Chaters, Gemma (2021) *An evaluation of the influence of livestock movements on the transmission, spread and persistence of infectious diseases in northern Tanzania.* PhD thesis.

http://theses.gla.ac.uk/82428/

An evaluation of the influence of livestock movements on the transmission, spread and persistence of infectious diseases in northern Tanzania

Gemma Chaters BVMS MSc MRCVS

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

University of Glasgow
Institute of Biodiversity, Animal Health and Comparative Medicine
College of Medical, Veterinary and Life Sciences

June 2021

Abstract

In northern Tanzania livestock are heavily relied upon as a main source of income, social status and nutritional security, especially by those living in the most impoverished communities (Covarrubias *et al.*, 2012; Government of Tanzania, 2017). The high dependence on livestock is accompanied by a high burden of infectious production-limiting and zoonotic pathogens circulating within the livestock population but poor access to veterinary services. Zoonotic pathogens can spill over to cause disease in people, which are often misdiagnosed and erroneously treated leading to worse patient outcomes (Crump *et al.*, 2013; Zhang *et al.*, 2016). For pathogens that cause disease in livestock alone, the economic returns from investing in disease control can far outweigh the costs (Jones *et al.*, 2016). Improved livestock health and productivity is widely recognised as a route out of and away from poverty for people living in the most marginalised communities (Randolph *et al.*, 2007). Funding and resources to invest in the livestock sector and livestock disease control are often lacking as the broad benefits to individuals, societies and economies are poorly documented and often overlooked (Rich and Perry, 2011; World Organisation for Animal Health (OIE), 2013). Policy makers need clear guidelines to develop efficient livestock disease control programmes that reduce livestock and zoonotic pathogen burden through minimal use of resources for maximal societal gains (FAO, 2009; Dutilly *et al.*, 2020).

Livestock movements are widely recognised as providing a route to move pathogens between populations (Fèvre *et al.*, 2006). These movements can drive large epidemic outbreaks of disease and also provide opportunities for pathogens with relatively low R_0 to persist in populations (Green, Kiss and Kao, 2006a; Prentice *et al.*, 2017). Where dense comprehensive data on livestock movements is available, this can be used by policy makers to guide effective disease control programmes (Kao *et al.*, 2006). However, information on livestock movements is not routinely collected and centrally recorded in Tanzania and is therefore not available to guide livestock disease control programmes.

Through this PhD, I gather and analyse data on three major types of permanent livestock movements in cattle and small ruminants across and beyond the Arusha, Manyara and Kilimanjaro regions in northern Tanzania. Permanent movements are those into or out of household herds and flocks with no plan to return the animal(s) to their origin. Using

household survey data in conjunction with livestock serological data, market survey data and government movement permit data, I evaluate how livestock movements contribute to epidemiological connectivity and disease risk. Movements to and from households, including market movements go largely unreported in the study area but can cover long distances up to 300 km in a single movement. I use the data to construct networks of livestock movements and use concepts from network analysis to identify sub-village and ward locations that can be targeted with efficient disease control and surveillance interventions.

My analysis shows that high risk locations for disease introduction are also those at high risk of onward transmission, and that locations at high risk for small ruminant pathogen transmission are also high risk for cattle pathogen transmission. Additionally, I show that locations at risk of introduction and onward transmission of less transmissible pathogens (e.g. *Brucella* spp.) are also high risk for epidemic-prone pathogens (e.g. Foot and Mouth disease virus) that are rapidly transmitted. The positive correlations identified between locations' risk ranks show that multi-species interventions which aim to prevent introduction and onward transmission of multiple pathogens could be an efficient use of disease control resources in northern Tanzania. Specifically, I show that household cattle introductions and sub-village betweenness are positively associated with cattle's risk of bovine viral diarrhoea virus (BVDV) exposure and this risk is ubiquitous across the study area. For other pathogens investigated, the associations between introductions and exposure are complex and vary depending on pathogen and the agro-ecological (A-E) class of the livestock keeping system. This complexity is likely driven by the varying transmission routes and shedding cycles of different pathogens, in combination with the different livestock mixing and contact rates between infectious and susceptible individuals in the different A-E systems. Nevertheless, risk factors relating to household and sub-village livestock introductions are identified and can be used to guide disease control interventions in different settings.

I also identify that livestock market movements are most often made on foot with increased risk of these livestock contacting local non-moving livestock and creating opportunities for pathogen dissemination across the landscape. Frequently travelled routes to and from market are therefore also identified as areas where an increase in active surveillance would benefit both local livestock and animals moving into the market system.

Findings from this work will be useful for policy makers in northern Tanzania who have minimal resources available to reduce livestock and zoonotic pathogen burden. High risk locations identified in this analysis can be made targets for knowledge exchange and information dissemination, active surveillance and multi-pathogen vaccination programmes. Additionally, the results from this study can be used to guide future research questions which address how temporary contacts between livestock from different herds and flocks might affect pathogen transmission in the area.

Table of Contents

3 Livestock introductions and the risk of infectious pathogen exposure in northern Tanzania 79

4 Describing livestock market movements in northern Tanzania and how these influence pathogen transmission 121

List of Tables

List of Figures

Glossary

Abbreviations used throughout this thesis

General

A-E	Agro-ecological
LMICs	Low and middle income countries

Network specific

GSCC	Giant strongly connected component
GWCC	Giant weakly connected component
Combined_E	Combined cattle and small ruminant network with link weights equal to the total number of expected cattle and small ruminants moved along them in one year
Combined_SR	Combined cattle and small ruminant network with links weighted heavily for small ruminants (0.9 * expected number of small ruminants moved in one year) and lightly for cattle (0.15 * expected number of cattle moved in one year)

Pathogens

BHV-1	Bovine herpes virus type 1
BVDV	Bovine viral diarrhoea virus
bTB	Bovine tuberculosis
PPRV	Peste des petits ruminants virus
RVF	Rift Valley fever
C. burnetii	*Coxiella burnetii*
Brucella spp.	Brucella melitensis, Brucella abortus
FMDV	Foot and mouth disease virus

Node and network properties brief definitions

Degree	Number of edges connected to a node (Newman, 2010) pg.133-46
In-degree	Number of inward connections in a directed network (Newman, 2010) pg.133-46
Out-degree	Number of outward connections in a directed network (Newman, 2010) pg.133-46
Eigenvector Centrality	Score of a node is proportional to the centrality scores of its neighbours (Newman, 2010)
Betweeness centrality	Frequency with which a node lies on the geodesic between other nodes (Newman, 2010)
Path	Sequence of nodes so that every consecutive pair in a path are connected by a link (Kiss, Green and Kao, 2006).
Geodesic	Shortest path between two nodes. If two nodes are not connected in a network they are in different 'Components' (Newman, 2010) pg. 139
Diameter	Shortest path (Geodesic) between the two most distant nodes on the network (Newman, 2010) pg. 139
Strongly Connected Component (SCC)	Nodes are in the same SCC if A connects to B and B connects to A via a directed path (Dorogovtsev, Mendes and Samukhin, 2001)
Giant Strongly Connected Component (GSCC)	Largest strongly connected component, used to estimate the lower bounds of an epidemic size on a directed network
Weakly Connected Component (WCC)	Nodes are in the same WCC if there is an undirected path between them (Newman, 2010)
Giant Weakly Connected Component (GWCC)	Largest weakly connected component, used to estimate the upper bounds of an epidemic size on a directed network
Local clustering	Proportion of neighbours of node *X* that are connected to each other (Newman, 2010)
Clustering coefficient	Density of triangles in a network (Newman, 2010) (Pg. 262)
Reciprocity	Fraction of edges in a network that are reciprocated (Newman, 2010)
Modularity	Measures the extent to which like connects to like in a network (Newman, 2010)
Assortative mixing	Nodes with similar characteristics (e.g. High degree) are more likely to connect to one another (Newman, 2010)
Disassortative mixing	Nodes connect to others with dissimilar characteristic values (Negative modularity) (Newman, 2010)
Cluster/Community	Many connections within, few connections between. (Newman, 2010)

Acknowledgements

I wish to gratefully acknowledge the support and guidance I have received from all members of my supervisory team over the past 4.5 years Paul Johnson, Rowland Kao, Will de Glanville, Louise Matthews and Kathryn Allen, without whom this PhD project would not have been possible. Also to Sarah, for introducing me into the field in Tanzania, for having confidence in me and being a permanent source of motivation and inspiration. Thank you for all believing in me, especially regarding the fieldwork, and for allowing me the opportunity to work as part of the SEEDZ modelling team. I have learned so much from all of you and will be forever grateful to you all for this PhD journey and the indelible mark it has made on my life. Extra special thanks to Paul for the grammar lessons, stats lessons and always making time to answer my R queries – especially from a distance during lockdown!

I wish to thank every member of the SEEDZ team I worked with in Tanzania, specifically Tito, Rigo and Sambeke for helping me organise my field work and Ray and Fadhilli for being excellent drivers and more! Special thank you to all household and market survey participants who gave up their valuable time to take part and help me answer my research questions. A heartfelt thanks to Dr Nyasebwa and Mr Msaki for your guidance and support for all fieldwork activities, for supporting the market survey work and sharing your many years of knowledge of livestock movements in Tanzania with me. This PhD would not be what it is without all of your input, support, trust and kindness.

Alicia and Jennika, thank you for sharing your wealth of anthropological knowledge and for the inter-disciplinary perspectives you brought to our discussions.

Anna, thank you for providing a home and family away from home, for sharing your knowledge and experiences and being a huge support and fantastic friend in Tanzania.

Suzan and Anthony, thank you for being a steadfast support throughout my PhD journey, providing another home in Glasgow and generally taking care of Molly, bump Elsie and I. Your kindness, hospitality and support over the years means the world to me and you have made completing a Tanzania-Glasgow based PhD possible for me.

Thank you to the funders at the BBSRC for funding the Zoonoses and Emerging Livestock Systems (ZELS) Initiative BB/L018926/1 and the associated studentships which made this project possible.

Elsie thank you for bringing so much joy and new perspective into the last two years of thesis writing. I can’t wait to take you to Tanzania to meet everyone. Rob, thank you for everything, for your patience, kindness, support, looking after our family and generally making the writing of this thesis enjoyable. Nanna, thank you for your quiet confidence and support and for helping me believe I could do something like this. Finally Mam and Dad, thank you for instilling a strong work ethic into me and for your unwavering support, patience, humour and encouragement throughout this PhD journey and everything that has lead up to it. I may not have lambed an elephant in Africa (ambition announced age 4) but I couldn’t be happier with the direction I have taken and couldn’t have done it without you. I dedicate this thesis to my family. Elsie, follow your dreams.

Author's declaration

I declare that this thesis and the research contained within it is my own work unless otherwise stated, and no part of it has been submitted as part of any other degree or qualification.

Gemma Chaters

June 2021

1 Introduction

1.1 Livestock reliance and infectious diseases

In northern Tanzania livestock are heavily relied upon as a main source of income, social status and nutritional security, especially by those living in the most marginalised communities (Covarrubias *et al.*, 2012; Government of Tanzania, 2017). Slow growth of the livestock agriculture sector in Tanzania has been linked directly to a slow rate of poverty and nutritional deficit reduction compared to what was expected in line with the country's overall economic growth (Pauw and Thurlow, 2011). The livestock agriculture sector is in need of rapid development to improve livestock productivity as this change will maximally benefit people living in the poorest communities (Christiaensen, Demery and Kuhl, 2011).

There is strong evidence to suggest that the high burden of infectious livestock and zoonotic pathogens circulating in the cattle and small ruminant livestock populations in northern Tanzania places a major constraint on the development of the livestock sector (Komba *et al.*, 2012; Sindato, Karimuribo and Mboera, 2012; Mathew *et al.*, 2017; Nandonde, Gebru and Stapleton, 2017; Haseeb *et al.*, 2019). The high livestock pathogen burden is problematic as it leads to high livestock losses through disease related morbidity and mortality, unstable reproduction and growth rates and financial and nutritional insecurity for the people and communities who are reliant upon livestock (Tempia *et al.*, 2010; Coppock, Desta and Gebru, 2012; Haseeb *et al.*, 2019). Healthy livestock, on the other hand, are associated with prevention of human disease, reduced malnutrition and improved household income and education (Thumbi *et al.*, 2015; Haseeb *et al.*, 2019).

Spill-over of zoonotic pathogens from animals (including livestock) to humans contribute to the majority of human infectious and emerging diseases globally (61% and 75% respectively) (Taylor, Latham and Woolhouse, 2001; Klous *et al.*, 2016). Endemic zoonoses however, receive little attention at the global scale because their overall impact on livelihoods are prone to underestimation due to insufficient access to cheap, reliable diagnostics and the consequent under reporting (Maudlin, Eisler and Welburn, 2009; Halliday *et al.*, 2012). In Tanzania, hospital based research studies have found zoonoses to be a leading cause of febrile illness in hospitalised patients but zoonoses are rarely listed as a potential or differential diagnosis (Biggs *et al.*, 2011; Crump *et al.*, 2013; Bodenham, Lukambagire, *et al.*, 2020). Limited awareness of zoonotic pathogens amongst both human and livestock health workers in Tanzania, along with limited resources for diagnosis, treatment and prevention mean that zoonotic infections are often misdiagnosed and erroneously treated which leads to worse patient outcomes (Crump *et al.*, 2013; Zhang *et al.*, 2016). The dual burden of zoonoses on individuals and societies in Tanzania motivates an integrated and collaborative approach between human and livestock health sectors for research, policy development and the design of disease control interventions (Morens, Folkers and Fauci, 2004; Coker *et al.*, 2011).

For pathogens that cause disease in livestock alone, the overall economic returns from control programmes can far outweigh the costs and thus pathogen control can directly contribute to improved livelihoods for farming communities (Jones *et al.*, 2016). Improvements to livestock- and public-health services can improve a country's health and economic wellbeing but the funding resources available for these remain limited (World Bank and TAFS Forum, 2011; World Organisation for Animal Health (OIE), 2013). Because of the limited governmental funding received by livestock sectors and veterinary services there is often scarce data available to clearly prioritise which pathogens should be targeted for control and where to focus control for maximal gains (Forman *et al.*, 2012; Brooks-Pollock *et al.*, 2015). Where there are multiple pathogens circulating in livestock populations, these are typically prioritised in government agendas based on their political profile and/or zoonotic potential, not necessarily on their cumulative impact on livestock productivity and livelihoods (Carslake *et al.*, 2011; Pieracci *et al.*, 2016). For example, Rift Valley fever virus (RVFV) is often prioritized due to its zoonotic potential and foot and mouth disease virus (FMDV) has a high political profile due to the estimated global costs of the disease and livestock trade restrictions placed on countries with circulating virus (Woods, 2004; Sinkala *et al.*, 2014; Munyua *et al.*, 2016; Pieracci *et al.*, 2016). In low and middle income countries (LMICs) specifically, improved diagnostic capacity, veterinary services, healthcare provider education and an increase in active surveillance are all required to improve livestock pathogen and zoonoses control (Christou, 2011; Sherman, 2011; Crump *et al.*, 2013; Zhang *et al.*, 2016; Cash-Goldwasser *et al.*, 2018). Pooling resources from multiple sectors to reduce the incidence of zoonoses in the livestock populations is recognised as the most economically efficient way to reduce the overall zoonoses burden in populations for maximum societal gains (Zinsstag *et al.*, 2007).

1.2 What influences pathogen transmission?

There are a wide range of pathogen- and host-specific factors that influence pathogen reproduction rates and transmission from infectious to susceptible individuals within populations (Woolhouse *et al.*, 1997; Blackburn *et al.*, 2019). In brief, pathogen shedding varies between hosts (Lloyd-Smith *et al.*, 2005) and between pathogens with some pathogens shed cyclically (e.g. *Brucella* spp.) (Lambert *et al.*, 2018), some intermittently (e.g. Bovine herpes virus type 1 (BHV-1))(Biswas *et al.*, 2013) and some persistently (e.g. bovine viral diarrhoea virus (BVDV))(Brownlie *et al.*, 1987). What is required for a successful transmission event between a susceptible and an infectious individual also varies with some pathogens requiring close contact (e.g. Peste des petits ruminants virus (PPRV) (Hammouchi *et al.*, 2012)) and others able to transmit via fomites or environmental contamination (e.g. *Leptospira* spp. (Barragan *et al.*, 2017) and *Coxiella* (Kersh *et al.*, 2013)). *Coxiella* is also an example of a pathogen which can remain infectious in the environment for several months (Seitz, 2014) whereas others such as *Brucella* spp. are unlikely to survive more than a few weeks in the environment (Aune *et al.*, 2012). The duration of infectious periods also varies with some pathogens, such as foot and mouth disease virus (FMDV), having infectious periods that last less than one week (Mardones *et al.*, 2010) and others such as *Leptospira* spp. having infectious periods that can last months or years (Bharti *et al.*, 2003). A further factor that affects transmission of pathogens is the involvement of vector species with examples such as Rift Valley fever (RVF) and Crimean-Congo haemorrhagic fever (CCHF) both transmitted through vectors to livestock and then directly from livestock to people (Ergönül, 2006; Aslam *et al.*, 2016; Métras *et al.*, 2020).

More broadly, pathogen transmission between hosts is influenced by population structure and contacts between susceptible and infectious individuals (Anderson and May, 1991). Intermittent direct and indirect contacts between livestock can provide opportunities for pathogen transmission events that allow chronic diseases with low transmission rates to persist in populations, provided the contacts occur during infectious periods (Cross *et al.*, 2005; VanderWaal, Gilbertson, *et al.*, 2017). With regards to population structure, a high variation in herd and flock sizes within a population can reduce the epidemic threshold for infectious pathogens (Caillaud, Craft and Meyers, 2013). In northern Tanzanian there is evidence for high variation in herd and flock sizes as the livestock population constitutes household herds and flocks that range from 1 to over 1,000 for each of cattle, sheep and goats (de Glanville *et al.*, 2020). Larger groups in a population also have larger infectious disease outbreaks within them that persist for longer periods of time. This means that in a population with more large groups, fewer movements between them are required for pathogen propagation because more infected individuals are moved (Cross *et al.*, 2005). There are multiple reasons that could explain the maintenance of endemic pathogens at low levels in northern Tanzanian livestock and the repeated epidemics seen from diseases such as FMD, RVF and PPR (Sindato, Karimuribo and Mboera, 2012; Kerfua *et al.*, 2018; Omondi *et al.*, 2019; Spiegel and Havas, 2019). It is very likely however, that transmission and maintenance of livestock pathogens is driven by livestock movements and contacts occurring between infectious and susceptible individuals as a result of these livestock movements (Prentice *et al.*, 2017; Kim *et al.*, 2021).

In northern Tanzania, a further concern regarding livestock disease transmission is that competition for natural resources is increasing in line with population growth and degradation of communal rangelands (Borjeson, Hodgson and Yanda, 2008). This will inevitably result in livestock herds and flocks and individuals being forced into closer proximity which could further increase the risk of pathogen transmission events (Lee and Barrett, 2001; Goldman and Riosmena, 2013; National Bureau of Statistics, 2013). Increased competition for resources could also negatively impact nutritional status which again might increase opportunities for pathogen transmission within and between species (Epstein, 2002; Patz and Khaliq, 2002; Cross *et al.*, 2004; Jones *et al.*, 2008).

Control measures that aim to reduce livestock and zoonotic pathogen burden within the livestock population need to be developed and implemented urgently, to prevent further and more pronounced livestock losses, human disease, social and nutritional insecurity.

1.3 Using livestock movements to target disease control measures

Targeted approaches to livestock disease control are an efficient way to use limited disease control resources (Kao *et al.*, 2006; Ortiz-Pelaez *et al.*, 2006; VanderWaal, Enns, *et al.*, 2017). Identifying targets for disease control interventions, however, requires epidemiological knowledge on which individuals, populations or geographic locations are at highest risk of disease introduction or onward transmission (Ortiz-Pelaez *et al.*, 2006; Fournie *et al.*, 2013; Rushmore *et al.*, 2014; Molia *et al.*, 2016; Mathew *et al.*, 2017). For pathogens that can be transmitted by livestock, livestock movements provide routes for transmission of infection between individuals and populations (Gilbert *et al.*, 2005; Fèvre *et al.*, 2006; Sherman, 2011). Livestock movements can occur at local, national and global scales which means that the risk of pathogen transmission and consequent disease risk also occur at these scales (Little, 2009). There are multiple motives for livestock movements but globally the major driver is trade to meet the growing demand for livestock produce (Williams, Spycher and Okike, 2006; Little, 2009; Mtimet *et al.*, 2021). The demand for livestock produce is often higher in urban areas compared to rural, where the majority of livestock rearing takes place (Covarrubias *et al.*, 2012; Fournié and Pfeiffer, 2013; Molia *et al.*, 2016; Nandonde, Gebru and Stapleton, 2017). Moving live animals from their point of origin to slaughter points close to the point of consumption is considered the most economical way to meet the demand for livestock meat products in urban areas in many LMICs. The cold chain in Tanzania, as in many other LMICs is insufficient to allow the safe movement of meat products between the points of origin and consumption (McClowry, 2014; Wilson, 2018). Live animal movements therefore occur over long distances and frequently cross international borders, both formally and informally through undocumented movements (Dean *et al.*, 2013; Motta *et al.*, 2017; Napp *et al.*, 2018).

Where extremely dense data on livestock movements exists this can be used to inform disease control programmes (Gibbens *et al.*, 2001). In some countries, often following outbreaks of an economically important livestock disease, national recording of livestock movements has been made compulsory to facilitate future tracing of infectious animals (McGrann and Wiseman, 2001; Derah and Mokopasetso, 2005; Bowling *et al.*, 2008; Smith *et al.*, 2008; Vernon, 2011; Cassius Moreki *et al.*, 2012; Ntokwane and Dibeela, 2016). A specific example of this is the National Animal Identification and Tracing programme set up in New Zealand following the *Mycoplasma bovis* outbreak (Ministry for Primary Industries, 2018). In countries where comprehensive data on livestock movements were lacking, multiple different methods have been used to gather this information (Tempia *et al.*, 2010; Pica-Ciamarra *et al.*, 2011; Flintan, 2012; Dean *et al.*, 2013; Fournie *et al.*, 2013; VanderWaal, Gilbertson, *et al.*, 2017). Data collection has often focussed on livestock markets as these have a high throughput of animals and can also have poor biosecurity, which in combination creates a high risk environment for pathogen transmission (Ortiz-Pelaez *et al.*, 2006; Kao *et al.*, 2007; Fournié *et al.*, 2011; Molia *et al.*, 2016; Júnior *et al.*, 2017). For example, the UK 2001 FMD epidemic was fuelled by a large dissemination event through sheep market sales (Gibbens *et al.*, 2001). Although livestock markets can increase disease transmission through widespread pathogen dissemination event, they do provide useful targets for implementing efficient disease control and surveillance programmes (Kao *et al.*, 2006; Ortiz-Pelaez *et al.*, 2006; Vallée *et al.*, 2013). In Mali and Hong Kong, data describing poultry movements were gathered at markets and used to identify key locations for surveillance and also the type of interventions that would reduce pathogen transmission with minimal trade disruption, such as market rest days (Fournié *et al.*, 2011; Molia *et al.*, 2016). In Togo, Mauritania, Senegal and Cameroon, studies have used market survey data in combination with movement permit data to identify epidemiological connections between multiple countries through cattle movements (Dean *et al.*, 2013; Motta *et al.*, 2017; Jahel *et al.*, 2020). Paper movement permits are often officially required to move livestock between locations in countries that don't have digitised livestock movement recording systems in place. These permits tend to contain information on movement origin, destination, number of animals moved, date of issue, health status of animals and a record of any tax paid for movement the movement. In Egypt, legal import permits for cattle and camels from Ethiopia and Sudan were used to identify high risk locations for Rift Valley fever (RVF) outbreaks and to determine changes that could be made to animal movement behaviour that would reduce risk (Napp *et al.*, 2018). In East Africa, from Somalia down to Tanzania, studies have used a combination of data from livestock market surveys, livestock trader surveys and GPS trackers to identify formal and informal trade routes and high risk locations for livestock and zoonotic pathogen outbreaks (Aklilu, Irungu and Reda, 2002; Mutua *et al.*, 2018; Mtimet *et al.*, 2021; Ng'asike, Hagmann and Wasonga, 2021).

In Ethiopia, Kenya, Sudan and Cameroon and in multiple West African countries, a hierarchical market structure has been identified. Livestock owners sell animals in small primary markets where livestock are batched together and moved on to secondary or tertiary markets by traders, closer to the point of consumption (Aklilu, Irungu and Reda, 2002; Dean *et al.*, 2013; Apolloni *et al.*, 2018; Motta *et al.*, 2019). Livestock traders therefore play an important role in the aggregation of livestock and movement between different locations, including across international boundaries, and along market chains (Motta *et al.*, 2017; Ng'asike, Hagmann and Wasonga, 2021).

In other countries and regions, the demand for livestock produce has been shown to vary throughout the year with increased demand often driven by religious festivals and holiday seasons. In West Africa, the trade of small ruminants and use of vehicles to move livestock are dramatically increased around the time of the Tabaski festival which changes the both the type and speed of pathogen spread (Nicolas *et al.*, 2018). In Ethiopia, poultry trade increases in line with festive periods so an increase in surveillance capacity is needed to accommodate the larger volume of animals traded (Vallée *et al.*, 2013). The trade of small ruminants from Somalia into Saudi Arabia also usually increases around the time of the Hajj pilgrimage, as trade restrictions between the countries are relaxed during this time to meet the increased demands for small ruminant produce (Mtimet *et al.*, 2021). The variation in demand that alters livestock mobility patterns should be captured where possible so that decisions about when and where pathogen surveillance or control programmes should be implemented account for seasonal changes.

The informal East African cross-border cattle trade is estimated to exceed a value of $60 million (Williams, Spycher and Okike, 2006; Little, 2009; Tempia *et al.*, 2010). An example of the cost of international livestock trade in the horn of East Africa is the trade between Somalia and Saudi Arabia which usually increases around the time of the Hajj pilgrimage festival. Current trade restrictions placed on livestock exports from Somalia into Saudi Arabia, driven by the global COVID-19 pandemic, are estimated to be worth US$770M over a 5-year period to Somalia. Over 54% of these losses will be felt by pastoralist livestock keepers (Mtimet *et al.*, 2021).

Livestock movements can also occur for reasons beyond those driven by economics through market chains. In northern Tanzania, as in many other extensive livestock keeping countries around the world, livestock can be moved in search of natural resources including pasture, water or salt, and/or transferred directly between households as gifts or borrowed (Coppolillo, 2000; Aktipis, Cronk and de Aguiar, 2011). In Kenya a combination of household surveys, GPS trackers and photograph recognition methods have been used to quantify livestock mobility patterns and identify high risk individuals, villages and times of year for increased disease risk (VanderWaal, Gilbertson, *et al.*, 2017; Ogola *et al.*, 2018; Floyd *et al.*, 2019). These types of household livestock movements provide opportunities for direct and indirect contacts between livestock from different herds and flocks which provide opportunities for pathogen transmission events.

Knowledge and quantification of livestock movements is important for policy makers and public and livestock health workers because this information can be used to identify which individuals and populations are high risk of disease exposure and onward transmission (Gibbens *et al.*, 2001; Kao *et al.*, 2006; Buhnerkempe *et al.*, 2014; Marquetoux *et al.*, 2016; Kim *et al.*, 2018). Within livestock populations, a small proportion of herds or individuals are often identified as responsible for the majority of movement activity (Woolhouse *et al.*, 1997; Volkova *et al.*, 2010). Heterogeneity in activity often translates into a heterogeneity in pathogen transmission risk and identification of high risk individuals, locations or populations means that these can be targeted by efficient disease control programmes (Kao *et al.*, 2006; Bansal, Grenfell and Meyers, 2007; Tempia *et al.*, 2010; Fournié *et al.*, 2011; Leventhal *et al.*, 2015; Guinat *et al.*, 2016). A major obstacle to implementing effective livestock disease control interventions in countries lacking centralised and comprehensive movement records is that these movements still need to be accounted for to prevent disease control programmes being undermined (Kivaria, 2003; Prentice *et al.*, 2017; Apolloni *et al.*, 2018).

1.4 Livestock movements in Tanzania

At the household level in northern Tanzania three major reasons for livestock movements have been identified: movement to access natural resources such as grazing and watering; movement between households as gifts, private sales, or loans; and sales into and through the market system (Aktipis, Cronk and de Aguiar, 2011; Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012). Despite the high dependence on livestock in northern Tanzania and multiple reasons for movements that could result in pathogen transmission events, there is no centralised system for identifying individuals and recording livestock movements within the country or to neighbouring countries (Mutua *et al.*, 2018). An estimated 100,000 cattle a year (valued at $15 million) are traded across the border into Kenya, but this trade is difficult to monitor and regulate despite its potential to facilitate pathogen transmission between distant, geographically distinct populations (Zaal *et al.*, 2006; Little, 2009). Movements to access natural resources can also cover long distances and cross international borders (Coppolillo, 2000; Bouslikhane, 2015). Unchecked long distance livestock movements provide hidden routes (Trojan livestock) for pathogen transmission into new human and livestock populations, reducing livestock and human resilience and increasing insecurity (Hotez *et al.*, 2009; Molyneux *et al.*, 2011; Rich and Perry, 2011; Bouslikhane, 2015).

In summary, there is a high burden of livestock and zoonotic pathogens in northern Tanzania, limited resources for diagnosis and control and multiple motives for undocumented livestock movements, contacts and thus pathogen transmission opportunities. The combination of these factors motivates data collection and analysis to establish a better understanding of livestock movements in northern Tanzania. More specifically, analysis should aim to quantify how different types of livestock movements at household and market levels contribute to pathogen propagation and maintenance. Improved knowledge of livestock movements and contacts can then be used to identify targets for implementation of efficient disease control programmes.

1.5 Network analysis

1.5.1 Constructing contact networks to represent underlying population epidemiological connections

Mathematical models, and more specifically network models and concepts from network analysis (NA), have been used increasingly over the past two decades in public health and veterinary epidemiology to describe and analyse population structure and livestock movements (Kao, 2002; Keeling and Eames, 2005; Danon *et al.*, 2011; Büttner *et al.*, 2013b; Craft, 2015; Enright and O’Hare, 2017; Enright and Kao, 2018). Social behaviour and contact structure in a population are important for pathogen transmission because it is the interactions between susceptible and infectious individuals, either directly or indirectly through vectors or fomites, that determine pathogen spread (Craft, 2015; Sah, Mann and Bansal, 2018). Heterogeneity in social contacts often naturally exists between individuals within a population and has profound implications for disease dynamics within a population (Bansal, Grenfell and Meyers, 2007). Behavioural heterogeneity translates into heterogeneity in the risk of acquiring and transmitting pathogens within and between groups (Corner, Pfeiffer and Morris, 2003; Böhm, Hutchings and White, 2009). Network models are a useful way of modelling heterogenous population structures and predicting disease spread through populations (Bansal, Grenfell and Meyers, 2007). Identifying the most active and thus high-risk individuals for targeted pathogen control is an efficient way to reduce overall pathogen burden in populations, and efficiency is particularly important when disease control resources are scarce (Kao *et al.*, 2006; VanderWaal, Enns, *et al.*, 2017).

To construct a network that represents an underlying study population, individuals, herds or locations can be represented as nodes in a network graph. A contact or movement between two nodes is represented by a link in the network graph which can be directed from one node to another or undirected. When designing a study, collecting data and constructing a network, what constitutes a node and a link between node pairs should be carefully considered and based on the biological process of interest (Keeling and Eames, 2005; Bansal *et al.*, 2010; Enright and O'Hare, 2017). Once constructed, network graphs that represent the underlying population contact structure can be used to explore how epidemiologically connected a study population is and how easily pathogens might transmit and persist in the population (Gross, D'Lima and Blasius, 2006; Craft, 2015). In some countries, all livestock movements are digitised and centrally recorded. This type of digitised and centralised data has been used to construct complete livestock movement networks for cattle, sheep and pigs in the UK, Sweden, Germany, Italy, France and Denmark to list a few (Natale *et al.*, 2009; Nöremark *et al.*, 2011; Rautureau, Dufour and Durand, 2011; Büttner *et al.*, 2013b; Smith, Cook and Christley, 2013; Ruget *et al.*, 2021). These networks have been used to identify high risk locations for disease control interventions such as active surveillance or movement bans and market closures. Where livestock movement records have not been digitised and centralised, household, trader and market survey data and official movement permit data have been used alone or in combination to construct cattle movement networks in many countries including Canada (Dubé *et al.*, 2008), Brazil (Júnior *et al.*, 2017), Cameroon (Motta *et al.*, 2017), Uruguay (VanderWaal *et al.*, 2016), Togo (Dean *et al.*, 2013) and Mauritania (Jahel *et al.*, 2020). Household, market and trader survey data have also been used to construct poultry movement networks in Ethiopia (Vallée *et al.*, 2013), Cambodia (Van Kerkhove *et al.*, 2009) and Madagascar (Rasamoelina-Andriamanivo *et al.*, 2014). These networks have described various resolutions of livestock market movements at regional, market and village levels and are considered representative subsamples of the true underlying networks. Finer scale networks of livestock movements between household herds and contacts between herds when they are moved to access natural resources have been created using household survey data and GPS tracker data in multiple studies in Kenya (VanderWaal, Gilbertson, *et al.*, 2017; Ogola *et al.*, 2018; Omondi *et al.*, 2021).

Once constructed, networks can be used to identify high risk individuals, populations or locations that are most central in the network, based on multiple node centrality metrics (further details in descriptive list of network measures and centrality metrics in glossary). The most central network nodes can be at highest risk of disease acquisition or onward transmission and therefore targets for efficient pathogen surveillance or control programmes (Bell, Atkinson and Carlson, 1999; Keeling and Eames, 2005; Kao *et al.*, 2006; Martínez-López, Perez and Sánchez-Vizcaíno, 2009b; Farine and Whitehead, 2015). Network models can be used to simulate the effectiveness of different pathogen control interventions that aim to reduce overall network connectedness such as removal of the highest risk (most central) nodes. Nodes that are identified as being at high risk of acquiring infection also have lower estimated time to infection during disease outbreaks (Christley *et al.*, 2005). A pre-existing knowledge of network structure and node's centrality metrics therefore enables plans to be made for the most efficient use of disease control resources and implementation of interventions in advance of disease outbreaks (Keeling and Eames, 2005; Colizza *et al.*, 2006; Salathé *et al.*, 2010; Motta *et al.*, 2017).

1.5.2 Using contact network structure to assess the extent of pathogen transmission in populations

Network structure is important because it affects the speed and extent of pathogen spread, which in a livestock movement network could represent the speed and extent of pathogen diffusion through the population (Newman, 2010). Networks can theoretically range from uniform and regular in structure to completely random (Newman, 2010). Between these two extremes exist small world networks. In small world networks most links connect local nodes, creating clusters of densely connected nodes, with a small number of long range links connecting more distant nodes that belong to different clusters (Watts, 1999; Mark D F Shirley and Rushton, 2005). This small world structure, with high clustering (neighbours of nodes are connected) but some long range links between distant nodes, creates short average path lengths which allow fast diffusion across the network (Watts and Strogatz, 1998; Latora and Marchiori, 2001; Cowan and Jonard, 2004; Guimerà *et al.*, 2005). If the movement or contact between clusters is temporary this can still be sufficient to allow pathogen transfer between clusters, provided the duration of contact is long enough for a transmission event to occur. This type of temporary contact can prevent local pathogen extinction even when the pathogen is not considered to be highly infectious and it has a low R_0 value (Keeling and Rohani, 2002).

Communities in networks are groups of highly connected nodes with many connections within the community and few between them. A network that has a strong community structure is susceptible to high rates of pathogen transmission within communities but slow transmission between them (Leventhal *et al.*, 2015). If networks have a high level of community structure this property can be utilised to tailor pathogen control programmes so that they reduce overall transmission whilst aiming to minimise disruption to normal activity by facilitating movement and trade within communities (Guinat *et al.*, 2016). Various methods and algorithms can be used to identify network communities (Newman and Girvan, 2004; Pons and Latapy, 2006; Martínez-López, Perez and Sánchez-Vizcaíno, 2009a). These work by optimising some definition of network modularity so that the density of network links is highest within communities compared to between (Blondel *et al.*, 2008).

