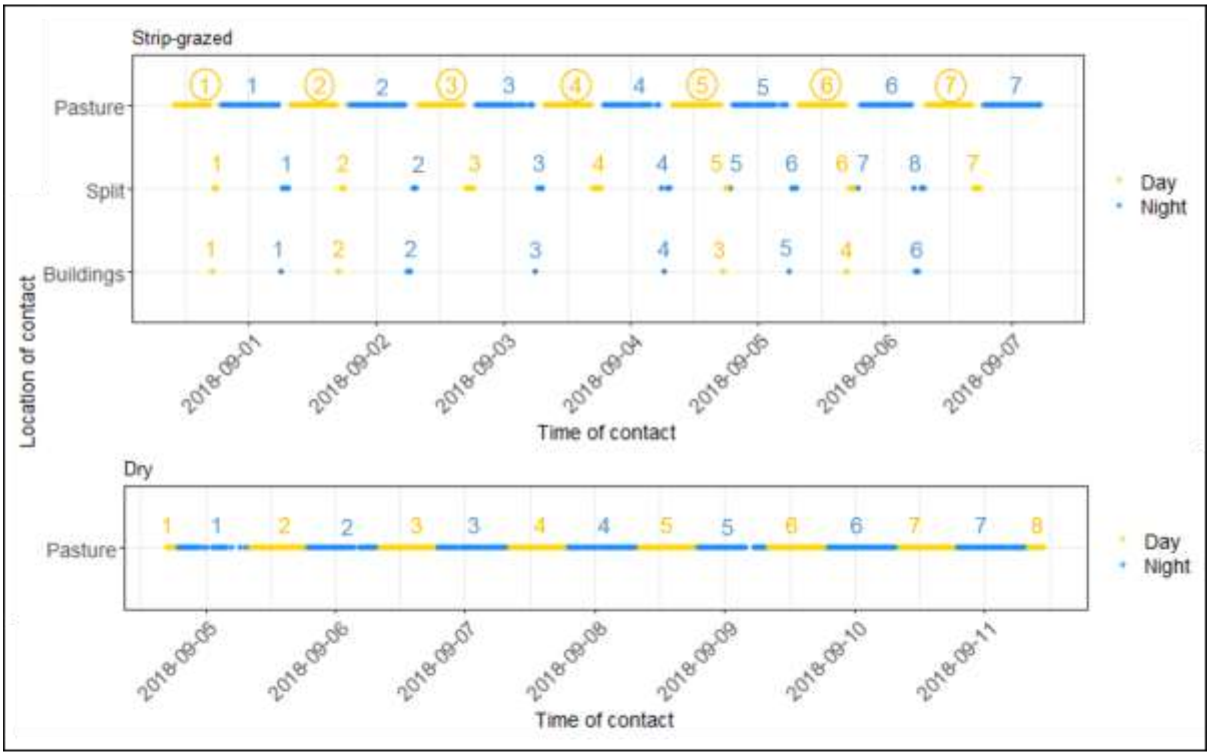


Figure E3: *Variation in the proportion of two-hour periods a dyad spent in contact over each 24-hour period of the study period, compared to the mean proportion of two-hour periods they spent in contact. Each point represents a dyad and is coloured by the P value comparing it to a random distribution, with P < 0.025 as triangles and P > 0.025 as crosses. Dyads more consistent than random exhibit a range of mean contact times.*

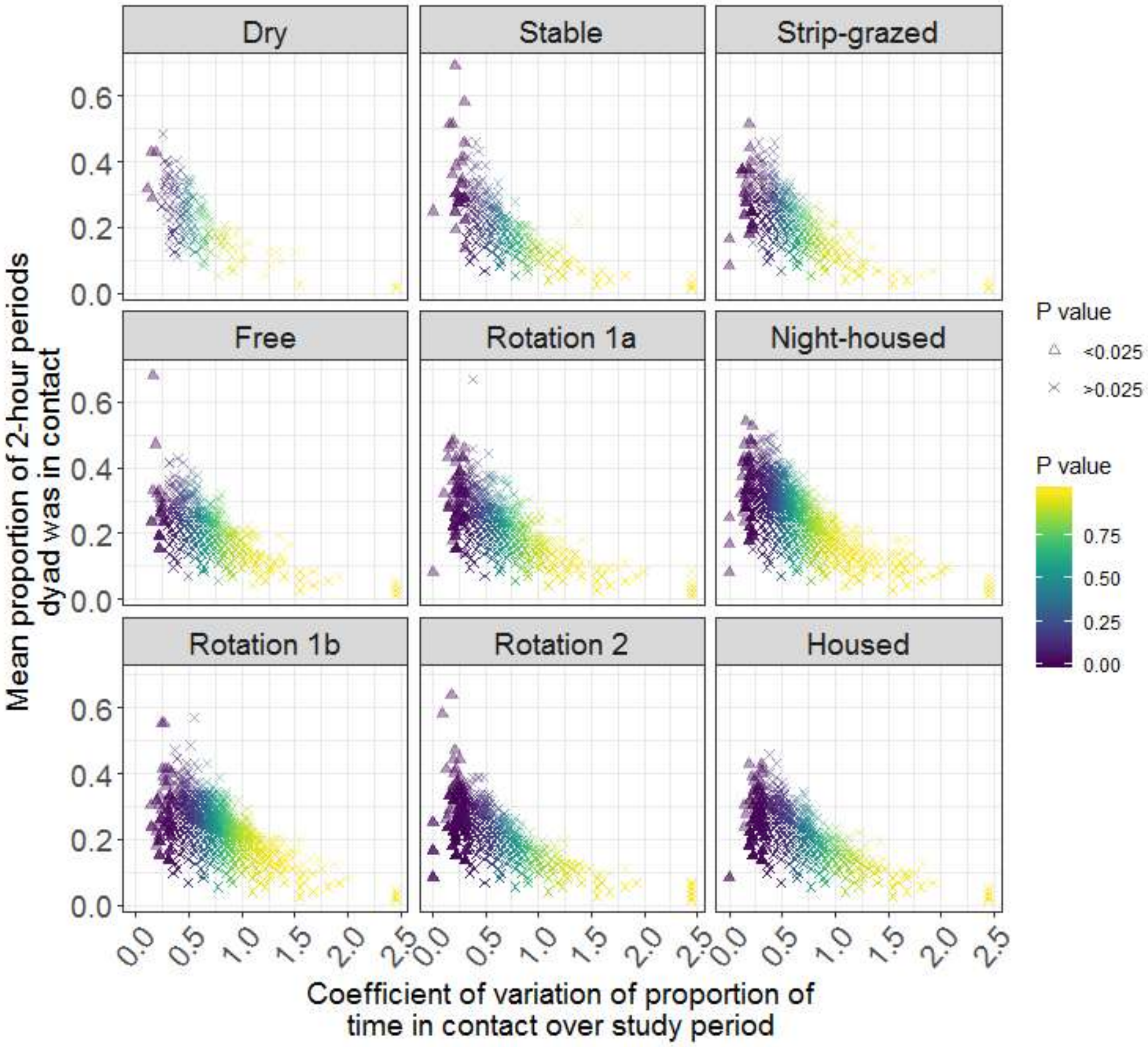


Figure E4. *Community structure for unweighted, unfiltered Full networks and Spatial networks (Buildings, Pasture and Split) for all study groups (NF) and for those networks filtered by removing edge weights below the 50th (F50), 75th, (F75) and 90th percentile (F90) of unfiltered edge weights. a) shows the number of communities detected by the fast-greedy algorithm (calculated in 'igraph') and b) shows the proportion of the group in the largest detected community. Study groups are in order of ascending group size.*

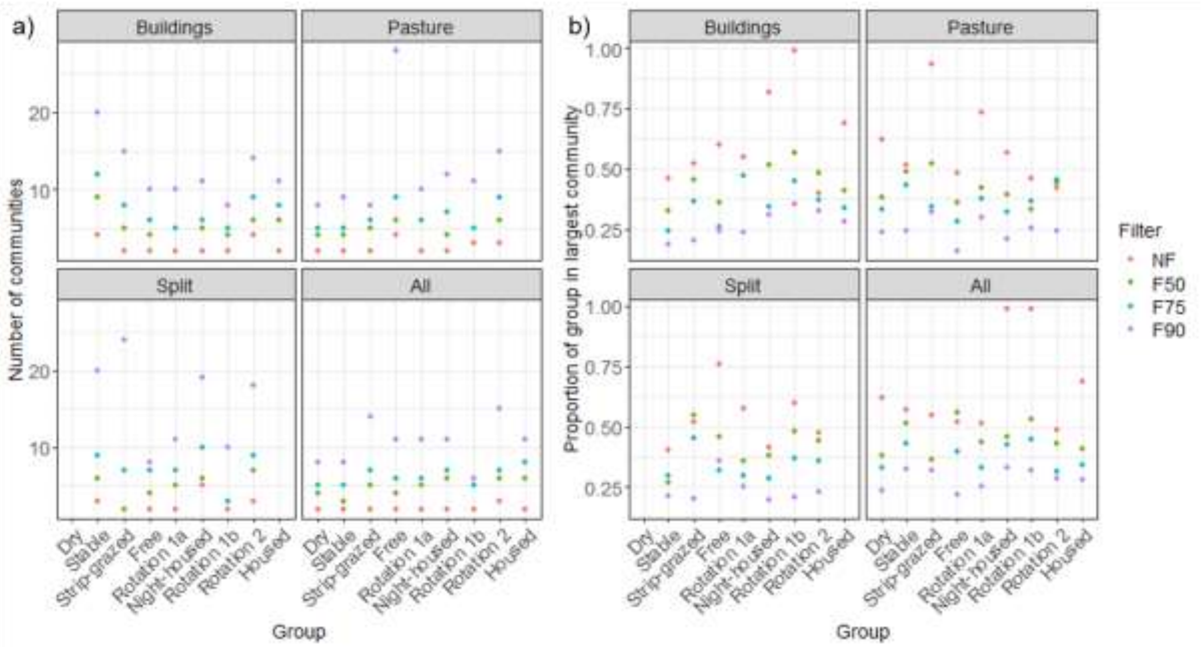


Figure E5. *Relative modularity values for each deployment for Full and Spatial networks (Buildings, Pasture and Split). The modularity was calculated for each of these additionally on three filtered networks removing edge weights less than the 50th (F50), 75th, (F75) and 90th percentile (F90) of unfiltered edge weights. Up and down facing triangles facing triangles respectively represent values higher and lower than 95% random values, crosses represent non-significance.*

Table E1. *Correlation values for comparing Spatial networks from quadratic assignment procedure analysis. Analysis was performed on unfiltered networks and networks filtered by removing edge weights less than the 50th (F50), 75th, (F75) and 90th percentile (F90) of unfiltered edge weights. Grey shading indicates correlation scores significantly different from random (P < 0.05).*

Farm	Metric	Networks compared											
		Pasture and Buildings				Pasture and Split				Buildings and Split			
	Threshold	NF	F50	F75	F90	NF	F50	F75	F90	NF	F50	F75	F90
Night-housed	*P* value	0.808	0.984	0.324	0.154	0.519	0.678	0.561	0.924	0.767	0.826	0.791	0.780
	R squared	0.0000	0.0000	0.0004	0.0007	0.0001	0.0001	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000
Strip-grazed	*P* value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	R squared	0.0972	0.0860	0.0574	0.0475	0.0942	0.0940	0.0649	0.0457	0.3096	0.2961	0.2289	0.1754
Free	*P* value	0.344	0.323	0.377	0.212	0.017	0.006	0.015	0.045	<0.01	<0.01	<0.01	0.005
	R squared	0.0013	0.0013	0.0009	0.0015	0.0092	0.0100	0.0081	0.0044	0.0580	0.0465	0.0268	0.0102
Rotation 1a	*P* value	<0.001	<0.001	0.002	0.041	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	R squared	0.0135	0.0154	0.0069	0.0020	0.0607	0.0549	0.0423	0.0169	0.0395	0.0374	0.0295	0.0355
Rotation 1b	*P* value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	R squared	0.0640	0.0587	0.0397	0.0225	0.0589	0.0540	0.0430	0.0315	0.1012	0.0912	0.0746	0.0510
Rotation 2	*P* value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.019	<0.001	<0.001	<0.001	<0.001
	R squared	0.0113	0.0110	0.0073	0.0028	0.0041	0.0040	0.0029	0.0018	0.0997	0.0935	0.0773	0.0560
Stable	*P* value	0.961	0.944	0.463	0.980	0.585	0.548	0.501	0.604	0.873	0.795	0.869	0.844
	R squared	0.0000	0.0000	0.0009	0.0000	0.0005	0.0006	0.0006	0.0003	0.0000	0.0001	0.0001	0.0001
Grey shading indicates statistical significance of *P* < 0.05													

Table E2. *Relative modularity values for communities detected by the fast-greedy algorithm on binary (unweighted) Full networks and Spatial networks. To analyse stronger ties, we filtered the networks by removing edge weights below the 50th (F50), 75th (F75), and 90th (F90) percentiles. NF indicates the unfiltered network. Dark grey shading of the cell indicates the observed relative modularity was lower than 2.5% lowest values of relative modularity calculated on 4999 randomised networks, light grey shading indicates observed values were higher than 97.5% of values on random networks.*

	Filter	NF	F50	F75	F90	NF	F50	F75	F90	NF	F50	F75	F90	NF	F50	F75	F90
	Location	Pasture				Buildings				Split				All			
Night-housed	LCI	0.0699	0.2487	0.4281	0.6751	0.0096	0.1772	0.3447	0.6128	0.1551	0.3356	0.5567	0.7896	0.0071	0.1767	0.3455	0.6106
	UCI	0.0964	0.3119	0.5156	0.7492	0.0199	0.2307	0.4231	0.6948	0.2063	0.4147	0.6436	0.8455	0.0177	0.2306	0.4236	0.6920
	Qrel	0.0651	0.2505	0.4296	0.7041	0.0103	0.2627	0.3984	0.6134	0.1738	0.3668	0.5821	0.7846	-0.0111	0.2722	0.4183	0.5948
Strip-grazed	LCI	0.0355	0.2517	0.4751	0.7294	0.0539	0.2676	0.4905	0.7331	0.0506	0.2768	0.4784	0.7336	0.0097	0.2379	0.4492	0.7150
	UCI	0.0599	0.3340	0.5782	0.8132	0.0825	0.3475	0.5939	0.8156	0.0786	0.3615	0.5820	0.8165	0.0335	0.3164	0.5542	0.7993
	Qrel	0.0231	0.2718	0.4561	0.7103	0.0501	0.2954	0.4285	0.6601	0.0636	0.3824	0.4978	0.3880	0.0130	0.2359	0.4579	0.6048
Free	LCI	0.2627	0.4816	0.6746	0.8677	0.0187	0.2278	0.4438	0.6988	-0.0203	0.2204	0.4259	0.6934	-0.0205	0.2176	0.4249	0.6898
	UCI	0.3438	0.5801	0.7610	0.9423	0.0379	0.3007	0.5440	0.7816	0.0199	0.2927	0.5256	0.7763	0.0055	0.2888	0.5259	0.7745
	Qrel	0.3155	0.5174	0.6451	0.7515	0.0085	0.2794	0.4320	0.6635	0.0204	0.2985	0.4497	0.6631	0.0056	0.2599	0.4787	0.6705
Rotation 1a	LCI	0.0261	0.2152	0.4039	0.6710	0.0498	0.2339	0.4295	0.6901	0.0961	0.3055	0.4876	0.7511	0.0172	0.2104	0.3929	0.6653
	UCI	0.0404	0.2795	0.4987	0.7511	0.0699	0.2994	0.5249	0.7663	0.1389	0.3846	0.5817	0.8179	0.0316	0.2756	0.4877	0.7466
	Qrel	0.0122	0.2348	0.3705	0.6908	0.0635	0.3502	0.5264	0.7516	0.1055	0.3179	0.5134	0.7537	0.0063	0.2035	0.4115	0.6509
Rotation 1b	LCI	0.0609	0.2330	0.3950	0.6609	-0.0115	0.1721	0.3392	0.6027	0.0363	0.2004	0.3644	0.6319	-0.0116	0.1699	0.3375	0.6051
	UCI	0.0803	0.2921	0.4803	0.7380	0.0024	0.2240	0.4168	0.6870	0.0489	0.2555	0.4462	0.7107	-0.0113	0.2233	0.4158	0.6898
	Qrel	0.0635	0.3050	0.4180	0.6683	-0.0094	0.2089	0.4103	0.6365	0.0297	0.2322	0.3598	0.6569	-0.0102	0.2238	0.3938	0.6021
Rotation 2	LCI	0.0628	0.2168	0.3855	0.6544	0.0401	0.1956	0.3528	0.6197	0.0973	0.2727	0.4339	0.7028	0.0299	0.1824	0.3462	0.6037
	UCI	0.0837	0.2710	0.4701	0.7294	0.0524	0.2477	0.4328	0.7012	0.1360	0.3379	0.5191	0.7729	0.0400	0.2337	0.4236	0.6867
	Qrel	0.0648	0.2630	0.4005	0.6369	0.0250	0.2137	0.3732	0.5727	0.1321	0.2973	0.4373	0.6542	0.0167	0.2096	0.3843	0.5286
Stable	LCI	0.0259	0.2706	0.4829	0.7263	0.3459	0.6126	0.7558	0.8786	0.2982	0.5478	0.7405	0.8771	0.0166	0.2639	0.4765	0.7296

