Infectious disease transmission and contact networks in wildlife and livestock

Meggan E. Craft

Department of Veterinary Population Medicine, University of Minnesota, St Paul, MN 55108, USA
MC, 0000-0001-5333-8513

The use of social and contact networks to answer basic and applied questions about infectious disease transmission in wildlife and livestock is receiving increased attention. Through social network analysis, we understand that wild animal and livestock populations, including farmed fish and poultry, often have a heterogeneous contact structure owing to social structure or trade networks. Network modelling is a flexible tool used to capture the heterogeneous contacts of a population in order to test hypotheses about the mechanisms of disease transmission, simulate and predict disease spread, and test disease control strategies. This review highlights how to use animal contact data, including social networks, for network modelling, and emphasizes that researchers should have a pathogen of interest in mind before collecting or using contact data. This paper describes the rising popularity of network approaches for understanding transmission dynamics in wild animal and livestock populations; discusses the common mismatch between contact networks as measured in animal behaviour and relevant parasites to match those networks; and highlights knowledge gaps in how to collect and analyse contact data. Opportunities for the future include increased attention to experiments, pathogen genetic markers and novel computational tools.

1. Introduction

Although there was optimism that the war on infectious diseases was over in the 1970s [1], humans are still faced with challenges of pandemics, as well as emerging and re-emerging diseases, many of which originate in animals [2]. Animal pathogens not only cause problems for humans (zoonoses), but infectious diseases also negatively affect the health of animals, including livestock and wild animals. Both livestock and wildlife can be infected by pathogens that infect multiple hosts; multihost pathogens have the potential to cause devastating effects on the livelihoods of livestock farmers, national economies and the environment via wildlife die-offs. Concern over multihost pathogens is increasing with compounding effects of land development, globalization and climate change. Infectious diseases in livestock can be exacerbated by intensive agriculture, as more individuals are concentrated in smaller spaces [3]. Additional infectious disease threats exist for wildlife, especially for small populations of conservation concern, which may be stressed owing to loss of habitat and are vulnerable to pathogens that ‘spill over’ from other more numerous hosts [4]. Larger human populations and more demand for space have led to increased contact between humans, wildlife and domestic animals, which creates more opportunities for pathogen transmission. The war on infectious diseases is clearly not over.

To understand how disease spreads through a population, we need to understand $\beta$, the transmission rate of infection. However, it is challenging to estimate $\beta$. Thus, it is commonly assumed that

$$\beta = \gamma + K,$$

where $\beta$ is the transmission rate of infection, $\gamma$ is the probability of pathogen transmission given a contact and $K$ is the contact rate. From the perspective of the population, this formula can be thought of as a disease transmission matrix between pairs of individuals ($\beta$; ‘who transmits to whom’), represented...
often represented as a contact network (\( \beta \)) scaled by a probability (\( \gamma \)). Black circles could represent individual animals, or groups of animals such as farms. Connections between circles represent pathogen transmission events in the transmission network (a) or contacts in the contact network (b). Figure adapted from the Serengeti lion networks [5].

Network modelling, or contact network epidemiology, gained popularity for studying sexually transmitted disease transmission in human populations [19] and later through the discovery of superspreaders during the Severe Acute Respiratory Syndrome (SARS) outbreak in 2003 [20,21]. Network modelling approaches can easily incorporate a heterogeneous contact structure and are most useful when each individual is in contact with a small fraction of the population [22]. Network models are especially useful for designing targeted interventions, either for particular types of members of a population or for specific individuals. Network models could be considered the most general modelling tool, where specific types of networks include compartmental SIR models (where all individuals are connected to each other) or lattice-based models (where individuals are connected only to their neighbours) [6]. Alternatively, network models could be considered an extension of an SIR model that incorporates more realistic contact structures, although network models are not as data hungry or computationally intensive as agent-based or individual-based spatial models [22]. The choice of model type is considered ‘a skill in itself’ and is dependent on the question asked, the quality of empirical data and the needed output [22]. Choosing the correct type of model is important; for example, compartmental SIR models can overestimate the speed of an outbreak if heterogeneous contacts are present but not captured in the model [23,24]. Although network modelling has been proved to be a useful tool to study transmission of infectious diseases in human populations [7,25–27], it has generally been an under-used approach for investigating disease dynamics in wildlife [6] and livestock populations [28,29].