The sizes of the largest connected components of a network can be used to estimate the upper and lower bounds of epidemics on networks (Dorogovtsev, Mendes and Samukhin, 2001; Kao *et al.*, 2006). The giant strongly connected component (GSCC) is the largest component of a network that is connected by directed links only and can be used to estimate the lower bounds of an epidemic's size on the network. The giant weakly connected component (GWCC) is the largest component of a network connected by undirected links and can be used to estimate the upper bounds of an epidemic's size (Newman, 2010).

1.5.3 Using network node centrality metrics to identify high risk locations for pathogen transmission and disease control interventions

The nodes in a network that connect communities are important for pathogen transmission dynamics because they act as bridges between large groups of nodes and reduce the number of steps required (average path length) for pathogen transmission across the network. These bridging nodes can be identified using the network node centrality measure betweenness. Betweenness centrality measures the frequency with which a node lies on the shortest path between all pairs of connected network nodes (Newman, 2010). When a pathogen is introduced onto a network, the higher the betweenness centrality of the seeding node, the greater the extent of the epidemic (Natale *et al.*, 2009; Sah, Mann and

Bansal, 2018). Removing just a small proportion (1-20%) of nodes with high betweenness centrality has been shown to rapidly fragment networks GSCCs in multiple settings, significantly reducing the potential for pathogen transmission across the network (Meyers *et al.*, 2003; Rautureau, Dufour and Durand, 2011; Marquetoux *et al.*, 2016; Motta *et al.*, 2017). Targeting nodes with high betweenness in advance of an epidemic, to raise awareness of the role they play in pathogen dissemination and control, has also been identified as an intervention that could improve preparedness to reduce the extent of future disease outbreaks (Rautureau, Dufour and Durand, 2011).

In addition to betweenness centrality, other measures of node centrality commonly used in disease epidemiology to identify high risk nodes include in-degree, out-degree, degree, geometric mean degree and eigenvalue centrality (De Arruda *et al.*, 2014; Bucur and Holme, 2020). More details on the relevance and use of these measures in network analysis are available elsewhere (Watts and Strogatz, 1998; Butts, 2008; Martínez-López, Perez and Sánchez-Vizcaíno, 2009b; Newman, 2010; Schimmer *et al.*, 2012; Farine and Whitehead, 2015). A basic description of the node centrality metrics commonly used to identify high risk nodes on static networks, and some descriptive network terms relevant to disease epidemiology that will be used in this thesis are included in the glossary.

In brief, in-degree counts the number of ingoing links to a node on a directed network and has been used to identify nodes (herds, markets and villages) at high risk of disease introduction in multiple countries (Lo Fo Wong *et al.*, 2004; Christley *et al.*, 2005; Natale *et al.*, 2009; Palisson *et al.*, 2016). Out-degree measures the number of outgoing links from a node and can be used to identify nodes that contribute most towards onward disease transmission (VanderWaal *et al.*, 2016). Degree is the sum of a node's in- and out-degree. Geometric mean degree is the square root of the product of in- and out-degree and quantifies a node's contribution to pathogen transmission on a network (Newman, 2010). A high level of correlation between in- and out-degree, for example livestock traders or markets, show that nodes at high risk of disease introduction are also high risk for onward transmission and thus have the potential to act as super spreaders on the network (Galvani and May, 2005; Kiss, Green and Kao, 2006; Natale *et al.*, 2009). Node eigenvalue centrality is calculated based on the sum of the centralities of its connected neighbours so it reflects a nodes connectedness in the network and thus its ability to transmit pathogens on to other nodes (Bonacich, 2007; De Arruda *et al.*, 2014).

Heterogeneity and right skew of centrality metric distributions are frequently seen in livestock movement networks globally, with most nodes having low numbers of connections (low degree centrality, low betweenness, low eigenvalue centrality etc) but few nodes having many connections (high degree centrality, high betweenness etc) (Bigras-Poulin *et al.*, 2006; Kiss, Green and Kao, 2006; Natale *et al.*, 2009; Nöremark *et al.*, 2011; Büttner *et al.*, 2013b; Lindström *et al.*, 2013; Smith, Cook and Christley, 2013; Motta *et al.*, 2017). Networks with right-skewed centrality distributions are vulnerable to rapid fragmentation of the large components through targeted removal of the most central nodes and these networks are less vulnerable to fragmentation if node removal is done randomly (Christley *et al.*, 2005; Shirley and Rushton, 2005; Nöremark *et al.*, 2011; Büttner *et al.*, 2013a; Smith, Cook and Christley, 2013). Successful and efficient fragmentation of networks into multiple small components that are not connected to one another rapidly reduces the potential for pathogen transmission across network.

More complex temporal dynamic networks that model the timing of movements and potential transmission events are also increasingly being used in disease epidemiology (Bajardi *et al.*, 2011; Masuda and Holme, 2013; Enright and Kao, 2016; Silk *et al.*, 2017). In countries such as the UK and Sweden with strong seasonal patterns to livestock movement, traditionally driven by birthing and growing seasons, temporal changes in network structure can be identified that will influence the speed and extent of pathogen spread (Kiss, Green and Kao, 2006; Nöremark *et al.*, 2011). Dynamic networks identify structural changes that can be otherwise masked by static networks but these are computationally expensive and require intensive data collection at appropriate time scales (Bansal *et al.*, 2010; Bajardi *et al.*, 2011; Farine and Whitehead, 2015). Static networks are still useful to identify high risk individual nodes when node behaviour does not vary substantially over time and or when disease transmission rates are low (Vernon and Keeling, 2009). In countries such as Ethiopia and Cameroon, no seasonal changes in trading behaviour of individuals were identified, so although the volume of livestock moved varied through the year, static networks were robust to temporal changes and could be used to identify high risk nodes (Nöremark *et al.*, 2011; Vallée *et al.*, 2013; Motta *et al.*, 2018).

1.6 Objectives and overview

In this thesis I aim to address the knowledge gaps around livestock movements in northern Tanzania and the impact that these movements have on the transmission of key livestock pathogens at the inter-household and market levels (Figure 1-1). I will collect market data and movement permit data across three regions (Arusha, Manyara and Kilimanjaro) and use these in combination with household survey data to quantify, for the different types of movement, numbers of cattle and small ruminants moved, distances travelled and variations in the modes of transport used and journey speeds. I will also assess how inter-household movements are associated with pathogen seroprevalence for multiple livestock and zoonotic pathogens in cattle and small ruminants. I will use the livestock movement data to construct contact networks and use concepts from network analysis to evaluate whether heterogeneity in movement activity exists at the inter-household network or market network levels. If heterogeneity exists, node centrality metrics calculated on the network models will be used to identify high risk locations for pathogen introduction or onward transmission and thus where disease control programmes should be targeted. I will also assess if the same locations are high risk for single or multi species pathogens by combining cattle and small ruminant networks.

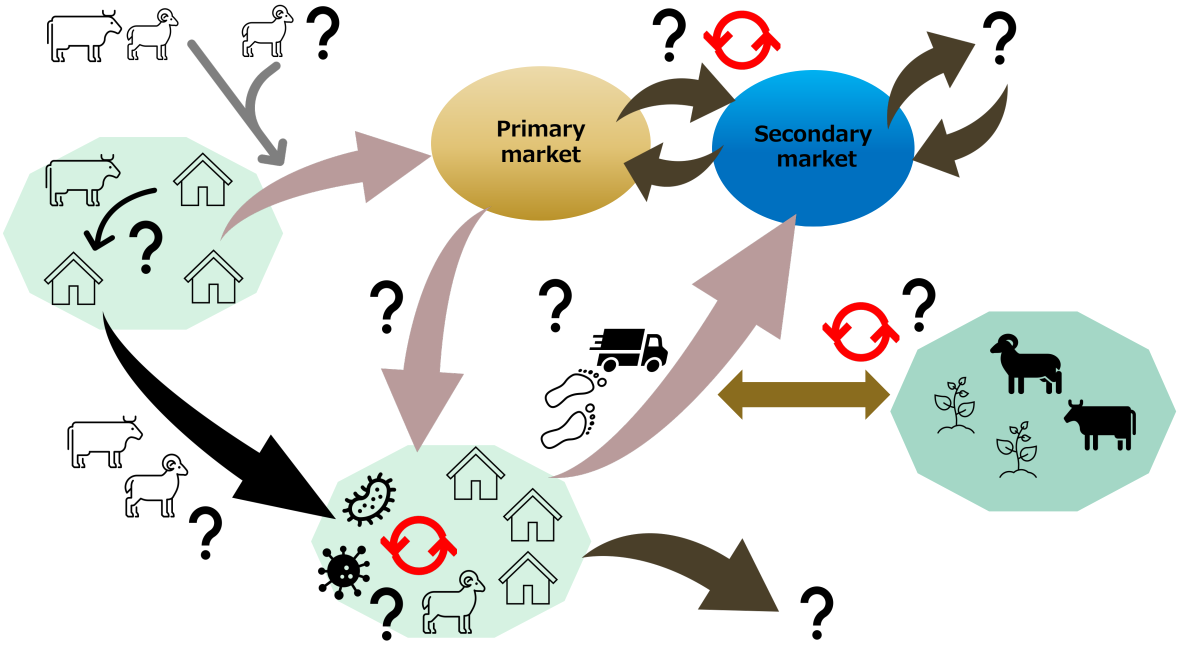


Figure 1-1: Schematic diagram showing movement types to be identified and quantified in this thesis

1.6.1 Chapter 2

In Chapter 2, household survey data from 404 household surveys collected from 47 sub-villages in Arusha and Manyara regions are used to describe movements of livestock between households and between households and markets. The household data is aggregated to the level of the sub-village (an administrative subunit within a village) and used to construct an inter-household movement network that represents a subsample of the underlying population inter-household movement network in the study area. Concepts from network analysis are used to assess study area connectivity through inter-household movements and assess if heterogeneity in movement activity exists between study sub-villages.

1.6.2 Chapter 3

In Chapter 3, household survey data and seroprevalence data for multiple livestock pathogens from cattle and small ruminants are used to assess whether livestock introductions and/or sub-village centrality metrics (calculated in Chapter 2) are associated with an increased risk of pathogen exposure.

1.6.3 Chapter 4

In chapter 4, market survey data collected from buyers and sellers at 22 markets across Arusha, Manyara and Kilimanjaro regions of northern Tanzania are used to describe and quantify the journeys of cattle and small ruminants to and from markets. The market survey data is also used to construct a network and assess how well-connected the study area is through the movements of cattle and small ruminants. The constructed network is used to calculate centrality metrics for the market and non-market nodes and evaluate whether heterogeneity exists between nodes which could be used to identify high risk locations for targeted pathogen surveillance programmes.

1.6.4 Chapter 5

In Chapter 5, Data from Tanzanian government livestock movement permits from Arusha, Manyara and Kilimanjaro regions are used to construct a static network of livestock trade movements within the study area. The network is used to evaluate if heterogeneity in movement activity exists between wards (administrative unit of around 12,000 people containing on average 4 villages). Node centrality metrics, calculated on the static network, are used to identify high risk wards and assess if removal of a small proportion (5%) of highest risk nodes, through implementation of targeted disease control interventions (vaccination vs movement ban) can significantly reduce pathogen transmission on the network.

1.6.5 Chapter 6

In Chapter 6, The same methods as in Chapter 5 are used to construct a small ruminant movement network from government movement permits data to quantify market trade movements of small ruminants to, from and between wards within the study area. The small ruminant and cattle networks are then combined and methods from network analysis are used to identify high risk locations for introduction and onward transmission of: (i) small ruminant only and (ii) combined cattle and small ruminant pathogens. Locations identified as high risk for small ruminant only or cattle and small ruminant pathogens are compared to determine if they are largely the same, or if locations identified as high risk vary depending on a pathogens host species. The comparison between the small ruminant only and combined cattle and small ruminant network models is made to evaluate whether multi-host, multi-pathogen disease control programmes can be efficiently and effectively implemented in northern Tanzania.

1.6.6 Summary

Overall, the data generated and used in this thesis address major data gaps regarding livestock movements in northern Tanzania and how these movements contribute to epidemiological connectivity and influence livestock and zoonotic pathogen transmission. The networks and results generated through this work can be used to identify if high risk locations for pathogen transmission exist at household, village and ward levels. The data can also be used to identify if the same locations are high risk for pathogens that are utilise different transmission routes and infect different livestock species. If there is positive correlation between risks within locations then these results can be used to develop multi-host multi-pathogen interventions for the most efficient use of publicly funded disease control resources.

2 The importance of largely undocumented household livestock movements on connectivity and pathogen transmission in northern Tanzania

Contribution

Data were collected as part of Social, Economic and Environmental Drivers of Zoonoses research project lead by Professor Sarah Cleaveland, University of Glasgow. Further details on this study are given within the chapter. Movement specific research questions which could be answered utilising available data were conceptualised by GC. Data cleaning, processing and analysis were done by GC.

2.1 Introduction

In Tanzania a large proportion of the population, particularly those in the most marginalised communities, rely heavily on livestock for a source of income, food security and social status (Covarrubias *et al.*, 2012; Davis and Sharp, 2020). The reliance on livestock is accompanied by a high prevalence of infections, including zoonotic pathogens, within the livestock (Hyera, Liess and Frey, 1991; Machang'u, Mgode and Mpanduji, 1997; Karimuribo *et al.*, 2007; Schoonman and Swai, 2010; Mdetele and Kassanga, 2014; Alonso *et al.*, 2016; Mathew *et al.*, 2017; Herzog *et al.*, 2019; Semango *et al.*, 2019). The pathogen burden leads to reduced productivity and excessive levels of uncertainty for people and societies whose livelihoods are reliant on livestock and their produce. Additionally, in the case of zoonoses, spill over of pathogens from livestock to humans leads to illnesses that are often misdiagnosed and improperly treated (Biggs *et al.*, 2011; Crump *et al.*, 2013; Allan *et al.*, 2015, 2018; Assenga *et al.*, 2015; Cash-Goldwasser *et al.*, 2018; Bodenham, Lukambagire, *et al.*, 2020). The relatively high prevalence of zoonotic pathogens in northern Tanzania has extensive negative health and socio-economic consequences that are exacerbated by a lack of awareness of zoonoses among livestock keepers and livestock and human health workers (Zhang *et al.*, 2016). Individuals are unable to prevent infection as there is limited epidemiological understanding of pathogen transmission routes and this problem is compounded by poor healthcare infrastructure and inadequate access to robust diagnostics and veterinary services (Chipwaza *et al.*, 2014; Zhang *et al.*, 2016; Bodenham, Mazeri, *et al.*, 2020). In any low-resource setting, knowledge exchange, active surveillance and disease control programmes need to be streamlined and targeted if control of endemic, zoonotic and emerging pathogens is to be both effective and sustainable (Heffernan and Misturelli, 2000; Maudlin, Eisler and Welburn, 2009; Hop *et al.*, 2011; Johnson *et al.*, 2020).

Movement of livestock between locations and/or populations provides a route for pathogen transmission. Where animal identification and movement recording are not routinely practised, tracing pathogen spread through a population becomes difficult and disease control efforts are hindered (Gibbens *et al.*, 2001). Livestock movement data, the use of this data to construct movement networks, and combining network analysis (NA) with epidemiological data have contributed to effective disease surveillance and control strategies globally (Christley *et al.*, 2005; Gilbert *et al.*, 2005; Ortiz-Pelaez *et al.*, 2006; Kao *et al.*, 2007; Volkova *et al.*, 2010; Nöremark *et al.*, 2011; Buhnerkempe *et al.*, 2014; Guinat *et al.*, 2016; VanderWaal *et al.*, 2016; Marquetoux *et al.*, 2016; Sah, Mann and Bansal, 2018; Kim *et al.*, 2018). There are many ways that livestock network data can be collected (Wongsathapornchai *et al.*, 2008; Nicolas *et al.*, 2013; Selby *et al.*, 2013; Buhnerkempe *et al.*, 2014; Motta *et al.*, 2017; VanderWaal, Gilbertson, *et al.*, 2017; Apolloni *et al.*, 2018). NA offers the opportunity to assess the effects of different types of livestock movements and the resulting epidemiological contacts on pathogen transmission through a population and this has been used increasingly in infectious disease epidemiology over the last two decades (Kao *et al.*, 2006; Kiss, Green and Kao, 2006; Martínez-López, Perez and Sánchez-Vizcaíno, 2009b; Natale *et al.*, 2009, 2011; Danon *et al.*, 2011; Rautureau, Dufour and Durand, 2011; Enright and Kao, 2018; Salvador *et al.*, 2018).

A contact network is composed of nodes that can represent individuals, populations or locations where livestock are kept, and these are connected by links which represent connections such as a contact or movement between nodes within the period being studied. Links are assigned numeric values and can be unweighted (binary), where the values 0 and 1 indicate the absence or presence of movement, or weighted (continuous), where the link values represent the number or rate of movements or contacts. Network links can be directed or un-directed and should be weighted depending on the question being asked, the timescale of the disease/event of interest and the level of data available (Craft, 2015). For a highly contagious pathogen such as African swine fever virus (ASFV) in pigs (Ferdousi *et al.*, 2019), Peste des petits ruminants (PPRV) in small ruminants (Fournié, Waret-Szkuta, Camacho, Laike M Yigezu, *et al.*, 2018) or foot and mouth disease virus

(FMDV) in livestock, any contact or movement between nodes might be epidemiologically important (Kao *et al.*, 2007). Thus, when constructing a network to model transmission of a highly infectious pathogen, network links can be binary so a link is present if any epidemiological contact between two nodes occurs. For a less contagious pathogen such as bovine tuberculosis (bTB) in cattle or scrapie in sheep the total number of contacts between two nodes and the duration of a contact or residency time becomes more important for transmission risk (Kao *et al.*, 2007). When constructing a network to model the risk of transmission of a less transmissible pathogen, links should be weighted by the total number of livestock movements that occur between all node pairs within a set time period.

In disease control epidemiology the aim of an intervention is to reduce or prevent pathogen transmission. Network fragmentation can be used to break a network into smaller components (or fragments), so there are fewer or zero routes for pathogens to transmit between nodes. This can be achieved by node or epidemiological link removal (Bienstock and Bonacich, 2003; Chen *et al.*, 2007; Chami *et al.*, 2017). Node removal can, for example, practically be achieved by vaccinating all livestock within it, increased active pathogen surveillance or preventing any movements into or out of the node, such as through market closures, rendering it inactive in the network (León *et al.*, 2006; Van Kerkhove *et al.*, 2009; Vallée *et al.*, 2013; Rasamoelina-Andriamanivo *et al.*, 2014; Molia *et al.*, 2016; Poolkhet *et al.*, 2016; Holme and Litvak, 2017). Construction and analysis of epidemiological contact networks, done with a pathogen(s) of interest in mind, allows node centrality metrics to be calculated. Centrality metrics are frequently used to identify the most central and influential nodes in contact networks where interventions can be targeted. Commonly used node centrality metrics that are used in this chapter include in-degree (number of ingoing links), out-degree (number of outgoing links), geometric mean degree (the square root of in-degree multiplied by out-degree) and betweenness centrality (most frequently lying on the shortest path between other nodes in the network) (Freeman, 1978; Newman, 2002). Identification of the most central and influential nodes (individuals, farms, markets, villages etc) in the network can identify those individuals or populations at greatest risk of acquiring and or propagating a pathogen(s) on the network (Palisson *et al.*, 2016). The most central nodes can then be targeted for removal for efficient fragmentation of the network with minimum disruption and minimal use of resources (Marquetoux *et al.*, 2016; Ferdousi *et al.*, 2019). Identifying highly central nodes also provides an opportunity to identify common characteristics of these which can then be used to identify other, similarly high risk nodes not captured by sub-sampled data (Rushmore *et al.*, 2014). High risk nodes can also be targets for active information dissemination and knowledge exchange programmes. Identification of the most central and influential network nodes is valuable in the fight to reduce disease burden when pathogen prevalence and the risk of emerging diseases are high but resources are scarce and knowledge of different pathogens and their transmission routes is limited (Maudlin, Eisler and Welburn, 2009; Colman *et al.*, 2019).

Where information on livestock movements is lacking, increasing the available knowledge on movements, including where livestock are moving to and from, over what distances and in what numbers, will help identify high risk areas for pathogen transmission and emergence (Bigras-Poulin *et al.*, 2006; Kiss, Green and Kao, 2006; Vanderwaal *et al.*, 2016). Quantifying the various types of livestock movement or contacts (such as shared resources or services) between locations, and the numerous potential routes for pathogen transmission is increasingly being identified as important in network epidemiology (Finn *et al.*, 2019; Porphyre *et al.*, 2020). A quantitative understanding of livestock movements and connectivity between host populations in any given area will help to inform the geographical scale needed for pathogen control programmes.

If different types of epidemiological contact between nodes in a population exist, multiple network layers that represent the different contact types can be combined to create a multiplex network representing the multiple pathogen transmission routes between nodes (Kinsley *et al.*, 2020). If multiple types of contact and thus methods for pathogen transmission between nodes exist on a network then the threshold required for large disease outbreak on the network is reduced (Zhao *et al.*, 2014; Zhang *et al.*, 2015).

In northern Tanzania there are multiple, complex socio-economic and environmental drivers for livestock movements which all have the potential to result in epidemiological contacts between individuals from different herds or populations (Covarrubias *et al.*, 2012; Goldman and Riosmena, 2013). Any data on livestock movements or identification, such as the recent national cattle branding programme, are scarce and not readily available (not digitised and or centralised). The predominant livestock movement types include movements to access natural resources such as grazing, water and salt points (Coppolillo, 2000), trade movements to, between and onward from markets (Chaters *et al.*, 2019), and movements between households. Inter-household livestock movements occur for many reasons, including the giving and receiving of gifts, payment for services, lending, and private sales and purchases. These inter-household movements are not necessarily captured by official government movement permits (used most frequently for movement of livestock batches onwards from markets (Chaters *et al.*, 2019) – further details in Chapter 4 and 5) as they are unlikely to represent large business transactions or be of substantial economic importance. Inter-household movements thus potentially remain undocumented and undetected at the regional and national levels and have received limited previous investigation. This results in uncertainty around the role of inter-household livestock movements on pathogen propagation and maintenance in the livestock, and by default for zoonoses, human populations in northern Tanzania.

The aim of this study is to address the gap in the knowledge of inter-household livestock movements in northern Tanzania and assess how these movements effect epidemiological connectivity in the study area. Household survey data will be used to construct an inter-household livestock movement network. Concepts from network analysis will then be used to quantify how these movements provide otherwise undocumented routes for pathogen transmission and maintenance in the livestock and human populations. The scale (distances and numbers moved) and extent (number of distinct locations that are connected) of epidemiological connectivity across the study area via inter-household movements will be assessed and the network used to identify if heterogeneities in movement activity exist between sampled locations. This information will be used to identify if highly central or influential locations exist within the inter-household livestock movement network that are at increased risk of pathogen exposure and onward transmission. Highly central locations, if identified and if the observed network structure lends itself to fragmentation with sufficiently minor disruption, can become targets for efficient implementation of surveillance, knowledge exchange and disease control programmes for both existing and emerging pathogens.

2.2 Methods

2.2.1 The study area and data collection

Data were collected as part of the Social, Economic and Environmental Drivers of Zoonoses (SEEDZ) research project cross-sectional household survey in Arusha and Manyara Regions of northern Tanzania between January and December 2016. Figure 2.1 shows a map of Tanzania with the two study regions highlighted in purple. The northern boundary of the study area is the international border with Kenya. Each region is made up of a number of districts. Within each district there are wards (an administrative unit of around 10,000 people) and within wards there are villages (*mean* = 3, median = 2, range 1-19 villages per ward). Villages are further divided into the smallest administrative unit the 'sub-village' (2-5 sub-villages per village). Multi-stage sampling was used to randomly select 20 villages from within the study areas (lists taken from national census data) with urban and Ngorongoro Conservation Area wards excluded. Two to three sub-villages were then randomly selected from each village and a central point determined within the sub-village (N = 47). All cattle, sheep and goat (livestock) keeping households in the sub-village were invited to attend the central point and register for participation in the study. Ten livestock-keeping households per sub-village were randomly selected for participation using a random number generator. Data were collected from a total of 404 households from 47 sub-villages by enumerators using Open Data Kit (ODK) Collect software (https://opendatakit.org/) on portable computer tablets in Kiswahili or Maa. Full details on the SEEDZ project, study design, implementation protocol and the household survey can be found elsewhere (Herzog *et al.*, 2019; de Glanville *et al.*, 2020).

In brief, data used in this study were generated from the survey questions 'How many cattle were introduced into this household in the last 12 months', followed by 'How many of these cattle were purchased?' and 'How many of these cattle were not purchased (e.g. Received as gifts or dowry)?' and the follow up questions: 'Where was the household most of the cattle were purchased from?' (asked up to 5 times to allow for introductions from multiple other households in different locations). The same questions were asked for all possible location types ('market', 'trader' and 'other') where introduced cattle may have originated (again up to 5 times for each location type). The same question structure was used to identify the numbers of, and destinations, for cattle exiting the household in the 12 months preceding the study. The same questions about household introductions and exits were asked for goats and small ruminants. The questionnaire sections and accompanying data dictionary used in this analysis are given in Appendix sections 9.2.2.1 and 9.2.2.2.

Livestock keeping households in the study area have been broadly categorized into three production system classes: pastoral, agropastoral and small holder/urban by de Glanville et al. using data from the SEEDZ survey (de Glanville *et al.*, 2020). Pastoral households generally have the largest herd sizes, are heavily reliant on livestock as the main source of income, largely report transhumance movement to access grazing land and practise limited crop production. Agropastoral households are reliant on livestock as well as crops and are generally in areas with the highest vegetation cover. Small holder households have the smallest numbers of cattle and goats, practise zero grazing around the homestead, are generally in denser human and livestock populated areas and are closest to the major road network. Urban households have no livestock and are in densely human populated towns (de Glanville *et al.*, 2020).

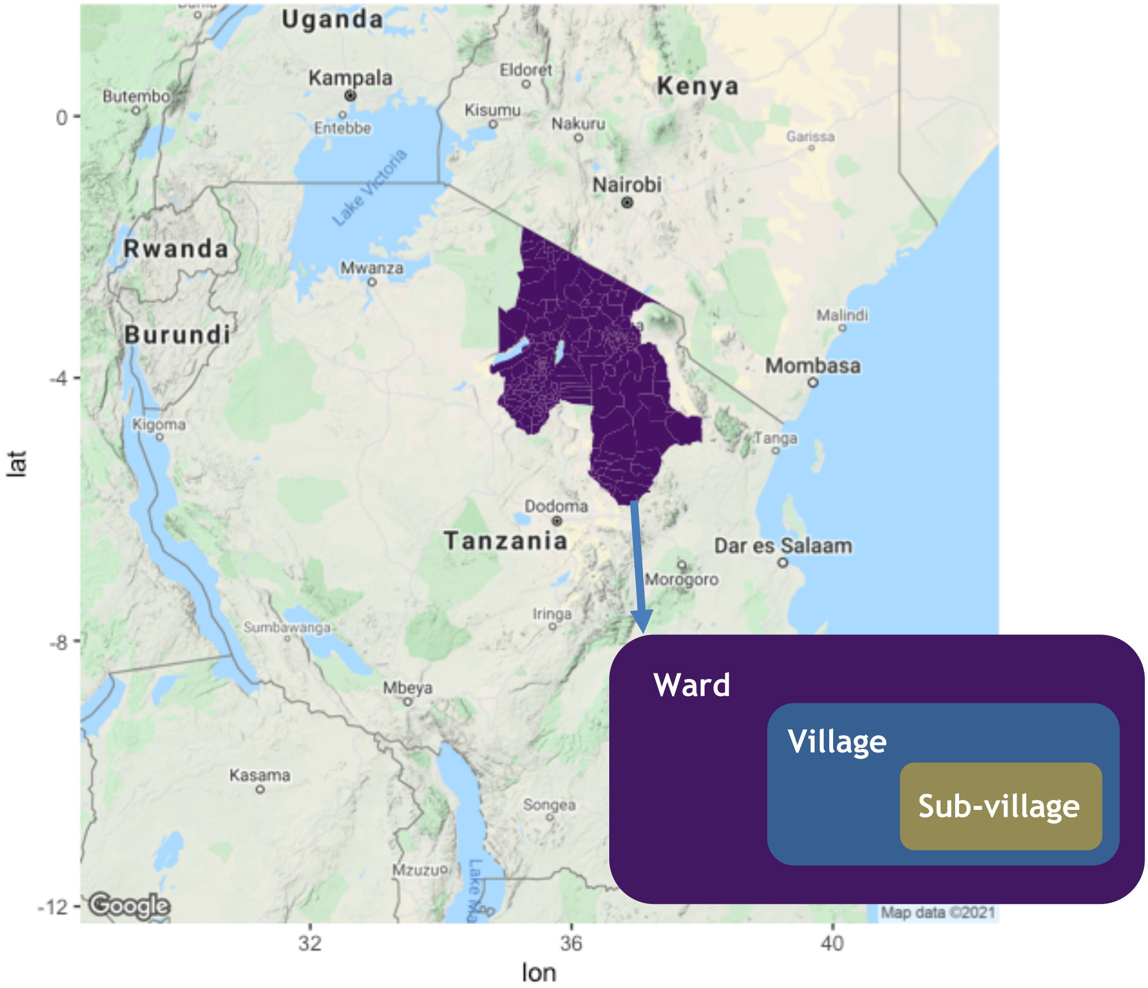


Figure 2-1: Arusha and Manyara study regions highlighted in map of Tanzania, sub-divisions show ward boundaries. Schematic diagram overlaid to depict the nested structure of administrative units in northern Tanzania - the smallest units sub-villages (sampled from in this study), are nested within villages which are nested within wards. Map created using Google Maps as base map.

2.2.2 Network construction

All reports of permanent livestock (cattle, sheep and goat) movements to and from the 404 surveyed household in the year preceding the study were aggregated at the sub-village level (N = 47). Household data are aggregated at the sub-village level because for this analysis it is assumed that mixing between livestock from different herds and flocks within sub-villages occurs throughout the study area due to the use of shared natural resources (de Glanville *et al.*, 2020). Herd and flock mixing provides opportunity for close contacts and pathogen transmission events between livestock from the same and different herds. The effect of livestock introductions into the sub-village on individuals' risk of pathogen exposure could therefore be assumed homogenous across the sub-village following an introduction event.

All household reports of permanent livestock movements into or out of the household herd and/or flock were used to construct a network of livestock movements. A directed link was created between a sub-village node and an origin or destination node for every reported movement. Each link represents the movement of a batch of livestock (at least one individual) and links are weighted by the number of times they are reported by all surveyed households from the sub-village.

There are three different types of permanent livestock movements used to construct the full inter-household movement network. These are gifting in and out movements, private sales (including for meat) and purchases, and market sales and purchases. Permanent movements of livestock between any two locations in Tanzania are supposed to be centrally recorded by use of a movement permit and a tax is supposed to be paid for each animal moved (Chaters *et al.*, 2019). To understand the scale of interhousehold movements potentially happening below the radar due to under/non-reporting, data were collected from households on whether the reported movement was centrally recorded by using a movement permit. If interhousehold movements are important for epidemiological connectivity but are also under reported centrally then this needs to be considered in the design of disease control interventions to prevent them being undermined by movement activity that is unaccounted for.

In addition to the full interhousehold movement network using all movement data, four subset networks were also constructed, one for each movement type and one which combines gifting and private sales/purchases. These four networks were analysed using the same methods used on the full network to assess the relative importance of each movement type on its contribution to study area connectivity and provision of undocumented pathogen transmission routes. The gifting and private sales/purchases movements were looked at in combination as practically these movements will be difficult to separate for disease control interventions and may require different interventions compared to movements centred around livestock markets.

All location nodes were assigned geographic coordinates, so that they could be plotted on a map for visualisation and so movement distances between nodes could be calculated. This was done using a stepwise process as naming and spelling of locations in the survey responses was variable. First, the surveyed sub-village nodes were assigned the sub-village central point co-ordinates (N matches = 47). The remaining non-sampled nodes (N = 100) name was matched first to the list of villages (N = 71) then wards (N = 12) and finally districts (N = 6) from the 2012 Tanzania NBS census data (the original list from which the study villages were selected)(National Bureau of Statistics, 2012) and coordinates from the geographic centre of the matched geographical area were assigned to the node. The remaining unmatched locations (N = 11) were manually checked by two local members of the study team (SM and RT) for potential spelling discrepancies, different naming for locations such as market names given rather than administrative unit, or when named locations fell out-with the study area. All 11 locations were successfully identified and assigned geographic central point coordinates using Google Earth to manually find the location and coordinates. Reported node types (market or other household/village), names and assigned coordinates were double checked by the same local members of the study team to minimise the risk of assigning erroneous locations in the case of a different or mis-spelt location name. For network plotting and the network analysis, sampled sub-village nodes are referred to as sub-village nodes. Nodes that were recorded as market locations due to the recorded movements to or from them being market sales or purchase are referred to as market nodes. Nodes that were recorded as other households or villages due to household private sales, purchases or gifting movements being to or from them are referred to as origin/destination location (OD location) nodes.

Data were imported, cleaned and analysed using R version 3.6.0 (R Core Team, 2019b) and the *igraph* package (Csardi and Nepusz, 2006) was used to construct and analyse the livestock movement networks. Spearman's ρ was calculated using the *cor.test* function in the R stats package(R Core Team, 2019a) to test the correlation between in-degree measures. For further details on data processing, methods and analysis see R scripts [HH_NW_1:5] descriptions and links in the Appendix section 9.1.1.

2.2.3 Movement distances

Link length (great circle distance between two nodes) was calculated to assess the scale at which the risk of pathogen transmission exists, via the recorded inter-household movements, within and out-with the study area. Links were categorised by species group (cattle or small ruminant which includes sheep and or goats), movement type (gift, private, or market), inward or outward movement, production system of outgoing and receiving nodes (agropastoral, pastoral or smallholder/urban), sub-village region (Arusha or Manyara) and production system type of sampled sub-village (agropastoral or pastoral) and were weighted by the number of livestock moved. A generalized linear mixed model (GLMM) was used to assess the effects of the aforementioned movement characteristic variables on movement distance with sampled sub-village included as a random effect. Distance was modelled on the log scale to satisfy model assumptions of normally distributed residuals. The model outcome estimate for movement distance was adjusted for Jensen's inequality to account for the potential uneven sampling between sub-villages and backwards log transformation of distance. All P-values presented in the results are Wald P-values unless otherwise stated. Model goodness of fit was assessed by visualising the distribution of model residuals and the model marginal and conditional R^2 were calculated to estimate the proportion of variation in movement distances explained by the model.

2.2.4 Network Analysis

Several concepts derived from network analysis were used to better understand the role of inter-household livestock movements in pathogen transmission. This included estimating the level of connectivity between populations in the study area via cattle and small ruminant movements, assessing the effects of different movement types on connectivity and identifying if any differences in connectedness exist between sub-villages.

The size of the giant strongly and weakly connected component (GSCC's and GWCC's) of a network can be used to evaluate how epidemiologically connected an area is and estimate the lower and upper bounds of epidemic sizes on a network, respectively. In the GSCC all nodes can reach one another via a directed link and in the GWCC links are assumed undirected so all nodes can reach each other via an undirected link. The size of the GSCC and GWCC were calculated using all recorded movement data and then separately on the subset market, gifting, private and combined private and gifting (from here on referred to as the private&gift)_networks.

The average path length, which is the average number of steps between all pairs of connected nodes in the network, and the diameter of the network, which is the shortest path between two most distant nodes on the network, were also calculated for all networks. These measures show how many sequential movements of infectious animals would be required to spread a pathogen on average between any two connected nodes and the two most distant nodes in the network respectively. Finally, the network fragmentation index, which is the proportion of pairs of nodes that are not able to reach each other, was calculated to assess how relatively difficult, or easy, it would be to fragment the networks through targeted node removal. Fragmentation breaks the network up into many small components so that as few nodes as possible are epidemiologically connected and so the routes available for pathogen transmission are minimised. The higher the fragmentation index of a network the fewer nodes or links need to be removed to fragment it. The aim of an efficient disease control programme is to fully fragment a network with as little disruption or use of resources as possible.