	Filter	NF	F50	F75	F90	NF	F50	F75	F90	NF	F50	F75	F90	NF	F50	F75	F90
	Location	Pasture				Buildings				Split				All			
	UCI	0.0517	0.3597	0.5885	0.8245	0.4477	0.7145	0.8513	1.0000	0.3915	0.6558	0.8347	1.0000	0.0441	0.3508	0.5841	0.8238
	Qrel	0.0213	0.3350	0.5605	0.7702	0.3742	0.5731	0.7007	0.8721	0.4334	0.6128	0.8084	0.8853	0.0223	0.3545	0.5919	0.7837
Dry	LCI	-0.0002	0.2831	0.5121	0.7132	–	–	–	–	–	–	–	–	-0.0023	0.2844	0.5134	0.7159
	UCI	0.0303	0.4056	0.6639	0.9158	–	–	–	–	–	–	–	–	0.0303	0.4060	0.6643	0.9192
	Qrel	0.0119	0.2624	0.4324	0.8054	–	–	–	–	–	–	–	–	0.0119	0.2624	0.4324	0.8054
Housed	LCI	–	–	–	–	0.0314	0.1789	0.3374	0.6004	–	–	–	–	0.0315	0.1794	0.3368	0.6001
	UCI	–	–	–	–	0.0413	0.2280	0.4134	0.6823	–	–	–	–	0.0416	0.2285	0.4121	0.6829
	Qrel	–	–	–	–	0.0120	0.1945	0.3521	0.5407	–	–	–	–	0.0120	0.1945	0.3521	0.5407
LCI = Lower bound of random networks (2.5%) UCI = Upper bound of random networks (97.5%) Qrel = relative modularity														Observed < LCI		Observed > UCI	

Appendix F (Chapter 6)

Figure F1. *Relationships between estimates of milk yields and estimates of somatic cell counts (SCC) taken at different time points; a) shows the mean daily milk yield during the study period in relation to a single milk yield recorded 1 week after the study period (r = 0.90, P < 0.001) in group HF1, showing that single values are highly correlated with mean values for milk yields on this farm. b) and c) show the relationship between milk yields (r = 0.59, P < 0.001) and SCCs (r = 0.67, P < 0.001) respectively, taken 16 days before and 11 days after the study period (34 days apart).*

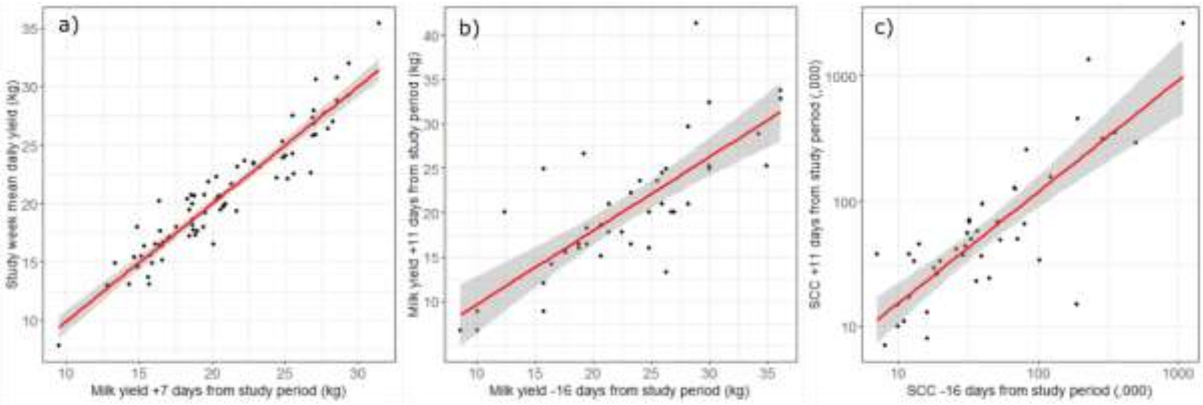


***Figure F2**. Posterior predictive distributions (n = 50) from Model 1, a Bayesian mixed effects model with milk yield of cows as the response variable, showing good model fit.*

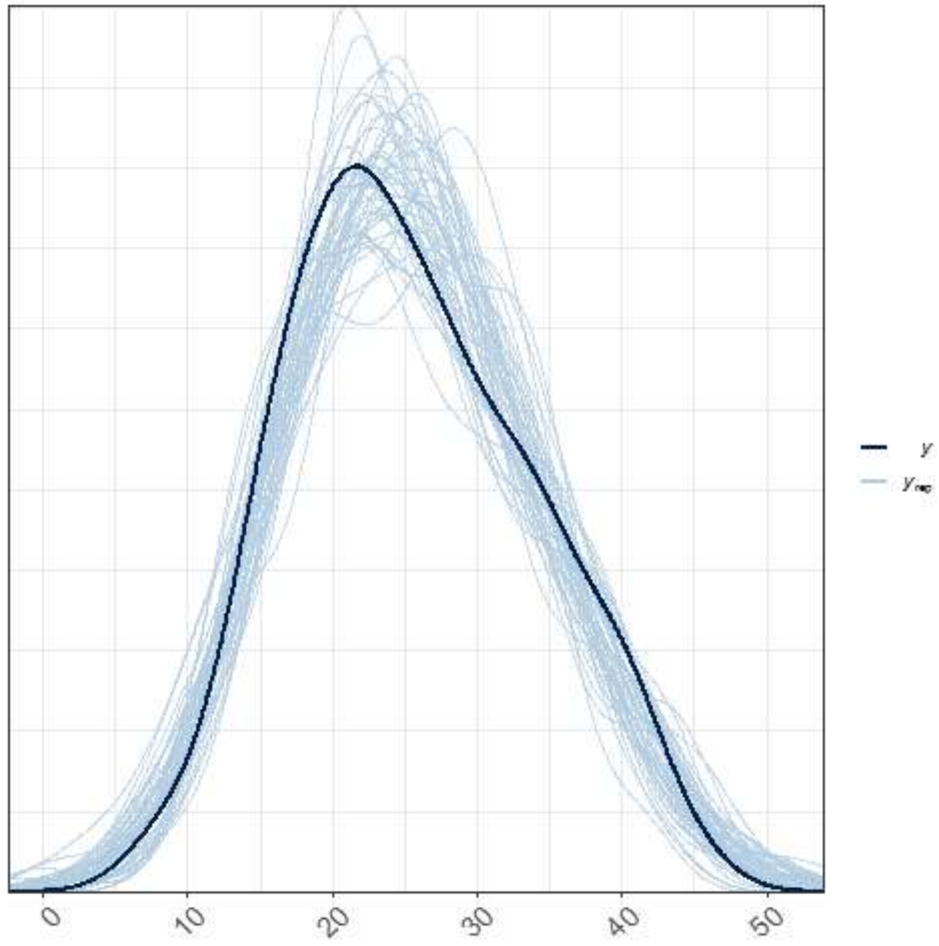


***Figure F3.** Posterior predictive distributions (n = 50) from Model 2, a Bayesian mixed effects model with log somatic cell count of cows as the response variable, showing good model fit.*

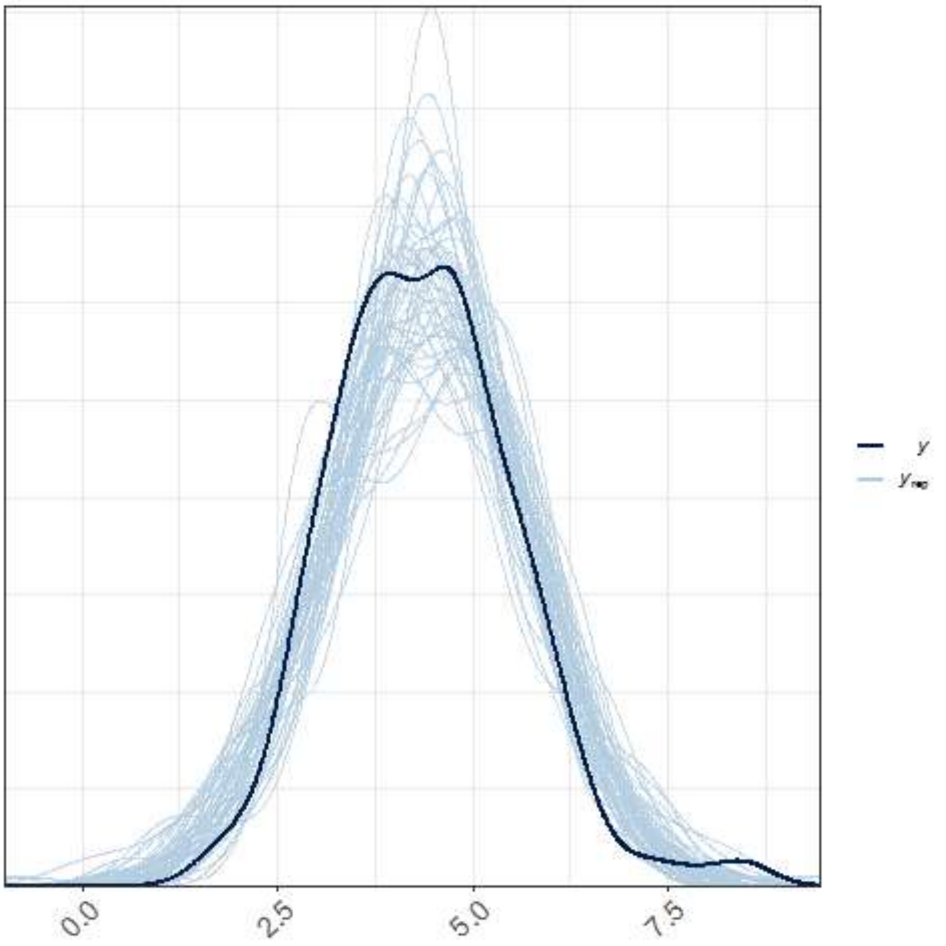


Table F1. *(overleaf). Table of effect sizes and 95% credible intervals from two Bayesian mixed effects models with the response variables of a) milk yield (kg; Model 1) and b) somatic cell count (an indicator of udder health; Model 2) from 197 cattle from three Cornish dairy herds. Fixed effects include mean association strength, mean time spent with 4 closest herd mates, number of 4 closest herd mates in the same lactation, lactation number, stage of lactation, and somatic cell count (only in Model 1) and milk yield (only in Model 2). Farm is included as a random effect in both models. Continuous variables are scaled to represent the effect size associated with a change in the explanatory variable from the 10th to the 90th percentile of the raw data. Effective sample sizes and Gelman-Rubin* $\hat{R}$ *values demonstrate good convergence of chains in both models.*

	Variable		Effect type	Estimate	Lower 95% credible interval	Upper 95% credible interval	10th percentile of raw data	90th percentile of raw data	Effective sample size	$\hat{R}$
Model 1: Response variable = milk yield per cow per day (kg)	Intercept			26.88	17.57	35.96	—	—	2716	1.00
	Mean association strength		Fixed, continuous	-0.36	-3.87	3.16	5.61	21.55	4121	1.00
	Mean top 4 contacts		Fixed, continuous	0.85	-2.92	4.57	27.38	108.22	4265	1.00
	Number of top 4 contacts in same lactation	1	Fixed, categorical	0.71	-1.21	2.63	—	—	4572	1.00
		2		-0.11	-2.41	2.24	—	—	4533	1.00
		3		-1.34	-4.09	1.47	—	—	4376	1.00
		4		-3.65	-9.59	2.20	—	—	4995	1.00
	Number of days in milk	71-170	Fixed, categorical	-2.53	-5.12	0.17	—	—	4002	1.00
		171-270		-5.98	-8.51	-3.37	—	—	3548	1.00
		271+		-13.11	-16.07	-10.22	—	—	3547	1.00
	Lactation number	2	Fixed, categorical	5.42	2.91	7.96	—	—	3750	1.00
		3		3.82	1.26	6.41	—	—	3223	1.00
		4		5.99	3.27	8.73	—	—	3352	1.00
		5+		6.54	4.06	9.03	—	—	3464	1.00
	Somatic cell count		Fixed, continuous	-0.94	-3.10	1.22	19.00	334.39	5909	1.00
	Farm	Ayrshire	Random	-4.99	-13.55	3.64	—	—	2713	1.00
		HF1		-1.00	-9.30	7.57	—	—		
		HF2		6.23	-2.08	14.81	—	—		
Model 2: Response variable = Log somatic cell count (thousand cells per ml)	Intercept			3.58	1.42	5.40	—	—	1544	1.00
	Mean association strength		Fixed, continuous	0.24	-0.41	0.90	5.61	21.55	3720	1.00
	Mean top 4 contacts		Fixed, continuous	0.17	-0.53	0.87	27.38	108.22	3274	1.00
	Number of top 4 contacts in same lactation	1		-0.05	-0.42	0.33	—	—	4254	1.00
		2		-0.11	-0.56	0.34	—	—	3846	1.00
		3		-0.04	-0.58	0.50	—	—	4346	1.00
		4		-0.05	-1.14	1.05	—	—	4479	1.00
	Number of days in milk	71-170	Fixed, categorical	0.54	0.02	1.06	—	—	3776	1.00
		171-270		0.90	0.40	1.41	—	—	2884	1.00
		271+		1.11	0.47	1.77	—	—	2652	1.00
	Lactation number	2	Fixed, categorical	-0.11	-0.62	0.41	—	—	2855	1.00
		3		-0.11	-0.63	0.38	—	—	2977	1.00
		4		0.04	-0.49	0.58	—	—	2827	1.00
		5+		0.51	0.02	0.99	—	—	2694	1.00
	Milk yield per cow per day (kg)		Fixed, continuous	-0.29	-0.88	0.32	15.46	37.38	3577	1.00
	Farm	Ayrshire	Random	-0.06	-1.74	1.96	—	—	5280	1.00
		HF1		0.37	-1.18	2.50	—	—		
		HF2		-0.12	-1.75	1.90	—	—		

References

Aagaard, C., Govaerts, M., Meikle, V., Gutiérrez-Pabello, J.A., McNair, J., Andersen, P., Suárez-Güemes, F., Pollock, J., Espitia, C., Cataldi, A., 2010. Detection of bovine tuberculosis in herds with different disease prevalence and influence of paratuberculosis infection on PPDB and ESAT-6/CFP10 specificity. Preventive Veterinary Medicine 96, 161–169. https://doi.org/10.1016/j.prevetmed.2010.06.007

Abernethy, D.A., Upton, P., Higgins, I.M., McGrath, G., Goodchild, a V, Rolfe, S.J., Broughan, J.M., Downs, S.H., Clifton-Hadley, R., Menzies, F.D., de la Rua-Domenech, R., Blissitt, M.J., Duignan, A., More, S.J., 2013. Bovine tuberculosis trends in the UK and the Republic of Ireland, 1995-2010. Veterinary Record 172, 312. https://doi.org/10.1136/vr.100969

Adkin, A., Brouwer, A., Simons, R.R.L.L., Smith, R.P., Arnold, M.E., Broughan, J., Kosmider, R., Downs, S.H., 2015. Development of risk-based trading farm scoring system to assist with the control of bovine tuberculosis in cattle in England and Wales. Preventive Veterinary Medicine 123, 32–38. https://doi.org/10.1016/j.prevetmed.2015.11.020

Agriculture and Horticutural Development Board, 2017. AHDB eFoodChainMap. https://www.emap.org.uk/HeatMap.aspx. Accessed on 27th September 2019.