This review article will focus on the use of contact networks to answer questions about infectious disease transmission in wildlife and livestock. I will describe how to combine animal contact data with network modelling, often with a cautionary tale. This review will: (i) highlight why wildlife and livestock often have a heterogeneous contact structure; (ii) illustrate the rising popularity of network approaches for understanding pathogen transmission in animal populations; (iii) discuss the common mismatch between contact networks measured in animal behaviour and pathogens that match the contact networks; (iv) highlight knowledge gaps; and (v) propose opportunities and ways forward for the field.

2. Heterogeneous contact patterns in wildlife and livestock

Infectious diseases spread through transmission routes that vary in contact intensity along a spectrum from indirect to direct transmission, where contact might be thought to increase in importance for transmission as one moves from environmental contamination to vector-borne, and then to aerosol, to close-contact, and finally to sexually transmitted pathogens [30]. Accordingly, the definition of what constitutes a contact relevant for disease transmission varies according to the pathogen of interest. For example, anthrax is spread via ingestion of spores from endemic environments, rabies is transmitted through the bite of a rabid animal, and sexually transmitted infections are spread via sexual interactions. The
extent of pathogen spread on a contact network is not only dependent on how a contact is defined, but is also dependent on both population and individual-level phenomena. Often there is variation, or heterogeneity, in the probability of contacts between individuals, leading to different macroscale network architecture between populations. Finally, it is worth noting that many pathogens have multiple transmission pathways; some sexually transmitted pathogens, for example, also spread vertically from mother to offspring.

At the population level, pathogens are expected to transmit faster and more extensively in populations that have less social or spatial structure (also known as modularity [15,31]), more variation in group size [32] and more movements or more connections between groups [33]. Wild animals often have a heterogeneous contact structure owing to their territorial nature, spatial distribution of individuals across heterogeneous landscapes or complex social systems [5]. Wild animal social systems can also vary between populations in different ecosystems owing to varying resource availability. For example, jackals live in small family groups unless they have access to ample food sources, in which case they live in larger communities [34,35].

For livestock, including poultry and managed fish populations, contact structure is also heterogeneous, but instead of being tied to social systems, contacts are typically dependent on group size, spatial structure and animal movements, which are frequently controlled by human management actions [28,36,37]. For example, swine in typical commercial systems in the US are grouped into rooms in buildings on farms, and there are scheduled movements of swine between rooms [38] and farms. Livestock movement networks (e.g. cattle, swine, poultry, farmed fish) are often directional and can be weighted by the number of animals moved, or the frequency of movements. Of course within a spatially restrained group, as defined by a farm, pen, pond or building, livestock could also have preferential associations owing to social structure [39,40]. Therefore, animal behaviour might be more important in driving within-group contacts than between-group contacts in livestock.

On an individual level, there is variation in the probability that an individual is exposed, gets infected and subsequently transmits a pathogen. Superspreaders are an extreme version of this variation. The concept of a superspreader has been expanded to include supershedders [41,42], and could be expanded further to include ‘supersusceptibles’ that easily get infected or ‘supermovers’ that might connect distant parts of a network. The cause of variation in these ‘super’ categories could be owing to a range of factors associated with the individual host, including host genotype, prior exposure to the pathogen (leading to an ability to rapidly produce antibodies), demographic class, behaviour and/or coinfection status. Demographics include age, sex and body size, which could also correlate with behaviour. Individual behaviour can vary along a continuum of behavioural types (personalities), such as shy versus bold (e.g. lions [43]), levels of aggression or activity, or dominance in group-living wildlife and livestock [44]. Individual behavioural types also affect interspecific contacts, as seen in badger–cattle interactions where higher-ranked cattle were more likely to have contact with badgers [45]. Major histocompatibility complex genotype, whether or not individuals are infected with other pathogens (i.e. co-infection), or their stress level can also determine how susceptible the individual is to infection, and how infectious the individual might become [46]. Heterogeneity in pathogen transmission at both the population and individual levels can create challenges for understanding the complex mechanisms influencing the spread of infectious diseases in populations, yet network modelling can be a useful tool to incorporate this complexity.