2.2.5 Node centralities

Node centrality metrics were calculated for sub-village nodes to ascertain if there is heterogeneity in the amount of livestock movement traffic, connectedness and disease risk, between sub-villages. Node centrality measures that are epidemiologically important for pathogen acquisition and onward transmission were calculated for the sampled sub-village nodes only. These include the in- and out-degree's, which are the sum of the ingoing and outgoing edges, the geometric mean degree which is the square root of the product of the in and out-degree's, and the betweenness centrality which is the number of times a node lies on the shortest path between two other nodes (Newman, 2010). Centrality metrics were not analysed in the same way for non-sampled market and OD location nodes because these were not sampled from in the same way as the sub-villages.

A sub-village node with high in-degree is at greater risk of bringing in pathogens from other nodes if infected livestock are moved along the network links (Palisson *et al.*, 2016). Sub-villages with high out-degree are at greater risk of infecting other nodes if they have pathogens circulating within the livestock population and infectious livestock move out of the node along the network links (Büttner *et al.*, 2013a). Geometric mean degree can be used to calculate the importance, or rank, of a node for its contribution to livestock movements on the network. The higher the geometric mean degree, the busier a node is (more in-going and out-going movements) thus the livestock in these nodes are both more likely to be exposed to infectious pathogens and transmit pathogens on to other nodes.

Betweenness measures the number of times a node lies on the shortest path between other nodes on the network (Newman, 2010). This represents a measure of a node's reachability, so high betweenness will lead to an increased risk of pathogen exposure as a node is more likely to be reached by livestock carrying pathogen(s) (Molia *et al.*, 2016). More specifically, livestock in nodes with a higher betweeness could be at higher risk of exposure to endemic pathogens because livestock repeatedly circulate through them.

Right skewed distributions of node centrality measures are common in many naturally occurring networks and this pattern has frequently been identified in livestock movement networks in other settings (Salathé *et al.*, 2010; Danon *et al.*, 2011; Büttner *et al.*, 2013a). A right skewed distribution of centrality measures show that few of the nodes in the network are responsible for much of the movement activity(Woolhouse *et al.*, 1997; Newman, 2002). The distribution of centrality measures can help to inform whether a network is susceptible to fragmentation through targeted removal of the most central nodes (Büttner *et al.*, 2013a). Successful fragmentation of a network diminishes the routes available for pathogens to transmit between nodes or groups of nodes with minimal disruption or resource use (Chami *et al.*, 2017).

Sub-village nodes that ranked highly for one centrality (in-degree, out-degree, betweenness and geometric mean degree) were evaluated to check if they also ranked highly for the others by calculating pairwise correlation coefficients using Spearman's rho (ρ). The P-value used to assess if the observed correlations are likely to be significant was calculated using the AS 89 algorithm (Best and Roberts, 1975). This analysis will show if nodes at high risk of pathogen acquisition are also those responsible for onward transmission and maintenance on the network. The amount of correlation between sub-village's risk of pathogen introduction or onward transmission will allow decisions to be made about how to implement disease control interventions most efficiently. For example, nodes with high in-degree at high risk of pathogen introduction may be good targets for vaccination but nodes with high betweenness might be good targets for active surveillance. If the centrality measures are highly correlated then the fixed costs of implementing different veterinary interventions can be shared and more high risk locations targeted with multiple intervention types using the available resources(Forman *et al.*, 2012).

2.2.6 Different methods for link weighting

Node centrality metrics are calculated to quantify the relative importance of nodes in a network. Calculated centrality metrics can depend on the weight of the links going into or out of nodes. For example, node in-degree is the sum of the links going into a node and this metric is used to gauge a node's risk of pathogen introduction. How links are weighted will therefore affect the calculated node centrality metrics and potentially where nodes rank when using centrality metrics to identify those at highest risk of pathogen introduction or onward transmission (Gates and Woolhouse, 2015a). How a link is weighted, when considering transmission of an infectious pathogen on a network, should depend on the infectiousness of the pathogen of interest, the granularity of the data available and needs to be carefully considered before constructing a network (Craft, 2015). For this study we are largely interested in the transmission of endemic zoonotic pathogens such as *Brucella spp*. and *Coxiella burnetii* that impair livestock productivity and cause disease in people (Hummel, 1976; Crump *et al.*, 2013; Shirima and Kunda, 2016). These pathogens can be transmitted through close contact between susceptible and infectious cattle and small ruminants, or contact between susceptible livestock and infectious reproductive fluids and or environmental contamination (Hummel, 1976; H. I.J. Roest *et al.*, 2011; Prabhu *et al.*, 2011; Shirima and Kunda, 2016). The primary analysis on the household survey movement data weights network links by the number of times a batch of at least one animal are moved between nodes as reported by all surveyed households within the sub-village. Each batch of at least one animal moved creates a link to another location with an associated risk of moving of at least one infectious animal and introducing pathogen to the sub-village where it can circulate directly and indirectly within and between herds and flocks. If interest was directed more towards a highly transmissible pathogen which is excreted and spread more rapidly between infectious and susceptible animals (e.g. foot and mouth disease virus (FMDV)), then binary link weighting could be used where any movement reported between nodes creates a single link (Green, Kiss and Kao, 2006a). Conversely, if this study was focussed on less transmissible pathogens (e.g. bovine tuberculosis (bTB) (VanderWaal, Enns, *et al.*, 2017)), links could be weighted by the total number of animals moved between nodes so the risk of pathogen introduction varies proportionally with the number livestock introductions.

The three different methods of link weighting, chosen depending on the transmissibility of a specific pathogen of interest, have the potential to identify different nodes as high or low risk. We wanted to assess if the results from the primary analysis in this study could be used to guide more generalised disease control programmes for pathogens of varying transmissibility. The two alternative methods for link weighting described above (binary or fully weighted) were therefore used to construct additional contact networks using all movement data so unweighted, partially weighted and fully weighted sub-village node centrality metric ranks could be compared. The level of correlation between the ranking of node metrics was assessed using Spearman’s rho (ρ) and the P-value used to check if the correlations are significant was calculated using the calculated using AS 89 algorithm (Best and Roberts, 1975). This was done to determine whether the same nodes ranked highly regardless of the method used to weight links and thus regardless of the transmissibility of the pathogen(s) of interest. If a high level of correlation exists between node ranks using the different methods of link weighting then pathogen control programmes that aim to reduce the burden of multiple pathogens with varying transmissibility could be developed. For further details see Appendix section 9.1.2.

2.2.7 Data rarefication

The network constructed using all household survey data represents a sub-sample of the true underlying inter-household livestock movement network in the study area. The 47 sampled sub-village nodes represent around 2% of sub-villages in the study area. These nodes belong to 20 (2.2%) randomly selected villages which in turn belong to 20 (8.2%) wards (1 village was sampled per ward) in the study area (Arusha and Manyara regions, northern Tanzania). If sub-villages are frequently connected to each other via permanent livestock movements then the sampling strategy should have captured many connections between sampled and non-sampled locations and the network will consist of a small number of large components. If sub-villages are not well connected via permanent livestock movements then the observed network will be heavily fragmented, made up of multiple components that are not connected to each other. We can check how connected sub-villages are by assessing the size and number of components in the observed network and then randomly removing proportions of the data and reassessing the network structure. If the observed network is not highly connected and it has numerous redundant links, then we would expect to see rapid fragmentation of the network through data removal. The level of data saturation that a study achieved using the sampling method can also be assessed by removing data and reassessing network structure. If data saturation, or close to this, has been achieved, then removing some data will not affect network structure. Conversely, if removing data does affect network structure, we can hypothesise that adding further data could increase the observed level of connectivity in the study area. The process of data rarefication will allow us to make transparent interpretations from any results and will indicate if sampling from more locations or extrapolating up from the observed network would increase the observed amount of connectivity in the study area.

To evaluate the level of connectivity and redundant links in the observed network and the level of data saturation 50% and 10% of data was removed randomly, using a random number generator within the *sample* function in R statistical software (R Core Team, 2019b) at two levels: (i) the sub-village and (ii) the village/ward level. The network was reconstructed (10 times for each scenario) using the remaining 50% and 90% of data. This was done at the sub-village level because these represent the sampled nodes in the network, and at the village/ward level because this was the primary level of sampling. If data saturation has been achieved, and the network is densely connected with many redundant links, then dropping data will not have a large effect on overall connectivity and size of the giant components. If data saturation has not been achieved then dropping half of the data will remove around half of the observations (nodes, edges) and greatly reduce connectivity measures and a similar pattern would be seen by dropping 10% of data.

In addition to assessing connectivity between sub-villages in the study area via inter-household movements, it is also important for us to assess how many wards are connected in the study area via these movements. This has practical relevance as the ward could be the smallest administrative unit that disease control programmes are implemented at. Knowledge of the level of ward connectivity via inter-household movements will also enable this work to be viewed in the context of others on livestock movements in the study area (Chaters *et al.*, 2019). To assess the extent of ward connectivity via inter-household movements in the study area the number of connected wards was calculated for the full network, the four subset networks constructed using different movement types and the 90% and 50% sub-sampled networks.

2.3 Results

2.3.1 Network construction

Data from the 404 surveyed households, from 47 sub-villages within 20 randomly selected villages, were aggregated at the sub-village level and used to construct an inter-household movement network. The constructed network using all data has 147 nodes. These are the 47 sampled sub-village nodes and 100 non-sampled origin/destination location nodes. Sixty-two of the origin/destination locations were active livestock markets (market nodes) and 38 origin/destination location nodes were households located in non-study sub-villages without markets, these location nodes are referred to as OD nodes. The full network has 816 links, each representing the movement of a batch of at least one animal. This network represents a static sub-sample of the underlying annual sub-village inter-household livestock movement network as we do not have information on the timing of movements throughout the year, or data from every household or every sub-village in the study area. Details on the size of the batches of livestock moved to and from sub-villages can be found in Table 2-1. Despite reporting a movement of some livestock between two locations, the exact number of animals moved was not reported for 55 of the 816 movements thus data on these movements are not included in Table 2-1. Number of links by species group (cattle or small ruminant) and the type of movement (gifting, private, market) are shown in Table 2-2.

Table 2-1: Frequency of batch sizes of reported livestock movements for all movements where batch size was recorded (N = 761) (55 movements had no batch size recorded)

Batch Size	Cattle frequency	Small ruminant frequency
1	82	51
2	64	57
3	33	43
4	31	30
5	31	46
6	15	22
7	7	12
8	12	6
9	2	0
10	23	33
11	3	0
12	4	11
13	2	3
14	0	3
15	3	19
16	1	3
17	0	1
18	3	0
20	4	27
21	0	1
22	1	0
24	0	1
25	0	4
26	0	2
27	0	1
28	0	1
30	7	10
32	2	0
35	0	1
40	1	6
42	0	1
45	0	1
48	0	1
50	2	8
54	2	0
55	0	1
60	4	3
78	2	0
80	3	1
92	0	1
100	0	2
104	0	1
234	0	2
Total edges	**344**	**417**

Table 2-2: Number of different edge types (Gift, Market and Private) for cattle and small ruminants, representing inward and outward movements in a static sub-sample of the underlying the northern Tanzanian inter-household livestock movement network

	In	Out	Total
Cattle Gift	8 (10%)	48 (16%)	56
Cattle Market	70 (86%)	217 (74%)	287
Cattle Private	3 (4%)	29 (10%)	32
Cattle Sub total	81	294	375
Small Ruminant Gift	2 (2%)	24 (7%)	22
Small Ruminant Market	98 (95%)	283 (84%)	381
Small Ruminant Private	3 (3%)	31 (9%)	34
Small ruminant Sub total	103	338	441
Total	184	632	816

The majority of household livestock introductions came from market purchases (86% cattle, 95% small ruminants) but only 32% (35/108) of households that introduced livestock from markets reported using a movement permit to record this. Receiving livestock as gifts accounted for 10% of cattle introductions and 2% of small ruminant introductions and private purchases accounted for fewest livestock introductions (4% of cattle and 3% of small ruminant). The predominant reason for livestock exits were market sales (74% of cattle exits and 85% of small ruminant exits) and again very few households (23% (56/240)) that took livestock to a market reported declaring this by using an official movement permit. The giving of livestock as gifts was the reason for 16% of cattle and 7% of small ruminant exits and private sales, including sales directly for meat, contributed to 10% of cattle and 9% of small ruminant household exits.

2.3.2 Movement distances

Movement distances were calculated as the great circle distance between origin and destination co-ordinates. The distribution of movement distances is given in Figure 2-2. There is right skew of the distribution of movement distances with most being short journeys (minimum 0.4 km) but some occurring between locations that are much farther apart (maximum 294 km). There is a slight bimodal distribution of movement distance with a slight increase in the frequency of movements between 100 and 200km (Figure 2-2). Results of the generalized linear mixed effects regression model (GLMM) to estimate movement distances are given in Table 2-3. The baseline movement distance (cattle, gifted into an agropastoral sub-village from an agropastoral origin) is estimated to be 19km (95% CI 12-29 km). Small ruminant movements were estimated to be 20% shorter than cattle movements (95% CI 0.70-0.88, P-value <0.001) which would make small ruminant gift movements into agropastoral sub-villages from an agropastoral origins an estimated 15km. Market and private movements were estimated to be 20% and 40% shorter than gifting movements respectively (market movements 95% CI 0.66-0.99 P-value 0.04 and private movements 95% CI 0.46-0.86, P-value <0.01). This means that cattle movements into households from markets (between agropastoral locations) are estimated to be around 15km and privately purchased cattle are estimated to be introduced from locations that are closer again, around 11km away. Outward movements were also estimated to be almost 20% shorter than inward movements (95% CI 0.69-0.99 P-value 0.04). There was strong evidence to suggest that the production system of the origin and destination nodes was associated with movement distance (Likelihood ratio Chi squared test P-value <0.001). Most notable were that movements between small holder/urban nodes and pastoral nodes were estimated to be around 10 times longer than agropastoral-agropastoral movements (To 95% CI 6.53-17.69, From 95% CI 5.31-17.72) which means that these movements are estimated to cover distances closer to 200km. Pastoral-agropastoral movements were also estimated to be three times longer than agropastoral-agropastoral movements (95% CI 2.02-4.78 P-value 0.01).

There was no evidence to suggest production system of the sampled sub-village node alone, the region of the sampled sub-village or the number of livestock moved were associated with movement distance (Likelihood ratio Chi squared test P-value > 0.3) so these variables were not included in the final model. The marginal and conditional R^2 values of the distance model show that the fixed effects explain an estimated 23% of variation in movement distance and the random effect of sampled sub-village explains a further 37% of the variation.

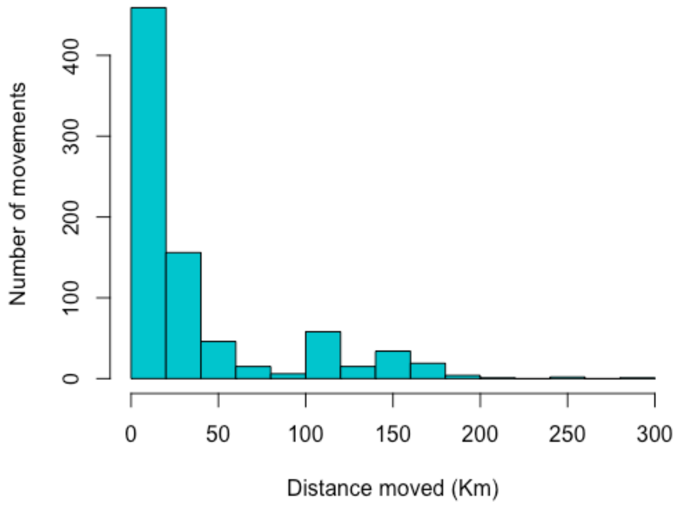


Figure 2-2: Distance of reported inter-household livestock movements in northern Tanzania

Table 2-3: Results of multivariable mixed effects regression model to estimate the distance livestock are moved to and from households. The estimated journey distance in kilometres for the baseline group (Cattle, moving into the household, as gifts, between agropastoral locations), adjusting for Jensen's inequality, is presented in the top row of the table. The multiplicative effect estimates of each variable on journey distance to and from the household with 95% confidence intervals are displayed in the rows below.

	Regression coefficient estimates (95% CI) (N_{obs} = 815)
Predicted distance in km (Baseline group: Cattle, Gift, Inward movement, agropastoral to agropastoral movement (ap-ap))	**19.00 (12.52-28.81)**
Small ruminant (vs cattle)	0.79 (0.70-0.88)***
Market movement (vs gift)	0.80 (0.66-0.99)*
Private movement (vs gift)	0.63 (0.46-0.86)**
Outward movement (vs inward)	0.81 (0.68-0.97) *
agropastoral to pastoral movement (vs ap-ap)	2.50 (1.62-3.86)***
agropastoral to smallholder/urban movement (vs ap-ap)	2.48 (1.62-3.86)*
pastoral to agropastoral movement (vs ap-ap)	3.12 (2.02-4.78)***
pastoral to pastoral movement (vs ap-ap)	2.83 (1.84-4.35)***
pastoral to smallholder/urban movement (vs ap-ap)	10.75 (6.53-17.69)***
smallholder/urban to agropastoral movement (vs ap-ap)	2.17 (1.05-4.49)*
smallholder/urban to pastoral movement (vs ap-ap)	9.70 (5.31-17.72)***
Sub-villages (N = 47) random effects variance	0.61
Marginal R squared	0.23
Conditional R squared	0.60
Note: Sub-village fitted as random effect	*p<0.05**p<0.01***p<0.01

2.3.3 Network analysis

The fully weighted network using all data is plotted over a map of the study area (Arusha and Manyara Regions in northern Tanzania) in Figure 2-3, Figure 2-4 and Figure 2-5 to visualise, at scale, a year of all reported inter-household livestock movements to and from 47 sub-villages (approximately 2% of all sub-villages in the study area). In Figure 2-3 link colour represents species moved, in Figure 2-4 link colour represents inward or outward movement and in Figure 2-5 link colour represents the production system of the receiving node. In all network plots link density is proportional to the number of animals moved. Where number moved was not reported the mean value based on link categories species, in or out movement and production system moved to was used (12 categories). Movements above the highlighted study area on the plots are those that cross the international Kenyan border.

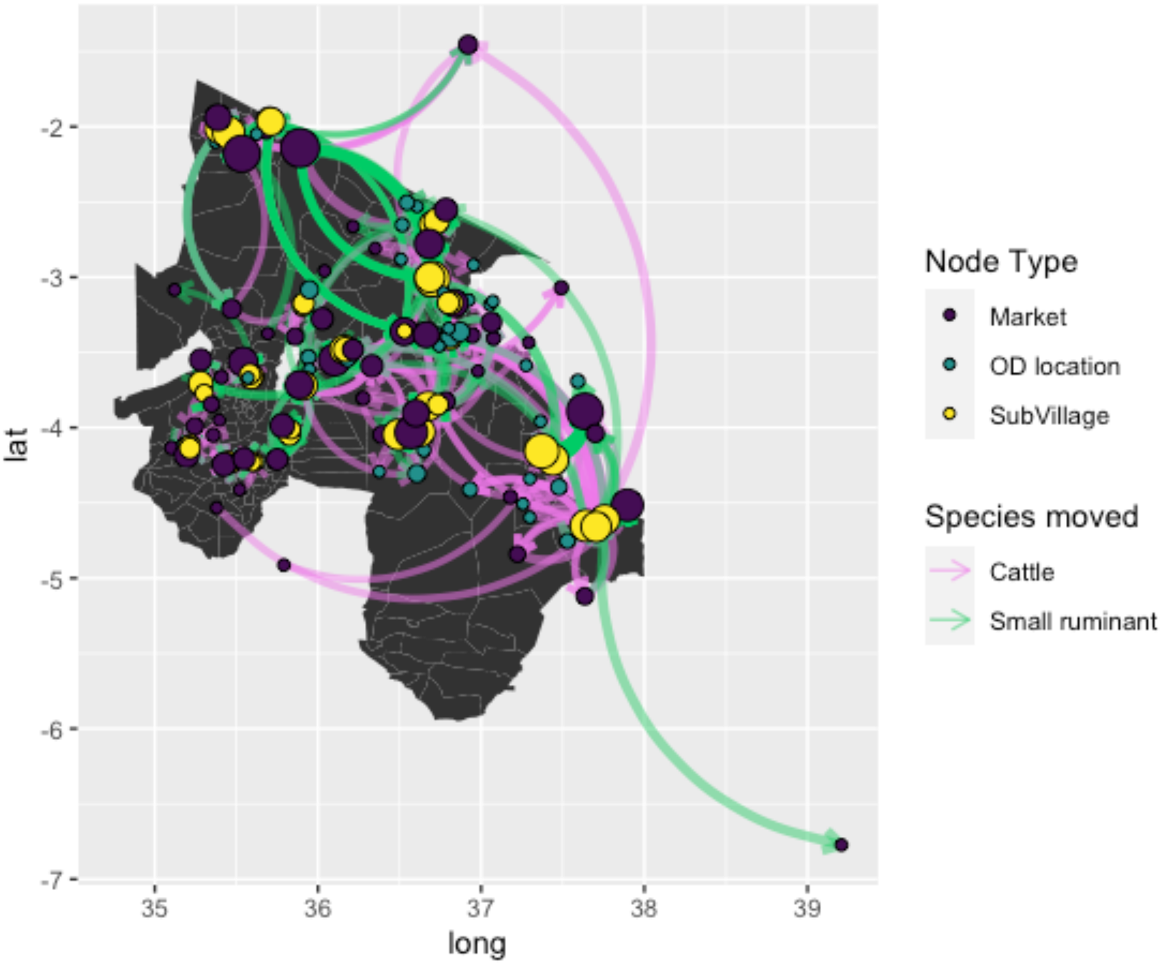


Figure 2-3: All household livestock movements from 404 surveyed households across 47 randomly selected sub-villages in northern Tanzania between January and December 2016. Movements represent reported gifts, sales and purchases to and from other households in other villages (OD locations) and market sales and purchases. Node size is proportional to its degree (busier nodes appear larger) and edge density is proportional to the total number of animals moved. Sub-village nodes are the sampled nodes. Link colour represents the species group moved.

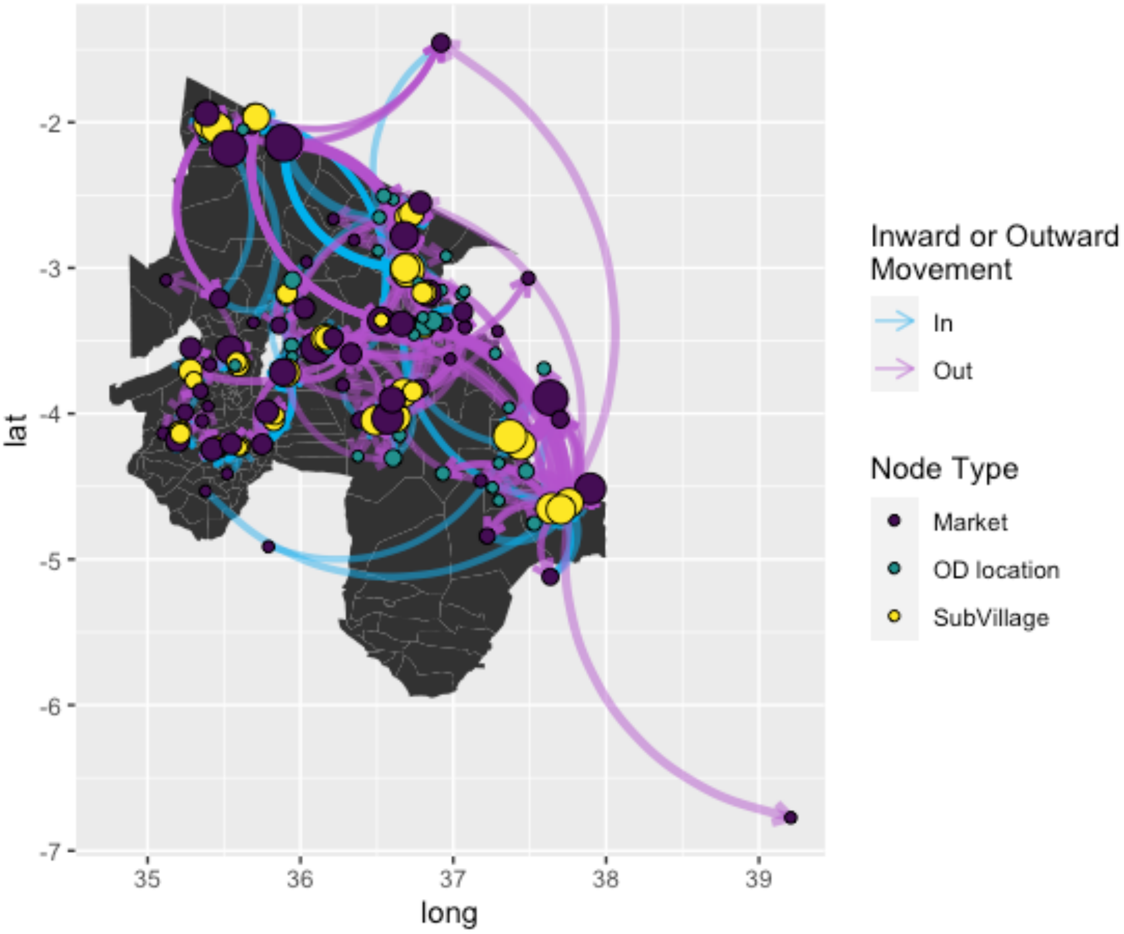


Figure 2-4: All household livestock movements from 404 surveyed households across 47 randomly selected sub-villages in northern Tanzania between January and December 2016. Movements represent reported gifts, sales and purchases to and from other households in other villages (OD locations) and market sales and purchases. Node size is proportional to its degree (busier nodes appear larger) and edge density is proportional to the total number of animals moved. Sub-village nodes are the sampled nodes. Link colour represents if the movement was reported as an inward or outward movement.

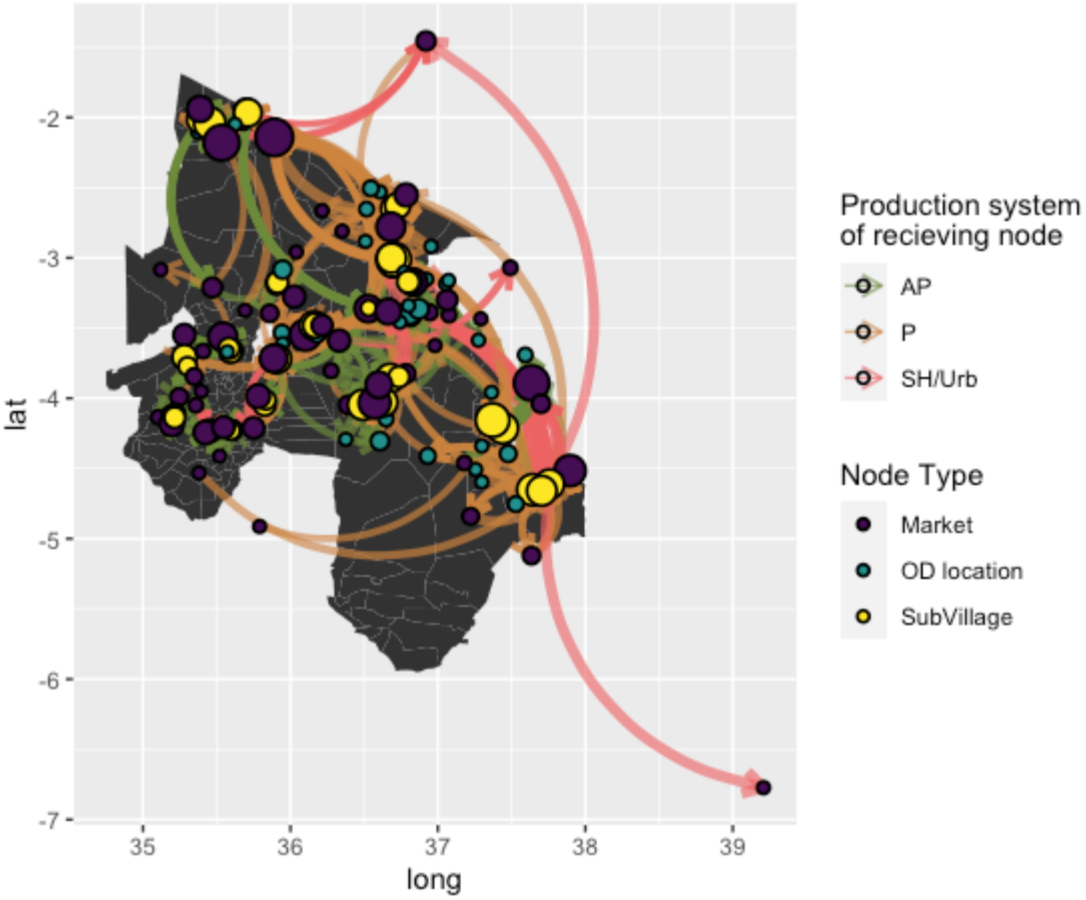


Figure 2-5: All household livestock movements from 404 surveyed households across 47 randomly selected sub-villages in northern Tanzania between January and December 2016. Movements represent reported gifts, sales and purchases to and from other households in other villages (OD locations) and market sales and purchases. Node size is proportional to its degree (busier nodes appear larger) and edge density is proportional to the total number of animals moved. Sub-village nodes are the sampled nodes. Link colour represents the agro-ecological class of the receiving node (Agropastoral (AP), Pastoral (P) or Small holder / Urban (SH/Urb))

The full network consists of 4 weakly connected components with the giant weakly connected component (GWCC) containing 137 nodes (93%). This means that, if links are bidirectional the network consists of 4 groups of connected nodes and 93% of nodes belong to the largest group. The strongly connected network, calculated assuming links are unidirectional and only travelled in the direction they are reported, consists of 10 connected components. The largest of the strongly connected components, the giant strongly connected component (GSCC), contains 29% of nodes (N=43). The size of the GSCC shows that despite sampling a small sub-sample of the underlying network and having missed some number of links, 30% of locations identified by this data are connected via a directed link and 93% by an undirected link.

The reciprocity of the network, which measures the proportion of links that are reported in both directions is 0.35, so 35% of the recorded movements in and out of sub-villages have happened in both directions between the two locations in the year preceding the study. The average path length to get between any pair of connected nodes in the network is 6.5 and the diameter is 16 (shortest path between the two most distant nodes in the connected network). The diameter (N = 16) could, in reality, represent the minimum number of sequential livestock movements required to move an infectious pathogen between the two most distant locations largest connected component of the network. This stepwise chain of transmission would require the duration of time between movements in and out of each node to be long enough for a transmission event to occur between infectious and susceptible individuals within the node.

The fragmentation index (proportion of pairs of nodes that cannot reach each other) of the full network is 0.91. This means that almost 10% of pairs of nodes in the network can reach each other via a directed path. The aim of a disease control programme would be to remove some of the most central nodes to fragment the network and reduce the size of the giant connected components so that fewer pairs of nodes can reach, and thus transmit pathogen to, each other. The full inter-household movement network also shows that despite sampling from sub-villages in only 8% of wards in the study area the observed network has movements that connect 34% of wards in the study area.

2.3.4 Subset networks

The data were seperated into the three movement types (private, gift and market) and used to construct four additional networks (private, gift, market and private&gift combined) to assess the extent to which each movement type contributes to study area connectivity. The three subset networks for the three distinct movement types (private, gift and market) are plotted in Figure 2-6 (a,b,c) and summary network statistics for the four subset networks (private, gift, market and private&gift combined) are given in Table 2-4). Private movements created the smallest network (48 nodes and 66 links). This network alone was completely fragmented (fragmentation index 1.00) with no GSCC and only 5 nodes (10%) in the GWCC. On its own this network might not be considered a cause for concern with regards to a rapidly spreading pathogen in an epidemic scenario, as these movements do not connect multiple locations. The movements do still pose a risk of introducing endemic pathogens that transmit slowly between individuals into populations. Market movements created the largest subset network (108 nodes and 669 links) with the largest GSCC (8 nodes (7%)) and GWCC (79 nodes (73%)). The gifting network sits between these two with a GSCC containing 3 nodes (5%), three other SCC each containing two nodes and a GWCC containing 33 nodes 52%).

The private&gift network was constructed to evaluate how these movements contribute to study area connectivity when combined, in case interventions cannot be implemented in a way that accounts for the differences between these movement types. In combination the gifting and private movements connect more nodes and wards than either movement type alone. This shows that these inter-household movements are fundamentally different when we consider the risk of pathogen transmission through these movements, because they create links with different locations. The combination of the gifting and private movements does not however create a network with a connected component as large as the one created through market movements. The fragmentation index of the market movements only network is 0.98 which is higher than the fragmentation index of the full network (0.91) and shows that only 2% of pairs of nodes in this inter-household market network can reach one another via a directed path. This indicates that if inter-household market movements were viewed alone, fewer nodes would have to be removed to fragment the network and disrupt pathogen transmission. When all inter-household movements are viewed in combination however, they reveal the additional links that exist between nodes that are not centred around livestock markets but could provide routes for pathogen transmission.

The four subset networks (private, gifting, market and private&gift) all consist of more smaller components compared to the full network, and are all closer to complete fragmentation. Despite being less connected than when assessed in combination the subset networks still include livestock movements to and from 11%, 17%, 27% and 22% of study area wards respectively when data were only collected from 8% of wards in the study area. Further details and links to the code for network analysis can be found in the Appendix section 9.1.1 (HH_NW_4 and 5).

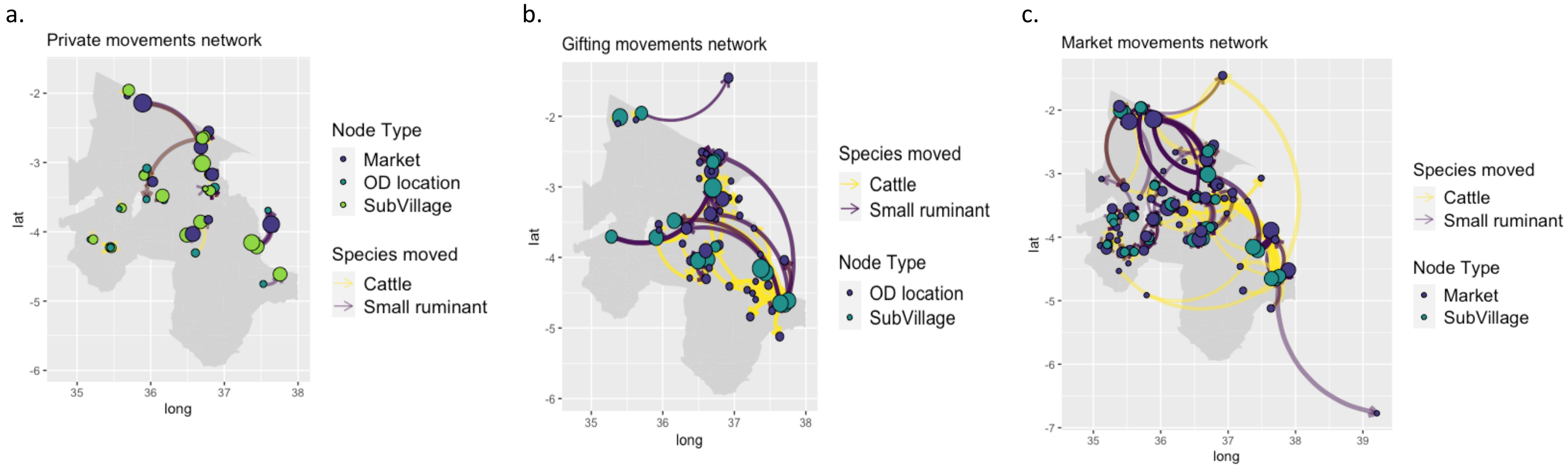


Figure 2-6: Subset networks showing connectivity across study area and beyond via private (a), gift (b) and market (c) movements of livestock. Node coloured by type and node size proportional to its degree. Link colour represents species moved and density is proportional to the number of livestock moved. Grey opaque area are three study regions Arusha, Manyara and Kilimanjaro. Northern straight-line boundary is the international border with Kenya. Sub-village nodes are sampled nodes, market nodes are origin/destination locations containing a market and OD locations are non-sampled origin/destination locations that do not contain a market.

Table 2-4: Network Summary statistics for the full inter-household livestock movement network constructed using all movement types, the gifting movements network, the private movements network, the market movements network and the gifting and private movements combined network.