AHDB Dairy, 2019a. Average UK milk yield, in: Agriculture and Horticulture Development Board. https://dairy.ahdb.org.uk/resources-library/market-information/farming-data/average-milk-yield/#.XT2Qt_JKiUk Accessed on 28th July 2019.

AHDB Dairy, 2019b. Average Herd Size, in: Agriculture and Horticulture Development Board. https://dairy.ahdb.org.uk/resources-library/market-information/farming-data/average-herd-size/#.XTINSPJKiUm. Accessed on 28th July 2019.

Allen, A.R., Skuce, R.A., Byrne, A.W., 2018. Bovine Tuberculosis in Britain and Ireland – A Perfect Storm? The Confluence of Potential Ecological and Epidemiological Impediments to Controlling a Chronic Infectious Disease. Frontiers in Veterinary Science 5, 109. https://doi.org/10.3389/fvets.2018.00109

Álvarez, J., Bezos, J., de la Cruz, M.L., Casal, C., Romero, B., Domínguez, L., de Juan, L., Pérez, A., 2014. Bovine tuberculosis: Within-herd transmission models to support and direct the decision-making process. Research in Veterinary Science 97, S61–S68. https://doi.org/10.1016/j.rvsc.2014.04.009

Álvarez, J., de Juan, L., Bezos, J., Romero, B., Sáez, J.L., Marqués, S., Domínguez, C., Mínguez, O., Fernández-Mardomingo, B., Mateos, A., Domínguez, L., Aranaz, A., 2009. Effect of paratuberculosis on the diagnosis of bovine tuberculosis in a cattle herd with a mixed infection using interferon-gamma detection assay. Veterinary Microbiology 135, 389–393. https://doi.org/10.1016/j.vetmic.2008.09.060

Álvarez, J., Perez, A.M., Bezos, J., Casal, C., Romero, B., Rodriguez-Campos, S., Saez-Llorente, J.L., Diaz, R., Carpintero, J., de Juan, L., Domínguez, L., 2012. Eradication of bovine tuberculosis at a herd-level in Madrid, Spain: study of within-herd transmission dynamics over a 12 year period. BMC Veterinary Research 8, 100. https://doi.org/10.1186/1746-6148-8-100

Anderson, R., Trewhella, W., 1985. Population dynamics of the badger (*Meles meles*) and the epidemiology of bovine tuberculosis (*Mycobacterium bovis*). Philosophical Transactions of the Royal Society B 310, 327–381.

Anderson, R.M., May, R.M., 1992. Infectious diseases of humans: dynamics and control. Oxford Scientific Press, Oxford, UK.

Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases: Part I. Nature 280, 455–461.

Andersson, M., Schaar, J., Wiktorsson, H., 1984. Effects of drinking water flow rates and social rank on performance and drinking behaviour of tied-up dairy cows. Livestock Production Science 11, 599–610. https://doi.org/10.1016/0301-6226(84)90074-5

Animal and Plant Health Agency, 2014a. Bovine tuberculosis : Infection status in cattle in England 2014 - Annual surveillance report. London, UK.

Animal and Plant Health Agency, 2014b. Approved Dedicated Sale for TB Restricted Cattle. https://www.gov.uk/government/publications/approved-dedicated-sale-for-tb-restricted-cattle Accessed on 2nd February 2018, London, UK.

Animal and Plant Health Agency, 2015a. Veterinary Investigation Diagnosis Analysis report - Yearly trends 2007 to 2014: Cattle. London, UK.

Animal and Plant Health Agency, 2015b. Pre- and post-movement testing of cattle in Great Britain. Report reference: AG-TBYHGB-04. London, UK.

Animal and Plant Health Agency, 2016. Bovine tuberculosis in Great Britain - Surveillance data for 2015 & historical trends. London, UK.

Animal and Plant Health Agency, 2017. Bovine tuberculosis in England in 2016 - Epidemiological analysis of the 2016 data and historical trends. London, UK.

Animal and Plant Health Agency, 2018a. Bovine tuberculosis in England in 2017: Epidemiological analysis of the 2017 data and historical trends. London, UK.

Animal and Plant Health Agency, 2018b. Pre-movement and post-movement testing of cattle in Great Britain. London, UK.

Animal and Plant Health Agency, 2018c. Exceptional private use of non-validated tests for TB on cattle in England, in: APHA Website: Http://Apha.Defra.Gov.Uk. http://apha.defra.gov.uk/vet-gateway/non-valid-tb-testing/ Accessed on 20th September 2019.

Animal and Plant Health Agency, Environmental Research Group Oxford, 2019. ibTB mapping tool for England and Wales. www.ibtb.co.uk.Accessed on 26th September 2019.

Arave, C.W., Albright, J.L., 1976. Social Rank and Physiological Traits of Dairy Cows as Influenced by Changing Group Membership. Journal of Dairy Science 59, 974–981. https://doi.org/10.3168/jds.S0022-0302(76)84306-8

Arave, C.W., Albright, J.L., Sinclair, C.L., 1974. Behavior, Milk Yield, and Leucocytes of Dairy Cows in Reduced Space and Isolation. Journal of Dairy Science 57, 1497–1501. https://doi.org/10.3168/jds.s0022-0302(74)85094-0

Arnott, G., Ferris, C.P., O’Connell, N.E., 2016. Review: welfare of dairy cows in continuously housed and pasture-based production systems. Animal 11, 261–273. https://doi.org/10.1017/S1751731116001336

Arruda, A.G., Friendship, R., Carpenter, J., Hand, K., Poljak, Z., 2016. Network, cluster and risk factor analyses for porcine reproductive and respiratory syndrome using data from swine sites participating in a disease control program. Preventive Veterinary Medicine 128, 41–50. https://doi.org/10.1016/j.prevetmed.2016.03.010

Atkins, P.J., 2000. The pasteurization of England: the science, culture and health implications of food processing, 1900-1950, in: Food, Science, Policy and Regulation in the Twentieth Century. University of Durham, UK., pp. 37–51.

Bansal, S., Grenfell, B.T., Meyers, L.A., 2007. When individual behaviour matters: Homogeneous and network models in epidemiology. Journal of the Royal Society Interface 4, 879–891. https://doi.org/10.1098/rsif.2007.1100

Bansal, S., Read, J., Pourbohloul, B., Meyers, L.A., 2010. The dynamic nature of contact networks in infectious disease epidemiology. Journal of Biological Dynamics 4, 478–489.

https://doi.org/10.1080/17513758.2010.503376

Barlow, N.D., Kean, J.M., Hickling, G., Livingstone, P.G., Robson, A.B., 1997. A simulation model for the spread of bovine tuberculosis within New Zealand cattle herds. Preventive Veterinary Medicine 32, 57–75. https://doi.org/10.1016/S0167-5877(97)00002-0

Barton, E., Donaldson, S., Ross, M., Albright, J., 1973. Social Rank and Social Index as Related to Age, Body Weight and Milk Production in Dairy Cows. Proceedings of the Indiana Academy of Science 83, 473–477.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Beggs, D.S., Jongman, E.C., Hemsworth, P.H., Fisher, A.D., 2018. Short communication: Milking order consistency of dairy cows in large Australian herds. Journal of Dairy Science 101, 603–608. https://doi.org/10.3168/jds.2017-12748

Beilharz, R.G., Butcher, D.F., Freeman, A.E., 1966. Social Dominance and Milk Production in Holsteins. Journal of Dairy Science 49, 887–892. https://doi.org/10.3168/jds.s0022-0302(66)87964-x

Bekara, M.E., Courcoul, A., Bénet, J.J., Durand, B., 2014. Modeling tuberculosis dynamics, detection and control in cattle herds. PLoS ONE 9, e108584. https://doi.org/10.1371/journal.pone.0108584

Bell, D.C., Atkinson, J.S., Carlson, J.W., 1999. Centrality measures for disease transmission networks. Social Networks 21, 1–21. https://doi.org/10.1016/S0378-8733(98)00010-0

Bennett, R.M., Cooke, R.J., 2006. Costs to farmers of a tuberculosis breakdown. Veterinary Record 158, 429–432. https://doi.org/10.1136/vr.158.13.429

Bernoulli, D., Blower, S., 2004. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it (review of 1760 analysis). Reviews in Medical Virology 14, 275–288. https://doi.org/10.1002/rmv.443

Bessell, P.R., Orton, R., White, P.C.L., Hutchings, M.R., Kao, R.R., 2012. Risk factors for bovine Tuberculosis at the national level in Great Britain. BMC Veterinary Research 8, 51. https://doi.org/10.1186/1746-6148-8-51

Bøe, K.E., Færevik, G., 2003. Grouping and social preferences in calves, heifers and cows. Applied Animal Behaviour Science 80, 175–190. https://doi.org/10.1016/S0168-1591(02)00217-4

Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., Aubert, A., 2007. Assessment of positive emotions in animals to improve their welfare. Physiology and Behavior 92, 375–397. https://doi.org/10.1016/j.physbeh.2007.02.003

Bouissou, M., Boissy, A., Le Neindre, P., Veissier, I., 2001. The social behaviour of cattle, in: Social Behaviour in Farm Animals. CABI, UK, pp. 113–145.

Boyland, N.K., Mlynski, D.T., James, R., Brent, L.J.N., Croft, D.P., 2016. The social network structure of a dynamic group of dairy cows: from individual to group level patterns. Applied Animal Behaviour Science 174, 1–10. https://doi.org/10.1016/j.applanim.2015.11.016

Bradbury, N. V., Probert, W.J.M., Shea, K., Runge, M.C., Fonnesbeck, C.J., Keeling, M.J., Ferrari, M.J., Tildesley, M.J., 2017. Quantifying the Value of Perfect Information in Emergency Vaccination Campaigns. PLoS Computational Biology 13, e1005318. https://doi.org/10.1371/journal.pcbi.1005318

Brakel, W.J., Leis, R.A., 1976. Impact of Social Disorganization on Behavior, Milk Yield, and Body Weight of Dairy Cows. Journal of Dairy Science 59, 716–721. https://doi.org/10.3168/jds.s0022-0302(76)84263-4

Brennan, M.L., Kemp, R., Christley, R.M., 2008. Direct and indirect contacts between cattle farms in north-west England. Preventive Veterinary Medicine 84, 242–260. https://doi.org/10.1016/j.prevetmed.2007.12.009

Brooks Pollock, E., de Jong, M.C.M., Keeling, M.J., Klinkenberg, D., Wood, J.L.N., 2015. Eight challenges in modelling infectious livestock diseases. Epidemics 10, 1–5. https://doi.org/10.1016/j.epidem.2014.08.005

Brooks Pollock, E., Keeling, M.J., 2009. Herd size and bovine tuberculosis persistence in cattle farms in Great Britain. Preventive Veterinary Medicine 92, 360–365. https://doi.org/10.1016/j.prevetmed.2009.08.022

Brooks Pollock, E., Roberts, G.O., Keeling, M.J., 2014. A dynamic model of bovine tuberculosis spread and control in Great Britain. Nature 511, 228–231. https://doi.org/10.1038/nature13529

Brooks Pollock, E., Wood, J.L.N., 2015. Eliminating bovine tuberculosis in cattle and badgers: insight from a dynamic model. Proceedings of the Royal Society B 282, 20150374. https://doi.org/10.1098/rspb.2015.0374

Broughan, J.M., Downs, S.H., Crawshaw, T.R., Upton, P.A., Brewer, J., Clifton-Hadley, R.S., 2013. *Mycobacterium bovis* infections in domesticated non-bovine mammalian species. Part 1: Review of epidemiology and laboratory submissions in Great Britain 2004-2010. Veterinary Journal 198, 339–345. https://doi.org/10.1016/j.tvjl.2013.09.006

Broughan, J.M., Judge, J., Ely, E., Delahay, R.J., Wilson, G., Clifton-Hadley, R.S., Goodchild, A. V., Bishop, H., Parry, J.E., Downs, S.H., 2016a. Review article a review of risk factors for bovine tuberculosis infection in cattle in the UK and Ireland. Epidemiology and Infection 144, 2899–2926. https://doi.org/10.1017/S095026881600131X

Broughan, J.M., Maye, D., Carmody, P., Brunton, L.A., Ashton, A., Wint, W., Alexander, N., Naylor, R., Ward, K., Goodchild, A. V., Hinchliffe, S., Eglin, R.D., Upton, P., Nicholson, R., Enticott, G., 2016b. Farm characteristics and farmer perceptions associated with bovine tuberculosis incidents in areas of emerging endemic spread. Preventive Veterinary Medicine 129, 88–98. https://doi.org/10.1016/j.prevetmed.2016.05.007

Brouwer, H., Bartels, C.J.M., Stegeman, A., van Schaik, G., 2012. No long-term influence of movement restriction regulations on the contact-structure between and within cattle holding types in the Netherlands. BMC Veterinary Research 8, 188. https://doi.org/10.1186/1746-6148-8-188

Brunton, L.A., Donnelly, C.A., O’Connor, H., Prosser, A., Ashfield, S., Ashton, A., Upton, P., Mitchell, A., Goodchild, A. V., Parry, J.E., Downs, S.H., 2017. Assessing the effects of the first 2 years of industry-led badger culling in England on the incidence of bovine tuberculosis in cattle in 2013–2015. Ecology and Evolution 7, 7213–7230. https://doi.org/10.1002/ece3.3254

Brunton, L.A., Nicholson, R., Ashton, A., Alexander, N., Wint, W., Enticott, G., Ward, K., Broughan, J.M., Goodchild, A.V., 2015. A novel approach to mapping and calculating the rate of spread of endemic bovine tuberculosis in England and Wales. Spatial and Spatio-temporal Epidemiology 13, 41–50. https://doi.org/10.1016/j.sste.2015.04.002

Buddle, B.M., Vordermeier, H.M., Chambers, M.A., de Klerk-Lorist, L.M., 2018. Efficacy and safety of BCG vaccine for control of tuberculosis in domestic livestock and wildlife. Frontiers in Veterinary Science 5, 259. https://doi.org/10.3389/fvets.2018.00259

Bürkner, P.-C., 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software 80. https://doi.org/10.18637/jss.v080.i01

Burnham, K.P., Anderson, D.R., 2003. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York. https://doi.org/10.1007/b97636

Büttner, K., Krieter, J., Traulsen, A., Traulsen, I., 2013. Efficient Interruption of

Infection Chains by Targeted Removal of Central Holdings in an Animal Trade Network. PLoS ONE 8, e74292. https://doi.org/10.1371/journal.pone.0074292

Büttner, K., Salau, J., Krieter, J., 2016a. Quality assessment of static aggregation compared to the temporal approach based on a pig trade network in Northern Germany. Preventive Veterinary Medicine 129, 1–8. https://doi.org/10.1016/j.prevetmed.2016.05.005

Büttner, K., Salau, J., Krieter, J., 2016b. Adaption of the temporal correlation coefficient calculation for temporal networks (applied to a real-world pig trade network). SpringerPlus 5, 165. https://doi.org/10.1186/s40064-016-1811-7

Butts, C.T., 2016. sna: Tools for Social Network Analysis. R package version 2.4.