### 3. Increasing interest in social network analysis and network modelling

The terminology surrounding ‘networks’ and its associated uses can be confusing because the word ‘network’ is used in both social network analysis (SNA) and network modelling. However, the two concepts are quite distinct. SNA is used to describe the contact structure of a population, whereas network modelling is a flexible tool by which to introduce and simulate a contagious process on a descriptive social, association or contact network [21,47]. SNA, also known as network analysis or contact analysis, has recently gained popularity in the field of animal behaviour to quantify animal social networks and to compare social networks within and between species [9,10,48]. SNA has also been used to investigate potential pathogen spread in wildlife. For example, SNA has been used to determine which individuals are at high risk of infection [49,50], identify traits of highly connected individuals [51], inform disease management policy [52] and correlate highly connected individuals with high rates of parasite infection [53].

SNA can also describe the heterogeneous network of livestock or poultry movements, often with the intention to determine the risk of disease spread in populations, identify high-risk herds or target surveillance [54–61]. In many cases, SNA has identified that a small percentage of farms contribute to the majority of livestock and poultry movements, and hence risk of disease transfer; these particular farms are the ones to target for increased surveillance or interventions [62–64]. Alternatively, SNA has shown that livestock movements are not a risk-factor for infectious diseases [65]. For fish, SNA has been used to identify aquaculture sites in a farmed-fish network that should be targeted for parasite control [66] and to explain virus infection pressure caused by water transport [67]. SNA typically represents interactions on static networks. However, the social network could be subdivided to represent contacts or movements occurring during different time periods: for example, by constructing different social networks in the breeding and non-breeding seasons [68], or between years with variable rainfall [69] or changes in the poultry and livestock trade network across seasons [70,71] and across years [58]. Descriptive network measures can also be used to compare competing models, such as for describing livestock movements dependent and/or independent of distance between farms [72].

Network modelling can build on the same data as SNA. The structure of the contact network influences how pathogens transmit on the model network [33,73]. For example, even with the same number of individuals and same number of effective contacts, when the configuration of the network is changed, it can affect both the rate at which individuals become infected and the final size of a disease outbreak [6]. Thus, it is important to get the most accurate contact structure possible, yet this must be weighed against the costs: collecting contacts can be time- and data-intensive. High-resolution data on contacts can be obtained from behavioural studies and new technology [6], including livestock tracking and reporting systems (e.g. active
and passive radio-frequency identification (RFID) tags). Network modelling is now more feasible for animal systems owing to an ability to collect high-resolution contact data and to integrate the contact data into new theoretical approaches.

Network models can be used to understand factors behind disease outbreaks, design effective surveillance and evaluate the consequences of disease outbreaks [74]. Contact network models have been used to predict the spread of exotic diseases between river catchment areas owing to anthropogenic movements of fish [75] and to recommend that the fish industry be designed around epidemiologically isolated management areas [37]. Network modelling has been used in wildlife systems to: identify the most connected individuals to target for vaccination [76]; illustrate that current vaccination levels may be inadequate to prevent large outbreaks [77]; identify the most important behaviour for disease transmission [78]; and incorporate seasonal-mixing preferences in models [79]. For livestock, network modelling and simulations have highlighted that disease transmission through livestock movements could explain the initial spread of the 2001 foot-and-mouth epidemic in the UK [80] and that network properties can direct targeted surveillance [28].

A combination of increasing exposure to network concepts and the wider application of novel technology has led to an increase in the use of SNA and network modelling for wildlife and livestock. According to a search run in the abstract and citation database of peer-reviewed literature, Scopus, with the search terms ‘contact network’ or ‘social network’ or ‘social-network’ or ‘association network’ or ‘contact analysis’ and ‘disease’ or ‘pathogen’ or ‘parasite’ in the title, abstract or keywords. The ‘wildlife’ category refers to publications with the search terms ‘wildlife’ or ‘wild animal’ or ‘social insect’, whereas ‘livestock’ refers to publications with the additional criteria of ‘poultry’ or ‘cattle’ or ‘cows’ or ‘swine’ or ‘pigs’ or ‘turkeys’ or ‘fishes’ or ‘livestock’ and not ‘Turkey (republic)’ and not ‘humans’. Search was run on 12 June 2014. This is a underrepresentation of publications, as many wild animals, such as mice, raccoons, apes and chimpanzees, did not show up in the ‘wildlife’ search.