	All movements	Private movements	Gifting movements	Market movements	Gifting and private combined
Nodes	147	48	64	108	93
Links	816	66	82	668	148
Number of strongly connected components > 1	10	0	4	15	3
Size of Giant Strongly connected component	43 (29%)	1	3 (5%)	8 (7%)	5 (5%)
Number of weakly connected components > 1	4	18	10	5	12
Size of Giant Weakly connected component	137 (93%)	5 (10%)	33 (52%)	79 (73%)	51 (35%)
Diameter	16	2	4	17	5
Fragmentation index	0.91	1.00	0.99	0.98	0.997
Average path length	6.5	1.09	1.74	4.94	1.98
Study area wards listed in the network (total N = 245)	83 (34%)	27 (11%)	41 (17%)	65 (27%)	54 (22%)

2.3.5 Node centrality metrics

A statistical summary of sub-village node centrality metrics in-degree, out-degree, geometric mean degree and betweenness is provided in Table 2-5. Histograms of the four node centrality metrics distributions are given in Figure 2-7(a, b, c, d). To visualise the level of correlation between sub-village in-degree and out-degree and sub-village in-degree and betweenness these are plotted in Figure 2-7 (e and f respectively).

The distributions of sub-village in-degree, out-degree, geometric mean degree and betweenness centrality metrics are right skewed (many nodes have a low value/connections, some have many). This is a common feature of many naturally occurring networks (Newman, 2002; Büttner *et al.*, 2013a). The variance to mean ratios are greater than 1 for node in-degree, out-degree and geometric mean degree, which along with the right skewed distributions, provides evidence to show that the node metric results are over-dispersed. Out-degree is less dramatically skewed than the other measures and there are more out-going links (N = 632, μ = 13.5) than in-going (N = 184 μ = 4) on the network.

Table 2-5: Summary statistics showing the distribution of sub-village node centrality metrics in-degree, out-degree, geometric mean degree and betweeness – network edges are weighted by the number of times they are reported by each sub-village node.

	Min	lower quartile	Median	Mean	upper quartile	Max	Variance	Variance: Mean ratio
In-degree	**0**	**1.5**	**4**	**4**	**6**	**13**	**9.7**	**2.4 (>1)**
Out-degree	**0**	**7**	**11**	**13.5**	**22**	**37**	**78.4**	**5.8 (>1)**
Geometric mean degree	**0**	**3.4**	**5.2**	**6.7**	**10.6**	**19.2**	**25.2**	**3.8 (>1)**
Betweenness	**0**	**5.5e-5**	**3.9e-3**	**2.0e-2**	**2.0e-2**	**1.5e-1**	**NA**	**NA**

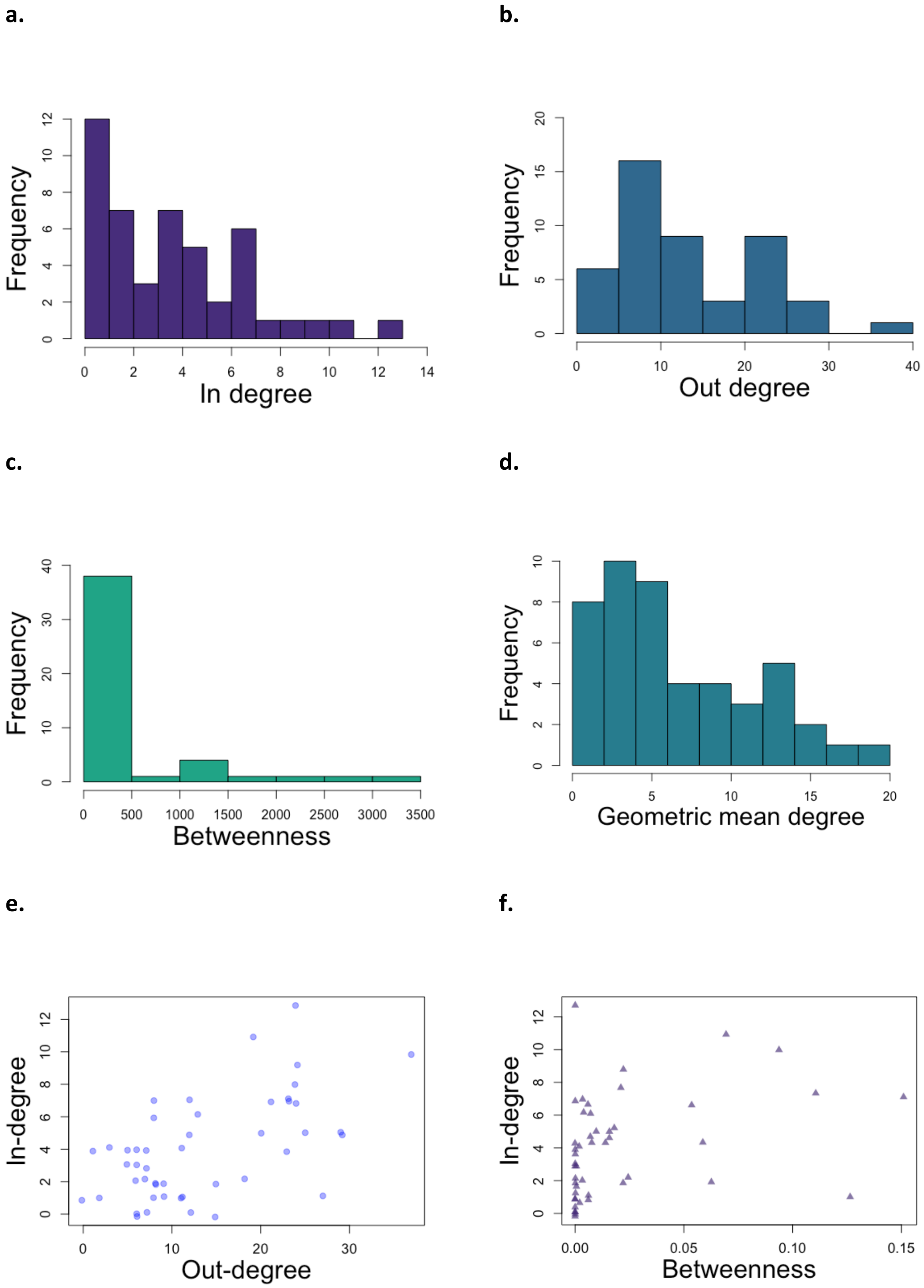


Figure 2-7: Sub-village (a) in-degree, (b) out-degree, (c) betweenness, (d) geometric mean degree, (e) sub-village in-degree plotted against out-degree (f) sub-village in-degree plotted against betweenness.

The Spearman's rank correlation coefficient was calculated to assess the level of correlation between sub-village node centrality metrics in-degree (risk of pathogen introduction), out-degree (measure of the risk of onward transmission), betweenness (risk of passing pathogen between other groups of nodes) and geometric mean degree (measures of the risk of pathogen propagation on the network). All pairwise correlations are positive (p-value <0.0001 calculated using AS 89 algorithm) for the four node centrality measures that can be used to identify central sub-village nodes, at high risk for pathogen introduction or propagation on the full network. Spearman's Rho correlation coefficients results are given in Table 2-6. In-degree and geometric mean degree are highly positively correlated but this is to some level expected as geometric mean degree is calculated in part from the in-degree (Spearman's ρ = 0.92).

Table 2-6: Spearman's Rho correlation coefficients showing pairwise correlations between four sub-village node centrality metrics. All pairwise correlation P-values less than 0.0001

	In-degree	Out-degree	Betweenness	Geometric mean degree
In-degree	-	-	-	-
Out-degree	0.51	-	-	-
betweenness	0.49	0.61	-	-
Geometric mean degree	0.92	0.76	0.64	-

2.3.6 Different methods for link weighting

The full network was reconstructed using two alternative methods to weight the links (binary so the links are present or absent and fully weighted by total number of animals reportedly moved). The additional networks were constructed to represent networks and associated node centralities that might be needed if the primary pathogens of interest for disease control were more or less transmissible than the endemic zoonotic pathogens of primary interest in this study. The centrality measure in-degree was re-calculated on the additional networks for all sub-village nodes and the three versions of in-degree (binary, partially weighted and fully weighted) were compared to assess how correlated they are. There was a high level of correlation between the three in-degree measures (Spearman's rho > 0.86 for all pairwise correlations P-value <0.0001) which suggests that, despite different link weighting methods the three measures are picking up the same information. A node identified as high risk of introducing of a highly infectious pathogen on the inter-household movement network is also at high risk of introducing a much less transmissible pathogen.

2.3.7 Data rarefication

Results from the data rarefaction are shown in Table 2-7. Summary statistics for network measures and number of wards connected via livestock movements are compared for networks constructed using all data and a random 50% and 90% of sub-village and 50% and 90% of sampled wards data. Removing 50% and 10% of the data at both sampling levels reduced all network measures. For the 50% data reduction the total number of links was halved at each level of rarefication. The total number of nodes in the network was least influenced by halving the number of sampled sub-villages (27% reduction) compared to wards (45% reduction). Both methods of removing 50% of data resulted in over 70% reduction in the size of the GSCC. Removing a random 10% of the data, at both levels also showed a relatively large reduction in the size of the GSCC (35% when data removed at sub-village level and 28% when data removed at the ward level).

Table 2-7: Network results (number of nodes, edges, size of the giant strongly and weakly connected components (GSCC and GWCC) and the number of study area wards included in the network) using; all SEEDZ household survey data (complete dataset), a randomly selected subset of data from 50% of sampled sub-villages and a randomly selected subset of data from 50% of sampled wards.

		Randomly removed 50% sampled sub-villages data		Randomly removed 50% sampled wards data		Randomly removed 10% sampled sub-villages data		Randomly removed 10% sampled wards data	
	All data	Min – Max (mean)	Mean % reduction	Min – Max (mean)	Mean % reduction	Min – Max (mean)	Mean % reduction	Min – Max (mean)	Mean % reduction
Nodes	147	86 – 100 (93)	27%	63 – 92 (81)	45%	124-139 (135)	8%	127-141 (136)	7%
Edges	816	371 – 458 (413)	49%	281 – 470 (409)	50%	673-769 (724)	11%	681-789 (740)	9%
GSCC	43	6 – 30 (10)	77%	7 -35 (12)	72%	8-43 (28)	35%	10-43 (31)	28%
GWCC	137	42 – 89 (64)	53%	18 – 84 (57)	58%	103-130 (122)	11%	108-133 (124)	9%
Wards	83	55 – 68 (61)	27%	39 – 59 (49)	41%	70-82 (79)	4%	73-81 (78)	6%

2.4 Discussion

2.4.1 General

The data described in this study show that inter-household livestock movements such as private purchase, gifts and market sales or purchases provide a largely undocumented level of epidemiological connectivity across northern Tanzania and beyond. The network of inter-household livestock movements provides a route for pathogen transmission between both geographically close (minimum 400m) and distant (maximum 294km) locations, including across the international border with Kenya. We show that disease control interventions in northern Tanzania should be discussed at a regional level if they are to avoid being undermined by movement activity and epidemiological connections that remain unaccounted for. The documented movements to and from Kenya provide a reminder that the international border here will not act as boundary to pathogens, as has been shown here and in other countries (Little, 2009; Dean *et al.*, 2013; Motta *et al.*, 2017).

The relatively high proportion of nodes contained in the giant components of the observed network (GSCC 30%, GWCC 93%) show that inter-household movements are responsible for creating epidemiological connections between a large proportion of locations. The minimum number of sequential movements taken to get between the two most distant locations on the observed network was 16. If more data gathered from other villages in the study area increased the density of network connections, which is what the data rarefication analysis suggests, the number of sequential movements required to cross the network would likely decrease. Either way, this number indicates that pathogens could be transmitted relatively quickly across the study area via inter-household movements if we consider the fact that livestock market movements were the most frequently reported and markets are active up to twice a week.

When viewed alone the three inter-household movements types (gifting, private and market) do not create particularly large or well-connected network graphs and the private network alone is completely fragmented. Alone therefore, the different movement types might not raise cause for concern with policy makers interested in developing pathogen control programmes, particularly when considering rapidly transmissible pathogens with epidemic potential. They may still however provide enough movement between populations to allow pathogens with low R_0 values to persist (Prentice *et al.*, 2017). Conversely, when the movements are combined into a single network, which we consider to be appropriate here as pathogens do not differentiate between the reasons for livestock movement, the network becomes relatively well connected (93% of nodes are connected to one another via an undirected link, 30% via a directed link). This means that the observed network and underlying population is susceptible to epidemic outbreaks of more infectious and rapidly transmissible pathogens (Kao *et al.*, 2006; Prentice *et al.*, 2017). This is observed despite the network data being collected from a very sparse sample of the underlying population (2% of sub-villages).

Although the observed network is well connected it does appear have a structure that lends itself to relatively easy fragmentation (FI 0.91) by targeted node removal, provided any data from additional unobserved nodes is similar to data from the sampled nodes. When 50% and 10% of the data are removed the network becomes much less connected and the giant components (GSCC and GWCC) shrink significantly. This shows that, if observed data are representative of the underlying network, few nodes would need to be removed to break apart the network to prevent pathogen transmission across it. The heterogeneity and right skew of sub-village node centrality measures (in-degree, out-degree, betweenness and geometric mean degree), in addition to the positive correlation between the different measures, shows that some high risk sub-village nodes can be identified and targeted for removal to fragment the network with minimal disruption (Marquetoux *et al.*, 2016; Chami *et al.*, 2017). The positive correlation between the different centrality measures shows that sub-village nodes at high risk of pathogen introduction are also high risk for onward pathogen transmission on the network (Rautureau, Dufour and Durand, 2011; Molia *et al.*, 2016; Palisson *et al.*, 2016). The high level of positive correlation between the different weighted networks (binary, partial and fully weighted) in-degree measures also show that interventions that target pathogens of varying transmissibility can be efficiently implemented at the high risk nodes identified on this network. The high level of correlation between all node centrality metrics evaluated here will make identifying high risk locations for implementing disease control interventions relatively straightforward. This is useful information for policy makers as it can be used to ensure harmonised disease control programmes are developed that focus on multiple pathogens, in the highest risk locations, for optimal cost efficiency and resource use.

Practically, node removal could be achieved by implementing movement restrictions, multi-pathogen active surveillance or vaccination programmes or knowledge exchange exercises so livestock keepers are aware of the risks associated with livestock introductions (Cowan and Jonard, 2004; Büttner *et al.*, 2013a; Vallée *et al.*, 2013; Colman *et al.*, 2019; Ferdousi *et al.*, 2019; Hidano, Gates and Enticott, 2019). Identifying the most central nodes in the study area, before being faced with a large-scale disease outbreak, will allow policy makers to be one step ahead in preparation for such an eventuality. Having established early warning surveillance platforms in the high risk locations will aid in the control of pathogens such as Rift Valley fever (Sindato, Karimuribo and Mboera, 2012) and peste des petits ruminants (Herzog *et al.*, 2019) that frequently plague the region causing significant socio-economic losses (Halliday *et al.*, 2012; Muma *et al.*, 2014). It will also be beneficial for surveillance of any new emerging livestock pathogens. Targeted control efforts are an efficient way to spend resources which is particularly important here in northern Tanzania where resources are scarce, funding for livestock disease control is limited and the burden of pathogens are high (Coker *et al.*, 2011; Rushton, Uggla and Magnusson, 2017). Targeted control measures can also be less disruptive and have less overall socio-economic consequences compared to blanket interventions such as movement bans in the face of a disease outbreak (Tildesley *et al.*, 2019).

Household livestock exits, that could result in onward pathogen transmission events, were reported more frequently (77% of reported movements) than introductions and network node centrality metrics show that out-degree was less skewed than other centrality metrics. This is likely to be a consequence of the broadly important and more frequent need for households to sell livestock to generate income (generating out-going edges), compared to the relatively rare practice of introducing livestock when an excess of cash capital is available. This finding shows the economic importance of animal movements at the household level and highlights that the drivers for these movements need to be carefully considered when designing disease control programmes (Haseeb *et al.*, 2019). Household needs and family responsibilities will still exist in the face of any livestock disease outbreak. A sweeping outright ban on livestock movements or market closures will remove a primary method used to generate household income in northern Tanzania and thus could be met with non-compliance and key stakeholder disengagement. The high reciprocity (35%) identified on the network indicates that there are factors involved in livestock movement decision making that encourage people to move animals between the same locations. Movements could have been to and from up to 900 villages in the study area and many more villages beyond this. The drivers that result in people from the same sub-villages often choosing to move livestock in both directions between the same locations within a year could be family or friendship ties or more economically focussed such as proximity to livestock markets. Regardless of the reason for high reciprocity on the network this is useful to have identified. The reciprocity could be taken advantage of to reinforce disease awareness message campaigns about risk of movements and pathogen transmission as information can flow between nodes along with livestock as people from each node interact. This could also be beneficial when implementing vaccination programmes as the high rate of livestock exchange between connected nodes (excluding markets) could ensure vaccine coverage does not drop, provided high risk vaccinated nodes largely connect to other vaccinated nodes. Harmonising disease control efforts to target multiple pathogens in locations identified as high risk for pathogen introduction and onward transmission, with an aim of reducing overall transmission rates and disease burden appears to be a sensible option for northern Tanzania.

The model fitted to estimate household movement distances showed that cattle movements generally extend about 20% further than small ruminant movements and gift movements are generally longer distance than market or private sale and purchase movements. This could be due to the fact that commercial movements have a lower opportunity cost if they are shorter whereas the motivations (family, clan, friendship ties) for gifting movements are not constrained by economics. The longer distances of gifting movements and their potential lack of economic predictability, again highlights the importance of considering these movements when developing disease control programmes. From an economics perspective we might also expect longer journeys to be made by larger batches of animals so that the fixed journey costs are shared and each animal journey then has a lower marginal cost (Tisdell and Adamson, 2017). The distance GLMM showed that there was no evidence to support this economic assumption in the household survey data. This finding highlights that there are motivations beyond economics for the inter-household livestock movements. These movements differ from commercial movements, with different drivers and so they are unlikely to be well predicted by economic models.

The production system of the giving and receiving nodes had the largest effect on estimated movement distance with the longest movements being those to and from small holder or urban locations (high livestock and human population density(de Glanville *et al.*, 2020)) from agropastoral or pastoral nodes (less densely populated, larger herd and flock sizes(de Glanville *et al.*, 2020)). These long-distance movements to and from more densely populated areas could explain the slightly bimodal nature of the movement distance distribution shown in Figure 2-2. It also supports findings from other studies that show livestock generally move from sparsely populated rural areas of production into densely populated areas for consumption (Motta *et al.*, 2018; Napp *et al.*, 2018). Pastoralist production systems have previously been associated with higher pathogen prevalence (Herzog *et al.*, 2019; Bodenham, Lukambagire, *et al.*, 2020). Identifying the existence of long-distance movements between the different production system types highlights that geographical distance from a high prevalence area will not render a node, and thus its population, immune to pathogen introduction. Ultimately it is important for public health officials and policy makers to be aware of these long-distance movements because, in the case of a disease outbreak, a pathogen could be rapidly moved into an area of high livestock and human population density from a distant location, without detection (Napp *et al.*, 2018).

Through the data rarefication process, we show that total number of nodes connected in the network was least influenced by removing data at the sub-village level compared to the ward level. This indicates that sub-villages within wards/villages, which are more spatially clustered, are more similar to each other and connect to similar locations. For future sampling to identify highly connected locations it would be more efficient to sample from different villages or wards rather than from multiple sub-villages within wards. Additionally, for the design and implementation of disease control programmes we can consider sub-villages within villages/wards to be more similar to each other. Sub-villages within villages/wards will thus benefit from similar interventions if these are targeted based on disease risk as a result of connectivity through inter-household livestock movements. Interventions can therefore be designed at the less granular spatial administrative level of the ward or village which is more practical and feasible compared to targeting sub-villages.

The data rarefication process also shows a less dramatic reduction in the number of wards listed in the 50% and 90% subsampled networks compared to the overall number of nodes. This finding shows that, if the observed network was constructed at the larger spatial scale of the ward, we would observe a more densely connected network with a number of redundant links. This data could therefore be used to calculate centrality measures of wards to identify the most influential wards at greatest risk of pathogen acquisition or onward transmission through household livestock movements in the study area. This is an important policy relevant point to note as practically the ward is likely to be the smallest administrative unit that a control programme is implemented at. If that is the case, further analysis is needed to identify key wards and thus locations to target control efforts. Identifying the most highly connected wards via inter-household movements would also allow these to be compared to the most central high risk wards that were identified using livestock movement permit data from the study area (Chaters *et al.*, 2019). Practically if the same nodes are identified as high risk this will make targeting pathogen control programmes even more efficient. If different wards are important for the different movement types, then careful consideration of the importance of each network layer will be needed.

In addition to the permit and household movement network layers mentioned above, there will be other types of movement or interaction between nodes in the study area, such as the lending of animals or the sharing of natural resources, that create epidemiological links. This could be particularly important for pathogens with indirect transmission routes such as *Coxiella Burnetii* (environmental contamination) or Rift Valley fever (mosquito transmission). We know that when multiple network layers are combined to create a multiplex with multiple routes available for pathogen transmission between nodes the minimum disease epidemic threshold is dramatically reduced compared to what is calculated on a single network layer (Zhao *et al.*, 2014; Zhang *et al.*, 2015). Having this network of inter-household livestock movements documented means that this data can be combined with other network layers from northern Tanzania in future to create a multiplex network of livestock movements and epidemiological connections (Chaters *et al.*, 2019). A multiplex network could then be analysed to identify high risk locations based on the multiple different types of livestock movement and epidemiological connections.

Data rarefication also shows that data saturation is unlikely to have been reached, and the observed network does not have a surplus of redundant links. Collecting more data from wards in the study area, or extrapolating up from our data, could further increase the observed level of connectivity via inter-household movements in the study area. Additional high-risk nodes might also be identified if sampling was more extensive. Care should be taken however to avoid unnecessary oversampling which would waste time and resources and further contribute to research fatigue. If additional data was collected it should be added to the network in a stepwise process and the network re-assessed, with random data removal, to evaluate at each stage if data saturation or close to it has being reached (no change in nodes centrality ranking and redundant links appearing). Through the data rarefication we have shown that the current data is unlikely to have identified the complete underlying network but we know that the value of adding new information to a network diminishes as the high risk nodes are identified (Colman *et al.*, 2019). Chami et al found using an acquaintance algorithm (Chami *et al.*, 2017) to be a more efficient method of network construction and identification of the most central nodes rather than oversampling from many more locations and wasting peoples' time. This method identifies the most connected locations by sampling from random locations and making the assumption that the most connected nodes in any network are likely to be those that are mentioned by the randomly selected others. Applying this method we could trial interventions such as active surveillance at the most central non-sampled nodes in the observed network and assess if removing these nodes efficiently fragments the underlying network. Given that the majority of household movements reported in this data are market movements this means that livestock markets are likely to be important non-sampled locations contributing significantly to epidemiological connectivity across northern Tanzania and beyond. The role of livestock markets in connectivity across Tanzania needs to be investigated further to evaluate if high risk locations can be identified and used as targets for implementation of efficient livestock disease control and surveillance programmes.

To improve the findings from this study and further understand the true impact of livestock movements on pathogen transmission risk in the study area, longitudinal data that details the timing of movements and disease incidence would also be valuable. This data could be used to construct a dynamic epidemiological transmission network model on which different intervention scenarios could be simulated to identify the most efficient type of interventions for maximum reduction of pathogen burden (Barnard *et al.*, 2018; Enright and Kao, 2018).

2.4.2 Data Limitations

The SEEDZ household survey was very detailed and took up to five hours to complete, with the questions regarding livestock husbandry and livestock movements asked towards the end of the survey. There is the potential for some interviewee and interviewer fatigue which could have resulted in a failure to recall all livestock movements in the previous year (reporting bias). The survey also asked the questions in relation to household livestock movements over the previous 12 months which in itself leads to the potential for under reporting due to relying on good memory/recall from the interviewee. Furthermore, the topic of livestock numbers and introductions is a sensitive subject in the study area because livestock numbers equate directly to wealth. For this reason, numbers of introductions could have been underreported or withheld. The location that animals are received from and sent to is a less sensitive subject which potentially made the inclusion of network measures as risk factors a more robust measure of livestock movement activity compared to absolute counts of livestock.

The movement permits that are supposed to be used to record all permanent livestock movements between two locations have a movement tax associated with them. The movement tax payable depends on the administrative boundaries crossed by the movement. An increasing tax is due if the movement crosses a ward, district or regional boundary (all administrative units increasing in size from ward to region). The person moving the livestock is responsible for getting the movement permit and paying the movement tax. This movement tax could be a major reason for underuse of permits and thus recording of livestock movements as people try to avoid paying it and thus avoid registering their livestock movements.

2.5 Conclusions

This study shows that inter-household movements of livestock add a layer of epidemiological connectivity in northern Tanzania that is otherwise undocumented. These movements link locations up to 300km apart and cross international boundaries. Alone and in combination with other types of livestock movements, inter-household movements should be a cause for concern for those charged with developing pathogen control programmes. The majority of reported movements were to and from markets but few households actually recorded these movements by using a government permit. Perhaps a method of movement reporting, that has minimal financial cost could be considered by the government if they can prioritise disease surveillance and control above tax collection.

The observed inter-household movement network exhibits properties that suggest it is susceptible to rapid fragmentation, by identification and removal of the most central nodes through targeted disease control interventions. This could be done via increased surveillance, vaccination or temporary movement restrictions. A high level of heterogeneity and right skew of the sub-village node centrality metrics and positive correlation between node centrality metrics means that targeted node removal to fragment the network could be done efficiently with minimal disruption to other livestock activity. This is an important finding as livestock movements are of vital socio-economic importance in northern Tanzania. The high level of correlation between in-degree centrality metrics calculated on the inter-household network using different methods to weight links (binary links weighted based on how one might represent risk of transmission of a rapidly transmissible pathogen such as FMDV and fully weighted links based on how one might represent risk of transmission of a less transmissible pathogen such as bovine Tuberculosis) show disease control programmes could target multiple pathogens of varying transmissibility in high risk locations. As society is moving towards improved efficiency of livestock production, it is important to identify where targeted interventions can be implemented, that do not require unnecessary culling of animals (wastage) or excessive drug use (pollution and increased resistance risk). A multi-pathogen approach focused on reducing overall pathogen burden in the livestock populations by reducing transmission between and within populations by regulating and reducing livestock movements appears to be a sensible and feasible option in this area. A control option that is focused on livestock movement with the potential to improve productivity and sustainability, while not costing as much financially, environmentally or socially as movement bans or drug use programmes should be considered in northern Tanzania. Finally however, if undocumented inter-household movements epidemiologically link regions beyond this study area, as this data suggests, we urge discussion of disease surveillance and control programmes at a regional scale to prevent all efforts being undermined.

3 Livestock introductions and the risk of infectious pathogen exposure in northern Tanzania

Contribution

Existing serological data from the SEEDZ and The Impact and Social Ecology of Bacterial Zoonoses in Tanzania (BacZoo) project, along with survey data from both studies were used in this chapter. Further details on these studies are given within the chapter. The movement specific research questions that could be answered utilising available qualitative and serological data, building on analysis results from the previous chapter in this thesis were conceptualised by GC. Data cleaning, processing and analysis, including choice of methods for analysis, were done by GC.

3.1 Introduction

Livestock movements create epidemiological links between locations and populations which provide direct routes for pathogen movement between populations (Fèvre *et al.*, 2006). Zoonotic pathogens such as *Brucella* spp., *Coxiella* and *Leptospira* spp. can be transmitted directly and indirectly between livestock (cattle and small ruminants) and from livestock to people. Close contact between people and livestock is identified as a major risk factor for zoonoses (Allan *et al.*, 2018; Maze *et al.*, 2018; Bodenham, Lukambagire, *et al.*, 2020; Johnson *et al.*, 2020). Zoonotic pathogens not only threaten human health directly but are also estimated to be responsible for 50% of global livestock losses by reducing livestock productivity (Molyneux *et al.*, 2011; World Bank and TAFS Forum, 2011). Thus, zoonotic pathogens, along with other production-limiting pathogens such as bovine viral diarrhoea virus (BVDV)(Lanyon and Reichel, 2013; Okumu *et al.*, 2019) and Bovine Herpes virus-1 (BHV-1)(Ata *et al.*, 2006; Mineo *et al.*, 2006; Mahmoud and Allam, 2013) cause reduced livestock productivity through impaired growth rates and milk production and excess reproductive and mortality losses. Poor livestock productivity has extensive negative human and livestock health and socio-economic consequences globally, particularly in communities that are heavily dependent upon livestock for socio-economic and nutritional security (Ashworth and Mainland, 1995; Kossaibati and Esslemont, 1997; Thompson *et al.*, 2002; Tomley and Shirley, 2009; Hendrik I J Roest *et al.*, 2011; Rich and Perry, 2011; Bouley *et al.*, 2012; Sindato, Karimuribo and Mboera, 2012; Vanderburg *et al.*, 2014; Wensman *et al.*, 2015; Lankester *et al.*, 2015; Njeru *et al.*, 2016; Guidoum *et al.*, 2020). The economic impacts of livestock diseases have long been recognised as keeping people in marginalised communities trapped in poverty (Steinfeld *et al.*, 2006; Perry and Grace, 2009; Rich and Perry, 2011). In a world with an expanding population, over 2 billion people living on the margins, prevalent macro and micronutrient deficiencies, environmental pressures and climate instability, there is a cumulative need to improve livestock productivity to ensure people have access to adequate nutrition (Upton, 2004; FAO, 2009; Nair, Augustine and Konapur, 2016; Myers *et al.*, 2017; Rushton *et al.*, 2018). Reducing the burden of livestock diseases is an obvious first step towards improving the efficiency of livestock production and global food security (WHO *et al.*, 2005; Hotez *et al.*, 2009; Molyneux *et al.*, 2011; Grace *et al.*, 2012; Muma *et al.*, 2014; Pradère, 2014).

In Tanzania a large proportion of the population depends on livestock as a source of income, food security and social status (Covarrubias *et al.*, 2012; Davis and Sharp, 2020). Resource-poor communities in Tanzania are also increasingly reliant on small ruminants for food security because these are considered more drought resilient and cheaper to replace (Smith *et al.*, 2004; Wodajo *et al.*, 2020). The reliance on livestock is accompanied by a high prevalence of infectious, including zoonotic, pathogens in the cattle and small ruminant livestock populations (Hyera, Liess and Frey, 1991; Machang'u, Mgode and Mpanduji, 1997; Karimuribo *et al.*, 2007; Schoonman and Swai, 2010; Mdetele and Kassanga, 2014; Alonso *et al.*, 2016; Mathew *et al.*, 2017; Herzog *et al.*, 2019; Semango *et al.*, 2019). This results in poor productivity and high levels of income uncertainty and food insecurity with up to 40% of household income lost through small ruminant drought and disease losses alone (Sherman, 2011; Covarrubias *et al.*, 2012; Haseeb *et al.*, 2019). Additionally, in the case of zoonoses, spill-over of pathogens from livestock to humans leads to illnesses that are often misdiagnosed and incorrectly treated (Biggs *et al.*, 2011; Crump *et al.*, 2013; Cash-Goldwasser *et al.*, 2018). Close contact with small ruminants has been identified as a significant contributor to the transmission of zoonotic pathogens *Coxiella burnetii* and *Brucella spp.* to people in Tanzania (Van den Brom *et al.*, 2015; Bodenham, Lukambagire, *et al.*, 2020). In addition to living with zoonotic pathogens, Tanzanian livestock lack routine pathogen testing, vaccination or control programmes for many endemic or production-limiting infections which is compounded by inadequate access to veterinary services (Komba *et al.*, 2012; Herzog *et al.*, 2019; Hughes *et al.*, 2019). In some European countries the detrimental effects of the production-limiting pathogen BVDV are considered so great in cattle that disease elimination campaigns have been established (Stott *et al.*, 2012; Metcalfe and Metcalfe, 2019). Prior work on the serological samples used in this study (Hodgkinson, 2017) suggests that BVDV and BHV-1 are circulating at a high prevalence in northern Tanzania compared to the zoonotic bacterial pathogens, *Brucella*, *Coxiella* and *Leptospira* (Hummel, 1976; Alonso *et al.*, 2016; Shirima and Kunda, 2016). There is variability in the shedding cycles of these five pathogens; *Brucella* and *Coxiella* can be shed cyclically after parturition, *Leptospira* and BHV-1 are shed intermittently and BVDV is shed continuously by persistently infected animals. These pathogens also utilise different transmission routes to spread between infectious and susceptible individuals. Close (nose to nose) contact between an infectious shedding animal and a susceptible can result in transmission events for all, *Leptospira* and *Coxiella* can be transmitted via contaminated environment and BHV-1 can utilise respiratory and sexual transmission routes (Brownlie *et al.*, 1987; Angelakis and Raoult, 2010; Poester, Samartino and Santos, 2013; Allan, 2016; Kipyego *et al.*, 2020). Regardless of the disparities in transmissibility between BVDV, BHV-1 and the zoonotic pathogens *Brucella*, *Coxiella* and *Leptospira*, they are all circulating and causing reduced livestock productivity while remaining largely unscreened for across northern Tanzania (Brackenbury, Carr and Charleston, 2003; Ata *et al.*, 2006; Chase, 2013; Okumu *et al.*, 2019).

Infection with BVDV in cattle can cause immunosuppression (Brownlie *et al.*, 1987; Chase, 2013). Coinfection between BVDV and bacterial zoonoses has been documented in cattle in neighbouring Kenya (Okumu *et al.*, 2019) but the potential immunosuppressive effects of belonging to a BVDV herd on cattle's risk of exposure to other infectious pathogens in Tanzania is unknown (Brackenbury, Carr and Charleston, 2003; Lanyon and Reichel, 2013). If BVDV presence in a herd enhances the risk of exposure to and infection with other endemic pathogens and the immunosuppressive effects make cattle more likely to shed other pathogens, this should be considered when designing livestock disease control programmes. This is particularly important if control programmes are to include vaccination campaigns that rely upon a healthy and responsive immune system in vaccinated livestock.

In low-resource settings, such as northern Tanzania, livestock disease control programmes need to be streamlined and targeted to make the control of endemic, zoonotic and emerging pathogens both effective and sustainable (Heffernan and Misturelli, 2000; Zinsstag *et al.*, 2007; Maudlin, Eisler and Welburn, 2009; Hop *et al.*, 2011). There is a need to identify common risk factors for pathogen exposure and whether these risk factors exist at the individual, household or wider population level, so that pathogen control programmes can be applied at an appropriate scale. Livestock movements and introductions have been identified as major risk factors for disease transmission in other populations, but the association between livestock movements and disease risk has received limited previous investigation in northern Tanzania (Gibbens *et al.*, 2001; Wilson and Mellor, 2009; Nöremark *et al.*, 2011; Segura-Correa *et al.*, 2016; Byrne *et al.*, 2017; Kaddour *et al.*, 2019; van Roon *et al.*, 2020).

Knowledge of livestock movements can help trace pathogen spread through a population and be used to slow, disrupt or prevent transmission (Gibbens *et al.*, 2001; Finn *et al.*, 2019; Porphyre *et al.*, 2020). Livestock movement data can be collected in multiple ways and can be used in combination with epidemiological data to construct contact networks (Wongsathapornchai *et al.*, 2008; Nicolas *et al.*, 2013; Selby *et al.*, 2013; Buhnerkempe *et al.*, 2014; Motta *et al.*, 2017; VanderWaal, Gilbertson, *et al.*, 2017; Apolloni *et al.*, 2018). Contact networks can then be analysed using concepts from NA to assess the effects of livestock movements on pathogen transmission and inform effective disease surveillance and control strategies (Kiss, Green and Kao, 2006; Gates and Woolhouse, 2015b; Motta *et al.*, 2017; Ruget *et al.*, 2021).

In northern Tanzania there are multiple, complex socio-economic and environmental drivers for livestock movements and contacts between herds (Covarrubias *et al.*, 2012; Goldman and Riosmena, 2013). Data on livestock movements or identification (such as the recent national cattle branding programme) are limited and not readily available (not digitised and/or centralised). Predominant livestock movement types include: movements to access shared natural resources such as grazing, water and salt points (Coppolillo, 2000); trade movements to, between and onward from markets (Chaters *et al.*, 2019); and movements between households. Inter-household livestock movements occur for many reasons, including the giving and receiving of gifts, payment for services, lending and private sales and purchases. Inter-household movements are not well captured by government movement permits, which are used most frequently for movement of livestock batches onwards from markets (Chaters *et al.*, 2019) (Chapter 1), so they remain undocumented at regional and national levels. Increasing the available knowledge on the role of inter-household livestock movements on pathogen exposure will help to identify high risk behaviours for both pathogen transmission and emergence and identify what and where disease control programmes are needed (Bigras-Poulin *et al.*, 2006; Kiss, Green and Kao, 2006; Vanderwaal *et al.*, 2016). As stated above, a targeted approach to pathogen control would be most useful in northern Tanzania where pathogen burden and the risk of emerging diseases are high but resources to implement broad scale control measures are scarce (Maudlin, Eisler and Welburn, 2009; Colman *et al.*, 2019).