Byrne, A.W., Graham, J., Brown, C., Donaghy, A., Guelbenzu-Gonzalo, M., McNair, J., Skuce, R.A., Allen, A., McDowell, S.W., 2018. Modelling the variation in skin-test tuberculin reactions, post-mortem lesion counts and case pathology in tuberculosis-exposed cattle: Effects of animal characteristics, histories and co-infection. Transboundary and Emerging Diseases 1–15. https://doi.org/10.1111/tbed.12814

Byrne, A.W., Guelbenzu-Gonzalo, M., Strain, S.A.J., McBride, S., Graham, J., Lahuerta-Marin, A., Harwood, R., Graham, D.A., McDowell, S., 2017. Assessment of concurrent infection with bovine viral diarrhoea virus (BVDV) and *Mycobacterium bovis* : A herd-level risk factor analysis from Northern Ireland. Preventive Veterinary Medicine 141, 38–47. https://doi.org/10.1016/j.prevetmed.2017.04.007

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P., Riddell, A., 2017. Stan: A probabilistic programming language. Journal of Statistical Software 76, 1–32. https://doi.org/10.18637/jss.v076.i01

Carrique-Mas, J.J., Medley, G.F., Green, L.E., 2008. Risks for bovine tuberculosis in British cattle farms restocked after the foot and mouth disease epidemic of 2001. Preventive Veterinary Medicine 84, 85–93. https://doi.org/10.1016/j.prevetmed.2007.11.001

Carter, S.P., Delahay, R.J., Smith, G.C., Macdonald, D.W., Riordan, P., Etherington, T.R., Pimley, E.R., Walker, N.J., Cheeseman, C.L., 2007. Culling-induced social perturbation in Eurasian badgers *Meles meles* and the management of TB in cattle: An analysis of a critical problem in applied ecology. Proceedings of the Royal Society B 274, 2769–2777. https://doi.org/10.1098/rspb.2007.0998

Cattuto, C., van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.F., Vespignani, A., 2010. Dynamics of person-to-person interactions from distributed RFID sensor networks. PLoS ONE 5, e11596. https://doi.org/10.1371/journal.pone.0011596

Charleston, B., Hope, J., Carr, B., Howard, C., 2001. Masking of two in vitro immunological assays for *Mycobacterium bovis* (BCG) in calves acutely affected with non-cytopathic bovine viral diarrhoea virus. Veterinary Record 149, 481–484.

Charlton, G.L., Rutter, S.M., 2017. The behaviour of housed dairy cattle with and without pasture access: A review. Applied Animal Behaviour Science 192, 2–9. https://doi.org/10.1016/j.applanim.2017.05.015

Charlton, G.L., Rutter, S.M., East, M., Sinclair, L.A., 2011. Effects of providing total mixed rations indoors and on pasture on the behavior of lactating dairy cattle and their preference to be indoors or on pasture. Journal of Dairy Science 94, 3875–3884. https://doi.org/10.3168/jds.2011-4172

Chase-Topping, M., Gally, D., Low, C., Matthews, L., Woolhouse, M.E.J., 2008. Super-shedding and the link between human infection and livestock carriage of *Escherichia coli* O157. Nature Reviews. Microbiology 6, 904–12. https://doi.org/10.1038/nrmicro2029

Chase-Topping, M., McKendrick, I.J., Pearce, M.C., MacDonald, P., Matthews, L., Halliday, J., Allison, L., Fenlon, D., Low, J.C., Gunn, G., Woolhouse, M.E.J., 2007. Risk factors for the presence of high-level shedders of *Escherichia coli* O157 on Scottish farms. Journal of Clinical Microbiology Microbiology 45, 1594–603. https://doi.org/10.1128/JCM.01690-06

Chen, S., Ilany, A., White, B.J., Sanderson, M.W., Lanzas, C., 2015. Spatial-temporal dynamics of high-resolution animal networks: What can we learn from domestic animals? PLoS ONE 10, e0129253. https://doi.org/10.1371/journal.pone.0129253

Chen, S., White, B.J., Sanderson, M.W., Amrine, D.E., Ilany, A., Lanzas, C., 2014. Highly dynamic animal contact network and implications on disease transmission. Scientific Reports 4, 04472. https://doi.org/10.1038/srep04472

Christensen, J., Gardner, I.A., 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. Preventive Veterinary Medicine 45, 83–106. https://doi.org/10.1016/S0167-5877(00)00118-5

Christley, R., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R., Turner, 2005a. Infection in Social Networks: Using Network Analysis to Identify High-Risk Individuals. American Journal of Epidemiology 162, 1024–1031. https://doi.org/10.1093/aje/kwi308

Christley, R., Robinson, S.E., Lysons, R., French, N.P., 2005b. Network Analysis of Cattle Movements in Great Britain. Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine 2005, 234–244.

Claridge, J., Diggle, P., McCann, C.M., Mulcahy, G., Flynn, R., McNair, J., Strain, S., Welsh, M., Baylis, M., Williams, D.J.L., 2012. *Fasciola hepatica* is associated with the failure to detect bovine tuberculosis in dairy cattle. Nature communications 3, 853. https://doi.org/10.1038/ncomms1840

Clark, B., Stewart, G.B., Panzone, L.A., Kyriazakis, I., Frewer, L.J., 2016. A Systematic Review of Public Attitudes, Perceptions and Behaviours Towards Production Diseases Associated with Farm Animal Welfare. Journal of Agricultural and Environmental Ethics 29, 455–478. https://doi.org/10.1007/s10806-016-9615-x

Clark, P.W., Ricketts, R.E., Krause, G.F., 1977. Effect on Milk Yield of Moving Cows from Group to Group. Journal of Dairy Science 60, 769–772. https://doi.org/10.3168/jds.S0022-0302(77)83933-7

Clauset, A., Newman, M.E.J., Moore, C., 2004. Finding community structure in very large networks. Physical Review 70, 066111. https://doi.org/10.1103/PhysRevE.70.066111

Clauset, A., Rohilla Shalizi, C., J Newman, M.E., 2009. Power-Law Distributions in Empirical Data. SIAM Review 51, 661–703. https://doi.org/10.1214/13-AOAS710

Cleaveland, S., Laurenson, M.K., Taylor, L.H., 2001. Diseases of humans and their domestic mammals: Pathogen characteristics, host range and the risk of emergence. Philosophical Transactions of the Royal Society B 356, 991–999. https://doi.org/10.1098/rstb.2001.0889

Clegg, T.A., Good, M., Duignan, A., Doyle, R., Blake, M., More, S.J., 2011. Shorter-term risk of *Mycobacterium bovis* in Irish cattle following an inconclusive diagnosis to the single intradermal comparative tuberculin test. Preventive veterinary medicine 100, 147–54. https://doi.org/10.1016/j.prevetmed.2011.02.015

Coad, M., Clifford, D., Rhodes, S.G., Hewinson, R.G., Vordermeier, H.M., Whelan, A.O., 2010. Repeat tuberculin skin testing leads to desensitisation in naturally infected tuberculous cattle which is associated with elevated interleukin-10 and decreased interleukin-1 beta responses. Veterinary Research 41, 14. https://doi.org/10.1051/vetres/2009062

Collis, K.A., 1976. An investigation of factors related to the dominance order of a herd of dairy cows of similar age and breed. Applied Animal Ethology 2, 167–173.

Collis, K.A., Kay, S.J., Grant, A.J., Quick, A.J., 1979. The effect on social organization and milk production of minor group alterations in dairy cattle. Applied Animal Ethology 5, 103–111. https://doi.org/10.1016/0304-3762(79)90082-8

Conlan, A., McKinley, T., Karolemeas, K., Brooks Pollock, E., Goodchild, A., Mitchell, A., Birch, C., Clifton-Hadley, R., Wood, J., 2012. Estimating the Hidden Burden of Bovine Tuberculosis in Great Britain. PLoS Computational Biology 8, e1002730. https://doi.org/10.1371/journal.pcbi.1002730

Conlan, A., Vordermeier, M., de Jong, M.C., Wood, J.L., 2018. The intractable challenge of evaluating cattle vaccination as a control for bovine Tuberculosis. eLife 7, e27694. https://doi.org/10.7554/eLife.27694

Conlan, A.J., Wood, J.L.N., 2016. Testing the dairy difference. The Veterinary Journal 217, 134–135. https://doi.org/10.1016/j.tvjl.2016.07.017

Conlan, A.J.K., Brooks Pollock, E., McKinley, T., Mitchell, A.P., Jones, G.J., Vordermeier, M., Wood, J.L.N., 2015. Potential benefits of cattle vaccination as a supplementary control for bovine tuberculosis. PLoS Computational Biology 11, e1004038. https://doi.org/10.1371/journal.pcbi.1004038

Corner, L.A.L., Pfeiffer, D.U., Morris, R.S., 2003. Social-network analysis of *Mycobacterium bovis* transmission among captive brushtail possums (Trichosurus vulpecula). Preventive Veterinary Medicine 59, 147–167. https://doi.org/10.1016/S0167-5877(03)00075-8

Courcoul, A., Ezanno, P., 2010. Modelling the spread of Bovine Viral Diarrhoea Virus (BVDV) in a managed metapopulation of cattle herds. Veterinary Microbiology 142, 119–128. https://doi.org/10.1016/j.vetmic.2009.09.052

Cox, D.R., Donnelly, C.A., Bourne, F.J., Gettinby, G., McInerney, J.P., Morrison, W.I., Woodroffe, R., 2005. Simple model for tuberculosis in cattle and badgers. Proceedings of the National Academy of Sciences of the United States of America 102, 17588–17593. https://doi.org/10.1073/pnas.0509003102

Craft, M.E., 2015. Infectious disease transmission and contact networks in wildlife and livestock. Philosophical Transactions of the Royal Society B 370, 20140107. https://doi.org/10.1098/rstb.2014.0107

Croft, D.P., Madden, J.R., Franks, D.W., James, R., 2011. Hypothesis testing in animal social networks. Trends in Ecology and Evolution 26, 502–507. https://doi.org/10.1016/j.tree.2011.05.012

Cross, P.C., Lloyd-Smith, J.O., Johnson, P.L.F., Getz, W.M., 2005. Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. Ecology Letters 8, 587–595. https://doi.org/10.1111/j.1461-0248.2005.00760.x

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal, Complex Systems, Version 1.2.2.

Dawkins, M.S., 2017. Animal welfare and efficient farming: Is conflict inevitable? Animal Production Science 57, 201–208. https://doi.org/10.1071/AN15383

Dawson, D.E., Farthing, T.S., Sanderson, M.W., Lanzas, C., 2019. Transmission on empirical dynamic contact networks is influenced by data processing decisions. Epidemics 26, 32–42. https://doi.org/10.1016/j.epidem.2018.08.003

de Freslon, I., Martínez-López, B., Belkhiria, J., Strappini, A., Monti, G., 2019. Use of social network analysis to improve the understanding of social behaviour in dairy cattle and its impact on disease transmission. Applied Animal Behaviour Science 213, 47–54. https://doi.org/10.1016/j.applanim.2019.01.006

de la Rua-Domenech, R., Goodchild, A.T., Vordermeier, H.M., Hewinson, R.G., Christiansen, K.H., Clifton-Hadley, R.S., 2006. Ante mortem diagnosis of tuberculosis in cattle: A review of the tuberculin tests, γ-interferon assay and other ancillary diagnostic techniques. Research in Veterinary Science 81, 190–210. https://doi.org/10.1016/j.rvsc.2005.11.005

Department for Environment Food and Rural Affairs, 2018. Quarterly publication of National Statistics on the incidence and prevalence of tuberculosis (TB) in Cattle in Great Britain – to end December 2017. London, UK.

Department for Environment Food and Rural Affairs, 2017a. Summary of badger control monitoring during 2017. London, UK.

Department for Environment Food and Rural Affairs, 2017b. Reducing the risk posed by inconclusive skin test reactors. http://www.tbhub.co.uk/tb-policy/england/reducing-the-risk-posed-by-inconclusive-skin-test-reactors/ Accessed on 15th June 2018.

Department for Environment Food and Rural Affairs, 2014. The Strategy for Achieving “Officially Bovine Tuberculosis- Free” Status for England. Report reference: PB14088. London, UK.

Department for Environment Food and Rural Affairs, 2013. Bovine Tuberculosis Evidence Plan. Report reference: PB13909. London, UK.

Dickson, D.P., Barr, G.R., Johnson, L.P., Wieckert, D.A., 1969. Social

Dominance and Temperament of Holstein Cows. Journal of Dairy Science 53, 904–907. https://doi.org/10.3168/jds.s0022-0302(70)86316-0

Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J., 1990. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology 28, 365–382. https://doi.org/10.1007/BF00178324

Dimitrov, I., Djorbineva, M., Sotirov, L., Tanchev, S., 2005. Influence of fearfulness on lysozyme and complement concentrations in dairy sheep. Revue de Medecine Veterinaire 156, 445–448.

Dohoo, I.R., Leslie, K.E., 1991. Evaluation of changes in somatic cell counts as indicators of new intramammary infections. Preventive Veterinary Medicine 10, 225–237. https://doi.org/10.1016/0167-5877(91)90006-N

Dommergues, L., Rautureau, S., Petit, E., Dufour, B., 2012. Network of Contacts between Cattle Herds in a French Area Affected by Bovine Tuberculosis in 2010. Transboundary and Emerging Diseases 59, 292–302. https://doi.org/10.1111/j.1865-1682.2011.01269.x

Donnelly, C., Nouvellet, P., 2013. The contribution of badgers to confirmed tuberculosis in cattle in high-incidence areas of England. PLOS Current Outbreaks. https://doi.org/10.1371/currents.outbreaks.097a904d3f3619db2fe78d24bc776098.Authors

Donnelly, C., Woodroffe, R., Cox, D.R., Bourne, J., Gettinby, G., Le Fevre, A.M., McInerney, J.P., Morrison, W.I., 2003. Impact of localized badger culling on tuberculosis incidence in British cattle. Nature 426, 834–837. https://doi.org/10.1038/nature02192

Downs, S.H.H., Broughan, J.M.M., Goodchild, A.V. V, Upton, P.A.A., Durr, P.A.A., 2016. Responses to diagnostic tests for bovine tuberculosis in dairy and non-dairy cattle naturally exposed to *Mycobacterium bovis* in Great Britain. The Veterinary Journal 216, 8–17. https://doi.org/10.1016/j.tvjl.2016.06.010

Drewe, J.A., 2010. Who infects whom? Social networks and tuberculosis transmission in wild meerkats. Proceedings of the Royal Society B 277, 633–642. https://doi.org/10.1098/rspb.2009.1775

Drosten, C., Lau, A.C.-W., Preiser, W., So, L.K.-Y., Yam, L.Y.-C., 2003. Chapter 4: Epidemiology, in: SARS Reference - 10/2003. www.sarsreference.com Accessed on 19th June 2018.