Figure 2. The number of publications in Scopus per year containing the search terms ‘contact network’ or ‘social network’ or ‘social-network’ or ‘association network’ or ‘contact analysis’ and ‘disease’ or ‘pathogen’ or ‘parasite’ in the title, abstract or keywords. The ‘wildlife’ category refers to publications with the search terms ‘wildlife’ or ‘wild animal’ or ‘social insect’, whereas ‘livestock’ refers to publications with the additional criteria of ‘poultry’ or ‘cattle’ or ‘cows’ or ‘swine’ or ‘pigs’ or ‘turkeys’ or ‘fishes’ or ‘livestock’ and not ‘Turkey (republic)’ and not ‘humans’. Search was run on 12 June 2014. This is an underrepresentation of publications, as many wild animals, such as mice, raccoons, apes and chimpanzees, did not show up in the ‘wildlife’ search.

4. Challenges of network modelling

(a) Mismatch between contact network and pathogen

Two types of missing data are common for network modelling: a lack of knowledge about contact patterns or a lack of information on infection patterns. Here, each of these is discussed in turn. The type and main cause of missing data might differ when considering wildlife or livestock.

(i) Lack of knowledge about contact patterns

Efforts to construct realistic network models can be hampered by a lack of detailed contact data. Although disease outbreaks in wildlife populations are notoriously difficult to observe and measure, in some rare situations outbreak data can be collected. This can occur when virulent pathogens cause high mortality in an easy-to-detect species (e.g. large marine mammals washing up on beaches or massive numbers of dead fish in lakes), or in well-monitored populations where individually identified animals disappear or carcasses are found [81,82]. Prevalence and incidence data are sometimes available for pathogens that cause clinical signs, such as tuberculosis in meerkats [83], mycoplasmal conjunctivitis in house finches [84] or rabies in carnivores [85]. In the uncommon situations where wildlife disease outbreaks have been observed, detailed contact data are often lacking, even to the point of not knowing the host abundance or social system. Without knowing the empirical social contacts, the wide choice of potential network structures makes the design of network models challenging. Thus, without contact patterns we are not able to tease apart competing hypotheses about the mechanisms of disease transmission for most systems, and as a consequence, are not able to implement targeted intervention and control measures.
One would think it would be easier to study and understand pathogen spread in livestock systems than in wildlife. Livestock farmers should know the numbers, movements, locations, birth rates, mixing patterns, etc., of each of their animals. However, there are different challenges in applying network modelling to livestock systems from those faced by wildlife managers. For livestock and managed fish populations, disease outbreaks can be observed, but detailed data on contacts are often missing. Some differences might have to do with the end result, as producers may not be interested in understanding basic ‘disease dynamics’ within a farm because depopulation and vaccination are often solutions that can locally eliminate a pathogen. On a regional level, challenges can revolve around data sharing and privacy issues. For example, in the USA, the basic fundamentals about the structure of regional farm and livestock movement networks are not known, including missing data about farm size, location and animal movements [86]; this is also true for fish industries [87].

(ii) Lack of information on infection patterns

The collection of contact data is getting easier. New remote bio-monitoring technologies, which include radiotelemetry, RFID, global positioning system tracking, tri-axial accelerometry and video tracking, are improving our ability to collect extensive contact data on study populations [6,88,89]. However, in the absence of an observable outbreak as discussed above, it remains a problem to match this contact data with relevant and appropriate parasites. For example, for wildlife species that are abundant, relatively easy to catch and handle, and where detailed data on contacts can be obtained using technology [90,91], often the circulating pathogens do not cause morbidity or mortality to the host. Because these pathogens do not necessarily cause disease or death, there might be a perceived need to simulate highly lethal pathogens on the contact network. If this is done, the choice of a pathogen to simulate should be made with caution, as discussed below.

For observational studies of wildlife populations with already-existing detailed contact data (e.g. [92,93]), there is a great opportunity to use network modelling owing to the detailed understanding of animal behaviour and the high resolution of observational data. However, behavioural researchers often lack data on pathogens with transmission modes and life histories that are relevant to the available contact data. There are at least two main reasons for this problem. First, detailed behavioural data are often collected on a single species, often in small populations of conservation concern (but see [94,95]). This is a problem because pathogens often infect multiple hosts, and researchers often lack information on contact patterns between various host species [96].