The objective of this study was to assess the relationship between livestock introductions and exposure to three zoonotic pathogens, *Brucella* spp. and *C. burnetii* in cattle and small ruminants and *Leptospira* serovar Hardjo in cattle, and two non-zoonotic, production-limiting pathogens, BHV-1 and BVDV, in cattle. Using seroprevalence to gauge exposure, livestock introductions at the household and sub-village levels (administrative sub-unit of a village), and sub-village centrality metrics in-degree and betweenness (calculated from the inter-household movement network in Chapter 2), were evaluated to see if these were associated with individual animal's risk of pathogen exposure. In-degree quantifies the number of livestock batches (at least one animal) from external locations a sub-village received in the year preceding the study. Betweenness quantifies the number of times a sub-village lies on the shortest path (acting as a bridge) between other pairs of location nodes in the connected network. The hypothesis tested here is that belonging to a household and/or a sub-village with high numbers of livestock introductions, or many epidemiological links from and between other locations via livestock movements (high in-degree and or betweenness centrality) is associated with an increased risk of pathogen exposure. Assessing the effects of introductions at the sub-village level allowed for the possibility that the risk of an introduction into a sub-village herd or flock would have consequences for all livestock in the sub-village due to between-herd mixing at shared resources (VanderWaal, Gilbertson, *et al.*, 2017). Belonging to a cattle herd with evidence of prior BVDV exposure was also assessed to see if this is associated with an increased risk of those cattle being exposed to other pathogens.

3.2 Methods

3.2.1 The study area

This study was conducted in three regions Arusha, Manyara and Kilimanjaro in northern Tanzania. Each region has 6 (Manyara) or 7 (Arusha and Kilimanjaro) districts, and within each district there are wards (an administrative unit of around 10,000 people). Wards are divided into villages (mean 3, median 2, range 1-19 per ward), and villages further sub-divided into smaller administrative units called sub-villages (2-5 sub-villages per village). Livestock-keeping households in the study area have been broadly categorized into three agro-ecological (A-E) classes: pastoral, agropastoral and smallholder (de Glanville *et al.*, 2020).

3.2.2 Data collection

3.2.2.1 Household survey data

Data were collected as part of two cross-sectional household surveys carried out in overlapping study areas in northern Tanzania. The two projects that generated the datasets used in this analysis are the Social, Economic and Environmental Drivers of Zoonoses (SEEDZ) project and the Impact and Social Ecology of Bacterial Zoonoses in Tanzania (BacZoo) project. Data were collected from a total of 65 sub-villages (47 SEEDZ, 18 BacZoo) across the study area (Figure 3-1).

3.2.2.1.1 SEEDZ

The primary data used for all parts of this analysis was generated from the SEEDZ research project (further details Chapter 2). Data used in this study included livestock numbers and numbers of introductions and exits, including the locations animals came from or went to. Location livestock were moved to or from, along with the number of livestock moved, were aggregated at the sub-village level and used to construct an inter-household movement network in Chapter 2. Sub-village node metrics in-degree and betweenness, calculated in Chapter 2, are assessed as potential risk factors for pathogen exposure in this analysis.

3.2.2.1.2 BacZoo

The BacZoo research project was carried out in the Kilimanjaro and Arusha Regions between September 2013 and March 2015. Multi-stage sampling was used to randomly select 18 wards from districts within the study areas (lists taken from national census data). One sub-village was then randomly selected from each ward and a minimum of five livestock-keeping households were randomly selected from within the sub-village for participation in the study. Data were collected from a total 121 households from the 18 sub-villages. Further details on the BacZoo project, study design and implementation can be found in (Bodenham, Mazeri, *et al.*, 2020). An important difference to note between the studies is that the BacZoo data does not specify a location from where livestock were acquired or sent to. Because there is no associated location data it could not be used to construct a movement network thus there are no sub-village network associated (in-degree and betweenness centrality) variables to test with these data.

Figure 3-1: Map of the three study regions (Arusha, Manyara and Kilimanjaro) in northern Tanzania with SEEDZ and BacZoo study sub-villages shown. Lines demarcate ward boundaries.

3.2.2.2 Serological data

Up to a maximum of 10 cattle, 10 sheep, and 10 goats for SEEDZ and 15 cattle, 15 sheep, and 15 goats for BacZoo (fewer if the household owned fewer than 10 or 15 respectively) were randomly selected from each participating household and a 10-ml blood sample taken for serological testing. All SEEDZ and BacZoo livestock serum samples were tested for the presence of antibodies to the zoonotic pathogens *Brucella spp*. (competitive ELISA, COMPELISA, APHA Scientific, UK) and *Coxiella burnetii* (Indirect ELISA, LSIVet Ruminant Q fever Serum/Milk ELISA kit, Thermofisher, USA). All study samples were tested in the same laboratory, Kilimanjaro Christian Research Institute (KCRI), Moshi, Tanzania. The SEEDZ cattle samples were also tested for the presence of antibodies to *Leptospira borgpetersenii* serovar Hardjo and/or *Leptospira interrogans* serovar Hardjo (*Leptospira* serovar Hardjo) (double sandwich ELISA, Linnodee Leptospira kit, Linnodee Animal Care, Ballyclare, Northern Ireland), bovine herpes virus 1 (BHV-1) (IBR Individual Antibody TEST, IDEXX test kit) and bovine viral diarrhoea virus (BVDV) (BVDV p80 Ab ELISA, IDEXX test kit).

3.2.3 Introductions

Household introductions and sub-village node centrality metrics are the primary risk factors of interest for pathogen exposure in this study. The household survey data were analysed to evaluate if the odds of the household having introduced any livestock in the 12 months preceding the study differed depending on household A-E class and herd or flock size. Logit-binomial generalised linear mixed models (GLMMs) were used to assess if there was any evidence of a relationship. Sub-village was included as a random effect and back transformed model predictions are adjusted for Jensen's inequality. Gaussian GLMMs were also used to assess if sub-village in-degree and betweenness metrics in the SEEDZ data were associated with sub-village estimated herd and flock size and A-E class.

3.2.4 Pathogen exposure risk

3.2.4.1 Model construction

Household survey and serological data from the household's cattle and small ruminants (sheep and goats combined) were used to assess if there was evidence of, and the strength and direction of, a relationship between livestock introductions (at the household and sub-village level) and exposure to five pathogens in cattle and two pathogens in small ruminants. Univariable and multivariable logit-binomial generalised linear mixed models (GLMMs) were fitted to identify potential risk factors for the binary pathogen exposure seropositivity outcomes (Chriél, Stryhn and Dauphin, 1999; Zuur *et al.*, 2009; Michel, Brun and Makowski, 2017). Seroprevalence estimates were calculated using GLMMs, accounting for the random effects of household and sub-village (Zuur *et al.*, 2009). Univariable models were fitted first to look for evidence of an association between all individual risk factors and potential confounding variables and the pathogen seropositivity outcomes, for each pathogen in both cattle and small ruminants. Multivariable models were then fitted, with all risk factor and potential confounding variables included, to calculate the adjusted odds of seropositivity in the exposed groups while controlling for the potential confounding effects of all other variables. Likelihood ratio tests were used to assess if a variable significantly ($P_{LRT} < 0.05$) improved models fit to the data. Non-significant variables were removed using backward stepwise elimination process, starting with the variable with the largest P-value, until all remaining variables in the models had a significant effect on the pathogen seropositivity outcome.

A summary of all model seropositivity outcomes in each livestock group, measured at the individual animal level, are given in Table 3-1. These are *Brucella* spp. and *C. burnetii* in both cattle and small ruminants using the SEEDZ and BacZoo data combined and *Leptospira* serovar Hardjo*,* BHV-1 and BVDV in cattle using the SEEDZ data only. "Any zoonoses seropositivity" is investigated as an outcome using both the cattle (combining *Brucella* spp., *C. burnetii* and *Leptospira* serovar Hardjo results (SEEDZ data only)) and small ruminant (combining *Brucella* spp. and *C. burnetii* results (using SEEDZ and BacZoo data)) data to evaluate if common risk factors for zoonoses exposure exist.

Table 3-1: Livestock group and pathogen exposures tested using SEEDZ only or combined SEEDZ and BacZoo datasets

Pathogen	**Cattle**	**Small Ruminants**
Brucella spp.	*SEEDZ and BacZoo*	*SEEDZ and BacZoo*
C. burnetii	*SEEDZ and BacZoo*	*SEEDZ and BacZoo*
Leptospira serovar Hardjo *(SEEDZ only)*	*SEEDZ only*	
Any zoonoses (SEEDZ only)	*SEEDZ only*	*SEEDZ and BacZoo*
BHV-1 (SEEDZ only)	*SEEDZ only*	
BVDV (SEEDZ only)	*SEEDZ only*	

3.2.4.2 Livestock introductions risk factor variables

The main risk factor exposure variables of interest, tested in the models using the combined SEEDZ and BacZoo data, were the number of household introductions (total number of cattle or small ruminants introduced to the household in the 12 months preceding the study) and the estimated number of sub-village introductions (total number of introductions of cattle or small ruminants in the 12 months preceding the study reported by all surveyed households from a sub-village, divided by number of surveyed households, multiplied by number of livestock keeping households in the sub-village). For the SEEDZ-only data the additional sub-village-level node centrality metrics in-degree and betweenness, calculated on the inter-household movement network, were also included as risk factor variables. Sub-village node centrality metrics were tested as risk factor variables to evaluate whether an individual’s risk of pathogen exposure can be better quantified by the connectedness of its sub-village rather than the total number of livestock introductions at household or sub-village level. More details on construction of the inter-household movement network and analysis of the network can be found in the Appendix section 9.2.1 and Chapter 2.

3.2.4.3 BVDV exposure risk factor variable

Herd BVDV exposure status was added as a risk factor variable to the SEEDZ cattle models to test the hypothesis that belonging to a herd with evidence of prior BVDV exposure will increase an individual’s risk of exposure to other pathogens due to the immunosuppressive effects of BVDV. Herds were assigned a positive status if any of the sampled cattle tested antibody positive for BVDV, under the assumption that this provides evidence to show there is, or has been, BVDV circulating locally.

3.2.4.4 Potential confounding variables

Potential confounding variables included in the models were age and sex at the individual animal level and herd or flock size at the household (total number of cattle or small ruminants reportedly owned by the household) and sub-village levels (mean number of livestock owned by all surveyed households in a sub-village multiplied by the number of livestock keeping households in the sub-village). Herd/flock size is important to include as it is possible that larger herds/flocks have a higher risk of both introductions and pathogen propagation once a pathogen is introduced, depending on how it transmits between infectious and susceptible individuals (de Jong, Diekmann and Heesterbeek, 1994; Begon *et al.*, 2002). Finally, A-E class, assigned at the sub-village level was also included as a potential risk factor variable. Sub-villages sampled during the SEEDZ and BacZoo studies were assigned an A-E class using a clustering method that was based largely upon household livestock keeping practices that include livestock densities, movement to access natural resources and herd sizes (de Glanville *et al.*, 2020). Pathogen exposure risk for livestock within households may be affected by some of the variables used in the classification process. For example, larger herds may have greater within-herd transmission once a pathogen is introduced and herds that move often to access natural resources might be at more risk of *de novo* exposure to different pathogens. Based on the possibility that pathogen transmission may vary within or between household and sub-village herds/flocks, based on some livestock husbandry factors used in the A-E classification process, all risk factor variables were assessed to see if they varied by A-E class (Begon *et al.*, 2002). Likelihood ratio tests ($P_{LRT} < 0.05$) were used to assess if there was evidence of an interaction between A-E class and the other risk factor variables in each multivariable model. Model results are presented with interaction terms included, along with the P_{LRT}, where there is evidence to suggest that including the interaction term improves the model fit. For any variable that has an interaction with A-E class, the interacting risk factor variable is centred on its geometric mean and the odds ratio for the effect of that variable are presented separately for each class, with each class fitted as the baseline group. The odds ratio effect of each A-E class is thus presented in the results table conditioned on the geometric mean value of the interacting risk factor variable(s). Further details on each of the variables used in the models are included in Table S2.1 in the Appendix section 9.2.2.

3.2.4.5 Guarding against collinearity

The presence of collinearity among covariates can result in misleading results from regression analysis models (Zuur, Ieno and Elphick, 2010). To guard against this correlation coefficients and variance inflation factors (VIFs) were calculated among all pairs of variables. The level of correlation between model variables was tested by calculating Spearman's ρ before they were all added to the multivariable disease models. If any two variables were strongly correlated, the author was aware of this before fitting both variables into the multivariable model as this could lead to multicollinearity and failure of model convergence. When variables were correlated, in the first instance both were included in the multivariable model and the variable with the highest VIF was removed in a backward stepwise process until all VIFs were less than 3 (James *et al.*, 2013). All variables with VIFs less than 3 remained in each multivariable mixed effects regression model for each disease and species unless they were removed during the backwards stepwise fitting process (Wang *et al.*, 2008).

3.2.4.6 Dealing with zero events in subgroups

For the *Brucella* spp. small ruminant models there were zero positive events in the smallholder households which meant these models could not be fitted using a maximum likelihood approach. Instead, a Bayesian approach was used to model this data. This resulted in a decision to combine smallholder and agropastoral data for the *Brucella* spp. small ruminant models because there was no evidence of a difference between the two groups. Further details are provided in Appendix section 9.2.3.

3.2.4.7 Model fitting

Data were imported into R statistical software version 3.6.0 (R Core Team, 2019b) and GLMMs were fitted by maximum likelihood using the *glmer* function in R package *lme4* (Bates *et al.*, 2015). The *vif* function in the *car* package (Fox and Weisberg, 2019) was used to check the model variable VIFs. Likelihood ratio tests were used to assess if there was sufficient evidence (level of acceptance $p < 0.05$) to support the effect each variable had on the coefficient. The d*rop1* function from R core package *stats* was used to conduct the likelihood ratio tests. All final multivariable model predictions estimates were adjusted for Jensen's inequality (Nakagawa, Johnson and Schielzeth, 2017). Further details on data preparation, analysis and links to the code can be found in the Appendix sections 9.2.1. Normally distributed random effects were fitted at the sub-village and household levels in all univariable and multivariable models to account non-independence of observations at these levels. The association between each exposure variable and outcome variable was tested using Wald P-values (P_W). Marginal R^2 (mR^2) was calculated to gauge the predictive power of the model, and conditional R^2 (cR^2) was calculated to assess the amount of unexplained variation at the levels of the random effects (Nakagawa and Schielzeth, 2013).

3.3 Results

3.3.1 Descriptive statistics

Sero-prevalence results which show individual exposure prevalence and 95% confidence intervals, for each pathogen in the SEEDZ and combined SEEDZ and BacZoo datasets broken down by agroecological class for cattle and small ruminants are in Table 3-2 and Table 3-3. Household exposure prevalence, which is the proportion of cattle or small ruminant owning households with any positive animals is also given for each pathogen, broken down by A-E class for cattle and small ruminants is shown in Table 3-2 and Table 3-3. A summary of all numeric risk factor variables used in the cattle and small ruminant disease models, presented as mean and standard deviation or number and percentage, are given in Table 3-4 and Table 3-5. The final data sets were comprised of 3,364 cattle samples (2,853 SEEDZ and 511 BacZoo) from 440 cattle owning households (79 BacZoo and 361 SEEDZ households) and 5,524 small ruminant samples (4,318 SEEDZ and 1,206 BacZoo) from 411 small ruminant owning households (106 BacZoo and 305 SEEDZ households). From the initial serology results, for all pathogens the household prevalence was greater than the individual animal prevalence in both cattle and small ruminants.

Table 3-2: Estimated pathogen prevalence in cattle for production systems in the SEEDZ and combined (SEEDZ and BacZoo) data sets, with summary household prevalence for each pathogen

	SEEDZ		Combined SEEDZ & BacZoo		Household prevalence (proportion of cattle owning households with any +ve animals)
Pathogen exposure	Number **+ve**	Prev. (95% CI)	Number **+ve**	Prev. (95% CI)	
***Brucella* spp**	**85**	**0.030 (0.018, 0.047)**	**136**	**0.040 (0.026-0.059)**	0.204 $^{\text{SEEDZ \& BacZoo}}$
Agropastoral	16	0.013 (0.006, 0.024)	17	0.014 (0.007, 0.026)	0.084
Pastoral	67	0.051 (0.033, 0.077)	110	0.062 (0.043-0.089)	0.414
Smallholder	2	0.009 (0.002, 0.039)	9	0.025 (0.011, 0.055)	0.076
C. burnetii	**110**	**0.038 (0.026, 0.054)**	**148**	**0.044 (0.032-0.061)**	0.241 $^{\text{SEEDZ \& BacZoo}}$
Agropastoral	51	0.036 (0.023, 0.058)	52	0.040 (0.025, 0.062)	0.223
Pastoral	57	0.044 (0.028, 0.067)	84	0.050 (0.034, 0.073)	0.343
Smallholder	2	0.012 (0.002, 0.044)	12	0.038 (0.019, 0.075)	0.109
***Leptospira* serovar Hardjo**	**911**	**0.326 (0.275, 0.380)**	-	-	0.778 $^{\text{SEEDZ}}$
Agropastoral	337	0.279 (0.229, 0.335)	-	-	0.718
Pastoral	541	0.386 (0.325, 0.451)	-	-	0.965
Smallholder	33	0.198 (0.126, 0.293)	-	-	0.429
Any zoonoses	1017	**0.361 (0.311, 0.413)**	-	-	0.820 $^{\text{SEEDZ}}$
Agropastoral	386	0.310 (0.261, 0.363)	-	-	0.785
Pastoral	597	0.435 (0.376, 0.495)	-	-	0.979
Smallholder	34	0.198 (0.129, 0.290)	-	-	0.452
BVDV	1,146	**0.376 (0.331, 0.423)**	-	-	0.803 $^{\text{SEEDZ}}$
Agropastoral	324	0.316 (0.253, 0.386)	-	-	0.661
Pastoral	732	0.458 (0.381, 0.536)	-	-	1.000
Smallholder	90	0.419 (0.315, 0.530)	-	-	0.762
BHV-1	1,890	**0.658 (0.613, 0.701)**	-	-	0.958 $^{\text{SEEDZ}}$
Agropastoral	838	0.619 (0.564, 0.672)	-	-	0.943
Pastoral	932	0.713 (0.660, 0.762)	-	-	1.000
Smallholder	120	0.588 (0.484, 0.685)	-	-	0.881

Table 3-3: Estimated pathogen prevalence in sheep and goats for production systems in the SEEDZ and combined (SEEDZ and BacZoo) data sets, with summary household prevalence for each pathogen

Pathogen exposure	**SEEDZ**		**Combined SEEDZ & BacZoo**		**Household prevalence (proportion of small ruminant owning households with any +ve animals)**
	Number **+ve**	**prevalence** (95% CI)	Number **+ve**	**prevalence** (95% CI)	
Brucella spp	**125**	***0.025 (0.012, 0.050)***	**169**	***0.025 (0.013, 0.049)***	***0.168***
Agropastoral/ smallholder	59	*0.015 (0.006, 0.033)*	59	*0.017 (0.007, 0.037)*	*0.086*
Pastoral	66	*0.033 (0.017, 0.063)*	110	*0.037 (0.019, 0.069)*	*0.305*
C. burnetii	**1059**	***0.202 (0.166, 0.243)***	**1280**	***0.223 (0.183, 0.270)***	***0.689***
Agropastoral	327	*0.172 (0.134, 0.218)*	331	0.159 (0.124, 0.200)	0.628
Pastoral	676	*0.267 (0.215, 0.326)*	858	0.266 (0.218, 0.319)	0.987
Smallholder	56	*0.200 (0.128, 0.294)*	91	0.124 (0.086, 0.173)	0.358
Any zoonoses	**1135**	***0.238 (0.193, 0.288)***	**1387**	***0.215 (0.176, 0.259)***	***0.686***
Agropastoral	366	0.195 (0.151, 0.246)	370	0.178 (0.140, 0.223)	0.628
Pastoral	*713*	0.275 (0.219, 0.337)	*926*	0.277 (0.226, 0.333)	0.981
Smallholder	*56*	0.215 (0.138, 0.315)	*91*	0.129 (0.089, 0.182)	0.349

Table 3-4: Summary statistics for all cattle specific variables used in cattle disease risk models for SEEDZ, BacZoo and combined data sets

		BacZoo		**SEEDZ**		**Combined**	
		Number / *Mean*	% / [*Std.dev]*	Number / *Mean*	% (95% CI) / [*Std.dev]*	Number / *Mean*	% (95% CI) / [*Std.dev]*
Number of households surveyed owning cattle		79	18% (row %)	361	82% (row %)	440	
Number of cattle sampled		511	15% (row %)	2,853	85% (row %)	3,364	
Sex	**Male**	140	27%	948	33%	1,088	32%
	Female	371	73%	1905	67%	2,276	68%
Age		***4.5***	***[3.10]***	***4.6***	***[3.10]***	***4.6***	***[3.10]***
Production system	**Agropastoral**	16	3%	1302	46%	1,318	39%
	Pastoral	369	72%	1318	46%	1,687	50%
	Smallholder	126	25%	233	8%	359	11%
Belong to a BVDV+ herd		-		2,399	84%		
Household cattle herd size		***35***	***[40]***	***67***	***[136]***	***62***	***[127]***
Household number of cattle introductions for all households		***0.69***	***[2.19]***	***1.18***	***[5.64]***	***1.09***	***[5.20]***
Household number of cattle introductions for households that reported any cattle introductions		***5.50***	***[3.50]***	***5.01***	***[10.82]***	***5.06***	***[10.28]***
Sub-village cattle herd size		***1332***	***[1375]***	***2537***	***[2460]***	***2354***	***[2367]***
Sub-village estimated number of cattle introductions		***32***	***[46]***	***57***	***[105]***	***54***	***[99]***
Degree		**-**		***19.8***	***[10.9]***		
In-degree		**-**		***4.6***	***[2.0]***		
Betweenness		**-**		***0.02***	***[0.04]***		

Table 3-5: Summary statistics for all small ruminant specific variables used in small ruminant disease risk models for SEEDZ, BacZoo and combined data sets

		BacZoo		**SEEDZ**		**Combined**	
		Number / ***Mean***	% **/** ***[Std.dev]***	Number / ***Mean***	% **/** ***[Std.dev]***	Number / ***Mean***	% **/** [***Std.dev]***
Number of households surveyed owning small ruminants		106	26% (row %)	305	74% (row %)	411	
Number of small ruminants sampled		1206	22% (row %)	4,318	78% (row %)	5,524	
Age		***3.1***	***[2.10]***	***3.4***	***[2.10]***	***3.4***	***[2.10]***
Sex	**Male**	363	30%	943	22%	1,306	24%
	Female	843	70%	3375	78%	4,218	76%
Production system	**Agropastoral**	39	3%	1743	40%	1,782	32%
	Pastoral	708	59%	2290	53%	2,998	54%
	Smallholder	459	38%	285	7%	744	13%
Household small ruminant flock size		***81***	***[90]***	***140***	***[270.00]***	***127***	***[243]***
Household number of small ruminant introductions for all households		***1.42***	***[6.71]***	***2.11***	***[7.94]***	***1.93***	***[7.64]***
Household number of small ruminant introductions for households introducing any small ruminants		***7.50***	***[14.16]***	***7.87***	***[13.82]***	***7.79***	***[13.82]***
Sub-village small ruminant flock size		***2895***	***[2855]***	***5818***	***[6520]***	***5180***	***[6038]***
Sub-village estimated number of small ruminant introductions		***93***	***[178]***	***131***	***[203]***	***123***	***[199]***
Degree				***19.8***	***[10.9]***		
In-degree				***4.6***	***[2.9]***		
Betweenness				***0.02***	***[0.04]***		

3.3.2 Introductions

Cattle introductions in the 12 months preceding the study date were reported by 22% of cattle owning households (85 SEEDZ and 10 BacZoo households), with the number introduced ranging from 1-94 (mean 6). Small ruminant introductions were reported by 25% (82 SEEDZ and 20 BacZoo) of small ruminant owning households, with the number introduced ranging from 1-110 (mean 10). Across all cattle owning households surveyed in the study area the mean number of cattle introduced was 1.1 [SD 5.2] and for small ruminant owning households the number of small ruminants introduced was 1.9 [SD 7.6]).

At the sub-village level 66% (37 SEEDZ, 5 BacZoo) of sub-villages reported cattle introductions with the estimated number introduced ranging from 4-822 (mean 67) and 78% of surveyed sub-villages (37 SEEDZ and 13 BacZoo) reported small ruminant introductions with the estimated number introduced ranging from 4 – 1,000 (mean 149).

There was no evidence of an association between the odds of cattle and small ruminant introductions and household herd or flock size (fitted on log scale) (Pw = 0.13 and Pw = 0.24) respectively. There was evidence to show the odds of a household introducing any cattle or small ruminants in the year preceding the study differed between A-E classes ($P_{LRT} < 0.05$). Pastoral households were more likely to introduce cattle compared to agro-pastoral (OR 2.27, 95% CI 1.28, 4.00, Pw = 0.005) and smallholder households were less likely to introduce cattle compared to agro-pastoral (OR 0.40, 95% CI 0.16, 0.96, Pw = 0.04). An estimated 8% (95% CI 4-17%) of smallholder households, 18% (95% CI 12-25%) of agro-pastoral and 32% (95% CI 25-41%) of pastoral households introduced cattle in the year preceding the studies. Pastoral households were also more likely to introduce small ruminants compared to agro-pastoral (OR 2.48, 95% CI 1.36, 4.54, Pw = 0.003). Model results estimate 14% (95% CI 9-24%) of smallholder households, 20% (95% CI 13-28%) of agropastoral and 37% (95% CI 28-46%) of pastoral households introduced small ruminants in the year preceding the studies.

A 10 fold increase in sub-village cattle herd size and or small ruminant flock size was associated with higher sub-village betweenness (OR 1.67, 95% CI 1.01-2.75, Pw 0.04 and OR 1.45, 95% CI 1.05-2.00, Pw 0.02 respectively) which shows larger sub-village herds and flocks act as bridges between multiple other locations. A 10 fold increase in sub-village cattle herd size was also associated with a higher in-degree (OR 2.34, 95% CI 1.46-3.76, Pw <0.001) so larger cattle herds at the sub-village level receive incoming livestock from a greater diversity of origins. There was no evidence of an association between sub-village flock size and in-degree (Pw = 0.12). Pastoral sub-villages had higher in-degree and betweenness compared to smallholder and agro-pastoral sub-villages in both the cattle data and small ruminant data so pastoral villages receive incoming livestock from a greater diversity of origins compared to other A-E classes (in-degree 95% CI 1.06-3.78, Pw = 0.003 and 95% CI 1.50-5.49, Pw = 0.002 respectively and betweenness OR 11.85, 95% CI 7.29-19.28, Pw <0.001 and OR 5.19 95% CI 3.66-7.36, Pw <0.001 respectively).

3.3.3 Disease risk models

3.3.3.1 Univariable analysis

All univariable model results for pathogen exposure risk factor variables and the potential confounding variables for cattle and small ruminants can be found in the Appendix Tables S2.2 and S2.3, along with a description in the Appendix results section 9.2.4. While several univariable comparisons suggest evidence of associations, these associations alter when the multivariable models are fitted.

3.3.3.2 Multivariable analysis

The results of the correlation tests are shown in the Appendix Table S2.4. The variable “sub-village number of small ruminant introductions” was removed from the SEEDZ small ruminant model due to its high level of correlation with other risk factor variables (e.g. sub-village herd size, in-degree and betweenness) and it having a VIF value > 3.00. No variables were removed from the cattle multivariable models based on their VIF values. Results from the final multivariable models for an individual’s odds of exposure to different pathogens for cattle and small ruminants, using the SEEDZ and BacZoo combined datasets and the SEEDZ only dataset, are given in Tables 3-6 to 3-9.

3.3.3.2.1 Cattle pathogen exposure - SEEDZ and BacZoo data model results

The results from the cattle multivariable models using the combined SEEDZ and BacZoo data are shown in Table 3-6. These models primarily assessed if household and sub-village cattle introductions were associated with an increase in the odds of exposure *to Brucella* *spp.* and *C. burnetii*. Increasing age was associated with an increased risk of exposure to both pathogens. Once age was adjusted in the combined data set there was no evidence of an association between any of the risk factor variables and exposure to *C. burnetii.*

For *Brucella spp.* there was evidence of an interaction between A-E class and household introductions ($P_{LRT} = 0.03$). In cattle from agro-pastoral households a twofold increase in household introductions was associated with an increased risk of exposure to *Brucella spp.* There was no evidence of an effect of household introductions on *Brucella spp* exposure in cattle from smallholder or pastoral households. Overall, belonging to household that was agro-ecologically classed as pastoral was associated with a fourfold increased risk of exposure to *Brucella spp*. Confidence limits around the model estimates for the effects of household introductions and A-E class are broad due to the low prevalence of *Brucella spp.* in the sampled cattle.

Table 3-6: Cattle disease multivariable mixed model results using combined SEEDZ and BacZoo data. The top row of the table shows predicted seroprevalence in the baseline group (Agropastoral cattle, age 0 belonging to households with the mean number of household introductions), adjusted for Jensen's inequality. Rows below show the odds ratio for the effect of each model covariate on pathogen exposure, relative to the baseline group, with 95% confidence limits.

	Odds ratio for pathogen exposure in cattle with 95% confidence interval	
	Brucella spp.	*C. burnetii*
Constant [a] (predicted seroprevalence for Agropastoral cattle, Age 0, mean number of household introductions)	**0.004 (0.002, 0.009)**	**0.029 (0.019, 0.045)**
Age (Years)	1.28 (1.19, 1.38) ***	1.10 (1.03, 1.17) **
Pastoral production system [b]	4.70 (2.39, 9.26)***	
Smallholder production system [b]	2.26 (0.60, 8.54)	
Household cattle introductions (log2)	*[0.03]*	
Agropastoral	2.26 (1.23, 4.17) **	
Pastoral [c]	0.89 (0.72, 1.10)	
Smallholder [d]	1.71 (0.21, 13.61)	
Random effects variance		
Household (440)	0.86	0.57
Sub-village (64)	0.35	0.55
Marginal R squared	0.05	0.003
Conditional R squared	0.09	0.049
Observations	3,357	3,330
Variable Wald P-value <0.05 * <0.01** <0.001***		

[a] Adjusted for Jensen's inequality

[b] Where number of household cattle introductions are centered on the geometric mean (The effect of agro-ecological class presented in the table is conditioned on the geometric mean value of the interacting risk factor variable(s)).

[c] Result given for model with pastoral group set as baseline

[d] Result given for model with smallholder group set as baseline

3.3.3.2.2 Cattle pathogen exposure - SEEDZ only data model results

The second set of cattle multivariable models were constructed using the SEEDZ data only (Table 3-7). The network centrality measures in-degree and betweenness and household herd BVDV status are included as additional risk factor variables in these models, along with household and sub-village introductions. The pathogen exposure (seropositivity) outcomes assessed in the SEEDZ cattle models are *Brucella sp., C. burnetii, Leptospira* serovar Hardjo*,* any zoonoses, BHV-1 and BVDV. There were only 2 disease events in the SEEDZ smallholder *Brucella spp*. and *C. burnetii* data thus smallholder data were excluded from these two models because their inclusion resulted in model instability. Across all pathogen models increasing age was associated with an increased odds of pathogen exposure and male cattle had increased odds of exposure to BHV-1.

In both the *Brucella spp.* and *C. burnetii* SEEDZ data models, pathogen exposure prevalence was relatively low and there was evidence to suggest an interaction between A-E class and household cattle introductions on the risk of pathogen of exposure ($P_{LRT} < 0.05$). In these models a twofold increase in household cattle introductions was associated with an increased odds of pathogen exposure in agro-pastoral cattle (*Brucella spp.* OR 2.30, CI 1.21 - 4.36, *C. burnetii* OR 1.62, CI 1.04 – 2.52). This is the same as the effect of household cattle introductions seen in the combined data model for *Brucella spp* exposure. A twofold increase in household cattle introductions was also associated with an increased risk of exposure to BVDV (OR 1.11, 95% CI 1.01-1.24) and this effect was the same across all three A-E classes. There was no evidence to suggest that household cattle introductions were associated with an increased risk of exposure to *Leptospira* serovar Hardjo*,* any zoonoses or BHV-1.

A twofold increase in the sub-village network centrality measure betweenness (node lies on shortest path between most other pairs of nodes in the network) was also associated with an increased risk of exposure to BVDV (OR 1.21, 95% CI 1.09-1.35) and again this effect was the same across all A-E classes.

There was strong evidence to suggest the effect of the sub-village network centrality measure in-degree (number of incoming batches of livestock from other locations) differed between A-E classes for exposure to *Leptospira* serovar Hardjo, any zoonoses and BHV-1 (P-value 0.005, 0.003 and 0.03 respectively). A twofold increase in in-degree was associated with increased odds of exposure to *Leptospira* serovar Hardjo and any zoonoses in smallholder cattle and BHV-1 in pastoral cattle. Despite evidence to support including these interaction terms the effect estimate confidence limits within classes were often overlapping and spanned 1.

The effect of belonging to larger household herds was uniform across A-E classes in the *Leptospira* serovar Hardjo, any zoonoses and BHV-1 models with a tenfold increase in household herd size positively associated with pathogen exposure. There was evidence (P_{LRT} 0.02) to suggest the effect of household herd size differed between A-E classes in the SEEDZ data *Brucella spp.* model. In this model only pastoral cattle belonging to a tenfold larger herd had an associated increased odds of exposure to *Brucella spp*. (OR 2.30, CI 1.21-4.36). There was no evidence of an effect of household herd size on *C. burnetii* or BVDV exposure.

In the SEEDZ data cattle models, belonging to a pastoral herd was associated with an increased risk of exposure to *Brucella spp.* (a sixfold increase in risk when household herd size and cattle introductions are at their geometric mean values), as in the combined cattle data model, and any zoonoses (a 63% increased risk of exposure when in-degree is set at its geometric mean value). Belonging to a pastoral or smallholder herd was associated with an increased risk of exposure to BVDV and smallholder cattle were associated with lower risk of exposure to BHV- 1 compared to agropastoral and pastoral cattle (when in-degree is set at its geometric mean value).

There was no evidence of an association between estimated sub-village herd size or estimated number of sub-village introductions and the odds of exposure to any cattle pathogens.

For the secondary hypothesis, the effects of belonging to a BVDV positive herd was associated with around a 60% increase in the odds of exposure to *Leptospira* serovar Hardjo (OR 1.64 95% CI 1.06, 2.05). This could be due to immunosuppression in the herds, but given the lack of evidence of a similar effect on the risk of exposure to other infectious pathogens it could also be a result of BVDV and *Leptospira* serovar Hardjo sharing similar transmission routes and other shared exposure risk factors.

Table 3-7: Cattle disease multivariable mixed model results using SEEDZ data only. The top row of the table shows predicted seroprevalence in the baseline group adjusted for Jensen's inequality. Rows below show the odds ratio for the effect of each model covariate on pathogen exposure, relative to the baseline group, with 95% confidence limits.