Dubé, C., Ribble, C., Kelton, D., McNab, B., 2011. Estimating potential epidemic size following introduction of a long-incubation disease in scale-free connected networks of milking-cow movements in Ontario, Canada. Preventive Veterinary Medicine 99, 102–111. https://doi.org/10.1016/j.prevetmed.2011.01.013

Dubé, C., Ribble, C., Kelton, D., McNab, B., 2009. A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development. Transboundary and Emerging Diseases 56, 73–85. https://doi.org/10.1111/j.1865-1682.2008.01064.x

Dubé, C., Ribble, C., Kelton, D., McNab, B., 2008. Comparing network analysis measures to determine potential epidemic size of highly contagious exotic diseases in fragmented monthly networks of dairy cattle movements in Ontario, Canada. Transboundary and Emerging Diseases 55, 382–392. https://doi.org/10.1111/j.1865-1682.2008.01053.x

Duncan, A.J., Gunn, G.J., Umstatter, C., Humphry, R.W., 2014. Replicating disease spread in empirical cattle networks by adjusting the probability of infection in random networks. Theoretical Population Biology 98, 11–18. https://doi.org/10.1016/j.tpb.2014.08.004

Duncan, A.J.J., Gunn, G.J.J., Lewis, F.I.I., Umstatter, C., Humphry, R.W.W., 2012. The influence of empirical contact networks on modelling diseases in cattle. Epidemics 4, 117–123. https://doi.org/10.1016/j.epidem.2012.04.003

Dutta, B.L., Ezanno, P., Vergu, E., 2014. Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period. Preventive Veterinary Medicine 117, 79–94. https://doi.org/10.1016/j.prevetmed.2014.09.005

Ellen, E.D., Bas Rodenburg, T., Albers, G.A.A., Elizabeth Bolhuis, J., Camerlink, I., Duijvesteijn, N., Knol, E.F., Muir, W.M., Peeters, K., Reimert, I., Sell-Kubiak, E., van Arendonk, J.A.M., Visscher, J., Bijma, P., 2014. The prospects of selection for social genetic effects to improve welfare and productivity in livestock. Frontiers in Genetics 5, 377. https://doi.org/10.3389/fgene.2014.00377

Enright, J., Kao, R.R., 2018. Epidemics on dynamic networks. Epidemics 24, 88–97. https://doi.org/10.1016/j.epidem.2018.04.003

Enright, J., Meeks, K., 2015. Deleting edges to restrict the size of an epidemic: A new application for treewidth, in: Lu, Z., Kim, D., Wu, W., Li, W., Du, D. (Eds.), Combinatorial Optimization and Applications. Lecture Notes in Computer Science. Springer, Cham. https://doi.org/https://doi.org/10.1007/s00453-017-0311-7

Enticott, G., 2014. Biosecurity and the Bioeconomy : The case of disease regulation in the UK & New Zealand, in: Morley, A., Marsden, T. (Eds.), Researching Sustainable Food: Building the New Sustainability Paradigm. Earthscan, London, UK, pp. 122–142. https://doi.org/10.4324/9780203083499

Erdös, P., Rényi, A, 1959. On random graphs. Publicationes Mathematicae 6, 290–297. https://doi.org/10.2307/1999405

European Economic Community, 1977. Council Directive 78/52/EEC of establishing the Community criteria for national plans for the accelerated eradication of brucellosis, tuberculosis and enzootic leukosis in cattle, Official journal of the European Communities. Brussels, Belgium.

Fauci, A.S., 2001. Infectious Diseases: Considerations for the 21st Century. Clinical Infectious Diseases 32, 675–685. https://doi.org/10.1086/319235

Faust, M.A., Kinsel, M.L., Kirkpatrick, M.A., 2001. Characterizing Biosecurity, Health, and Culling During Dairy Herd Expansions. Journal of Dairy Science 84, 955–965. https://doi.org/10.3168/jds.S0022-0302(01)74554-7

Ferguson, N.M., Donnelly, C.A., Anderson, R.M., 2001a. The Foot and Mouth epidemic in Great Britain: Pattern of spread and impact of interventions. Science 292, 1061020. https://doi.org/10.1038/35097116

Ferguson, N.M., Donnelly, C.A., Anderson, R.M., 2001b. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. Nature 413, 542–548. https://doi.org/10.1038/35097116

Fielding, H.R., McKinley, T.J., Silk, M.J., Delahay, R.J., McDonald, R.A., 2019. Contact chains of cattle farms in Great Britain. Royal Society Open Science 6, 180719. https://doi.org/dx.doi.org/10.1098/rsos.180719

Fike, J.H., Staples, C.R., Sollenberger, L.E., Macoon, B., Moore, J.E., 2010. Pasture Forages, Supplementation Rate, and Stocking Rate Effects on Dairy Cow Performance. Journal of Dairy Science 86, 1268–1281. https://doi.org/10.3168/jds.s0022-0302(03)73711-4

Foris, B., Zebunke, M., Langbein, J., Melzer, N., 2018. Comprehensive analysis of affiliative and agonistic social networks in lactating dairy cattle groups. Applied Animal Behaviour Science 210, 60–67. https://doi.org/10.1016/j.applanim.2018.10.016

Frössling, J., Nusinovici, S., Nöremark, M., Widgren, S., Lindberg, A., 2014. A novel method to identify herds with an increased probability of disease introduction due to animal trade. Preventive Veterinary Medicine 117, 367–374. https://doi.org/10.1016/j.prevetmed.2014.07.013

Frössling, J., Ohlson, A., Björkman, C., Håkansson, N., Nöremark, M., 2012. Application of network analysis parameters in risk-based surveillance –

Examples based on cattle trade data and bovine infections in Sweden. Preventive Veterinary Medicine 105, 202–208. https://doi.org/10.1016/j.prevetmed.2011.12.011

Fu, Y., Huang, C., Sun, C., 2015. Identifying Super-Spreader Nodes in Complex Networks. Mathematical Problems in Engineering 2015, 675713. https://doi.org/10.1155/2015/675713

Gardner, I.A., Willeberg, P., Mousing, J., 2002. Empirical and theoretical evidence for herd size as a risk factor for swine diseases. Animal Health Research Reviews 3, 43–55. https://doi.org/10.1079/AHRR200239

Gardy, J.L., Johnston, J.C., Sui, S.J.H., Cook, V.J., Shah, L., Brodkin, E., Rempel, S., Moore, R., Zhao, Y., Holt, R., 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. New England Journal Of Medicine 364, 730–739. https://doi.org/DOI: 10.1056/NEJMoa1003176

Garske, T., Rhodes, C.J., 2008. The effect of superspreading on epidemic outbreak size distributions. Journal of Theoretical Biology 253, 228–237. https://doi.org/10.1016/j.jtbi.2008.02.038

Gates, M., 2014. Characteristics of replacement breeding cattle trade in Great Britain. Veterinary Record 175, 67. https://doi.org/10.1136/vr.102351

Gates, M., Humphry, R.W., Gunn, G.J., Woolhouse, M.E.J., 2014. Not all cows are epidemiologically equal: quantifying the risks of bovine viral diarrhoea virus (BVDV) transmission through cattle movements. Veterinary Research 45, 110. https://doi.org/10.1186/s13567-014-0110-y

Gates, M., Volkova, V., Woolhouse, M.E.J., 2013a. Risk factors for bovine tuberculosis in low incidence regions related to the movements of cattle. BMC Veterinary Research 9, 225. https://doi.org/10.1186/1746-6148-9-225

Gates, M., Volkova, V., Woolhouse, M.E.J., 2013b. Impact of changes in cattle movement regulations on the risks of bovine tuberculosis for Scottish farms. Preventive Veterinary Medicine 108, 125–136. https://doi.org/10.1016/j.prevetmed.2012.07.016

Gates, M., Woolhouse, M.E.J., 2015. Controlling infectious disease through the targeted manipulation of contact network structure. Epidemics 12, 11–19. https://doi.org/10.1016/j.epidem.2015.02.008

Ghani, A., Garnett, G., 2000. Risks of Acquiring and Transmitting Sexually Transmitted Diseases in sexual partner networks. Sexually Transmitted Diseases 27, 579–587.

Gibbens, J.C., Sharpe, C.E., Wilesmith, J.W., Mansley, L.M., Michalopoulou, E., Ryan, J.B., Hudson, M., 2001. Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. Veterinary Record 149, 729–743. https://doi.org/10.1136/vr.149.24.729

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. https://doi.org/10.1038/nature03548

Gilchrist, M.J., Greko, C., Wallinga, D.B., Beran, G.W., Riley, D.G., Thorne, P.S., 2007. The potential role of concentrated animal feeding operations in infectious disease epidemics and antibiotic resistance. Environmental Health Perspectives 115, 313–316. https://doi.org/10.1289/ehp.8837

Godfray, C., Donnelly, C.A., Hewison, G., Winter, M., Wood, J.L.N., 2018. Bovine TB Strategy Review. London, UK.

Godfray, H.C.J., Donnelly, C.A., Kao, R.R., Macdonald, D.W., McDonald, R.A., Petrokofsky, G., Wood, J.L.N., Woodroffe, R., Young, D.B., McLean, A.R., 2013. A restatement of the natural science evidence base relevant to the control of bovine tuberculosis in Great Britain. Proceedings of the Royal Society B 280, 20131634. https://doi.org/10.1098/rspb.2013.1634

Godfrey, S.S., Bull, C.M., James, R., Murray, K., 2009. Network structure and parasite transmission in a group living lizard, the gidgee skink, *Egernia stokesii*. Behavioral Ecology and Sociobiology 63, 1045–1056. https://doi.org/10.1007/s00265-009-0730-9

Goodchild, T., Clifton-Hadley, R., Thoen, C., Steele, J., Gilsdorf, M., 2008. The fall and rise of bovine tuberculosis in Great Britain., in: *Mycobacterium bovis* Infection in Animals and Humans. pp. 100–116.

Gopal, R., Goodchild, A., Hewinson, G., de la Rua Domenech, R., Clifton-Hadley, R., 2006. Introduction of bovine tuberculosis to north-east England by bought-in cattle. Veterinary Record 159, 265–271. https://doi.org/10.1136/vr.159.9.265

Gorden, P.J., Plummer, P., 2010. Control, management, and prevention of bovine respiratory disease in dairy calves and cows. Veterinary Clinics of North America - Food Animal Practice 26, 243–259. https://doi.org/10.1016/j.cvfa.2010.03.004

Green, D.M., Gregory, A., Munro, L.A., 2009. Small- and large-scale network structure of live fish movements in Scotland. Preventive Veterinary Medicine 91, 261–269. https://doi.org/10.1016/j.prevetmed.2009.05.031

Green, D.M., Kao, R.R., 2007. Data quality of the Cattle Tracing System in Great Britain. Veterinary Record 161, 439–443. https://doi.org/http://dx.doi.org/10.1136/vr.161.13.439

Green, D.M., Kiss, I.Z., Kao, R.R., 2006. Modelling the initial spread of foot-and-mouth disease through animal movements. Proceedings of the Royal Society B 273, 2729–2735. https://doi.org/10.1098/rspb.2006.3648

Green, D.M., Kiss, I.Z., Mitchell, A.P., Kao, R.R., 2008. Estimates for local and movement-based transmission of bovine tuberculosis in British cattle. Proceedings of the Royal Society B 275, 1001–1005. https://doi.org/10.1098/rspb.2007.1601

Green, D.M., Werkman, M., Munro, L.A., Kao, R.R., Kiss, I.Z., Danon, L., 2011. Tools to study trends in community structure: Application to fish and livestock trading networks. Preventive Veterinary Medicine 99, 225–228. https://doi.org/10.1016/j.prevetmed.2011.01.008

Green, L.E., Cornell, S.J., 2005. Investigations of cattle herd breakdowns with bovine tuberculosis in four counties of England and Wales using VETNET data. Preventive Veterinary Medicine 70, 293–311. https://doi.org/10.1016/j.prevetmed.2005.05.005

Griffin, D., Chengappa, M.M., Kuszak, J., McVey, D.S., 2010. Bacterial pathogens of the bovine respiratory disease complex. Veterinary Clinics of North America - Food Animal Practice 26, 381–394. https://doi.org/10.1016/j.cvfa.2010.04.004

Griffin, J.M., Hahesy, T., Lynch, K., Salman, M.D., McCarthy, J., Hurley, T., 1993. The association of cattle husbandry practices, environmental factors and farmer characteristics with the occurence of chronic bovine tuberculosis in dairy herds in the Republic of Ireland. Preventive Veterinary Medicine 17, 145–160. https://doi.org/10.1016/0167-5877(93)90025-O

Gutmann, A.K., Špinka, M., Winckler, C., 2015. Long-term familiarity creates preferred social partners in dairy cows. Applied Animal Behaviour Science 169, 1–8. https://doi.org/10.1016/j.applanim.2015.05.007

Gygax, L., Neisen, G., Wechsler, B., 2010. Socio-spatial relationships in dairy cows. Ethology 116, 10–23. https://doi.org/10.1111/j.1439-0310.2009.01708.x

Hadrich, J.C., Wolf, C.A., Lombard, J., Dolak, T.M., 2018. Estimating milk yield and value losses from increased somatic cell count on US dairy farms. Journal of Dairy Science 101, 3588–3596. https://doi.org/10.3168/jds.2017-13840

Hand, K.J., Godkin, A., Kelton, D.F., 2012. Milk production and somatic cell counts: A cow-level analysis. Journal of Dairy Science 95, 1358–1362. https://doi.org/10.3168/jds.2011-4927

Handcock, R.N., Swain, D.L., Bishop-Hurley, G.J., Patison, K.P., Wark, T., Valencia, P., Corke, P., O’Neill, C.J., 2009. Monitoring animal behaviour and environmental interactions using wireless sensor networks, GPS collars and satellite remote sensing. Sensors 9, 3586–3603. https://doi.org/10.3390/s90503586

Harrell, F., 2001. Regression Modeling Strategies. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4757-3462-1

Hasegawa, N., Nishiwaki, A., Sugawara, K., Iwao, I., 1997. The effects of social exchange between two groups of lactating primiparous heifers on milk production, dominance order, behavior and adrenocortical response. Applied Animal Behaviour Science 51, 15–27. https://doi.org/10.1016/S0168-1591(96)01082-9

Haskell, M., Rennie, L., Bowell, V., Bell, M., Lawrence, A., 2006. Housing system, milk production, and zero-grazing effects on lameness and leg injury in dairy cows. Journal of Dairy Science 89, 4259–4266. https://doi.org/10.3168/jds.S0022-0302(06)72472-9

Henderson, D.A., 2011. The eradication of smallpox - An overview of the past, present, and future. Vaccine 29S, D7–D9. https://doi.org/10.1016/j.vaccine.2011.06.080

Holme, P., 2005. Network reachability of real-world contact sequences. Physical Review E 71, 046119. https://doi.org/10.1103/PhysRevE.71.046119

Holme, P., Saramäki, J., 2012. Temporal networks. Physics Reports 519, 97–125. https://doi.org/10.1016/j.physrep.2012.03.001

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression. Wiley, New York. https://doi.org/10.1002/0471722146

Houlihan, M.G., Dixon, F.W., Page, N.A., 2008. Outbreak of bovine tuberculosis featuring anergy to the skin test, udder lesions and milkborne disease in young calves. Veterinary Record 163, 357–361. https://doi.org/10.1136/vr.163.12.357

Hussein, A.N., Al-Marashdeh, O., Bryant, R.H., Edwards, G.R., 2016. Relationship between social dominance and milk production of dairy cows grazing pasture. Proceedings of the New Zealand Society of Animal Production 76, 69–72.

Hyun, Y., Ellis, M., Johnson, R.W., 1998. Effects of Feeder Type, Space Allowance, and Mixing on the Growth Performance and Feed Intake Pattern of Growing Pigs. Journal of Animal Science 76, 2771–2778. https://doi.org/10.2527/1998.76112771x

James, R., Croft, D.P., Krause, J., 2009. Potential banana skins in animal social network analysis. Behavioral Ecology and Sociobiology 63, 989–997. https://doi.org/10.1007/s00265-009-0742-5

Jezierski, T.A., Podluzny, M., 1984. A quantitative analysis of social behaviour of different crossbreds of dairy cattle kept in loose housing and its relationship to productivity. Applied Animal Behaviour Science 13, 31–40.