Second, researchers often lack data on host infection status in wildlife populations. This situation is especially common when infected individuals cannot be identified without capturing the hosts, which can be logistically challenging, unethical or even impossible. Thus, studies that use behavioural data to construct a contact network often do not have infection data; although this can be overcome with simulation of pathogen spread through network modelling. However, we must be careful to simulate pathogen spread on appropriately matched contact networks. For example, we should simulate infectious diseases with short infectious periods (e.g. weeks) on networks where the contacts used to create the network were aggregated over weeks (versus years). Long-term aggregation of the contacts would likely create a network that is substantially more connected than the network that the pathogen actually experiences [97]. Caution should be used whenever social networks are used to infer disease transmission, as the method of data collection also affects the construction of the social network [90]. If infectious diseases are going to be simulated on social networks, edge construction needs to incorporate parasite life cycles and reflect the type of exposure needed for pathogen transmission [98].

Although the health status of livestock is more easily collected compared to wildlife, infection status of individuals or of herds is not typically shared on a regional level owing to privacy issues, even if it is known. Unless animal movements and animal testing are mandated and regulated, regional studies of disease spread in agricultural settings are challenging. In general, it might be easier to overcome the lack of either contact or pathogen data in agricultural systems than in wildlife systems. In summary, there is the potential to have a mismatch between a contact network and a pathogen for both wildlife and livestock; missing data are often the root cause of the mismatch. Often, we have either observed an outbreak, but do not know the contact structure, or know the contact structure but lack an appropriate pathogen. In order to combine animal contact data with network modelling, and to avoid a contact network–pathogen mismatch, a simple thought process could be implemented (box 1).

(b) Questions to consider in the design of network models

Numerous challenges exist when attempting to use data on host contacts to construct infectious disease models, or when trying to infer contact structure from outbreak data. Below are a few questions to consider in the design of network modelling studies, some of which are unresolved.

(i) How to define a ‘contact’?

Defining a contact is challenging (box 1, step (iii), [99]). We often do not know how a pathogen is transmitted between hosts and therefore cannot define a contact sufficient for disease transmission [51,100]. For directly transmitted pathogens, ideally a contact would be defined as a susceptible and infectious host being within a specified distance for a specified length of time; however, the distance and time period are normally unknown for the host species and pathogen strain of interest. The definition of a contact is even more challenging for indirectly transmitted pathogens owing to environmental persistence (e.g. macroparasites spread by faecal–oral transmission, anthrax spread by infection of spores, fungi infecting bees or bats) or the importance of vectors (e.g. ticks and mosquitoes). Because of these challenges and with the exception of a few studies on indirectly transmitted pathogens, network modelling has mostly been used to model directly transmitted diseases [101].

(ii) How does infection with a pathogen change the contact structure of a population?

Individual behaviour, and hence the contact network, can change once an individual is infected [102]; it is rare that the change in behaviour is quantified. Healthy animals could avoid or potentially be attracted to sick animals, and sick
animal behaviour can be ‘attracted to’ sick animals; lions generally separate themselves from their pride mates when they are sick, but, in the case of canine distemper virus infection, lions that had convulsions in the presence of their companions were sometimes attacked (Craig Packer 2014, personal communication). As another example of behaviour change, animals infected with rabies can either experience ‘furious’ rabies and likely have more contacts, while others exhibit ‘dumb’ rabies and have fewer contacts [77]. Similarly, toxoplasmosis can cause behavioural changes in the intermediate host (e.g. rat) to enhance predation by the definitive host (e.g. cat) to complete the parasite’s life cycle [104,105]. In an experimental setting, the grooming behaviour of ants changes upon fungal exposure; exposed ants increase individual self-grooming rates and reduce allogrooming of conspecifics [106], with consequent changes on the structure of the ant contact networks.

Ideally one would have empirical data describing how a network changes once infection occurs and spreads (e.g. [107]), although simulating a change in behaviour can be a useful tool to work around the lack of data on behavioural changes caused by infection. Data collection for step (vi) of box 1 could occur in a population preferably before, during, and after an outbreak, so any changes in behaviour could be documented on a population level. If empirical data are available on behavioural changes, dynamic networks are tools to incorporate temporal changes in contact network structure.