		Odds ratio (OR) for pathogen exposure in cattle 95% confidence intervals (CI)					
		Brucella sp.	*C. burnetii*	*Leptospira* serovar Hardjo	Any zoonoses	BHV-1	BVD
Constant [a] (predicted seroprev. for baseline group in all covariate categories)		**0.002 (0.0004, 0.005)**	**0.02 (0.01, 0.04)**	**0.08 (0.05, 0.13)**	**0.12 (0.08, 0.17)**	**0.10 (0.06, 0.15)**	**0.55 (0.35, 0.74)**
Age (Years)		1.32 (1.20, 1.44) ***	1.13 (1.06, 1.21) **	1.18 (1.14, 1.83) ***	1.20 (1.16, 1.23) ***	1.67 (1.60, 1.75) ***	1.08 (1.05, 1.12) ***
Sex (Male)						1.57 (1.26, 1.96) ***	
Production system Pastoral [b]		6.20 (2.10, 18.25) **	1.03 (0.58, 1.82)	1.45 (0.96, 2.18)	1.63 (1.11, 2.38) *	1.03 (0.69, 1.54)	1.69 (1.11, 2.56) *
Production system Smallholder [b]		-	-	0.67 (0.37, 1.21)	0.59 (0.33, 1.04)	0.52 (0.29, 0.92) *	1.87 (1.13, 3.09) *
Household herd size (log 10)		[0.02]#		1.37 (1.03, 1.83) *	1.39 (1.05, 1.83) *	2.04 (1.47, 2.82) ***	
	Agro-pastoral	0.16 (0.02, 1.10) .					
	Pastoral [c]	1.65 (0.92, 2.95) *					
	Small-holder [d]	-	-				
Household cattle introductions (log 2)		[0.01]#	[0.05]#				1.11 (1.01, 1.24) *
	Agro-pastoral	2.30 (1.21, 4.36) *	1.62 (1.04, 2.52) *				
	Pastoral [c]	0.90 (0.72, 1.13) **	1.02 (0.83, 1.24)				
	Small-holder [d]	-	-				
Herd BVD status				1.64 (1.06, 2.55) *			-
In-degree				*[0.005]#*	*[0.003]#*	*[0.03]#*	

(log2)							
	Agro-pastoral			0.72 (0.53, 0.99) *	0.77 (0.58, 1.03) .	0.75 (0.56, 1.01) .	
	Pastoral [c]			1.03 (0.79, 1.35)	1.03 (0.81, 1.32)	1.24 (0.96, 1.59) *	
	Small-holder [d]			2.16 (1.10, 4.25) **	2.29 (1.20, 4.38) **	1.00 (0.62, 1.62)	
Betweenness (log2)							1.21 (1.09, 1.35) ***
Random effects variance							
Household (361)		0.74	0.12	0.34	0.33	0.34	0.22
Sub-village (46)		-	0.47	0.30	0.23	0.20	0.91
Marginal R squared		0.05	0.01	0.10	0.12	0.35	0.12
Conditional R squared		0.07	0.03	0.21	0.22	0.42	0.31
Observations		2,618	2,617	2,848	2,853	2,848	2,848
Variable Wald P-value <0.05 * <0.01** <0.001***							

[#] Interaction term likelihood ratio test P-value

[a] Adjusted for Jensen's inequality

[b] Conditional on any variable with an interaction effect being centred around, and set at, its geometric mean

[c] Result given for model with pastoral group set as baseline

[d] Result given for model with smallholder group set as baseline

Note. Smallholder data were removed from the *Brucella spp.* and *C. Burnetii* models as there are only 2 disease events in each group which made the multivariable models unstable with large uninformative confidence intervals for the smallholder groups

3.3.3.2.3 Cattle data models variation explained by fixed and random effects

In the combined SEEDZ and BacZoo dataset cattle models, the random effect of 'household' captures much of the variation for the *Brucella spp.* and *C. burnetii* models (Random effects variance 0.86 and 0.57 respectively). This means that animals within households are more similar to each other compared to animals from other households within the sub-village. For the *C. burnetii* SEEDZ only data model this finding was not repeated as the sub-village variance was larger than the household variance which indicates that in the SEEDZ villages cases, and thus *C. burnetii* exposure risk appears to cluster at the sub-village level. The mR^2 value for both the combined data and SEEDZ only data *C. burnetii* models was particularly low with the models fixed effects explaining <1% of the variation in the data. The mR^2 for both *Brucella spp.* model was slightly higher at 5% which indicates that losing observations to test for the effects of the network centrality measures and herd BVDV status did not compromise, nor improve, this models predictive power. For the SEEDZ only data cattle *Brucella spp.* model there was no additional variance at the sub-village level so this was not included as a random effect. The additional variation explained by the random effects in the combined SEEDZ and BacZoo dataset and SEEDZ only data cattle *Brucella spp.* models was also small with an increase in the mR^2 to the cR^2 of 4% and 2% respectively.

In the BVDV model the variance results show there was more clustering of cases at the sub-village level and the mR^2 (12%) and cR^2 (31%) show the models random effects explain more of the variation in exposure prevalence than the fixed.

The variance results from the *Leptospira* serovar Hardjo, any zoonoses and BHV-1 models show that unexplained risk of exposure to these pathogens was more clustered at the household level rather than the sub-village level. For *Leptospira* serovar Hardjo and any zoonoses the proportion of variation explained by the models fixed and random effects are similar and for BHV-1 the models fixed effects explain more of the variation in the data than the random effects. These results show that there are different livestock management and husbandry risk factors for pathogen exposure at the level of both the household and sub-village, depending on the pathogen of interest, not captured by these models.

3.3.3.2.4 Small ruminant pathogen exposure - all model results

Results from the small ruminant multivariable models, fitted using the combined SEEDZ and BacZoo data to assess the effect of household and sub-village introductions on small ruminants risk of *Brucella spp*., *C. burnetii* and ‘any zoonoses’ exposure are shown in Table 3-8. The results from the small ruminant models fitted using the SEEDZ data only, with the additional network centrality measures in-degree and betweenness included as risk factor variables in the analysis, are given in Table 3-9. Agro-pastoral and smallholder data are combined in both *Brucella spp.* models as there were no *Brucella spp.* seropositive events in smallholder small ruminants and no evidence of a difference in risk between the smallholder and agropastoral groups. For further details see Appendix section 9.2.3.

For all small ruminant models, increasing age and belonging to a tenfold larger household flock was associated with an increased risk of pathogen exposure across all A-E classes. In both *C.burnetii* models*,* male small ruminants had a lower odds of exposure compared to females and belonging to a pastoral flock was associated with an increased odds of exposure.

With the effects of age, sex, flock size (household and sub-village) and A-E class controlled for in all small ruminant models there is no evidence of an association between household and sub-village introductions of small ruminants and exposure to *Brucella spp*. or *C.burnetii.*

In the SEEDZ data only small ruminant *Brucella spp.* model there is evidence of an interaction between the network centrality measure in-degree and A-E class (P_{LRT} = 0.03) on the risk of pathogen exposure. In pastoral villages a twofold increase in sub-village in-degree is associated with an approximately twofold increase in the odds of small ruminant exposure to *Brucella spp* (OR 2.47, 95% CI 1.03-5.89).

In both the combined and SEEDZ-only small ruminant *Brucella spp.* models there was good evidence to suggest the effects of sub-village flock size on the risk of pathogen exposure differed between A-E classes (P_{LRT} = 0.03 and 0.01 respectively). Small ruminants from agropastoral and smallholder households had an increased odds of exposure to *Brucella spp.* with increasing sub-village flock size .

Table 3-8: Small ruminant Brucella and Coxiella multivariable mixed model results - BacZoo and SEEDZ data. The top row of the table shows predicted seroprevalence in the baseline group adjusted for Jensen's inequality. Rows below show the odds ratio for the effect of each model covariate on pathogen exposure, relative to the baseline group, with 95% confidence limits.

	Odds ratio for pathogen exposure in small ruminants with 95% confidence interval		
	Brucella spp.	*Coxiella burnetii*	Any Zoonoses
Constant [a] (predicted seroprev. for baseline group in all covariate categories))	**0.002 (0.0002, 0.009)**	**0.03 (0.02, 0.06)**	**0.03 (0.02, 0.05)**
Age (Years)	1.26 (1.14, 1.38) ***	1.28 (1.23, 1.33) ***	1.29 (1.24, 1.34) ***
Sex (Male)		0.60 (0.49, 0.74) ***	0.61 (0.50, 0.74) ***
Production system Pastoral [b]	0.74 (0.25, 2.24)	1.83 (1.26, 2.66) **	
Production system Smallholder [b]	NA	0.89 (0.54, 1.47)	
Household herd/flock size (log 10)	4.59 (1.61, 13.06) **	1.47 (1.12, 1.94) **	1.90 (1.45, 2.48) ***
Sub-village herd/flock size (log 10)	[0.03] [#]		
Agropastoral/smallholder	5.90 (1.09, 32.02) *		
Pastoral [c]	0.39 (0.06, 2.41)		
Random effects variance			
Household (411)	3.02	0.25	0.29
Sub-village (64)	1.33	0.72	0.92
Marginal R squared	0.04	0.09	0.07
Conditional R squared	0.13	0.21	0.23
Observations	5,524	5,500	5,524

Wald p-value <0.05 * <0.01** <0.001***

[#] Interaction term likelihood ratio test P-value

[a] Adjusted for Jensen's inequality

[b] Conditional on any variable with an interaction effect being centered around, and set at, its geometric mean

[c] Result given for model with pastoral group set as baseline

Table 3-9: Small ruminant Brucella sp. and C.burnetii multivariable mixed model results - SEEDZ data only. The top row of the table shows predicted seroprevalence in the baseline group adjusted for Jensen's inequality. Rows below show the odds ratio for the effect of each model covariate on pathogen exposure, relative to the baseline group, with 95% confidence limits.

	Odds ratio (OR) for pathogen exposure in small ruminants with 95% confidence intervals (CI)		
	Brucella spp.	*Coxiella burnetii*	Any Zoonoses
Constant (predicted seroprev. for baseline group in all covariate categories)[a]	**0.001 (0.0001, 0.010)**	**0.03 (0.02, 0.06)**	**0.04 (0.02, 0.06)**
Age (Years)	1.27 (1.14, 1.42) ***	1.33 (1.27, 1.39) ***	1.34 (1.28, 1.39) ***
Sex (Male)		0.66 (0.53, 0.84) ***	0.69 (0.55, 0.86) **
Production system Pastoral [b]	0.42 (0.12, 1.41)	1.67 (1.12, 2.50) *	
Production system Smallholder [b]	-	1.42 (0.78, 2.59)	
Household flock size (log 10)	5.05 (1.61, 15.87) **	1.45 (1.07, 1.95) *	1.68 (1.25, 2.25) ***
Sub-village herd/flock size (log 10)	*[0.01]*[#]		
Agropastoral/Small holder	4.56 (0.68, 30.23)		
Pastoral [c]	0.08 (0.008, 0.87) *		
Sub-village in-degree (log 2)	*[0.03]*[#]		
Agropastoral/Small holder	0.72 (0.33, 1.58)		
Pastoral [c]	2.47 (1.03, 5.89) *		
Random effects Variance			
Household ID (N=305)	3.02	0.25	0.29
Sub-village (N=46)	1.04	0.76	0.92

Marginal R squared	0.03	0.09	0.08
Conditional R squared	0.12	0.22	0.24
Observations	4,318	4,316	4,318
Variable Wald p-value <0.05 * <0.01** <0.001***			

[#] Interaction term likelihood ratio test P-value

[a] Adjusted for Jensen's inequality

[b] Conditional on any variable with an interaction effect being centered around, and set at, its geometric mean

[c] Result given for model with pastoral group set as baseline

3.3.3.2.5 Small ruminant data models variation explained by fixed and random effects

The mR^2 (variation explained by the model's fixed effects only) value in the combined data *Brucella spp., C. burnetii* and 'any zoonoses' models is increased from 4%, 9% and 7% respectively, to the cR^2 of 13%, 21% and 23% (amount of variation explained by the combination of fixed and random effects) respectively. Similarly, in the SEEDZ data models the mR^2 values are similar at 3%, 9% and 8% and these are increased to the cR^2 values 12%, 22% and 24% for the *Brucella spp.*, *C. burnetii and* any zoonoses models respectively. These results show that for all small ruminant models the random effects of household and sub-village explain more of the variation in exposure prevelance than the model fixed effects. The small ruminant brucella models have a particularly low mR^2 values compared to the other models with only 3-4% of the observed variation being explained by the fixed effects. The slight decrease in the mR^2 between the combined data model and the SEEDZ only data model shows that losing some observations to test for the effects of the network centrality measures on *Brucella* spp. exposure did not improve this model's fit to the data. The larger proportion of variation explained by the random effects compared to the fixed effects in all small ruminant models shows that there are factors at both the household and sub-village levels that explain exposure prevalence which are not included in these models. In the *Brucella spp.* small ruminant models a large amount of the variation in prevalence is captured by the random effects household ID (both models 3.02) and sub-village (combined 1.33, SEEDZ 1.04) which shows a high level of clustering of exposure at both the household and sub-village levels (Table 3-8). In both the combined data and the SEEDZ-only data *C.burnetii* and 'any zoonoses' models there is also substantial variation and clustering of exposure among sub-villages (*C.burnetii* sub-village model variance's 0.72 and 0.76 and any zoonoses sub-village model variance's 0.92 and 0.92 respectively) but less evidence of clustering at the household level. This means that the sub-village the small ruminant belongs to affects its risk of pathogen exposure to a greater extent than the household. There are therefore livestock management and husbandry practices, or environmental contamination which affect exposure risk, clustered at the sub-village level and not captured by the variables used in these models.

3.4 Discussion

3.4.1 General

Livestock movements are widely recognised as providing a route for pathogen transmission between otherwise well-separated populations (Green, Kiss and Kao, 2006a; Palisson *et al.*, 2016). High in-degree and betweenness have also frequently been identified as risk factors for disease introduction in other livestock keeping systems (Palisson *et al.*, 2016; Ferdousi *et al.*, 2019). The primary objective of this study was to investigate the extent to which different types of permanent livestock introductions at the household and sub-village levels, which go largely unreported in northern Tanzania, affect livestock's risk of exposure to five infectious pathogens.

Through this study, analyses show the association between seroprevalence and household introductions and sub-village in-degree differ for pathogens of varying transmissibility between the different A-E systems and between species. When other important or potentially confounding risk factors are controlled for, my analysis shows that:

- Increasing numbers of household introductions are associated with increased exposure to BVDV (all A-E classes), and *Brucella* spp. (AP only) and *C. burnetii* (AP only) in cattle.
- At the sub-village level, increasing betweeness centrality is associated with increased BVDV exposure across all A-E classes.
- Increasing in-degree at the sub-village level is also associated with (i) increased exposure to *Brucella spp.* in pastoral small ruminants, (ii) BHV-1 in pastoral cattle and (iii) *Leptospira* serovar Hardjo and any zoonoses in smallholder cattle.

The pathogens used to test the hypothesis in this study have different shedding patterns (cyclical vs intermittent vs persistent) and utilise multiple different transmission routes (close contact, sexual, respiratory and environmental). In addition to the variations in transmissibility of the pathogens studied, cattle and small ruminant stocking density, mixing patterns and use of shared resources vary between A-E classes (de Glanville *et al.*, 2020). The combination of varying pathogen transmissibility and opportunities for effective contacts between infectious and susceptible individuals in the different A-E classes likely explain the inconsistent pattern observed here between introductions and seroprevalence. The variation in risk associated with household and sub-village level introductions and exposure to the different pathogens needs careful consideration in the design of effective control programmes for production-limiting and zoonotic pathogens in this population.

3.4.2 Household introductions

In the 12 months preceding the study 22% and 25% of cattle and small ruminant owning households respectively reported livestock introductions with a high variation in the numbers introduced (mean of 5 [SD 10.3] cattle and 8 [SD 13.8] small ruminants).

These results show that the majority of household herds and flocks in the study area, especially those in agropastoral and smallholder villages, operate in one sense as closed units reporting no introductions. If household introductions are the major risk factor for exposure to pathogens in this population then few households will experience the increased risk each year. On the other hand, with over 20% of households introducing potentially infectious animals a year this could be sufficient to sustain transmission between households and allow pathogens to persist in the population even when their R_0 values are low (Prentice *et al.*, 2017). If other contacts between livestock, such as those that occur through the use of shared resources, are more important for pathogen transmission and maintenance in this population then the potentially beneficial effect of belonging to a somewhat ‘closed’ herd will be less evident.

3.4.3 Sub-village introductions

At the sub-village level, introductions were more frequent with 66% and 78% of surveyed sub-villages experiencing cattle and small ruminant introductions respectively in the year preceding the studies. Pastoralist sub-villages and all sub-villages with larger cattle herds were connected to more other locations through incoming livestock movements (higher in-degree) and were more likely to act as bridges between other locations (higher betweenness). Sub-villages with larger small ruminant flocks across all A-E classes also had higher betweenness. This means that if the risk of pathogen exposure is increased through connectivity at the sub-village level then livestock in pastoral sub-villages and sub-villages with larger herds and flocks will be at increased risk of exposure. Additionally, because introductions are more frequent at the sub-village level compared to household level, if the risk of pathogen exposure is experienced at the sub-village level then a higher proportion of livestock are at risk.

3.4.4 Disease risk

3.4.4.1 Household level

At the household level household introductions were positively associated with BVDV exposure across all A-E classes and *Brucella spp.* and *C. burnetii* exposure in agro-pastoral cattle.

For BVDV, this result indicates that once BVDV is introduced into a household herd the frequency of within herds contacts that result in within herd transmission are the same across all A-E classes. For *Brucella spp.* and *C.burnetii* these results provide no definitive answer as to why household introductions act as a risk factor for exposure in agro-pastoral cattle but not in the other A-E classes. One possibility is that the type of introductions vary between A-E classes with market introductions more common in agro-pastoral herds but gifts more common in pastoral herds (Aktipis, Cronk and de Aguiar, 2011). Additionally the management of recently calved animals, most likely to be shedding pathogen, may vary between A-E classes which will have a direct effect on opportunities for transmission events.

3.4.4.2 Sub-village

At the sub-village level increasing in-degree was associated with increased risk of *Brucella* spp. exposure in pastoral small ruminants, BHV-1 in pastoral cattle and *Leptospira* serovar Hardjo in smallholder cattle. The variation in the effects of increasing in-degree in the different A-E classes for different pathogens in different species groups is likely due to the differences in the pathogen transmission routes and livestock husbandry practices between A-E classes (Otte and Chilonda, 2003; Schoonman and Swai, 2010; Allan *et al.*, 2018).

For *Brucella* spp. in pastoral small ruminants and BHV-1 in pastoral cattle perhaps the concept of defined individual herds and flocks in these settings is less important, specifically when considering the transmission routes used by these pathogens (close contact, reproductive and respiratory). There may be more opportunities for mixing and effective contacts between susceptible and infectious individuals from different herds and flocks at shared grazing and water points in pastoral settings that can be utilised as transmission routes by these pathogens (Valle *et al.*, 1999; Palisson *et al.*, 2016; Rossi *et al.*, 2019). This hypothesis is supported by findings from Böhm et al. who show the most social within herd individuals are most likely to interact with individuals outside the herd making them potential hubs for transmission between infectious and susceptible individuals (Böhm, Hutchings and White, 2009). This could explain why in pastoral settings introductions at the sub-village level are a risk factor for some pathogens that require close contact for transmission events. The opportunities for close contact interactions between flocks from different households in pastoral settings might also be different to the opportunities available to cattle. For example, cattle may be taken away to distant grazing sites whereas all small ruminants may stay closer to the homesteads and thus have more opportunities for interactions between post-partum, infectious shedding animals and susceptibles from other flocks.

For *Leptospira* serovar Hardjo there is an environmental component to transmission and also the involvement of vectors such as rats. These factors could explain why introductions into more densely populated smallholder sub-villages specifically act as a risk factor for *Leptospira* serovar Hardjo.

Also at the sub-village level increasing betweenness (number of times a sub-village lies on the shortest path between other locations) was associated with increased risk of BVDV across all A-E classes (BVDV is shed persistently by infected animals (PIs) and close contact between infectious and susceptible can result in transmission event (Brownlie *et al.*, 1987; Qi *et al.*, 2019)). This result shows that transmission of BVDV between herds within sub-villages is ubiquitous across all A-E classes and indicates that some within-sub-village transmission routes exist for BVDV across all A-E classes that are less utilisable by other pathogens investigated here. This finding alone is noteworthy as little is currently known about the transmission of BVDV within northern Tanzania (Hyera, Liess and Frey, 1991; Hodgkinson, 2017). The high level of within sub-village transmission of BVDV once introduced, and the higher overall sero-prevalence in pastoral and smallholder cattle could be due to BVDV transmission being more density dependent rather than frequency dependent (Begon *et al.*, 2002) (smallholder cattle densely populated and pastoral herds both large and densely confined in bomas overnight).

More specifically the relationship between BVDV and betweenness could be explained by transmission events for highly transmissible vs less transmissible (specifically those that are cyclically or intermittently shed) pathogens being less dependent on the duration of a contact or stay of an infectious animal (in this case a PI) in a herd (Kao *et al.*, 2007). The reverse causality option is also possible, in that BVDV PIs might under-perform and frequently be sold on and replaced, thus driving up sub-village betweeness (Hidano and Gates, 2019). Both of these possible explanations should be interpreted with caution as they make an assumption that it is the same animals which are moving in and out of sub-villages.

Either way, the relationship between increasing sub-village betweenness centrality and increasing risk of exposure to highly transmissible pathogens should be explored further. This could be done by assessing whether the same effect is observed for exposure to other highly transmissible livestock pathogens such as PPRV (Jones *et al.*, 2016), contagious caprine pleuropneumonia (CCPP) (Swai and Nesella, 2010) and FMD (Green, Kiss and Kao, 2006a). PPRV in northern Tanzania is already associated with pastoralist livestock keepers and, as shown by the results from this study, pastoralist sub-villages do tend to have higher betweenness centrality than sub-villages in the other A-E classes (Herzog *et al.*, 2019).

In addition to the increased risk of exposure to highly transmissible pathogens, sub-villages with high betweenness link most other locations in a contact network and thus are high risk for passing pathogens on to many other locations. The combination of increased risk of highly transmissible pathogen introduction, propagation and onward transmission means that sub-villages with high betweenness are ideal targets for interventions that will break chains of transmission. These could include increased veterinary knowledge exchange on the specific effects of BVDV in herds and how to prevent it or reduce the burden, early warning surveillance platforms or sub-village vaccination programmes.

Any zoonoses exposure was analysed as an outcome in case strong risk factors could be identified that resulted in livestock, and thus the people in close contact with them, being at high risk of zoonoses exposure. The disparate prevalence and transmission routes of the pathogens used to create the *any zoonoses* measure mean the risk factors for the most prevalent pathogens drive the model outcomes. This makes the any zoonoses category less practically useful than if all zoonotic pathogens had similar prevalence and transmission routes.

3.4.5 BVDV exposure

In this study we also tested the hypothesis that belonging to a BVDV positive herd increased the risk of exposure to other pathogens in cattle. A herd was considered positive for BVDV circulation if one or more animals from the herd tested positive for BVDV antibody. This assumption falls down if the cattle sampled are PIs and therefore test sero-negative or the cattle test antibody positive due to historic exposure but were only recently added to the sampled herd and thus do not represent circulating infection. We do not have data available on the timing of any household introductions or the duration of stay in the herd of sampled cattle so results from this analysis must be interpreted with caution.

Leptospira serovar Hardjo was the only pathogen tested that was associated with BVDV herd positivity. This could be due to the immunosuppressive effects of the virus at the herd level and perhaps the effect is weak so only picked up for pathogens with a high prevalence. It also possible that *Leptospira* serovar Hardjo and BVDV exposure have other common risk factors as they both circulate endemically in this cattle population. The high herd prevalence of BVDV in the study area (80%) means there are few sero-negative herds to compare to the sero-positive herds and thus show evidence of an effect of herd seropositivity on risk of exposure to other pathogens.

Despite not identifying a significant relationship between BVDV herd seropositivity and exposure to other pathogens, the high herd prevalence of BVDV is important to note. This acts as a reminder that livestock disease control programmes can be inefficient and potentially ineffective if they tackle individual diseases in isolation (Mahmoud and Allam, 2013). If a herd is immunosuppressed due to circulating BVDV (Chase, 2013) then the effectiveness of vaccines against other pathogens will be reduced. Similarly, if an animal or herd tests positive for one pathogen following a clinical syndrome such as an abortion, this is not necessarily the only pathogen causing disease as coinfections occur and can have more severe consequences on individual and herd productivity (Mahmoud and Allam, 2013; Mathew *et al.*, 2017). Furthermore the impact of BVDV on pathogen shedding and thus zoonotic risk posed by animals that are infected with *Brucella spp.* and *C. burnetii* is not clear but there is reason to believe there may be an association if shedding increases with immunosuppression (Mathew *et al.*, 2017; Okumu *et al.*, 2019).

3.4.6 Control measures

The effects of livestock introductions on individual animals' risk are not homogenous across species, pathogens and A-E classes. Introductions should therefore not be relied upon as the sole guide for when and where to implement pathogen surveillance and control programmes in this livestock population as they have been in other countries with disparate livestock husbandry practices (VanderWaal, Enns, *et al.*, 2017; Salvador *et al.*, 2018). Nonetheless, these results can be used to guide the design of optimal surveillance and control programmes for the pathogens studied here and potentially for others with similar shedding patterns and transmission routes within the different A-E classes. Targeted pathogen control interventions, at either the household or sub-village level in locations identified as high risk for exposure, could greatly benefit communities in northern Tanzania by improving overall livestock productivity and public health (Lanyon and Reichel, 2013; Haseeb *et al.*, 2019).

3.4.6.1 Target interventions based on household risk

If the risk factor for exposure exists at the household level then the responsibility of the risk belongs at that level – i.e. to the person who chooses to import the livestock. In this scenario livestock owners should have access to veterinary knowledge so they can weigh up the pros and cons of introducing livestock and come to an informed decision about disease risk. Sharing results of the risk of bacterial zoonoses exposure associated with household introductions with livestock keepers in agropastoral villages will give them greater agency to make informed decisions about livestock introductions based on the potential risks posed by introducing animals. Similarly, cattle keepers in villages across all A-E classes should be made aware of the risk of introducing BVDV to their household and sub-village through cattle introductions. Cattle keepers could implement risk reduction strategies such as spending a small amount of money on testing, quarantining, practicing zero grazing for a short time and or buying from known disease-free sources. The cost of preventing pathogen introduction would be offset by the both herd and community benefits of reduced pathogen burden in the livestock and human population (Jones *et al.*, 2016; Cash-Goldwasser *et al.*, 2018). The relatively low rate of household introductions and generally small numbers of cattle and small ruminants introduced means that the interventions listed above, to reduce pathogen introduction into the household via these routes, could be focussed and achievable.

3.4.6.2 Target interventions based on sub-village risk

When the risk of pathogen exposure through livestock introductions exists at the sub-village level then interventions that can be publicly or collectively funded and enacted are needed because the livestock and public health benefits are a collective public good (Rich and Perry, 2011; Forman *et al.*, 2012; Pradère, 2014; Thumbi *et al.*, 2015). Examples of collective targeted interventions could be: vaccination against rapidly transmissible pathogens such as BVDV, development and maintenance of sub-village test and quarantine facilities, and provision of improved veterinary services in sub-villages with high betweenness. Veterinary knowledge exchange interventions that aim to reduce high risk behaviour and decision making can be introduced rapidly at a low financial and resource cost, unlike control programmes that require drug administration or surveillance platforms. The disadvantages of knowledge exchange as a disease control intervention is that even when disease knowledge is considered good, complex relationships can exist between livestock keepers' behaviour and biosecurity practices (Rich and Perry, 2011; Hidano, Gates and Enticott, 2019). These complexities will need to be considered when deciding how to measure the success of knowledge exchange programmes.

3.4.7 Limitations

The topic of livestock numbers and introductions is a sensitive subject in the study area because livestock numbers equate directly to wealth. For this reason, numbers of introductions could have been underreported or withheld. The location animals are received from and sent to is a less sensitive subject which potentially made the inclusion of network measures as risk factors a more robust measure of livestock movement activity compared to absolute counts of livestock.

Sero-prevalence results, used here as a proxy for overall risk of pathogen exposure, also have their limitations. Sero-positivity is not correlated with pathogen shedding, for example cattle persistently shedding BVDV will never test sero-positive for antibodies and the relationship between shedding and seropositivity for *C. burnetii* and *Leptospira* serovar Hardjo is variable (Brownlie *et al.*, 1987; de Cremoux *et al.*, 2012; Allan, 2016). A further limitation to using serology results is that owing to the moderate power of each individual test for seropositivity, and the large number of tests, false positives are possible and false negatives within the dataset are likely. Despite the limitations associated with using sero-prevalence as a proxy for pathogen exposure these two overlapping studies carried out in northern Tanzania were large with robust study design and implementation. This means that the general results presented here will be difficult to improve upon.

The influence of the timing of events is another limitation we need to be aware of when interpreting the results. Many livestock could have been added to the herd in previous years/months yet none in the year which the study asks about or the reverse. Because the measure of disease used in this study is point prevalence seropositivity data, infection could have been introduced to the herd in previous years, or may have been so recently introduced that few animals are currently infected and thus it remains undetected. To understand the true impact of livestock introductions on individual disease status, a longitudinal study would be needed, gathering data on timing of movements and time of pathogen exposure (Barnard *et al.*, 2018; Enright and Kao, 2018). A further limitation to the approach taken in this study is that there is no account made for potential transmission events, either direct or indirect, between livestock and wildlife that could be acting as reservoirs for multiple livestock pathogens in the study area. Details on direct and indirect contact with wildlife, through the use of shared resources, could be included in future longitudinal studies that aim to identify key risk factors for pathogen transmission events and introductions into herds and flocks.

3.5 Conclusion

The relationship between livestock introductions at the herd or flock and sub-village level and onward pathogen transmission within herds and flocks is complex. It varies depending on the A-E class of the management systems under which the livestock are kept and the pathogen characteristics which likely include the shedding pattern, infectiousness and transmission route(s) utilised. The overarching hypothesis tested in this work, that introductions cause pathogen exposure, embeds an underlying assumption about the relationship between introduction of disease, residency times, mixing and seroprevalence. Within these results there likely lies an interaction between the risk of pathogen introduction and risk of persistence as pathogens are transmitted within and between herds and flocks through the use of shared resources.

The relationships identified in this data could be investigated further using data on other more transmissible small ruminant pathogens, or other serology data from studies with similar differences in livestock husbandry practices between different communities. Further work in northern Tanzania would ideally use a longitudinal cohort study design to identify the differences in the risk of pathogen introduction, propagation and exposure within and between herds and flocks (Bajardi *et al.*, 2011; Enright and O'Hare, 2017).

The results from this study show that livestock movements between households provide routes for pathogen transmission and maintenance between populations in northern Tanzania. Where the increased risk of pathogen exposure is due to livestock introductions at the sub-village level, network node centrality metrics combined with A-E class data can be used to identify high risk villages to target with disease surveillance and control programmes. These interventions should aim to break transmission chains and prevent pathogen propagation within sub-villages. Sub-village level interventions should be publicly funded because they provide wider community benefits by reducing livestock and zoonotic pathogen burden, improving livestock health and productivity and thus improving household income, food security and public health (Salvador *et al.*, 2018).

This work shows that multiple zoonotic and socio-economically important pathogens have similar livestock movement exposure risk factors despite their varying shedding patters and transmission routes. This builds on the theory that pathogens should not be viewed in isolation with regards to control campaigns, particularly in populations such as northern Tanzania where multiple pathogens circulate at high levels and control resources are scarce. This work provides data to show that targeted approaches to disease control (albeit nuanced depending on the different A-E systems), that focus on reducing pathogen incursion and onwards transmission via livestock movements, could encompass multiple pathogens and target multiple species. A single pathogen vaccination programme, such as the one aimed for by GALVmed with the Brucella vaccine development challenge (GALVmed, 2018), is unlikely to reduce the all-encompassing impacts of zoonoses in northern Tanzania's human and livestock populations. Where public or private funded facilities are set up to test or vaccinate for one pathogen, multiple pathogens that share exposure risk factors and/or cause similar clinical disease (for example livestock abortion) can and should be included. This multi-pathogen approach to disease control will not only maximise the cost efficiency of control programmes but will have a more noticeable clinical effect on the ground. Where we should be moving towards improved efficiency of livestock production it is important to identify where targeted interventions can be implemented. A multi-pathogen approach, that focuses on reducing overall burden in the livestock populations by reducing transmission between and within populations by regulating (for example by introducing pre- and post-movement quarantines) and reducing livestock movements (through knowledge exchange to make people aware of the disease risks incurred at household and sub-village level with livestock introductions) appears to be a sensible and feasible option in this area.

Where we are still faced with a lack of standardised processes to quantify the true socio-economic burden of animal diseases (productivity losses and zoonotic burden) it is difficult to secure funding for large control and elimination programmes (Hotez *et al.*, 2009; Molyneux *et al.*, 2011). To this end it remains imperative that through future work, we improve knowledge and awareness of specific risk factors for livestock disease, for livestock managed in different ways. Having more specific information on the risk of disease introduction and onward transmission will allow livestock keepers to act to mitigate these risks in an effort to protect the health of their herds, families and wider communities.

4 Describing livestock market movements in northern Tanzania and how these influence pathogen transmission

Contribution

Conceptualisation of research questions, study design, survey design, data collection, processing and analysis were all done by GC. Data collection was supported by a field team made up of three individuals trained by GC.

4.1 Introduction

The northern zone of Tanzania contains 15% of Tanzania's 30 million cattle and around 25% of the 19 million goats and 5.5 million sheep, with many households heavily reliant on livestock as a major or sole source of income (Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012; National Bureau of Statistics and Tanzania Government, 2016). The high reliance on livestock is accompanied by a high pathogen prevalence in both cattle and small ruminant livestock with limited access to robust diagnostics and treatment and limited resource and funding for disease control programmes (Chipwaza *et al.*, 2014; Alonso *et al.*, 2016; Shirima and Kunda, 2016; Mathew *et al.*, 2017; Allan *et al.*, 2018). Many of the livestock pathogens are also zoonotic and spill over to cause disease that is regularly misdiagnosed and incorrectly treated with profound negative consequences for human health and wellbeing (Crump *et al.*, 2013; Bodenham, Lukambagire, *et al.*, 2020). In northern Tanzania it is essential to reduce pathogen transmission within the livestock population to reduce the socio-economic losses and zoonoses that result from unhindered transmission (Molyneux *et al.*, 2011; Halliday *et al.*, 2012).

Movement of livestock between locations create epidemiological links that provide routes for pathogen transmission. The geographical scale and speed of livestock movements therefore influences the scale and speed of pathogen transmission (Cowan and Jonard, 2004; Fèvre *et al.*, 2006; Green, Kiss and Kao, 2006a; Buhnerkempe *et al.*, 2014; Palisson *et al.*, 2016). Market movements of livestock have frequently been identified as important for pathogen amplification and propagation within and between distinct populations (Robinson and Christley, 2007; Fournié *et al.*, 2011; Rautureau, Dufour and Durand, 2011; Dean *et al.*, 2013; Fournié and Pfeiffer, 2013; Fournie *et al.*, 2013). The well-travelled routes livestock take to reach markets have also been identified as important for pathogen transmission in countries where livestock are kept extensively and market journeys are made on foot (Tempia *et al.*, 2010; Jahel *et al.*, 2020). In northern Tanzania many types of livestock movement exist and these include the movement of livestock to, from and between markets. This creates a network of movements between and across geographical areas and these can cross, sometimes multiple, international boundaries (Watts and Strogatz, 1998; Kao *et al.*, 2006; Dean *et al.*, 2013; Motta *et al.*, 2017). A relevant example of this is the recent description of movements from northern Tanzania into Kenya as part of the Kenyan beef supply chain (Mutua *et al.*, 2018).

Globally livestock markets have been identified as highly central hubs that attract sellers and buyers from multiple and potentially distant locations (Gibbens *et al.*, 2001; Molia *et al.*, 2016; Motta *et al.*, 2017). Markets are therefore frequently identified as important locations contributing to pathogen spread, but also where targeted disease control programmes could be most effective (Ortiz-Pelaez *et al.*, 2006; Robinson and Christley, 2007; Molia *et al.*, 2016). Targeted control programmes might include active surveillance, knowledge exchange or vaccination programmes or under extreme circumstances movement restrictions or market closures (Vallée *et al.*, 2013).