Johnston, W.T., Vial, F., Gettinby, G., Bourne, F.J., Clifton-Hadley, R.S., Cox, D.R., Crea, P., Donnelly, C.A., McInerney, J.P., Mitchell, A.P., Morrison, W.I., Woodroffe, R., 2011. Herd-level risk factors of bovine tuberculosis in England and Wales after the 2001 foot-and-mouth disease epidemic. International Journal of Infectious Diseases 15, e833–e840. https://doi.org/10.1016/j.ijid.2011.08.004

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., Daszak, P., 2008. Global trends in emerging infectious diseases. Nature 451, 990–993. https://doi.org/10.1038/nature06536

Jordan, D., McEwen, S. a, 1998. Herd-level test performance based on uncertain estimates of individual test performance, individual true prevalence and herd true prevalence. Preventive Veterinary Medicine 36, 187–209. https://doi.org/10.1016/S0167-5877(98)00087-7

Kao, R., 2011. Risk-based surveillance for tuberculosis in cattle (bTB). Report No. UGW/003/10. The Scottish Government, Edinburgh.

Kao, R.R., 2002. The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. Trends in Microbiology 10, 279–286. https://doi.org/10.1016/S0966-842X(02)02371-5

Kao, R.R., Green, D.M., Johnson, J., Kiss, I.Z., 2007. Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the UK. Journal of the Royal Society Interface 4, 907–916. https://doi.org/10.1098/rsif.2007.1129

Kappeler, P.M., Cremer, S., Nunn, C.L., 2015. Sociality and health: Impacts of sociality on disease susceptibility and transmission in animal and human societies. Philosophical Transactions of the Royal Society B 370, 20140116. https://doi.org/10.1098/rstb.2014.0116

Kay, S.J., Collis, K.A., Anderson, J.C., Grant, A.J., 1977. The effect of intergroup movement of dairy cows on bulk-milk somatic cell numbers. Journal of Dairy Research 44, 589–593. https://doi.org/10.1017/S0022029900020549

Keeling, M.J., 2005. Models of foot-and-mouth disease. Proceedings of the Royal Society B 272, 1195–1202. https://doi.org/10.1098/rspb.2004.3046

Keeling, M.J., 1999. The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society B 266, 859–867.

Keeling, M.J., 1997. Modelling the persistence of measles. Trends in Microbiology 5, 513–518. https://doi.org/10.1016/S0966-842X(97)01147-5

Keeling, M.J., Eames, K.T.D., 2005. Networks and epidemic models. Journal of The Royal Society Interface 2, 295–307. https://doi.org/10.1098/rsif.2005.0051

Keeling, M.J., Woolhouse, M.E., 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, 1065973. https://doi.org/10.1126/science.1065973

Keeling, M.J., Woolhouse, M.E.J., May, R.M., Davies, G., Grenfell, B.T., 2003. Modelling vaccination strategies against foot-and-mouth disease. Nature 421, 136–142. https://doi.org/10.1038/nature01343

Kermack, W.O., McKendrick, A.G., 1927. Contributions to the Mathematical Theory of Epidemics - 1 (reprinted in Bull. Math. Biol. 1991. 53, 33-55). Proceedings of the Royal Society 115A, 700–721.

Kilgour, R.J., 2012. In pursuit of “normal”: A review of the behaviour of cattle at pasture. Applied Animal Behaviour Science 138, 1–11. https://doi.org/10.1016/j.applanim.2011.12.002

Kilpatrick, A.M., Gillin, C.M., Daszak, P., 2009. Wildlife-livestock conflict: The risk of pathogen transmission from bison to cattle outside Yellowstone National Park. Journal of Applied Ecology 46, 476–485. https://doi.org/10.1111/j.1365-2664.2008.01602.x

Kiss, I.Z., Green, D.M., Kao, R.R., 2006. 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. https://doi.org/10.1098/rsif.2006.0129

Klovdahl, A.S., 1985. Social networks and the spread of infectious diseases: The AIDS example. Social Science and Medicine 21, 1203–1216. https://doi.org/10.1016/0277-9536(85)90269-2

Konschake, M., Lentz, H.H.K., Conraths, F.J., Hövel, P., Selhorst, T., 2013. On the Robustness of In- and Out-Components in a Temporal Network. PLoS ONE 8, e55223. https://doi.org/10.1371/journal.pone.0055223

Kosmala, M., Miller, P., Ferreira, S., Funston, P., Keet, D., Packer, C., 2016. Estimating wildlife disease dynamics in complex systems using an

Approximate Bayesian Computation framework. Ecological Applications 26, 295–308. https://doi.org/https://doi.org/10.1890/14-1808

Krawczel, P.D., Klaiber, L.B., Butzler, R.E., Klaiber, L.M., Dann, H.M., Mooney, C.S., Grant, R.J., 2012. Short-term increases in stocking density affect the lying and social behavior, but not the productivity, of lactating Holstein dairy cows. Journal of Dairy Science 95, 4298–4308. https://doi.org/10.3168/jds.2011-4687

Krebs, J., Anderson, R., Clutton-Brock, T., Morrison, I., Young, D., Donnelley, C., 1997. Bovine Tuberculosis in Cattle and Badgers. MAFF, London, UK.

Laister, S., Stockinger, B., Regner, A.M., Zenger, K., Knierim, U., Winckler, C., 2011. Social licking in dairy cattle-Effects on heart rate in performers and receivers. Applied Animal Behaviour Science 130, 81–90. https://doi.org/10.1016/j.applanim.2010.12.003

Lanzas, C., Brien, S., Ivanek, R., LO, Y., Chapagain, P.P., Ray, K.A., Ayscue, P., Warnick, L.D., Gröhn, Y.T., 2008. The effect of heterogeneous infectious period and contagiousness on the dynamics of *Salmonella* transmission in dairy cattle. Epidemiology and Infection 136, 1496–1510. https://doi.org/10.1017/S0950268807000209

Lass, S., Hudson, P.J., Thakar, J., Saric, J., Harvill, E., Albert, R., Perkins, S.E., 2013. Generating super-shedders: co-infection increases bacterial load and egg production of a gastrointestinal helminth. Journal of the Royal Society Interface 10, 20120588. https://doi.org/10.1098/rsif.2012.0588

Lawes, J.R., Harris, K.A., Brouwer, A., Broughan, J.M., Smith, N.H., Upton, P.A., 2016. Bovine TB surveillance in Great Britain in 2014. Veterinary Record 178, 310–315. https://doi.org/10.1136/vr.i1616

Lentz, H.H.K., Koher, A., Hövel, P., Gethmann, J., Sauter-Louis, C., Selhorst, T., Conraths, F.J., 2016. Disease spread through animal movements: a static and temporal network analysis of pig trade in Germany. PLoS ONE 11, e0155196. https://doi.org/10.1371/journal.pone.0155196

Li, S., 2016. Evaluating surveillance strategies for bovine tuberculosis in Scotland (Doctoral thesis). University of Edinburgh.

Lindstrom, T., Lewerin, S.S., Wennergren, U., 2012. Influence on disease spread dynamics of herd characteristics in a structured livestock industry. Journal of The Royal Society Interface 9, 1287–1294. https://doi.org/10.1098/rsif.2011.0625

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438, 355–359. https://doi.org/10.1038/nature04153

Mansley, L.M., Donaldson, A.I., Thrusfield, M. V, Honhold, N., 2011. Destructive tension: mathematics versus experience--the progress and control of the 2001 foot and mouth disease epidemic in Great Britain. Scientific and Technical Review of the Office International des Epizooties (Paris) 30, 483–498. https://doi.org/10.20506/rst.30.2.2054

Marcé, C., Ezanno, P., Seegers, H., Pfeiffer, D.U., Fourichon, C., 2011. Within-herd contact structure and transmission of *Mycobacterium avium* subspecies *paratuberculosis* in a persistently infected dairy cattle herd. Preventive Veterinary Medicine 100, 116–125. https://doi.org/10.1016/j.prevetmed.2011.02.004

Mardones, F.O., Perez, A.M., Valdes-Donoso, P., Carpenter, T.E., 2011. Farm-level reproduction number during an epidemic of infectious salmon anemia virus in southern Chile in 2007–2009. Preventive Veterinary Medicine 102, 175–184. https://doi.org/http://dx.doi.org/10.1016/j.prevetmed.2011.07.005

Martin, S.W., Shoukri, M., Thorburn, M.A., 1992. Evaluating the health status of herds based on tests applied to individuals. Preventive Veterinary Medicine 14, 33–43. https://doi.org/10.1016/0167-5877(92)90082-Q

Martin, V., Zhou, X., Marshall, E., Jia, B., Fusheng, G., FrancoDixon, M.A., de Haan, N., Pfeiffer, D.U., Soares Magalhães, R.J., Gilbert, M., 2011. Risk-based surveillance for avian influenza control along poultry market chains in South China: The value of social network analysis. Preventive Veterinary Medicine 102, 196–205. https://doi.org/10.1016/j.prevetmed.2011.07.007

Martínez-López, B., Perez, A.M., Sánchez-Vizcaíno, J.M., 2009. Social network analysis. Review of general concepts and use in preventive veterinary medicine. Transboundary and Emerging Diseases 56, 109–120. https://doi.org/10.1111/j.1865-1682.2009.01073.x

Mathers, C., Stevens, G., Hogan, D., 2017. Global and regional causes of death: patterns and trends 2000-2015, in: Disease Control Priorities, Third Edition (Volume 9): Improving Health and Reducing Poverty. The International Bank for Reconstruction and Development / The World Bank, pp. 43–65. https://doi.org/10.1596/978-1-4648-0527-1_ch3

Matthews, L., Reeve, R., Woolhouse, M.E.J., Chase-Topping, M., Mellor, D., Pearce, M., Allison, L., Gunn, G., Low, J., Reid, S., 2009. Exploiting strain diversity to expose transmission heterogeneities and predict the impact of targeting supershedding. Epidemics 1, 221–229. https://doi.org/10.1016/j.epidem.2009.10.002

May, R.M., 2006. Network structure and the biology of populations. Trends in Ecology and Evolution 21, 394–399. https://doi.org/10.1016/j.tree.2006.03.013

May, R.M., Anderson, R.M., 1987. Transmission dynamics of HIV infection. Nature 326, 137–42. https://doi.org/10.1038/326137a0

McDonald, R.A., Delahay, R.J., Carter, S.P., Smith, G.C., Cheeseman, C.L., 2008. Perturbing implications of wildlife ecology for disease control. Trends in Ecology and Evolution 23, 53–56. https://doi.org/10.1016/j.tree.2007.10.011

McKinley, T., Cook, A.R., Deardon, R., 2009. Inference in Epidemic Models without Likelihoods. The International Journal of Biostatistics 5, 24. https://doi.org/10.2202/1557-4679.1171

McKinley, T., Lipschutz-Powell, D., Mitchell, A.P., Wood, J., Conlan, A.J.K., 2018. Risk factors and variations in detection of new bovine tuberculosis breakdowns via slaughterhouse surveillance in Great Britain. PloS ONE 13, e0198760.

Mclennan, K.M., 2013. Social bonds in dairy cattle: the effect of dynamic group systems on welfare and productivity. Doctoral Thesis. The University of Northampton.

Mekonnen, G.A., Ameni, G., Wood, J.L.N., The ETHICOBOTS Consortium, Berg, S., Conlan, A.J.K., 2019. Network analysis of dairy cattle movement and its implication on Bovine Tuberculosis spread and control in emerging dairy belts of Ethiopia. BMC Veterinary Research 15, 262. https://doi.org/10.1186/s12917-019-1962-1

Mellor, D.J., 2015. Positive animal welfare states and encouraging environment-focused and animal-to-animal interactive behaviours. New Zealand Veterinary Journal 63, 9–16. https://doi.org/10.1080/00480169.2014.926800

Mellor, D.J., Beausoleil, N.J., 2015. Extending the “Five Domains” model for animal welfare assessment to incorporate positive welfare states. Animal Welfare 24, 241–253. https://doi.org/10.7120/09627286.24.3.241

Melmer, D., Sullivan, O., Terri, L., 2018. A descriptive analysis of swine movements in Ontario (Canada) as a contributor to disease spread. Preventive Veterinary Medicine 159, 211–219. https://doi.org/10.1016/J.PREVETMED.2018.09.021

Mench, J.A., 2002. Broiler breeders: feed restriction and welfare. World’s Poultry Science Journal 58, 23–29. https://doi.org/10.1079/wps20020004

Meyers, L.A., Pourbohloul, B., Newman, M.E.J., Skowronski, D.M., Brunham, R.C., 2005. Network theory and SARS: predicting outbreak diversity. Journal of Theoretical Biology 232, 71–81. https://doi.org/10.1016/j.jtbi.2004.07.026

Milwid, R.M., O’Sullivan, T.L., Poljak, Z., Laskowski, M., Greer, A.L., 2019. Comparing the effects of non-homogenous mixing patterns on epidemiological outcomes in equine populations: A mathematical modelling study. Scientific Reports 9, 3227. https://doi.org/10.1038/s41598-019-40151-2

Miranda-de la Lama, G.C., Mattiello, S., 2010. The importance of social behaviour for goat welfare in livestock farming. Small Ruminant Research 90, 1–10. https://doi.org/10.1016/j.smallrumres.2010.01.006

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. Animal Science 80, 265–273. https://doi.org/10.1079/ASC50020265

Molia, S., Boly, I.A., Duboz, R., Coulibaly, B., Guitian, J., Grosbois, V., Fournié, G., Pfeiffer, D.U., 2016. Live bird markets characterization and trading network analysis in Mali: Implications for the surveillance and control of avian influenza and Newcastle disease. Acta Tropica 155, 77–88. https://doi.org/10.1016/j.actatropica.2015.12.003

Monies, R.J., 2000. Tuberculous pneumonia and BVD in housed calves. Cattle Practice 8, 119–125.

Morgan, N., Prakash, A., 2006. International livestock markets and the impact of animal disease. Scientific and Technical Review of the Office International des Epizooties (Paris) 25, 517–528.

Morris, R.S., Wilesmith, J.W., Stern, M.W., Sanson, R.L., Stevenson, M.A., 2001. Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain, 2001. Veterinary Record 149, 137–144. https://doi.org/10.17660/ActaHortic.2001.566.43

Mweu, M.M., Fournié, G., Halasa, T., Toft, N., Nielsen, S.S., 2013. Temporal characterisation of the network of Danish cattle movements and its implication for disease control: 2000-2009. Preventive Veterinary Medicine 110, 379–387. https://doi.org/10.1016/j.prevetmed.2013.02.015

Napolitano, F., Knierim, U., Grass, F., De Rosa, G., 2009. Positive indicators of cattle welfare and their applicability to on-farm protocols. Italian Journal of Animal Science 8, 355–365. https://doi.org/10.4081/ijas.2009.s1.355

Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G., Calistri, P., 2009. Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. Preventive Veterinary Medicine 92, 341–350. https://doi.org/10.1016/j.prevetmed.2009.08.026

Neill, S.D., Bryson, D.G., Pollock, J.M., 2001. Pathogenesis of tuberculosis in cattle. Tuberculosis 81, 79–86. https://doi.org/10.1054/tube.2000.0279

Neill, S.D., Hanna, J., Mackie, D.P., Bryson, T.G.D., 1992. Isolation of *Mycobacterium bovis* from the respiratory tracts of skin test-negative cattle. Veterinary Record 131, 45–47.