### Box 1. Implementation: how to combine animal contact data with network modelling.

Whether the contact data come from new technology or from observations, the following steps will be useful in combining relevant contact data with a pathogen:

(i) Define pathogen of interest.
(ii) Identify the life cycle of the pathogen, including host range (e.g. are there multiple hosts or vectors or environmental stages?)
(iii) Identify type of contact needed for pathogen transmission. Evaluate whether this type of transmission can be captured by measurements of social contact.
(iv) Identify the infectious period for the pathogen.
(v) Determine the definition of a node based on the minimum level at which it is important and feasible to collect data on heterogeneous contacts (e.g. will the nodes be individuals, social units, rooms in farms, farm premises or habitat patches?)
(vi) Collect the number of contacts between nodes on a time scale relevant to (likely equivalent to or less than) the infectious period of the pathogen. These will be the edges.
(vii) If outbreak data exist, estimate the transmission rate of the pathogen. The network model can then be used for hypothesis testing or ‘what if’ scenarios.
(viii) If no outbreak data exist, use epidemiological parameters obtained from the literature to simulate outbreaks from a range of parameters and test hypotheses.
(ix) Test control strategies on the network.

New technologies, such as proximity contact loggers, have the potential to capture the contacts of an entire population, but they are expensive. Behavioural observations are also expensive in terms of person-hours; even then, not all animals are observed simultaneously. Therefore, studies of animal contacts typically focus on a subset of a larger population. Because we are interested in disease dynamics in a biologically meaningful geographical area (such as a protected area or a region of high livestock production), we need additional tools to enable us to ‘scale up’ to populations of relevant sizes. Methods for scaling up empirical networks are being developed, but on a case-by-case basis [68,77,106,111].

### (v) How to capture contacts on multiple host species?

Pathogens that infect multiple hosts can be challenging to model for both well mixed or network models, as normally we do not have within- and between-species contact data for multiple hosts. Making behavioural observations for individually identified multiple host species is challenging. Novel technology, such as automated proximity loggers or even camera traps, seems to be a promising route to better document

transmission of foot-and-mouth disease (short infectious period) or bovine tuberculosis (long infectious period) in cattle herds, do we need to include contact heterogeneity between individuals within a farm, or just movements between farms, sub-regions or regions? While the framework for modelling heterogeneity both within and between groups is available [32], it is unclear at which level capturing within-group heterogeneity is necessary. The scale likely depends on an interaction between host contacts and infectious period [97], and the relative importance of direct versus indirect transmission [109]. In a study on US cattle movements, heterogeneities in cattle shipments only emerged at the county level, not at the state level; thus, analysis of county-level networks was sufficient to understand movements in this system [110].

### (iv) How to scale up the observed network to a biologically relevant size?

While it is unclear at what scale interactions should be studied [108], the level of resolution will likely depend on the host, pathogen and system of interest, with consideration of what is computationally feasible [22,99]. For example, to model
contacts [45,112]. However, automated proximity loggers and camera traps are not feasible when we lack information on which hosts are involved and where interspecific contacts might take place, therefore informing where to place the loggers and traps. Likely owing to the inability to understand the contact structure of multiple populations simultaneously, studies of the dynamics of zoonoses [113] or multihost pathogens at the livestock–wildlife interface are rare [114].

5. Opportunities for understanding disease transmission using contact networks

(a) Experiments and manipulations

A promising area for further research includes increased attention to experiments and manipulations. Both hosts and pathogens can be experimentally manipulated through removals or additions, and consequent changes on the contact network can be documented [107]. Abundant and small wild animals are useful study systems for manipulations of free-ranging animals in natural systems. Experimental manipulations can be done on these common species to see how varying the host–pathogen environment influences different network measures and pathogen spread [107,115,116]. In addition, studies of common pathogens in common species are highly useful for two reasons. For applied research questions, pathogens found in abundant wildlife species are often important for their role in ‘spilling over’ into humans, agricultural animals or wildlife of conservation concern [117]. Second, for basic research questions, network studies on common wild animal species are extremely useful and important for testing mechanistic principles of transmission on networks.