There is currently little known about the characteristics of market movements in northern Tanzania (Chaters *et al.*, 2019). This includes a lack of information on how livestock travel to market (in vehicle or on foot), and the risk of pathogen transmission events along market journeys through journey sharing or other close contacts made with local livestock en route (Pica-Ciamarra *et al.*, 2011). It was legislated in 2000 that vehicles must be used to transport livestock to and from markets to protect land, the health of the livestock moved and to reduce pathogen transmission between moving and local livestock (Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012). This legislation is difficult to enforce and the level of compliance with vehicle transport to and from markets remains unknown. The additional cost of vehicle transport to move animals to market might also be beyond what poorer households can afford and force them to make difficult decisions about market movements. This could include moving livestock to and from markets on foot, increasing the risk of journey pathogen transmission events, or having to sell livestock to traders rather than in the markets themselves and losing out on opportunities to achieve the best price (Pica-Ciamarra *et al.*, 2011). If livestock being moved are infectious and have contact with others along their market journey this could lead to pathogen dissemination beyond movement origins and destinations. If this is the case it will make surveillance more difficult and makes targeted surveillance at markets less effective as a lower proportion of overall infectious-susceptible contacts and observed infections will be caused at markets. The routes livestock take to markets and local livestock living along them would need to be considered in the design of effective surveillance programmes, similar to what was suggested for locations along major livestock trade routes between Somalia and Ethiopia (Tempia *et al.*, 2010).

There is also a lack of information on the socio-economic drivers for livestock market movements and who the key stakeholders are that are involved in market movements in northern Tanzania. In other LMICs chains of livestock trade movements are driven by a mismatch between the high volume of livestock in rearing locations and low volume in locations where demand for livestock produce is highest (Nicolas *et al.*, 2018; Mtimet *et al.*, 2021). Livestock traders can be prominent actors in the livestock market chain, often trading in multiple countries as they buy, move and sell livestock in different locations, moving along a price gradient to make a profit (Dean *et al.*, 2013). Livestock traders can also be high risk for widespread pathogen dissemination events as they gather and distribute livestock from and to multiple locations (Robinson and Christley, 2007). Knowledge of the reasons for livestock sale or purchase at markets and prices available provides information about the major drivers for livestock market movements and who the key stakeholders are that would be affected by market movement restrictions. The socio-economic factors surrounding livestock market movements in northern Tanzania should be reflected in the design of disease control programmes. The aim would be to ensure stakeholders remain engaged and that interventions are equitable, compensate for socio-economic losses and have minimal negative externalities (Beileh, 2006; Rich and Perry, 2011).

More generally, if livestock movement data is recorded it can be used to construct a network which can then be analysed using concepts from network analysis to assess the level of connectivity and scale of pathogen transmission risk in an area or population. A network constructed from livestock market movements would be made up of nodes that represent markets and other origin and destination locations that livestock have moved from or are moving to, connected by links that represent the movement of animals between nodes.

Once constructed a network structure can be analysed using concepts from network analysis to identify how epidemiologically connected an area is. The giant strongly connected components (GSCC) and weakly connected components (GWCC) of a network can be used to estimate the upper and lower bounds of epidemics in the connected population (Newman, 2010). Pathogen control programmes often aim to fragment networks, breaking them up into smaller disconnected components by removing epidemiological connections while minimizing overall network disruption, to reduce the potential size of epidemics (Marquetoux *et al.*, 2016; Chami *et al.*, 2017). The diameter of a network is the minimum number of steps taken to get between the two most distant nodes in the largest connected component (Newman, 2010). The diameter can be used to estimate of how many sequential livestock movements of infectious animals would be needed to transmit a pathogen between the two most epidemiologically distant nodes on a connected network (Danon *et al.*, 2011). The average path length on a network is the average number of movements needed to get between any two connected nodes (Newman, 2010). This can be useful when thinking about the speed of pathogen movement and subsequent speed needed to implement disease control interventions in the case of an outbreak. Network nodes reciprocity is also useful in disease epidemiology as it measures the proportion of connections on the network that are reciprocated (reported in both directions) and thus how often a pathogen can flow in both directions (Lindström *et al.*, 2013).

Node centrality metrics can be calculated on a network and used to identify whether heterogeneity in movement activity exists between nodes (Kao *et al.*, 2006; Danon *et al.*, 2011). Nodes with high centrality metrics can be those at greatest risk of pathogen introduction or onward transmission and can be made targets for resource efficient disease control programmes (Ortiz-Pelaez *et al.*, 2006; Molia *et al.*, 2016). Node centrality metrics that are often the most informative for epidemiological networks, to identify high risk locations for targeted disease control interventions are; in- and out-degree (number of ingoing or outgoing links); geometric mean degree (square route of product of in-degree and out-degree); and betweenness centrality (number of times a node lies on the shortest path between other pairs of nodes, acting as a bridge)(Dubé *et al.*, 2008; Lichoti *et al.*, 2016; Bucur and Holme, 2020). A detailed description of these centrality metrics and their epidemiological relevance can be found in the introduction to this thesis.

If node centrality metrics have a right skewed distribution this indicates that the network will be vulnerable to fragmentation by using targeted disease control interventions to remove the most central nodes (Büttner *et al.*, 2013a; Chami *et al.*, 2017). Additionally, if a high level of correlation between the different centrality metrics is identified, on a network these can be targeted with multipurpose interventions that can reduce both pathogen introduction and onward transmission.

Two major types of livestock market, primary and secondary, exist in northern Tanzania (Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012) (further details in The Appendix Section 9.3.1). Primary markets are smaller and widespread across the country, run by local governments with a market tax payable by buyers to local governments. Secondary markets are owned by the national government and are larger, busier and generally close to major towns and road networks with a tax payable by buyers to central government. There was no published data available on the differences between primary and secondary livestock markets in northern Tanzania at the time of writing this document. Movement of livestock between markets in the study area has been identified using government movement permit data and this is described in Chapter 5 (Chaters *et al.*, 2019) and Chapter 6. Movement permits are officially required to move livestock between any two locations in Tanzania. Household survey data from northern Tanzania however shows that compliance with permit use is low for household movements to and from markets and they are predominantly used for movements between markets or onward from markets to end point slaughter destinations (Chapter 1 and further details on livestock movement permits in Appendix section 9.3.2). Because of the low rate of permit use at the household level and because livestock have not historically been formally identified nor movements centrally recorded, the large gap in the knowledge around livestock movements between households and markets remains.

It is important that we improve the overall quantitative knowledge around market movements in northern Tanzania so they can be appropriately accounted for in the design of efficient disease control programmes (Beileh, 2006; Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012; Chaters *et al.*, 2019). Policy makers require information on the scale needed for effective disease control programmes, where these can be most efficiently implemented and how to prevent enforcing interventions that are met with resistance and non-compliance. Having detailed knowledge of the characteristics of livestock market movements that will affect pathogen transmission is particularly important in areas with limited resources, high pathogen burden, and a limited awareness of transmission routes (Maudlin, Eisler and Welburn, 2009; Crump *et al.*, 2013; Chipwaza *et al.*, 2014; Colman *et al.*, 2019)

The aim of this study is to use market survey data to address four major knowledge gaps around livestock market movements in northern Tanzania. First we need to improve the knowledge available on how market movements might influence pathogen transmission along a market journey beyond potentially moving a pathogen between distinct locations. To do this we will quantify what proportion of journeys by vehicle and on foot are shared by livestock from different batches and how likely livestock moving to markets are to have contact with local, non-moving livestock, along their journey. Second, a better understanding of the socio-economic drivers behind livestock market movements and the key stakeholders involved with market movements is needed. This would help determine potential drivers of pathogen transmission and who is affected by any restrictions. To achieve this the major reasons for sale and purchase of different species in the different market types will be identified, along with any difference in sale prices between market types. Third, the more quantitative market movement characteristic data gaps that may be associated with the geographical area at risk, and rate of pathogen transmission through market movements will be assessed. This will be achieved by evaluating how many livestock are moved on market journeys, over what distance and the duration of time livestock spend travelling to and from markets. Fourth, a static network of livestock market movements will be constructed to assess how epidemiologically connected northern Tanzania is via market movements which include household movements not necessarily recorded through use of a movement permit. The network will then be used to calculate market and non-market node centrality metrics. Centrality metrics will be evaluated for evidence of heterogeneity in connectivity and thus pathogen acquisition and onward transmission risk and for correlation between metrics to see if high risk locations can be targeted with multi-purpose disease control interventions.

4.2 Methods

4.2.1 Study area and livestock market selection

The data collected for this study were from livestock markets across Arusha, Manyara and Kilimanjaro Regions of northern Tanzania. At the time of this study there were 85 active livestock markets across the three study regions. This included 3 secondary and 82 primary markets. Livestock market data were generated from the Zonal Veterinary Centre in Arusha town, Arusha, Tanzania. To achieve samples from buyers and sellers that were representative of those moving livestock to and from markets in the study regions we sampled a substantial portion (N = 22, 26%) of these (Figure 4-1). To make inferences about the differences between market types and markets across all three regions this required sampling of all three secondary markets in the study area and all four primary markets from Kilimanjaro. The remaining 15 markets consisted of seven primary markets from Manyara Region and eight primary markets from Arusha Region which were selected randomly from the lists of functioning primary markets using the *sample* function in the *base* package in R version 3.6.0 (R Core Team, 2019b). To minimize seasonal variation in observed market characteristics, market visits were carried out within a 7-week period between the 25th April and 16th May 2017. The analysis presented here therefore do not account for seasonal variation in livestock market movements in northern Tanzania. Two markets from the original list of 15 randomly selected primary markets could not be accessed due to logistical constraints during the 7-week study period so these were replaced by randomly selected replacements from the same region. Further descriptive details on livestock markets can be found in the Appendix Sections 9.3.3 and 9.3.4.

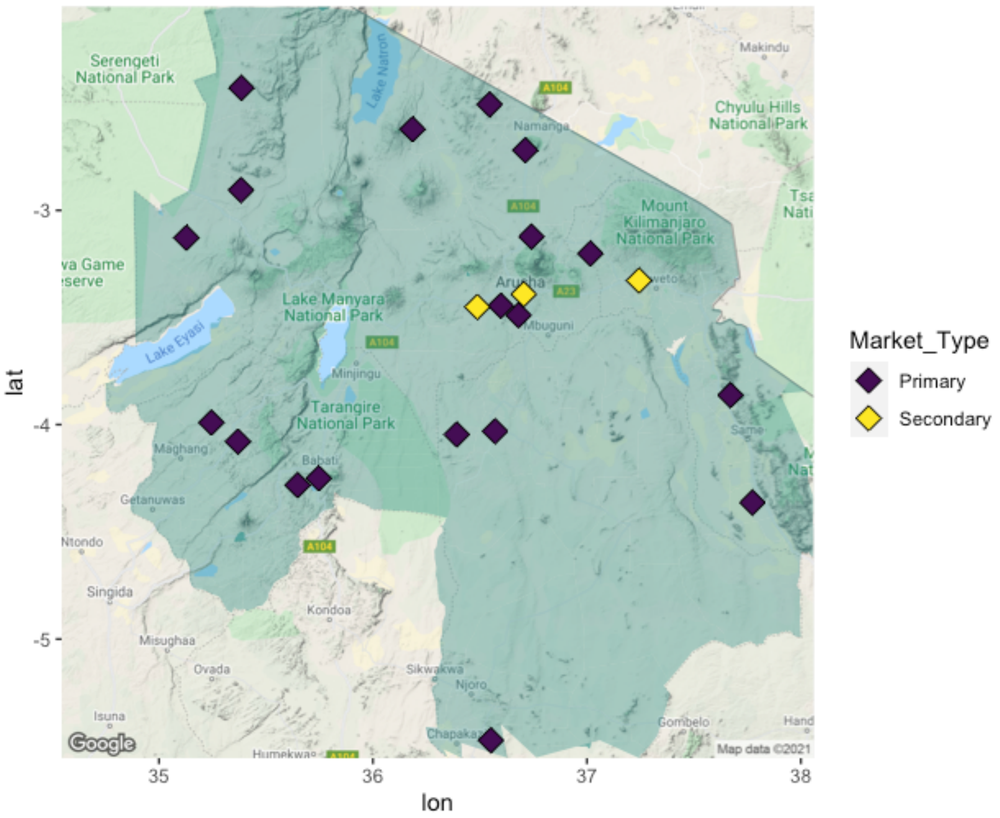


Figure 4-1: Arusha, Manyara and Kilimanjaro study regions highlighted on map of northern Tanzania with sampled markets plotted and point colour representing market type.

4.2.2 Survey

A field team of three interviewers conducted the market survey at all 3 secondary markets and 19 (23%) of 82 primary markets in the study area. The survey was created and implemented using Open Data Kit (ODK) Collect software (https://opendatakit.org/) on portable computer tablets in Kiswahili or Maa. The market survey can be found in Appendix section 9.3.5

4.2.3 Ethics

All participants provided written informed consent. The protocols, questionnaire tools and consent and assent procedures were approved by the ethics review committees of the Kilimanjaro Christian Medical Centre (KCMC/832) and National Institute of Medical Research (NIMR/2028) in Tanzania, and in the UK by the ethics review committee of the College of Medical, Veterinary and Life Sciences at the University of Glasgow (39a/15). Approval for study activities was also provided by the Tanzanian Commission for Science and Technology (COSTECH) and by the Tanzanian Ministry of Livestock and Fisheries.

4.2.4 Participant selection

The three interviewers arrived together at the start of the market day and each had a target of completing a minimum of 10 surveys per market visited. An effort was made to ensure an equal number of sellers and buyers were surveyed by asking one seller and then one buyer, but this was not always possible as buyers move around and can be difficult to locate. Where markets had both small and large ruminants for sale, interviewers split themselves evenly between the separate market sections during the visit spending half their time in each section.

A central point was selected in the market and the three interviewers set out at the same time in different directions following a spiral or zig-zag direction of movement, depending on the shape of the market area, to ensure all corners of each market area were covered. Multiple entry points into the market necessitated the spiral or zig-zag strategy to capture people arriving from different locations and directions, and congregating in different areas of the market.

Depending on the size of the market, number of buyers and sellers present and buyer and seller willingness to participate, every seller, every other seller or every third seller was interviewed to ensure the sample was representative of the buyer and seller population. Sellers generally stand with their livestock until sold, buyers were more difficult to identify and locate within the markets. Buyers were therefore interviewed if they were within easy conversation distance of the interviewer when a seller’s interview was finished. If a buyer was not near they were sought out between interviews with sellers by the local livestock field officer or market master who was familiar with and could recognise the buyers present, and asked them if they were willing to participate in the study.

4.2.5 Statistical analysis

4.2.5.1 Descriptive summary

We summarised market movement characteristics that could influence the extent and rate of pathogen transmission and the opportunity for transmission events during market journeys (Table 4-1). The data were stratified by market type (primary and secondary) and species group (cattle or small ruminants) to compare characteristics of market movements (and therefore pathogen transmission risk) between the four groups. Binomial and Gaussian generalised linear mixed models (GLMMs) are fitted to the descriptive journey characteristics data, depending on whether the outcome of interest is binary or continuous, with market fitted as a random effect. The assumptions of linearity and normally distributed residuals were assessed visually by plotting the residuals against the fitted values. Residuals for outcomes numbers sold and purchased, journey distance and journey duration required log-transformation. Where model outcomes were log transformed (including the binomial models), any effect estimates calculated using the models were adjusted for Jensen's inequality due to the back transformation of the logged outcome variable and in case of uneven sampling across markets (Nakagawa, Johnson and Schielzeth, 2017). For model outcomes that relate to contacts made on the journey, either through journey sharing or close contacts with local livestock at grazing and watering points en route, mode of transport was also included as a covariate in the descriptive models.

Results of the binomial GLMMs are given as odds ratios. The results given in the tables and text are the multiplicative effect of each variable on the baseline value. For the Gaussian GLMM's estimating price data, where the continuous outcomes is not log transformed, the effect estimate results presented are additive.

All data processing and analysis was done using R statistical software (R Core Team, 2019b). The Binomial and Gaussian GLMMs were fitted using the *lme4* package (Bates *et al.*, 2015). The validity of model assumptions was judged by visualising the model residuals to assess if they exhibited homoscedasticity using the *plotResiduals* function in the *DHARMa* package (Hartig, 2020). The marginal R^2 (variance explained by the fixed effects) and conditional R^2 (variance explained by both fixed and random effects) of the GLMMs are calculated using the *rsquared* function in the *piecewiseSEM* package (Lefcheck, 2016). All plots were made using the *ggplot2* package (Wickham, 2016). All P-values shown in the results are Wald P-values unless otherwise stated.

Table 4-1: List of market journey characteristics described in this data

	Journey Characteristic	Binary or continuous	Stratified by
Journey contacts and transmission opportunities	Transport to market (Vehicle or foot)	Binary	Market type, Species group
	Transport from market (Vehicle or foot)	Binary	Market type, Species group
	Journey to market shared with others	Binary	Market type, Species group
	Journey from market shared with others	Binary	Market type, Species group, Transport used
	Contacts made with local livestock on journey to market	Binary	Market type, Species group, Transport used
Socio-economic drivers	Reason for sale (As part of a livestock trading business or for private household needs)	Binary	Market type, Species group
	Origin	Binary	Market type, Species group
	Price	Continuous	Market type, Species group
Quantitative factors	Number sold/purchased	Continuous	Market type, Species group
	Journey distance	Continuous	Market type, Species group,
	Journey duration	Continuous	Market type, Species group

4.2.5.2 Transport

Binomial GLMMs were used to estimate the proportion of livestock transported to and from market by vehicle and assess whether this differed between cattle and small ruminants and market type. If mode of transport differs for the different market journeys this will affect both the speed of pathogen spread and the potential for pathogen dissemination along market routes.

4.2.5.3 Shared journey contacts

Binomial GLMMs were used to estimate the proportion of livestock batches sold or purchased that shared (or would share) their market journey with livestock from other households (or batches), adjusting for market type, species group, and mode of transport. If shared journeys are more common for livestock transported by vehicles there will be greater risk of pathogen transmission between these animals as they are kept in close physical confinement and share the same air space.

4.2.5.4 Local livestock journey contacts

Binomial GLMMs were used to assess if there was evidence of a difference in close contacts made with livestock from other households at grazing or watering points during journeys to market, adjusting for market type, species group and mode of transport. If there are high rates of close contact with local livestock along market journeys these could provide pathogen transmission opportunities and should be accounted for in the design of disease control programmes.

4.2.5.5 Reason for sale or purchase

Binomial GLMMs were used to assess whether there was evidence of a difference in the reason for sale (traders selling livestock for business reasons, as part of a commercial enterprise or private sellers selling their own animals to directly generate household income) between cattle and small ruminant batches sold at the different market types. This will help identify who the key stake holders are that will be affected by livestock market centred interventions or movement restrictions and thus who needs to be engaged in the development of equitable interventions to improve both compliance and effectiveness.

4.2.5.6 Origins

Sellers were asked what type of origin of the livestock they were selling came from, the response options to select from were *home*, *another household/village* or *another market* or *other*. The number and type of livestock origins will influence the risk of pathogen introduction into the market system. Improving the understanding of where livestock that enter the market system come from will help ascertain the scale needed for disease control interventions. This information will identify where knowledge exchange campaigns can be most effectively targeted, especially if livestock entering the market chain are primarily privately owned in contrast to sold by traders and gathered from multiple locations.

4.2.5.7 Price

Gaussian GLMMs were used to assess if there was evidence of a difference in price wanted or price achieved, for cattle and small ruminants sold at primary and secondary markets. If price differs between locations and market types this could affect the distances people are willing to travel and how much they are willing to spend on using a vehicle to transport their livestock. The distances travelled and ability to spend money on using a vehicle to transport the livestock to market could both influence the speed and extent of pathogen dissemination through market centred livestock movements. Price disparity between locations would also identify if and where improved market access is needed by livestock keepers to ensure they can access the best prices.

4.2.5.8 Numbers sold and purchased

Gaussian GLMMs were used to assess if there was evidence of a difference in the number of livestock being bought and sold at each market type for cattle and small ruminants. Livestock managed in larger batches might have a higher risk of pathogen spread within the group if one moving animal is infectious.

4.2.5.9 Journey distance and duration

Gaussian GLMM's were used to assess if there was evidence of a relationship between the outcomes journey distance (calculated as the great circle distance between origin/destination co-ordinates and market co-ordinates of all reported livestock market movements captured by the survey data) and duration (reported by seller/buyer) for both the buyer and seller data, with multiple covariates fitted in each model. The distance and speed of livestock movements to and from markets will be directly associated with the speed and geographical area at risk of pathogen transmission as a result of these movements. The explanatory variables included in both the distance and duration models for buyer and seller data were market type (primary or secondary), species (cattle or small ruminant), number in batch, transport type (vehicle or foot), reason for sale or purchase (private or business), shared journey and journey contacts made. Variables that showed no evidence of an association with any of the outcome variables (LR test P-value <0.05) were removed from the models using backwards stepwise selection.

4.2.6 Network analysis

4.2.6.1 Network construction and analysis

Livestock origin and destination data collected from sellers and buyers was used to construct a market-centred static livestock movement network. The number and size of the strongly and weakly connected components (all nodes in the component can reach each other by a directed or undirected path respectively) of the network were calculated to assess the extent to which market-centred movements connect geographically distinct locations. The network diameter, average path length and reciprocity of the market network were also calculated.

Central point coordinates from sampled markets and origins or destinations livestock travelled from or to were used to represent network nodes. Links were formed between these nodes, weighted by the number of livestock moved, each time a movement was reported. Unweighted node centrality metrics in-degree, out-degree, betweeness and geometric mean degree were calculated for all market nodes. The same node metrics were calculated for non-sampled origin/destination location nodes but this was done on the weighted network with links weighted by the total number of animals reportedly moving to and from the destinations. For market nodes the centrality metrics were used to assess whether heterogeneity existed between markets in their contribution connectivity, and thus pathogen transmission risk, in the study area. Unweighted unique links were used to calculate the market nodes centrality metrics in an effort to adjust for any bias in sampling size and proportion of sellers and buyers sampled at the different markets. A node would have an in-degree of 1 if 10 sellers surveyed were all selling cattle and had all travelled from the same location. If all survey respondents listed a different origin then the markets in-degree would be 10. The larger a markets in-degree the greater diversity in origins the livestock in the market are sampled from. Weighted links were used to analyse non-market nodes to identify if locations outside the sampled markets have a large amount of livestock traffic and therefore a high risk of pathogen introduction or onward transmission.

4.2.6.2 Node centrality metrics

The non-market nodes were analysed separately, to assess if there was evidence of heterogeneity in their centrality metrics and if any of these could be made targets for resource efficient disease control interventions. Other studies have found targeting contacts or ‘friends’ of randomly sampled nodes to be an efficient way to identify highly connected network nodes and fragment networks by removing these highly connected nodes, if there is again evidence of heterogeneity in node centrality measures.

Correlation between centrality metrics was estimated using Spearman’s rho separately for the market and non-market nodes and the P-value to test the significance of any observed correlation calculated using the AS 89 algorithm (Best and Roberts, 1975). A high level of positive correlation between centrality metrics would allow the most influential nodes in the network to be identified. These could then be targeted with multipurpose disease control interventions that aim to both reduce pathogen introduction and onward transmission. Multi-purpose interventions would have the benefit of a shared initial fixed cost of community engagement and infrastructure development so authorities would save resources by not having to implement different types of interventions in different locations.

Northern Tanzania has four major agro-ecological classes that villages have been divided into which are pastoral, agropastoral, small holder and urban. In pastoral areas the main household income is from livestock, in agropastoral areas household income is from both livestock and crops, in small holder areas livestock are kept tethered at the household in small numbers and these are more densely human and livestock populated areas. Urban areas are classified as those where no large livestock are kept by the households. Further details on agro-ecological classifications in the study area can be found in (de Glanville *et al.*, 2020). The agro-ecological class of market locations might affect the type of movement required to access it and the probability of connection to the individual locations so market node centrality metrics should be assessed to see if they differ between agro-ecological classes of market locations.

For the market nodes, Gaussian GLMMs were used to assess the relationship between the two centrality metrics *in-degree* and *betweenness* (fitted as continuous outcomes) and the market characteristic covariates; market type (primary and secondary) and agro-ecological class of the market area (pastoral, agropastoral, small holder/urban). Market was fitted in each model as a random effect. This was done to evaluate if the study results could be generalised from across the study area and beyond if markets identified as high risk based on their centrality metrics also had other defining characteristics. The network construction and analysis was done in R statistical software (R Core Team, 2019b) using the *igraph* package (Csardi and Nepusz, 2006). Livestock market movement network plots were created using the *ggplot2* package (Wickham, 2016).

4.3 Results

4.3.1 Number of surveys

Data were collected from 507 sellers and 216 buyers at 19 primary markets and three secondary markets. The numbers of buyers and sellers surveyed stratified by market type and those selling cattle or small ruminants are given in Table 4-2. The majority of sellers and buyers surveyed at primary markets were dealing with small ruminants whereas the majority in secondary markets were dealing with cattle. Data from five sellers and seven buyers dealing with both species groups and one small ruminant seller (primary market) and two cattle sellers (secondary market) who had used mixed transport were excluded from descriptive analysis as these group sizes are too small to analyse. Raw survey data for all variables analysed are given in Appendix Table S3.1 stratified by market type and species group. All model estimates described in these results and the corresponding model summary statistics which include number of observations, random effects, variance and marginal and conditional R^2 values are given in this document in Table 4-3 and in the Appendix Table S3.9 respectively.

Table 4-2: Number of buyers and sellers surveyed at different market types stratified by species group being sold or purchased

		Primary market (N = 19)	Secondary market (N = 3)	Total
Sellers	**All**	**425**	**82**	**507**
	Cattle	137 (32%)	66 (80%)	203
	Small ruminant	283 (67%)	16 (20%)	299
	Cattle and small ruminants	5 (1%)		5
Buyers	**All**	**185**	**31**	**216**
	Cattle	46 (25%)	22 (71%)	68
	Small ruminant	132 (71%)	9 (29%)	141
	Cattle and small ruminants	7 (4%)		7

4.3.2 Journey transport and contacts

4.3.2.1 Mode of transport (vehicle or on foot) used to and from market

Overall, the majority of livestock journeys to and from markets are made on foot (Figure 4-2). Sellers at secondary markets were more likely to have moved their livestock to market by vehicle compared to those who had travelled to primary markets (OR 6.47, 95% CI 1.12-37.46, P = 0.04). Forty-two percent (95% CI 17-71%) of journeys to secondary markets were made by vehicle compared to only 15% (95% CI 8-23%) of journeys made to primary markets (Table 4-3). Mode of transport did not differ between cattle and small ruminant groups sellers (P = 0.30). Buyers of small ruminants were more likely to use a vehicle for onward transport (OR 4.23, 95% CI 1.51-11.85, P = 0.006) compared to cattle buyers; this represents 50% of small ruminant buyers (95% CI 36-63%) compared to 28% of cattle buyers (95% CI 15-45%) (Table 4-3). There was no difference in mode of transport used for onward journeys between livestock purchased at primary or secondary markets (P = 0.31).

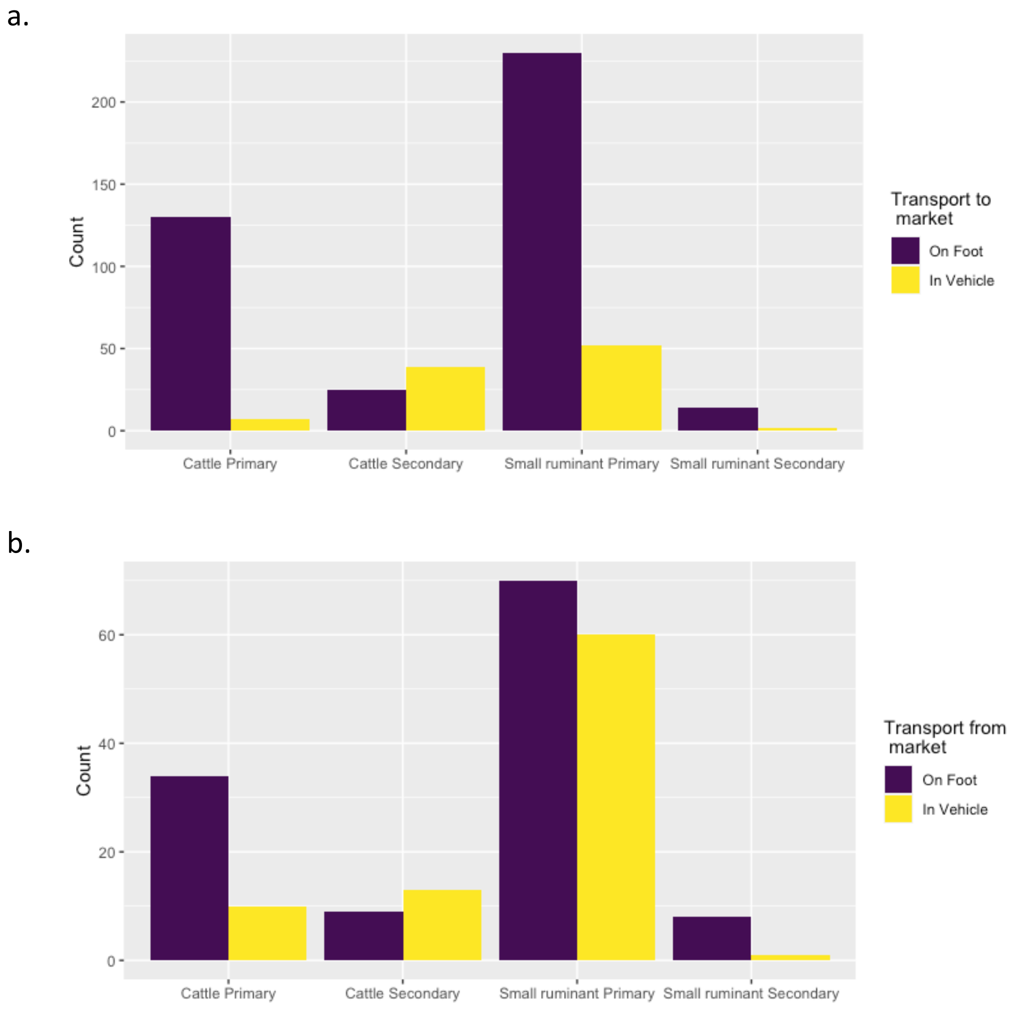


Figure 4-2: Mode of transport used to get cattle and small ruminants to (a) and from (b) primary and secondary markets

4.3.2.2 Journey contacts with livestock from other households

The majority of livestock batches being sold at both primary and secondary markets had some contact with livestock from another household during their journey to market (Figure 4-3). The proportions of contact type were similar regardless of species group and destination market type.

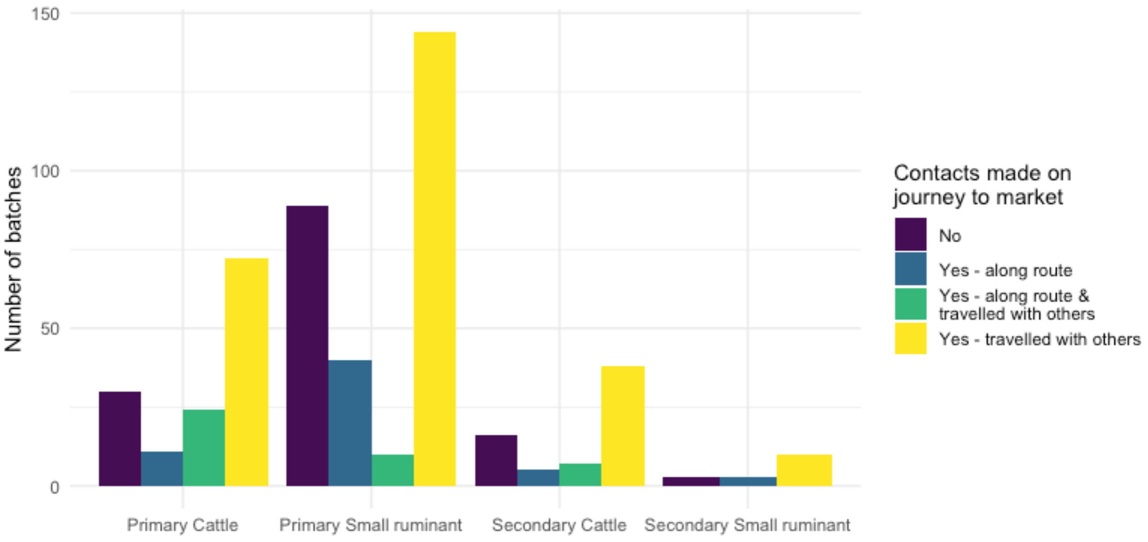


Figure 4-3: Types of journey contacts made by batches of livestock as they are moved to market, stratified by market type

Livestock moved to market on foot were more likely (OR 2.4, 95% CI 1.15-5.07, P = 0.02) to make contact with livestock from other households at grazing and watering points along their journey compared to livestock moved to market by vehicle. Twenty-one percent (95% CI 17-27%) of livestock batches moved to market by foot had some close contact with livestock from other non-moving, local households at grazing and watering points compared to 10% when moved by vehicle (95% CI 5-19%) (Table 4-3). There was no evidence to suggest that species group or the type of market being moved to had any effect on contacts made at grazing and water points during the journey (P = 0.07 and 0.35 respectively).

4.3.2.3 Journey sharing to and from markets

Livestock transported to and from market by vehicle were more likely to be transported with livestock that belonged to other households or batches compared livestock moved to and from market on foot (OR 4.19, 95% CI 2.37-7.40 P <0.001 and OR 6.56 95% CI 2.78-15.48, P <0.001 respectively). Cattle were more likely to share their journey to and from market with livestock from another household or batch compared to small ruminants (OR 1.53, 95% CI 1.03-2.28, P = 0.03 and OR 2.75, 95% CI 1.23-6.12, P = 0.01 respectively). There was no evidence of an effect of market type being travelled to or from on the odds of the market journey being shared with others (P = 0.91 and 0.65 respectively). Model estimates for the proportion of cattle and small ruminants that share their market journey with others, for journeys to and from markets, both by vehicle and on foot, are given in Table 4-3.

4.3.3 Socio-economic factors

4.3.3.1 Reason for sale and money use

Livestock sold in secondary markets were more likely to be sold by traders listing business as the major reason for sale, compared to livestock being sold in primary markets (OR 7.1, 95% CI 1.9-26.4, P = 0.003). An estimated 47% (95% CI 24-70%, P = 0.003) of sales in secondary markets were made by livestock traders rather than private sellers compared to only 14% of sales in primary markets (95% CI 9%-21%)(Table 4-3). There was no difference in the proportion of sellers listing business as the primary reason for livestock sale (business or private) between those dealing with cattle or small ruminants (P = 0.57).

A summary of the number of responses given for reasons for livestock sale are given in the Appendix table S3.2. Some of the more common reasons listed for the sale of livestock include ‘to generate money to buy other (non-livestock) things’ 199/502 (40%) and ‘for household purchases/problem at home’ 41/502 (8%). In response to the question about what the money generated from the livestock sales will be used for, the most common response was ‘for a household purchase’ (223/502 (44%)). Other common responses given to this question include ‘school fees’ 154/502 (30%), ‘treating illness (livestock or human)’ 125/502 (25%), ‘buying food’ 68/502 (13%) and for ‘business’ 96/502 (19%). A full list of responses about the what the money generated from livestock sales will be used for, and the number of respondents listing these reasons, are given in the Appendix table S3.3. Further details on who sellers are hoping to sell to and what they plan to do with non-sold livestock are given in Appendix tables S3.4 and S3.5. Eight percent of livestock sellers (all at secondary markets) planned to leave any unsold livestock at the market site until the next active market day, 78% planned to take any unsold livestock home.

4.3.3.2 Reason for purchase

The most commonly listed primary reason for choosing to buy specific livestock was ‘price’ 128/216 (59%). Other responses given for *primary reason for choosing to purchase these animals* were ‘health of livestock’ (39% (N=85)), followed by ‘quality of animal’ (36% (N=78)), ‘amount of meat on carcase’ (26% (N=56)) and finally for ‘milk production’ (2.2% (N=5)).

4.3.3.3 Price

Cattle sellers at secondary markets were hoping to achieve, and generally did achieve, a higher price for the cattle they were selling compared to cattle sellers at primary markets (57% and 42% more, respectively) (Figure 4-4 and Table 4-3). For small ruminants there was no evidence of a difference between the price per head wanted or achieved at primary and secondary markets (P = 0.42 and 0.34 respectively). Summary data on price wanted and price achieved, stratified by species group and market type is in Appendix Tables S3.6 and S3.7.