Nigsch, A., Costard, S., Jones, B.A., Pfeiffer, D.U., Wieland, B., 2013. Stochastic spatio-temporal modelling of African swine fever spread in the European Union during the high risk period. Preventive Veterinary Medicine 108, 262–275. https://doi.org/10.1016/j.prevetmed.2012.11.003

Nöremark, M., 2010. Infection through the farm gate. Doctoral Thesis. Swedish University of Agricultural Sciences.

Nöremark, M., Håkansson, N., Lewerin, S.S., Lindberg, A., Jonsson, A., 2011. Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk based surveillance. Preventive Veterinary Medicine 99, 78–90. https://doi.org/10.1016/j.prevetmed.2010.12.009

Nöremark, M., Widgren, S., 2014. EpiContactTrace: an R-package for contact tracing during livestock disease outbreaks and for risk-based surveillance. BMC Veterinary Research 10, 71. https://doi.org/10.1186/1746-6148-10-71

Nuñez-Garcia, J., Downs, S.H., Woolliams, J.A., Sharp, M., Gunn, J., More, S.J., Cook, A.J., Welsh, M., Rolfe, S., Parry, J.E., Clifton-Hadley, R.S., Rhodes, S., Upton, P.A., de la Rua-Domenech, R., Abernethy, D.A., Watson, E., Goodchild, A. V., Broughan, J.M., Cameron, A.R., Whelan, A.O., Vordermeier, H.M., Greiner, M., 2018. Meta-analyses of the sensitivity and specificity of ante-mortem and post-mortem diagnostic tests for bovine tuberculosis in the UK and Ireland. Preventive Veterinary Medicine 153, 94–107. https://doi.org/10.1016/j.prevetmed.2017.02.017

Office for National Statistics, 2018. Labour in the agriculture industry, UK. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/labourintheagricultureindustry/2018-02-06 Accessed on 20th September 2019.

Oltenacu, P., Broom, D., 2010. Impact of Genetic Selection for Increased Milk Yield on the Welfare of Dairy Cows. Animal Welfare 19, 39–49.

Orsel, K., Bouma, A., Dekker, A., Stegeman, J.A., de Jong, M.C.M., 2009. Foot and mouth disease virus transmission during the incubation period of the disease in piglets, lambs, calves, and dairy cows. Preventive Veterinary Medicine 88, 158–163. https://doi.org/10.1016/j.prevetmed.2008.09.001

Orton, R.J., Deason, M., Bessell, P.R., Green, D.M., Kao, R.R., Salvador, L.C.M., 2018. Identifying genotype specific elevated-risk areas and associated herd risk factors for bovine tuberculosis spread in British cattle. Epidemics 24, 34–42. https://doi.org/10.1016/j.epidem.2018.02.004

Palisson, A., Courcoul, A., Durand, B., 2016. Role of cattle movements in bovine tuberculosis spread in France between 2005 and 2014. PLoS ONE 11, e0152578. https://doi.org/10.1371/journal.pone.0152578

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic spreading in scale-free networks. Physical Review Letters 86, 3200–3203. https://doi.org/10.1103/PhysRevLett.86.3200

Patison, K.P., Swain, D.L., Bishop-Hurley, G.J., Pattison, P., Robins, G., 2010. Social companionship versus food: The effect of the presence of familiar and unfamiliar conspecifics on the distance steers travel. Applied Animal Behaviour Science 122, 13–20. https://doi.org/10.1016/j.applanim.2009.11.001

Perkins, S.E., Cagnacci, F., Stradiotto, A., Arnoldi, D., Hudson, P.J., 2009. Comparison of social networks derived from ecological data: Implications for inferring infectious disease dynamics. Journal of Animal Ecology 78, 1015–1022. https://doi.org/10.1111/j.1365-2656.2009.01557.x

Phillips, C.J.C., Foster, C.R.W., Morris, P.A., Teverson, R., 2002. Genetic and management factors that influence the susceptibility of cattle to *Mycobacterium bovis* infection. Animal Health Research Reviews 3, 3–13. https://doi.org/10.1079/AHRR200236

Polikarpus, A., Kaart, T., Mootse, H., Rosa, G. De, Arney, D., 2015. Influences of various factors on cows’ entrance order into the milking parlour. Applied Animal Behaviour Science 166, 20–24. https://doi.org/10.1016/j.applanim.2015.02.016

Pollock, J.M., Neill, S.D., 2002. *Mycobacterium bovis*: Infection and Tuberculosis in Cattle. The Veterinary Journal 163, 115–127. https://doi.org/10.1053/tvjl.2001.0655

Poutanen, S.M., Low, D.E., Henry, B., Finkelstein, S., Rose, D., Green, K., Tellier, R., Draker, R., Adachi, D., Ayers, M., Chan, A.K., Skowronski, D.M., Salit, I., Simor, A.E., Slutsky, A.S., Doyle, P.W., Krajden, M., Petric, M., Brunham, R.C., McGeer, A.J., 2003. Identification of severe acute respiratory syndrome in Canada. New England Journal of Medicine 348, 1995–2005. https://doi.org/10.1056/NEJMoa030634

Pradhan, A.K., Mitchell, R.M., Kramer, A.J., Zurakowski, M.J., Fyock, T.L., Whitlock, R.H., Smith, J.M., Hovingh, E., Van Kessel, J.A.S., Karns, J.S., Schukken, Y.H., 2011. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis* in a longitudinal study of three dairy herds. Journal of Clinical Microbiology 49, 893–901. https://doi.org/10.1128/JCM.01107-10

Probst, C., Freuling, C., Moser, I., Geue, L., Köhler, H., Conraths, F.J., Hotzel, H., Liebler-Tenorio, E.M., Kramer, M., 2011. Bovine tuberculosis: making a case for effective surveillance. Epidemiology & Infection 139, 105–112. https://doi.org/10.1007/s13398-014-0173-7.2

Proudfoot, K., Habing, G., 2015. Social stress as a cause of diseases in farm animals: Current knowledge and future directions. The Veterinary Journal 206, 15–21. https://doi.org/10.1016/j.tvjl.2015.05.024

Proudfoot, K.L., Jensen, M.B., Weary, D.M., von Keyserlingk, M.A.G., 2014. Dairy cows seek isolation at calving and when ill. Journal of Dairy Science 97, 2731–2739. https://doi.org/10.3168/jds.2013-7274

Proudfoot, K.L., Weary, D.M., von Keyserlingk, M.A.G., 2012. Linking the social environment to illness in farm animals. Applied Animal Behaviour Science 138, 203–215. https://doi.org/10.1016/j.applanim.2012.02.008

Pryce, J., Veerkamp, R.F., 2001. The incorporation of fertility indices in genetic improvement programmes, in: Fertility in the High Producing Dairy Cow. BNAS Occasional Publication. BNAS, pp. 237–250. https://doi.org/10.1017/S0263967X00033711

QGIS Development team, 2019. QGIS. Open Source Geospatial Foundation.

R Core Team Version 3.5.3, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rathore, A.K., 1982. Order of cow entry at milking and its relationships with milk yield and consistency of the order. Applied Animal Ethology 8, 45–52. https://doi.org/10.1016/0304-3762(82)90131-6

Rault, J.L., 2012. Friends with benefits: Social support and its relevance for farm animal welfare. Applied Animal Behaviour Science 136, 1–14. https://doi.org/10.1016/j.applanim.2011.10.002

Raussi, S., Boissy, A., Delval, E., Pradel, P., Kaihilahti, J., Veissier, I., 2005. Does repeated regrouping alter the social behaviour of heifers? Applied Animal Behaviour Science 93, 1–12. https://doi.org/10.1016/j.applanim.2004.12.001

Rautureau, S., Dufour, B., 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. https://doi.org/10.1111/j.1865-1682.2010.01187.x

Read, J.M., Eames, K.T.D., Edmunds, W.J., 2008. Dynamic social networks and the implications for the spread of infectious disease. Journal of the Royal Society Interface 5, 1001–1007. https://doi.org/10.1098/rsif.2008.0013

Reinhardt, V., Reinhardt, A., 1981. Cohesive Relationships in a Cattle Herd (*Bos indicus*). Behaviour 77, 121–151.

Reinhardt, V., Reinhardt, A., 1975. Dynamics of Social Hierarchy in a Dairy Herd. Zeitschrift für Tierpsychologie 38, 315–323. https://doi.org/10.1111/j.1439-0310.1975.tb02007.x

Reynolds, D., 2006. A review of tuberculosis science and policy in Great Britain. Veterinary Microbiology 112, 119–126. https://doi.org/10.1016/j.vetmic.2005.11.042

Riolo, C.S., Koopman, J.S., Chick, S.E., 2001. Methods and measures for the description of epidemiological contact networks. Journal of Urban Health 78, 446–456.

Robinson, S.E., Christley, R.M., 2007. Exploring the role of auction markets in cattle movements within Great Britain. Preventive Veterinary Medicine 81, 21–37. https://doi.org/10.1016/j.prevetmed.2007.04.011

Robinson, S.E., Christley, R.M., 2006. Identifying temporal variation in reported births, deaths and movements of cattle in Britain. BMC Veterinary Research 2, 11. https://doi.org/10.1186/1746-6148-2-11

Roeder, P.L., 2011. Rinderpest: The end of cattle plague. Preventive Veterinary Medicine 102, 98–106. https://doi.org/10.1016/j.prevetmed.2011.04.004

Rossi, G., De Leo, G.A., Pongolini, S., Natalini, S., Zarenghi, L., Ricchi, M., Bolzoni, L., 2017a. The Potential Role of Direct and Indirect Contacts on Infection Spread in Dairy Farm Networks. PLoS Computational Biology 13, 1–19. https://doi.org/10.1371/journal.pcbi.1005301

Rossi, G., Smith, R.L., Pongolini, S., Bolzoni, L., 2017b. Modelling farm-to-farm disease transmission through personnel movements: From visits to contacts, and back. Scientific Reports 7, 1–11. https://doi.org/10.1038/s41598-017-02567-6

Rozins, C., Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., McDonald, R.A., Weber, N., Boots, M., 2018. Social structure contains epidemics and regulates individual roles in disease transmission in a group-living mammal. Ecology and Evolution 8, 12044–12055. https://doi.org/10.1002/ece3.4664

Sah, P., Leu, S.T., Cross, P.C., Hudson, P.J., Bansal, S., 2017. Unraveling the disease consequences and mechanisms of modular structure in animal social networks. Proceedings of the National Academy of Sciences USA

114, 4165–4170. https://doi.org/10.1073/pnas.1613616114

Salvador, L.C.M., Deason, M., Enright, J., Bessell, P.R., Kao, R.R., 2018. Risk-based strategies for surveillance of tuberculosis infection in cattle for low-risk areas in England and Scotland. Epidemiology and Infection 146, 107–118. https://doi.org/10.1017/S0950268817001935

Sambraus, H., 1969. Zur sozialen Rangordnung von Rindern. Zeitschrift für Tierzüchtung und Züchtungsbiologie 86, 240–257.

Santos, N., Almeida, V., Gortázar, C., Correia-Neves, M., 2015. Patterns of *Mycobacterium tuberculosis*-complex excretion and characterization of super-shedders in naturally-infected wild boar and red deer. Veterinary Research 46, 129. https://doi.org/10.1186/s13567-015-0270-4

Sart, S., Bencini, R., Blache, D., Martin, G., 2004. Calm Ewes Produce Milk With More Protein Than Nervous Ewes. Animal Production in Australia 25, 307.

Sato, S., 1984. Social licking pattern and its relationships to social dominance and live weight gain in weaned calves. Applied Animal Behaviour Science 12, 25–32. https://doi.org/10.1016/0168-1591(84)90093-5

Sato, S., Tarmuizu, K., Sonoda, T., 1991. Social, behavioural and physiological functions of allo-grooming in cattle, in: Applied Animal Behaviour: Past, Present and Future. Proceedings of the international congress Edinburgh, pp. 77–78.

Scanlon, M., Quinn, P., 2000. Inactivation of *Mycobacterium bovis* in cattle slurry by five volatile chemicals. Journal of Applied Microbiology 89, 854–861. https://doi.org/10.1046/j.1365-2672.2000.01190.x

Schärrer, S., Widgren, S., Schwermer, H., Lindberg, A., Vidondo, B., Zinsstag, J., Reist, M., 2015. Evaluation of farm-level parameters derived from animal movements for use in risk-based surveillance programmes of cattle in Switzerland. BMC Veterinary Research 11, 149. https://doi.org/10.1186/s12917-015-0468-8

Schein, M., Fohrman, M., 1955. Social dominance relationships in a herd of dairy cattle. The British Journal of Animal Behaviour 3, 45–55.

Schukken, Y.H., Whitlock, R.H., Wolfgang, D., Grohn, Y., Beaver, A., Vankessel, J., Zurakowski, M., Mitchell, R., 2015. Longitudinal data collection of *Mycobacterium avium* subspecies *paratuberculosis* infections in dairy herds: the value of precise field data. Veterinary Research 46, 65. https://doi.org/10.1186/s13567-015-0187-y

Sevi, A., Taibi, L., Albenzio, M., Muscio, A., Dell’Aquila, S., Napolitano, F.,

2001. Behavioral, adrenal, immune, and productive responses of lactating ewes to regrouping and relocation. Journal of Animal Science 79, 1457–1465. https://doi.org/10.2527/2001.7961457x

Shen, Z., Ning, F., Zhou, W., He, X., Lin, C., Chin, D.P., Zhu, Z., Schuchat, A., 2004. Superspreading SARS Events, Beijing, 2003. Emerging Infectious Diseases 10, 256–260. https://doi.org/10.3201/eid1002.030732

Shirley, M.D.F., Rushton, S.P., 2005. The impacts of network topology on disease spread. Ecological Complexity 2, 287–299. https://doi.org/10.1016/j.ecocom.2005.04.005

Shultz, S., Dunbar, R., 2010. Bondedness and sociality. Behaviour 147, 775–803. https://doi.org/10.1163/000579510x501151

Sibley, R., 2010. Biosecurity in the dairy herd. In Practice 32, 274–280. https://doi.org/10.1136/inp.c3913

Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., Boots, M., Weber, N., McDonald, R.A., 2017a. Using social network measures in wildlife disease ecology, epidemiology, and management. BioScience 67, 245–257. https://doi.org/10.1093/biosci/biw175

Silk, M.J., Croft, D.P., Delahay, R.J., Hodgson, D.J., Weber, N., Boots, M., McDonald, R.A., 2017b. The application of statistical network models in disease research. Methods in Ecology and Evolution 8, 1026–1041. https://doi.org/10.1111/2041-210X.12770

Silva, P.R.B., Moraes, J.G.N., Mendonça, L.G.D., Scanavez, A.A., Nakagawa, G., Fetrow, J., Endres, M.I., Chebel, R.C., 2013. Effects of weekly regrouping of prepartum dairy cows on metabolic, health, reproductive, and productive parameters. Journal of Dairy Science 96, 4436–4446. https://doi.org/10.3168/jds.2012-6464

Simm, G., Oldham, J., Coffey, M., 2001. Dairy cows in the future. BSAP Occasional Publication 26, 1–18. https://doi.org/doi:10.1017/S0263967X00033553

Sintayehu, D.W., Prins, H.H.T., Heitkönig, I.M.A., de Boer, W.F., 2017. Disease transmission in animal transfer networks. Preventive Veterinary Medicine 137, 36–42. https://doi.org/10.1016/j.prevetmed.2016.12.017

Skuce, R., Allen, A.R., Mcdowell, S.W.J., 2011. Bovine Tuberculosis (TB): A Review Of Cattle-To-Cattle Transmission, Risk Factors and Susceptibility, in: Report. Agri-food and Biosciences Institute; http://www.dardni.gov.uk.