Livestock systems are a great system to test the theory of infectious disease spread [118,119], because producers actively manage host demographics and contact rates, and can be willing to test new methods for disease control, including targeted control measures. In addition, pathogen surveillance and detection in managed populations are more feasible since livestock are often more easily handled than wildlife.

(b) Pathogen genetic markers

In order to overcome some of the challenges in measuring contacts between members of the host population, parasite genetic markers, microbial genetics, whole-genome sequencing, pathogen phylogenetics and phylodynamics are increasingly used to infer pathogen transmission networks [120,121]. Instead of focusing on the host, these methods focus on the pathogen to infer transmission networks. Ideally, genetic studies should be conducted in parallel with traditional detailed contact studies in order to define which interactions are important for transmission [99].

It is important to find the correct pathogen to use as a genetic marker. Pathogens that are spread via direct contact with conspecifics and that experience rapid mutation rates relative to the host genome are useful for evaluating transmission of directly transmitted pathogens in populations [121]. RNA viruses are particularly well suited as markers of host contact, because their fast mutation rate and short generation time result in unique genetic signatures that can be compared and related across individuals [122–128]. However, isolating virus typically requires invasive sampling to collect blood, tissue or saliva.

Non-invasive sampling (e.g. collection of faeces, hair) is a preferred method over invasive sampling; non-invasive sampling is especially appealing for collecting pathogens from wild animals. Bacteria can be collected from faeces and can be used as indicator organisms to infer transmission or contact networks. Two bacteria found in faeces, *Escherichia coli* and *Salmonella enterica*, have been used to deduce if contact networks (based on association networks) correspond to the transmission networks of *E. coli* and *Salmonella* (as inferred by similarity in genetic subtypes). Bacterial transmission networks have been correlated with social association patterns in wild lizards and giraffe populations [129,130], but not in elephants [131], potentially owing to extensive space use of elephants. Pathogen transmission networks can also be used to shed light on contact patterns between multiple host species. For example, *E. coli* infects a wide range of hosts and has been used to infer superspreading species and deduce multihost association patterns and habitat overlap between humans, livestock and wild animals [132,133].

Another non-invasive proxy of transmission that could be used to look at social contacts is the similarity in the gut microbial community. 16S rRNA marker-based studies provide a new opportunity to obtain and analyse high-depth representations of microbial communities at a low cost. Profiling of an entire microbial population has been used extensively over the past few years in the study of human conditions, revealing links between the microbiome, nutrition and diseases [134,135]. It has also been used in livestock to study the effects of management changes on the gastrointestinal microbiome [136–138]. The power of this approach over the use of indicator organisms such as *E. coli* is the high level of discriminatory power, since it relies on entire microbial populations instead of the genetic background of a single organism. Social relationships likely influence the composition of the gut microbiome [93,139–141], and similarity in the gut microbiome could be used to infer social contacts to simulate the spread of pathogens on network models, although shared habitat and behaviours must be controlled for. It should be noted that host-associated microbial communities are often linked with health benefits [142], rather than health costs. So instead of using microbial communities as proxies to understand disease dynamics, contact networks could be used to model epidemics of beneficial organisms (i.e. the spread of ‘healthy’) on social networks. This avenue of future research could shed light on the intersection between sociality and health, both in terms of health costs and benefits.

(c) Computational tools and methods

Contact network studies began with what were fairly simple descriptive tools (e.g. SNA) and have now advanced to more sophisticated modelling techniques. New computational tools and methods for understanding disease transmission continue to evolve. These include Bayesian inference for partial contact tracing [143], pathogen whole-genome sequencing [144], and new statistical frameworks to bridge epidemiological models and coalescent-based inference methods [145]. There is a push for further development of robust analytical methods for network epidemic models [146], and for methods to make efficient phylogenetic inferences on the large amounts of sequence data currently being produced [147].

Because of the improvement in technology to collect detailed temporal data on contacts, empirical data to parametrize contact
networks now exist. However, large datasets can create bioinformatic challenges. While new technologies, such as proximity loggers or RFID tags, improve our ability to collect very detailed data on contact structures in host populations, analysing these data presents a computational challenge, especially in a dynamic network framework where the network structure is changing. Currently, the methods for modelling dynamic networks are being developed [8,148,149] and will need to continue to develop to keep up with the growing size and temporal resolution of the contact data.

6. Conclusion

The incorporation of contact networks into the study of infectious diseases has advanced our understanding of pathogen transmission and dynamics, with applications for wild animals and livestock. New network modelling techniques build on the foundations of SNA and traditional SIR modelling, often blending the two methods. Network modelling approaches are well suited to capture heterogeneous contact patterns within and between animal groups, to simulate disease spread and to test hypotheses across a wide range of wild and domestic animal social systems. Network models can use detailed data on behavioural contacts, collected from either observational behavioural studies or from new technology. It is imperative, however, that the type of contact matches the transmission route of the pathogen of interest.

In addition to furthering our understanding of basic research on pathogen dynamics, contact networks can produce useful applied outcomes for animal populations. Network methods are ideally suited to help design targeted surveillance and control programmes; targeted measures can be more efficient and effective than random surveillance and control. Targeted control measures for individuals, or groups of individuals, can include proactive or reactive vaccination, culling or depopulating and/or quarantine of selected individuals. Targeted surveillance and control can also focus on distinct time periods, as elucidated by changes in contact patterns between different periods of time. Because of these varied benefits and despite inconsistencies in ‘network’ terminology—the use of SNA and network modelling is increasing.

However, some critical gaps remain in our knowledge of contact networks. The field is wrestling with how best to define a contact, capture contacts between multiple host species and scale up observed networks to biologically relevant sizes. In addition, we do not know how infection with a pathogen changes individual behaviour (and consequently the contact structure at the population level) or the minimum level of detail needed to accurately capture important features of contact structure.

Some of these challenges can be solved using new approaches. Great potential exists for increased use of experiments in either common wild animal populations or livestock systems. There is excitement about using new pathogen genetic markers from viruses, bacteria and the gut microbial community to infer transmission networks. The transmission networks could be used to infer contact networks, which could be used to simulate pathogen spread with similar modes of transmission. Finally, new computational tools are being developed to keep up with the influx of large contact datasets, and new modelling methods are being designed to incorporate the importance of including temporal changes on the contact network into dynamic network models. Many of these novel methods are being developed on an individualized case-by-case basis; hence no off-the-shelf and user-friendly software currently exists for these needs.

While the progression towards using new genetic markers and developing computational techniques for contact networks can provide powerful methods for understanding disease dynamics in animal populations, the increasing complexity of methods could present an obstacle in terms of accessibility. Behavioural ecologists, producers or veterinarians would likely find some of the cutting-edge contact network modelling tools too mathematical and theoretical. However, access to many of these new methods is possible through collaborations with physicists, bioinformaticians or molecular analysts who are developing and using the new methods. Collaboration between disease modellers and scientists who know the data and the field system well can shed light on animal epidemiology [150,151] and can make headway on discerning both theoretical and applied questions. Applied questions could become more relevant if network models were constructed with an appropriate policy analysis as the output (e.g. evaluate vaccination versus do nothing, and the cost/benefit of each approach), as is being done in the field of public health. In addition, collaboration will likely expand the utility of contact network models in shedding light on the linkages between sociality (e.g. contact networks) and health outcomes (e.g. risk of infectious pathogen exposure or acquisition of beneficial bacteria). Contact network models could be used to compare the effects of increased sociality on health at a population level, and the role of social individuals on an individual level, including the role of social factors in infectious disease spread.

Acknowledgements. Thanks to P. Kappeler, C. Nunn, E. Enos, A. Kissling, K. VanderWaal, L. White, two anonymous reviewers and University of Minnesota’s Veterinary Medical Library staff for invaluable input and assistance.

Funding statement. National Science Foundation (DEB-1413925), the University of Minnesota’s Institute on the Environment, Agricultural Experiment Station General Agricultural Research Funds, and College of Veterinary Medicine (Animal Health Formula Funds and Population Systems Signature Program) provided funding. This material is based upon work supported by the Cooperative State Research Service, US Department of Agriculture, under Project nos. MINV 62-044 and 62–051. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the view of the US Department of Agriculture.

Conflict of interests. I have no competing interests.

References

4. Smith K, Acevedo-Whitehouse K, Pedersen A. 2009 The role of infectious diseases in biological...