Skuce, R.A., Allen, A.R., McDowell, S.W.J., 2012. Herd-level risk factors for bovine tuberculosis: A literature review. Veterinary Medicine International

2012, 621210. https://doi.org/10.1155/2012/621210

Small, M., Tse, C.K., Walker, D.M., 2006. Super-spreaders and the rate of transmission of the SARS virus. Physica D 215, 146–158. https://doi.org/10.1016/j.physd.2006.01.021

Smith, G.C., 2001. Models of *Mycobacterium bovis* in wildlife and cattle. Tuberculosis 81, 51–64. https://doi.org/10.1054/tube.2000.0264

Soffié, M., Thinès, G., Marneffe, G.D.E., 1976. Relation between milking order and dominance value in a group of dairy cows. Applied Animal Ethology 2, 271–276.

Sottysiak, T., Nogalski, Z., 2010. The effects of social hierarchy in a dairy cattle herd on milk yield. Polish Journal of Natural Sciences 25, 22–30.

Sowerby, M., Polan, C., 1978. Milk Production Response to Shifting Cows Between Intraherd Groups. Journal of Dairy Science 61, 455–460. https://doi.org/10.3168/jds.S0022-0302(78)83620-0

Springer, A., Kappeler, P.M., Nunn, C.L., 2017. Dynamic vs. static social networks in models of parasite transmission: predicting *Cryptosporidium* spread in wild lemurs. Journal of Animal Ecology 86, 419–433. https://doi.org/10.1111/1365-2656.12617

Stein, R.A., 2011. Super-spreaders in infectious diseases. International Journal of Infectious Diseases 15, e510–e513. https://doi.org/10.1016/j.ijid.2010.06.020

Stoffel, M.A., Nakagawa, S., Schielzeth, H., 2017. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods in Ecology and Evolution 8, 1639–1644. https://doi.org/10.1111/2041-210X.12797

Stoye, S., Porter, M.A., Dawkins, M.S., 2012. Synchronized lying in cattle in relation to time of day. Livestock Science 149, 70–73. https://doi.org/10.1016/j.livsci.2012.06.028

Sumner, K.M., McCabe, C.M., Nunn, C.L., 2018. Network size, structure, and pathogen transmission: A simulation study comparing different community detection algorithms. Behaviour 155, 639–670. https://doi.org/10.1163/1568539X-00003508

Swain, D.L., Bishop-Hurley, G.J., 2007. Using contact logging devices to explore animal affiliations: Quantifying cow-calf interactions. Applied Animal Behaviour Science 102, 1–11. https://doi.org/10.1016/j.applanim.2006.03.008

Swift, B.M.C., Convery, T.W., Rees, C.E.D., 2016. Evidence of *Mycobacterium* *tuberculosis* complex bacteraemia in intradermal skin test positive cattle detected using phage-RPA. Virulence 7, 779–788. https://doi.org/10.1080/21505594.2016.1191729

Swinton, J., Tuyttens, F., Macdonald, D., Nokes, D.J., Cheeseman, C.L., Clifton-Hadley, R., 1997. A comparison of fertility control and lethal control of bovine tuberculosis in badgers: The impact of perturbation induced transmission. Philosophical Transactions of the Royal Society B 352, 619–631. https://doi.org/10.1098/rstb.1997.0042

Takeda, K., Sato, S., Kazuo, S., 2000. The number of farm mates influences social and maintenance behaviours of Japanese Black cows in a communal pasture. Applied Animal Behaviour Science 67, 181–192. https://doi.org/10.1016/S0168-1591(99)00124-0

Takeda, K.I., Sato, S., Sugawara, K., 2003. Familiarity and group size affect emotional stress in Japanese Black heifers. Applied Animal Behaviour Science 82, 1–11. https://doi.org/10.1016/S0168-1591(03)00039-X

Talebi, A., von Keyserlingk, M.A.G., Telezhenko, E., Weary, D.M., 2014. Reduced stocking density mitigates the negative effects of regrouping in dairy cattle. Journal of Dairy Science 97, 1358–1363. https://doi.org/10.3168/jds.2013-6921

te Beest, D.E., Hagenaars, T.J., Stegeman, J.A., Koopmans, M.P.G., van Boven, M., 2011. Risk based culling for highly infectious diseases of livestock. Veterinary Research 42, 81. https://doi.org/10.1186/1297-9716-42-81

The Government Office for Science, 2011. Foresight. The Future of Food and Farming. London, UK.

Thompson, D., Muriel, P., Russell, D., Osborne, P., Bromley, A., Rowland, M., Creigh-Tyte, S., Brown, C., 2002. Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001. Scientific and Technical Review of the Office International des Epizooties (Paris) 21, 675–687. https://doi.org/10.20506/rst.21.3.1353

Thurmond, M.C., 2003. Conceptual foundations for infectious disease surveillance. Journal of Veterinary Diagnostic Investigation 15, 501–514. https://doi.org/10.1177/104063870301500601

Tildesley, M.J., Bessell, P.R., Keeling, M.J., Woolhouse, M.E.J., 2009. The role of pre-emptive culling in the control of foot-and-mouth disease. Proceedings of the Royal Society B 276, 3239–48. https://doi.org/10.1098/rspb.2009.0427

Tildesley, M.J., Keeling, M.J., 2009. Is R_0 a good predictor of final epidemic size: Foot-and-mouth disease in the UK. Journal of Theoretical Biology 258, 623–629. https://doi.org/10.1016/j.jtbi.2009.02.019

Tildesley, M.J., Savill, N.J., Shaw, D.J., Deardon, R., Brooks, S.P., Woolhouse, M.E.J., Grenfell, B.T., Keeling, M.J., 2006. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. Nature 440, 83–86. https://doi.org/10.1038/nature04324

Tresoldi, G., Weary, D.M., Filho, L.C.P.M., von Keyserlingk, M.A.G., 2015. Social licking in pregnant dairy heifers. Animals 5, 1169–1179. https://doi.org/10.3390/ani5040404

Tuomisto, L., Huuskonen, A., Jauhiainen, L., Mononen, J., 2019. Finishing bulls have more synchronised behaviour in pastures than in pens. Applied Animal Behaviour Science 213, 26–32. https://doi.org/10.1016/j.applanim.2019.02.007

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. Journal of Theoretical Biology 254, 45–54. https://doi.org/10.1016/j.jtbi.2008.05.007

van Pinxteren, L.A.H., Ravn, P., Agger, E.M., Pollock, J., Andersen, P., 2000. Diagnosis of Tuberculosis Based on the Two Specific Antigens ESAT-6 and CFP10. Clinical and Vaccine Immunology 7, 155–160. https://doi.org/10.1128/CDLI.7.2.155-160.2000

VanderWaal, K.L., Ezenwa, V.O., 2016. Heterogeneity in pathogen transmission: mechanisms and methodology. Functional Ecology 30, 1606–1622. https://doi.org/10.1111/1365-2435.12645

VanderWaal, K.L., Picasso, C., Enns, E.A., Craft, M.E., Alvarez, J., Fernandez, F., Gil, A., Perez, A., Wells, S., 2015. Network analysis of cattle movements in Uruguay: Quantifying heterogeneity for risk-based disease surveillance and control. Preventive Veterinary Medicine 123, 12–22. https://doi.org/10.1016/j.prevetmed.2015.12.003

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vernon, M.C., 2011. Demographics of cattle movements in the United Kingdom. BMC Veterinary Research 7, 31. https://doi.org/10.1186/1746-6148-7-31

Vernon, M.C., Keeling, M.J., 2012. Impact of regulatory perturbations to disease spread through cattle movements in Great Britain. Preventive Veterinary Medicine 105, 110–117. https://doi.org/10.1016/j.prevetmed.2011.12.016

Vernon, M.C., Keeling, M.J., 2009. Representing the UK's cattle herd as static and dynamic networks. Proceedings of the Royal Society B 276, 469–476. https://doi.org/10.1098/rspb.2008.1009

Vial, F., Miguel, E., Johnston, W.T., Mitchell, A., Donnelly, C.A., 2015. Bovine Tuberculosis Risk Factors for British Herds Before and After the 2001 Foot-and-Mouth Epidemic: What have we Learned from the TB99 and CCS2005 Studies? Transboundary and Emerging Diseases 62, 505–515. https://doi.org/10.1111/tbed.12184

Vidondo, B., Voelkl, B., 2018. Dynamic network measures reveal the impact of cattle markets and alpine summering on the risk of epidemic outbreaks in the Swiss cattle population. BMC Veterinary Research 14. https://doi.org/10.1186/s12917-018-1406-3

Vijayakumar, M., Park, J.H., Ki, K.S., Lim, D.H., Kim, S.B., Park, S.M., Jeong, H.Y., Park, B.Y., Kim, T. Il, 2017. The effect of lactation number, stage, length, and milking frequency on milk yield in Korean Holstein dairy cows using automatic milking system. Asian-Australasian Journal of Animal Sciences 30, 1093–1098. https://doi.org/10.5713/ajas.16.0882

Villarroel, A., Dargatz, D.A., Lane, V.M., McCluskey, B.J., Salman, M.D., 2007. Suggested outline of potential critical control points for biosecurity and biocontainment on large dairy farms. Journal of the American Veterinary Medical Association 230, 808. https://doi.org/10.2460/javma.235.8.937

Volkova, V. V, Howey, R., Savill, N.J., Woolhouse, M.E., 2010a. Potential for transmission of infections in networks of cattle farms. Epidemics 2, 116–122. https://doi.org/10.1016/j.epidem.2010.05.004

Volkova, V. V, Howey, R., Savill, N.J., Woolhouse, M.E.J., 2010b. Sheep movement networks and the transmission of infectious diseases. PloS ONE 5, e11185. https://doi.org/10.1371/journal.pone.0011185

von Keyserlingk, M.A.G., Olenick, D., Weary, D.M., 2008. Acute Behavioral Effects of Regrouping Dairy Cows. Journal of Dairy Science 91, 1011–1016. https://doi.org/10.3168/jds.2007-0532

Warren, K.S., Anderson, Roy M., Capasso, V., Cliff, A.D., Dietz, K., Fenner, F., Fiennes, R.N., Grossman, Z., Knolle, H., Mann, P.G., Molineaux, L., Schad, G., Schenzle, D., 1982. Transmission: patterns and dynamics of infectious diseases group report, in: Anderson, R.M., May, R.M. (Eds.), Population Biology of Infectious Diseases. pp. 67–85.

Waters, W.R., Buddle, B.M., Vordermeier, H.M., Gormley, E., Palmer, M. V., Thacker, T.C., Bannantine, J.P., Stabel, J.R., Linscott, R., Martel, E., Milian, F., Foshaug, W., Lawrence, J.C., 2011. Development and evaluation of an enzyme-linked immunosorbent assay for use in the detection of bovine tuberculosis in cattle. Clinical and Vaccine Immunology 18, 1882–1888. https://doi.org/10.1128/CVI.05343-11

Waters, W.R., Palmer, M. V., Buddle, B.M., Vordermeier, H.M., 2012. Bovine tuberculosis vaccine research: Historical perspectives and recent advances. Vaccine 30, 2611–2622. https://doi.org/10.1016/j.vaccine.2012.02.018

Webb, C.R., 2006. Investigating the potential spread of infectious diseases of sheep via agricultural shows in Great Britain. Epidemiology and Infection 134, 31–40. https://doi.org/10.1017/S095026880500467X

Weber, N., Bearhop, S., Dall, S.R.X., Delahay, R.J., McDonald, R.A., Carter, S.P., 2013a. Denning behaviour of the European badger (*Meles meles*) correlates with bovine tuberculosis infection status. Behavioural Ecology and Sociobiology 67, 471–479. https://doi.org/10.1007/s00265-012-1467-4

Weber, N., Carter, S.P., Dall, S.R.X., Delahay, R.J., McDonald, J.L., Bearhop, S., McDonald, R.A., 2013b. Badger social networks correlate with tuberculosis infection. Current Biology 23, R915–R916. https://doi.org/10.1016/j.cub.2013.09.011

Weycker, D., Edelsberg, J., Elizabeth Halloran, M., Longini, I.M., Nizam, A., Ciuryla, V., Oster, G., 2005. Population-wide benefits of routine vaccination of children against influenza. Vaccine 23, 1284–1293. https://doi.org/10.1016/j.vaccine.2004.08.044

White, P.C., Harris, S., 1995. Bovine tuberculosis in badger (*Meles meles*) populations in southwest England: An assessment of past, present and possible future control strategies using simulation modelling. Philosophical Transactions of the Royal Society B 349, 415–432. https://doi.org/10.1098/rstb.1995.0127

Wilkinson, D., Smith, G.C., Delahay, R.J., Rogers, L.M., Cheeseman, C.L., Clifton-Hadley, R.S., 2000. The effects of bovine tuberculosis (*Mycobacterium bovis*) on mortality in a badger (*Meles meles*) population in England. Journal of Zoology 250, 389–395. https://doi.org/10.1111/j.1469-7998.2000.tb00782.x

Williams, R.S., Hoy, W.A., 1930. The viability of B. tuberculosis (Bovinus) on pasture land, in stored faeces and in liquid manure. Journal of Hygiene 30, 413–419. https://doi.org/10.1017/S0022172400010561

Wilson-Aggarwal, J.K., Ozella, L., Tizzoni, M., Cattuto, C., Swan, G.J.F., Moundai, T., Silk, M.J., Zingeser, J.A., McDonald, R.A., 2019. High-resolution contact networks of free-ranging domestic dogs Canis familiaris and implications for transmission of infection. PLOS Neglected Tropical Diseases 13, e0007565. https://doi.org/10.1371/journal.pntd.0007565

Wilson, A.D.M., Krause, S., Dingemanse, N.J., Krause, J., 2013. Network position: A key component in the characterization of social personality types. Behavioral Ecology and Sociobiology 67, 163–173. https://doi.org/10.1007/s00265-012-1428-y

Wood, M.T., 1977. Social grooming patterns in two herds of monozygotic twin dairy cows. Animal Behaviour 25, 635–642. https://doi.org/10.1016/0003-3472(77)90114-2

Woodbine, K.A., Medley, G.F., Moore, S.J., Ramirez-Villaescusa, A.M., Mason, S., Green, L.E., 2009. A four year longitudinal sero-epidemiological study of bovine herpesvirus type-1 (BHV-1) in adult cattle in 107 unvaccinated herds in south west England. BMC Veterinary Research 5, 5. https://doi.org/10.1186/1746-6148-5-5

Woodroffe, R., Donnelly, C.A., Ham, C., Jackson, S.Y.B., Moyes, K., Chapman, K., Stratton, N.G., Cartwright, S.J., 2016. Badgers prefer cattle pasture but avoid cattle: implications for bovine tuberculosis control. Ecology Letters 19, 1201–1208. https://doi.org/10.1111/ele.12654

Woodroffe, R., Donnelly, C.A., Jenkins, H.E., Johnston, W.T., Cox, D.R., Bourne, F.J., Cheeseman, C.L., Delahay, R.J., Clifton-Hadley, R.S., Gettinby, G., Gilks, P., Hewinson, R.G., McInerney, J.P., Morrison, W.I., 2006. Culling and cattle controls influence tuberculosis risk for badgers. Proceedings of the National Academy of Sciences USA 103, 14713–14717. https://doi.org/10.1073/pnas.0606251103

Woods, A., 2011. A historical synopsis of farm animal disease and public policy in twentieth century Britain. Philosophical Transactions of the Royal Society B 366, 1943–1954. https://doi.org/10.1098/rstb.2010.0388

Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Liu, W.C., Mellor, D.J., Thomas, M.R., 2005. Epidemiological implications of the contact network structure for cattle farms and the 20-80 rule. Biology Letters 1, 350–352. https://doi.org/10.1098/rsbl.2005.0331

Zwald, A., Shaver, R., 2012. Effect of pen change on milk yield by dairy cows in 2 commercial herds. The Professional Animal Scientist 28, 569–572.