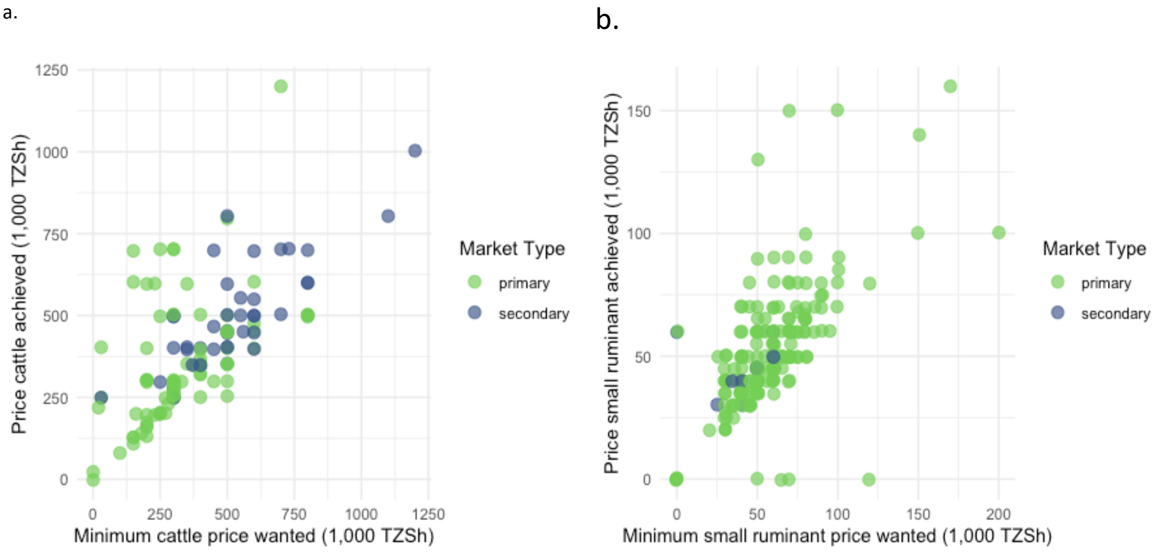


Figure 4-4: Minimum price wanted and price achieved for (a) cattle and (b) small ruminants being sold at primary and secondary markets in northern Tanzania. Data displayed from N= 131 cattle sellers and N = 239 small ruminant sellers who were able to responded to both questions.

4.3.3.4 Livestock origins

The majority of sellers in both market types had brought the livestock they were selling from their home. Other types of location livestock were reported to have originated from include 'other markets' and 'other villages' (Figure 4-5). Cattle sellers in secondary markets were less likely to have brought the animals from home compared to cattle sellers in primary markets and any small ruminants sellers (OR 0.21, 0.08-0.55, P = 0.001) (Figure 4-6). An estimated 80% of sellers at primary markets (cattle 95% CI 71-88% and small ruminants 95% CI 77-88%) and small ruminant sellers at secondary markets (95% CI 62-97%) were selling livestock brought from home compared to only 50% of cattle sellers at secondary markets (95% CI 33-67%). A summary of the raw data for responses to livestock origin and if the livestock sold originally belonged to the seller is given in Appendix Tables S3.1 and S3.8 respectively.

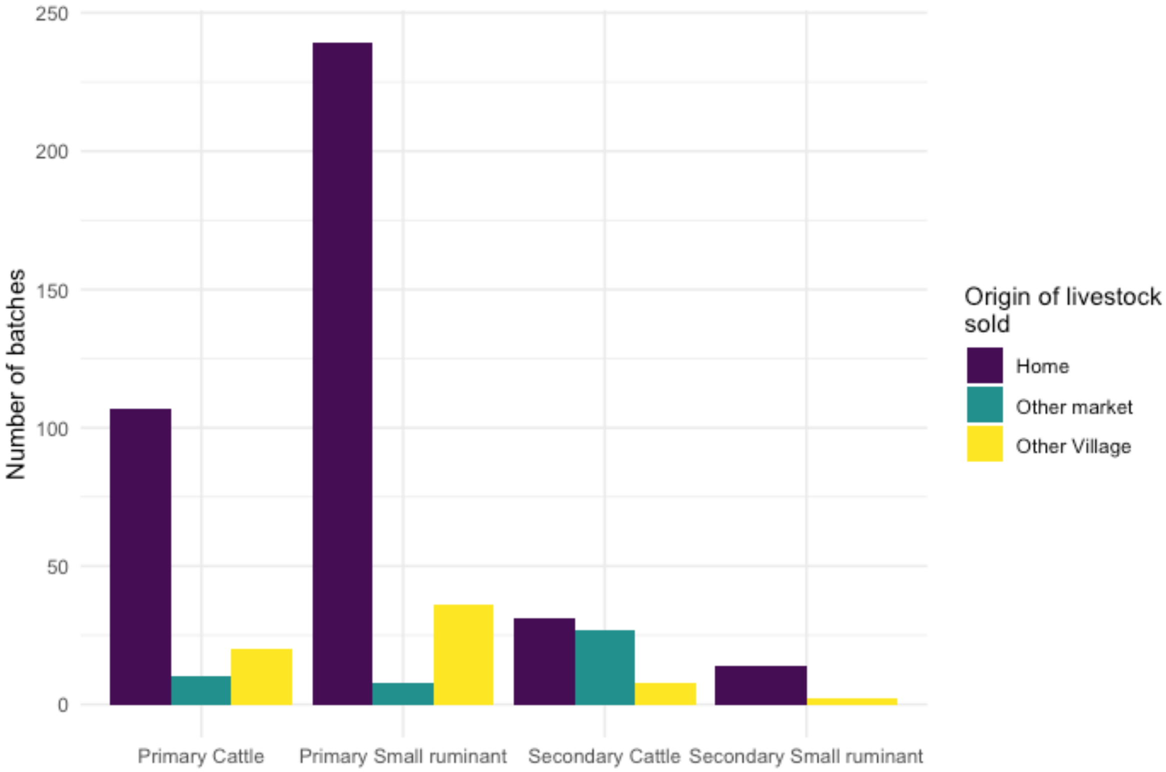


Figure 4-5: Type of origin cattle and small ruminants sold in primary and secondary markets came from

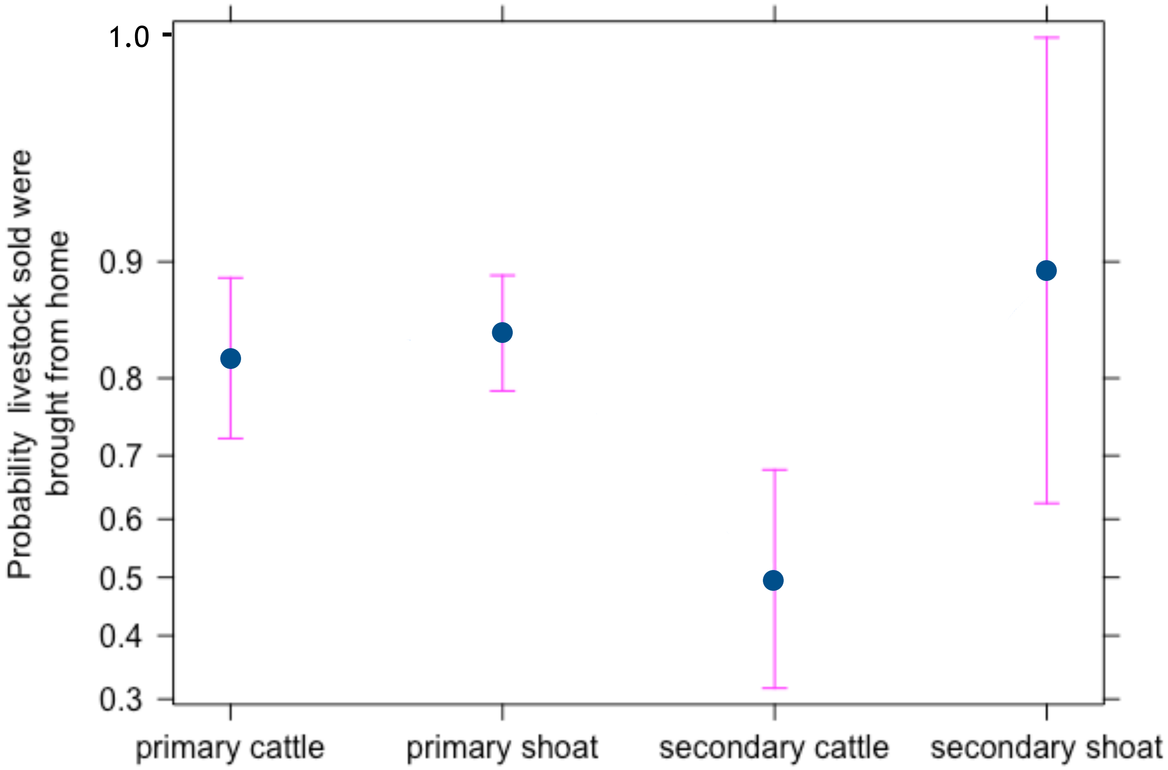


Figure 4-6: Model estimates for probability that a seller is selling livestock brought from home rather than some other location in the primary and secondary markets, for cattle and small ruminant (shoat) sellers

4.3.3.5 Livestock destinations

Other market was the most commonly reported destination for livestock batches purchased at primary markets (48% cattle and 44% small ruminants) and small ruminants purchased at secondary markets (66%) (Table 4-3 and Figure 4-7). The effect of market type differed between species groups for destination choice *meat* and *other market* so an interaction was fitted between market type and species group in these models (P = 0.05). Cattle batches purchased at secondary markets were most likely to be taken direct to slaughter (OR 7.21, 95% CI 1.73-30.0, P = 0.007) and had a lower odds of being taken on to other markets (OR 0.11, 95% CI 0.02-0.81, P = 0.03) compared to small ruminants purchased at secondary markets and both cattle and small ruminants purchased at primary markets. Livestock purchased at secondary markets were less likely to be taken home compared to those purchased at primary markets (OR 0.19, 95% CI 0.04-0.89, P = 0.03) and there was no evidence to suggest this effect differed between species groups (P = 0.47). Small ruminants purchased at primary and secondary markets were most likely to be taken on to another market (44% from primary markets and 66% purchased at secondary markets). Model estimates for the proportion of purchased livestock at each market going to the three most commonly reported destination types is given in Table 4-3 and a summary of the raw data numbers of buyers taking each livestock species to each destination are given in the Appendix Table S3.1.

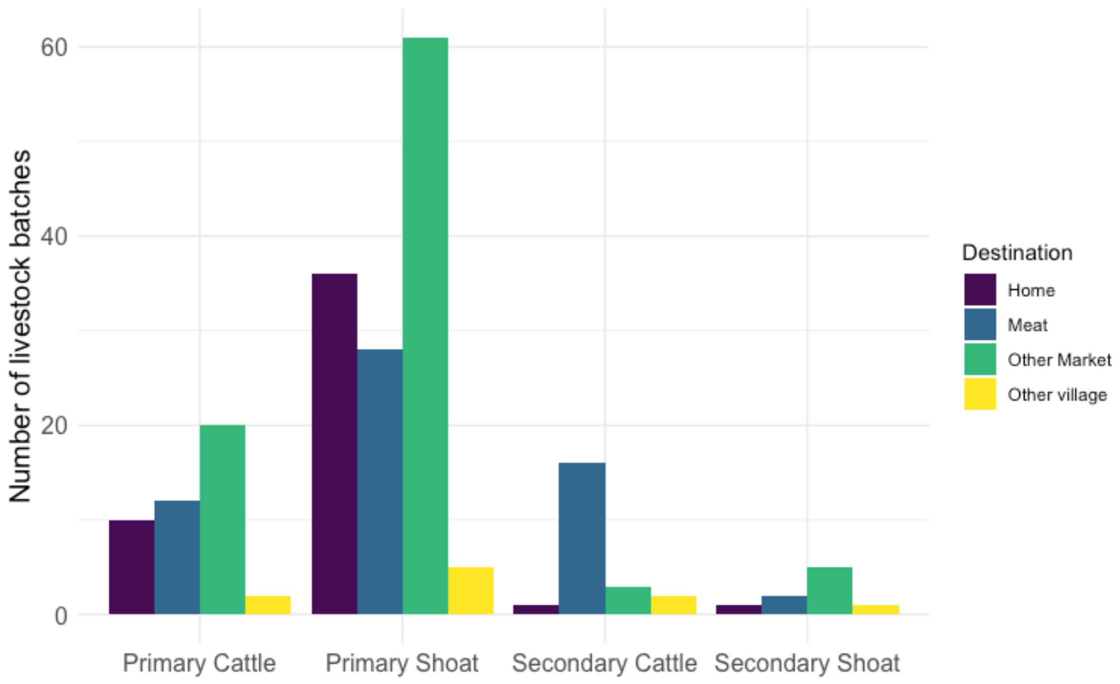


Figure 4-7: Number of buyers taking the livestock they purchased to their home, for slaughter, to another market or to another village location.

Table 4-3: Model estimated proportion (95% CI) of livestock batches in each variable category, by market type and species, estimated from binomial GLMMs

(see methods for details)

	Primary (N = 19)		Secondary (N = 3)	
	Cattle	Small ruminants	Cattle	Small ruminants
Estimated proportion of livestock Transported to market in vehicle	15% (8-23%)		42% (17-71%)	
Estimated proportion of livestock Transported from market in vehicle	28% (15-45)	50% (36-63)	28% (15-45)	50% (36-63)
Estimated proportion of livestock that contacted local livestock at grazing or watering, of those moved to market on **foot**	21% (17-27%)			
Estimated proportion of livestock that contacted local livestock at grazing or watering, of those moved to market in a **vehicle**	10% (5-19%)			
Estimated proportion of livestock sharing their journey to market with others, for those transported by **foot**	59% (51-67)	39% (30-49)	59% (51-67)	39% (30-49)
Estimated proportion of livestock sharing their journey to market with others, for those transported by **vehicle**	79% (60-90)	73% (57-85)	79% (60-90)	73% (57-85)
Estimated proportion of livestock sharing their onward journey from market with others, for those transported by **vehicle**	90% (77-96)	78% (65-87)	90% (77-96)	78% (65-87)
Estimated proportion of livestock sharing their onward journey from market with others, for those transported on **foot**	61% (46-75)	39% (27-53)	61% (46-75)	39% (27-53)
Proportion of livestock sold for Business	14% (9-21)		47% (24-70)	
Estimated Price wanted (1,000 TzSh)	329 (273– 384)	59 (55-63)	516 (412-621)	58 (55-63)
Estimated Price achieved (1,000 TzSh)	343 (276-408)	46 (30-62)	486 (365-607)	46 (30-62)
Estimated proportion of livestock sold that originated from the home	81% (71-88)	83% (77-88)	50% (33-67)	88% (62-97)
Estimated proportion of livestock purchased that were destined for another market	48% (32-64)	44% (33-56)	12% (3-40)	66% (26-92)
Estimated proportion of livestock purchased that were destined for meat/consumption	26% (14-43)	20% (13-30)	69% (43-87)	20% (4-59)
Estimated proportion of livestock purchased that were destined for the home	23% (12-40)	27% (20-36)	5% (1-27)	11% (1-52)

4.3.4 Quantitative journey characteristics

4.3.4.1 Batch sizes

There is strong evidence of a difference in estimated batch sizes sold between cattle and small ruminants sold at primary and secondary markets (Table 4-4 and Figure 4-8). Cattle sold at secondary markets were sold in batches that were around 5 times larger than batches sold at primary markets (95% CI 3.05 – 7.87, P <0.001). Small ruminants sold at primary and secondary markets were also sold in larger batches than cattle at primary markets (P <0.001) but there was no evidence of a difference in batch sizes of sold small ruminants between the different market types (Table 4-4).

Small ruminants were purchased in larger batches than cattle (95% CI 1.12-2.30, P = 0.009) (Table 4-4 and Figure 4-9). There was no evidence of an effect of market type on the batch size of purchased cattle or small ruminants (P = 0.60) though the number of buyers surveyed at secondary markets (N=22 cattle buyers and N=9 small ruminant buyers) are too small for meaningful differences to be excluded.

Table 4-4: Model estimates (95% CI) from Gaussian GLMMs for batch size of cattle and small ruminants sold and purchased at primary and secondary markets

	Primary (N = 19)		Secondary (N = 3)	
	Cattle	Small ruminants	Cattle	Small ruminants
Batch Size Sold	2.64 (2.15-3.23)	3.53 (2.96-4.22)	9.36 (6.11-14.35)	4.69 (2.74-8.03)
Batch size Purchased	8.24 (5.90 – 11.63)	13.32 (10.23 – 17.34)	8.24 (5.90 – 11.63)	13.32 (10.23 – 17.34)

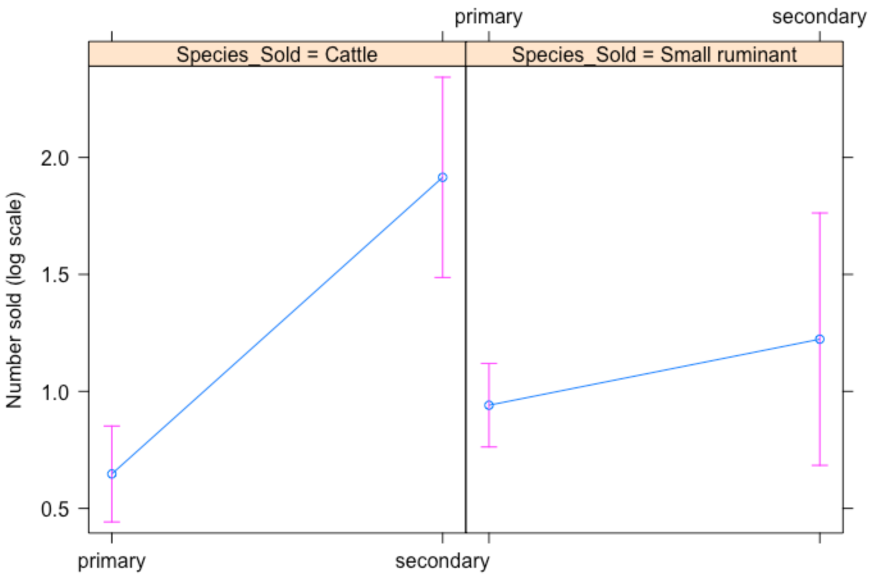


Figure 4-8: Batch sizes for cattle and small ruminants sold at primary and secondary markets

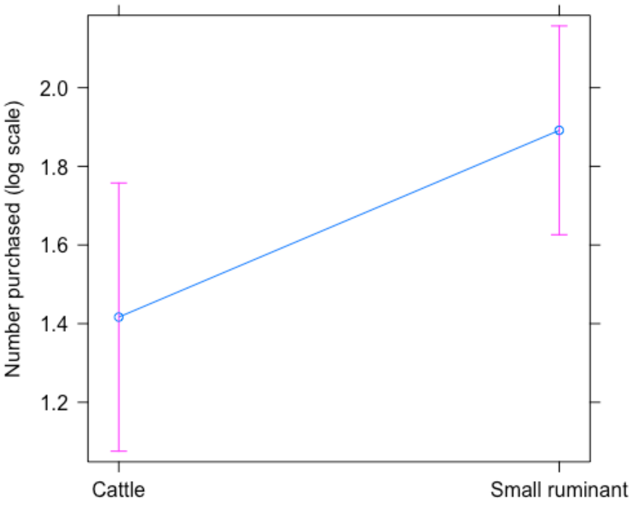


Figure 4-9: Batch sizes for purchased cattle and small ruminants

4.3.4.2 Journey duration

When all other variables are controlled for, cattle spent longer travelling to market than small ruminants (4.2 hours (95% CI 3.5-5.2) vs 2.1 hours (95% CI 1.8-2.5)) and journeys to secondary markets were 68% longer duration (95% CI 19-138% P = 0.007) than journeys to primary markets. Journeys that were primarily for business reasons took 65% longer than private sale journeys (95% CI 28-113%, P <0.001) and livestock that experienced close contact with local livestock at grazing or watering points along their journey had travelled for longer compared to those that had no contact with local livestock (31% longer journeys, 95% CI 5-64% P = 0.02). Journey duration also increased by an estimated 28% for every twofold increase in the size of the batch of livestock being moved to market (95% CI 18-38%, P <0.001).

For onward journeys from market, the estimated journey duration for one livestock being transported on foot from primary markets, without being grouped with others, is 4.75 hours (95% CI 3.0-7.5 hours). There was no evidence of a difference in the duration of onward journey between cattle and small ruminants (P = 0.86). A twofold increase in batch size was associated with a 24% increase in estimated onward journey duration (95% CI 12-38% P <0.001). An increase in onward journey duration was also positively associated with livestock being purchased based on them being a good price and a plan to share the onward journey with other batches of livestock (60% (95% CI 13-126% P = 0.01) and 175% (95% CI 87-303% P <0.001) increase respectively). Onward journeys from secondary markets and those that used a vehicle were both associated with shorter journey duration times (0.44 (95% CI 0.20-0.97, P = 0.05) and 0.39 (95% CI 0.26-0.59, P <0.001) respectively). Results from both journey-to-market and journey-from-market duration models are plotted in Figures 4-10 and 4-11 given in Table 4-5.

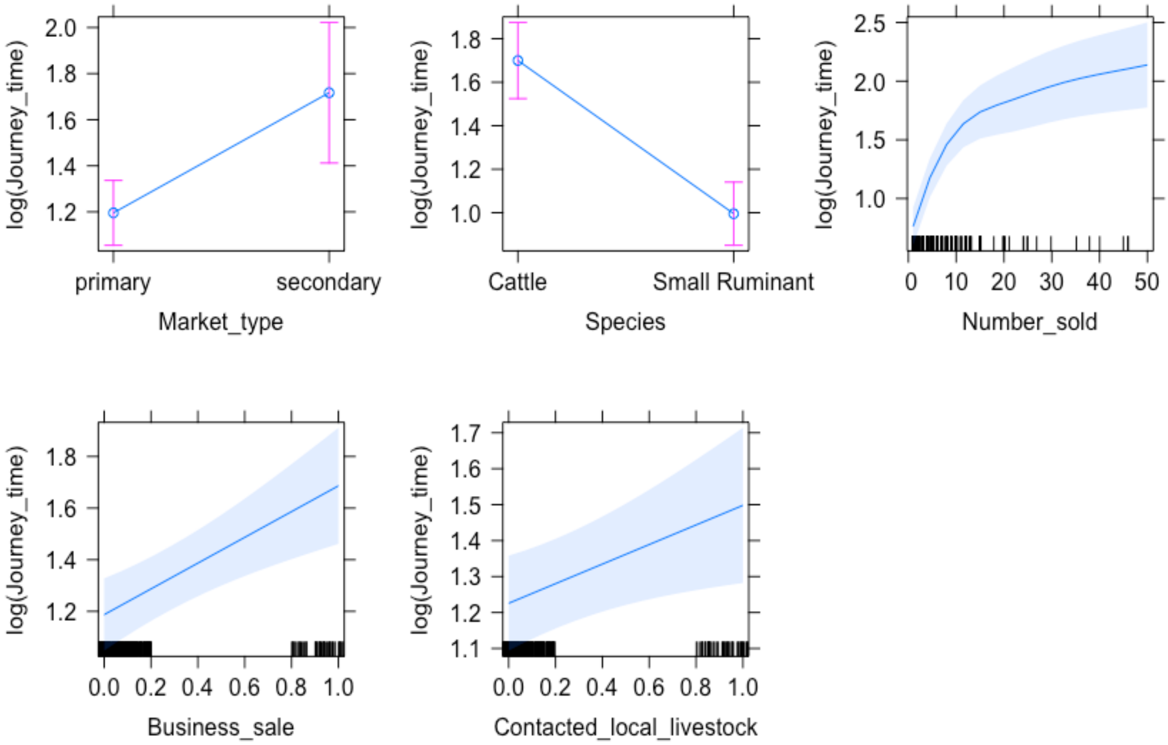


Figure 4-10: GLMM results for effects of covariates on livestock journey to market duration

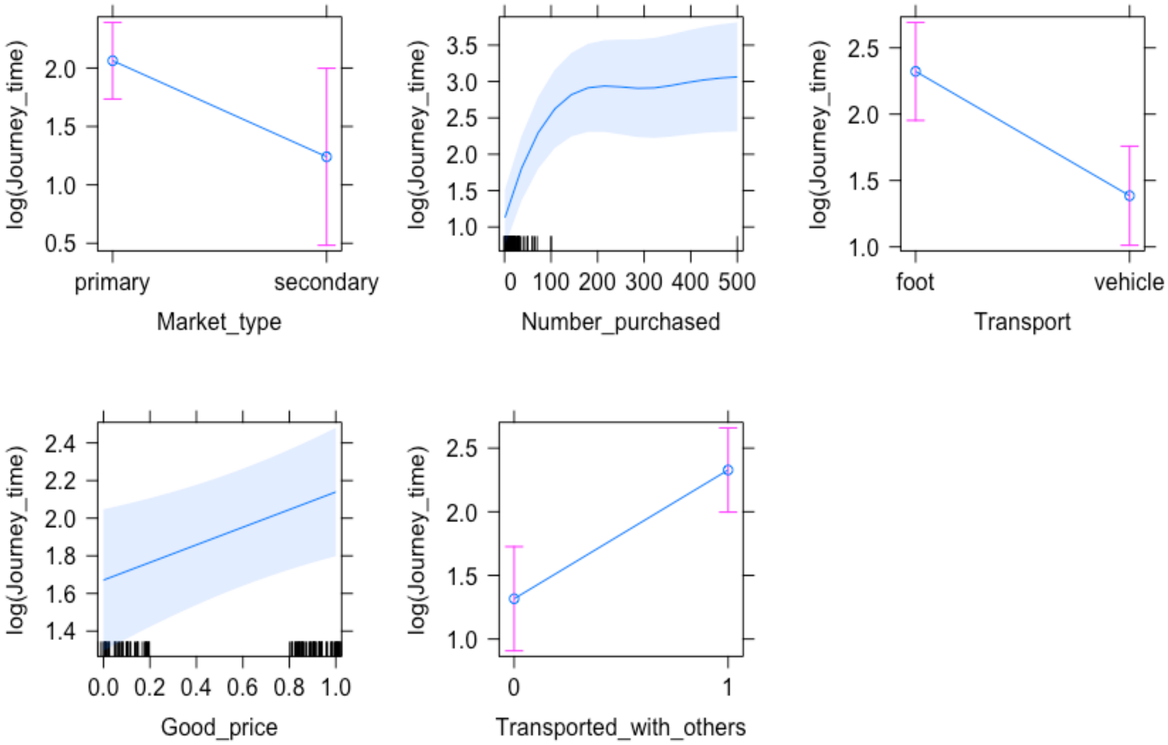


Figure 4-11: GLMM effects of covariates on journey from market duration

Table 4-5: Results from generalised linear mixed model, with journey duration (time in hours) fitted on the log scale to satisfy model assumption of normally distributed outcome. In the top row of the table the results show the estimated journey duration in hours for the baseline group (1 cattle, going to/from primary market on foot as a private sale and having no contacts with local livestock en route), adjusting for Jensen's inequality. Rows below show the multiplicative effect estimates of each variable on journey duration to and from market with 95% confidence intervals.

		Regression coefficient estimate (95% CI)	
		Journey to market (N_{obs} = 499)	*Journey from market* (N_{obs} = 205)
Predicted duration in hours (#1 Cattle, Primary market, On foot, private sale, no contacts with local livestock on journey)		**4.24 (3.45, 5.20)**	**4.75 (3.01, 7.50)**
Multiplicative effect estimate:	Secondary market	1.68 (1.19, 2.38) **	0.44 (0.20, 0.97) *
	Small ruminants	0. 50 (0.41, 0.60) ***	-
	Log2 Number in batch	1.28 (1.18, 1.38) ***	1.24 (1.12, 1.38) ***
	Vehicle Transport	-	0.39 (0.26, 0.59) ***
	Journey close contacts with local livestock	1.31 (1.05, 1.64) *	-
	Business sale	1.65 (1.28, 2.13) ***	-
	Reason for purchase Good Price	-	1.60 (1.13, 2.26) *
	Transported onwards with other livestock	-	2.75 (1.87, 4.03) ***
Markets (N = 22) random effects variance		0.03	0.28
Marginal R^2		0.35	0.32
Conditional R^2		0.37	0.44
Note:			$^{*}p<0.05^{**}p<0.01^{***}p<0.001$ Market fitted as random effect

4.3.4.3 Journey distance

Livestock journey distance was calculated as the great circle distance between the origin (to) and destination (from) coordinates and the market central point coordinates. Summary data of market distances travelled to and from markets are given in Appendix Table S3.1. The longest recorded distance of a single market movement was for cattle that were moved 511 km to a secondary market. The longest recorded onward journey distance was 297 km for small ruminants moved from a primary market.

The effect of *market type* on journey distance differed between the species groups in both *journey distance to* and *from* models (Likelihood ratio test P <0.001 and 0.05 respectively) so an interaction was fitted between species group and market type and the model results are presented separately for each group (cattle primary, cattle secondary, small ruminant primary and small ruminant secondary)(Table 4-5). Cattle travelling to secondary markets travelled twice as far as those travelling to primary markets and cattle travelling onwards from secondary markets had about a third of a journey distance ahead of them compared to cattle purchased in primary markets. Small ruminants travelled shorter distances to and from primary and secondary markets compared to cattle and their journey distances were not affected by market type.

Estimated distance travelled to market for the baseline group of one cattle, travelling on foot to primary markets was 22.2km (95% CI 16.1-31.6km). A twofold increase in number of livestock being moved was associated with a 20% increase in journey distance (95% CI 14-27%) and use of a vehicle to travel to market was associated with a twofold increase in distance travelled (2.15 (95% CI 1.79-2.58 P <0.001)). Livestock that had close contacts with local livestock on their journey had also travelled an estimated 21% further than those who made no contacts with local livestock at grazing and watering points on their journey (95% CI 4-41% P = 0.008).

Estimated distance travelled from markets for one cow (or bull) moved onwards on foot from primary markets was 36.2km (95% CI 25.0-52.4km). A twofold increase in number of livestock purchased was associated with a 19% increase in onward journey distance (95% CI 11-28%, P <0.001) and livestock that were transported onwards by vehicle were estimated to be moving over twice as far compared to those moved onwards on foot (2.22, 95% CI 1.69-2.92 <0.001). Livestock that were purchased primarily for meat production had shorter onward journeys compared to livestock that weren't purchased primarily for meat (0.62, 95% CI 0.52, 0.90 P = 0.007).

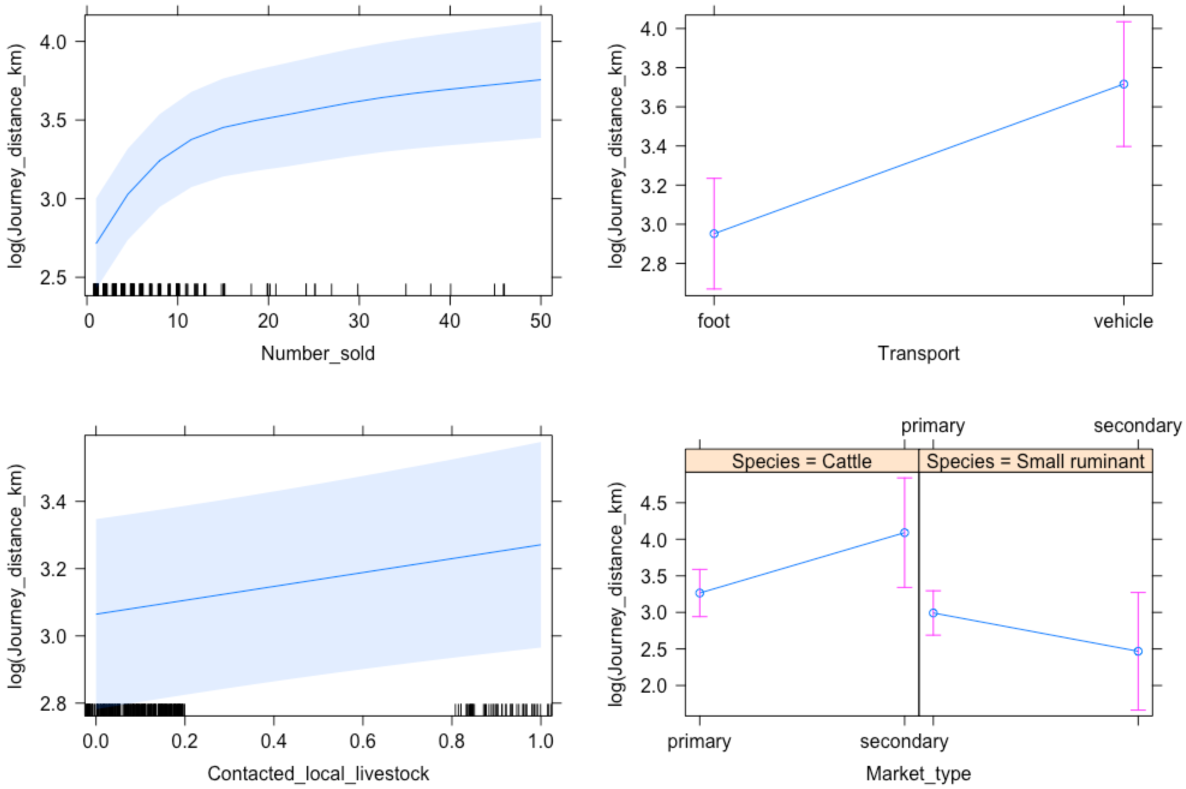


Figure 4-12: Effect of GLMM covariates on distance livestock travelled to market

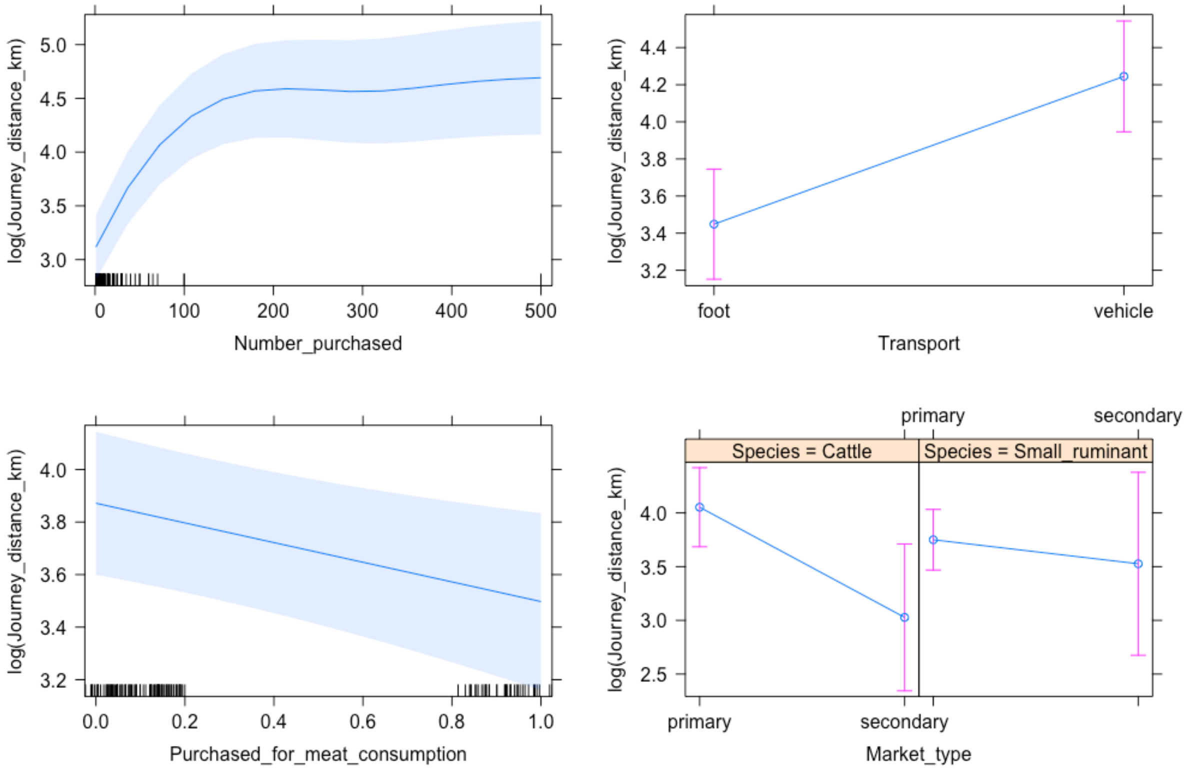


Figure 4-13: Effect of GLMM covariates on distance livestock travel onwards from markets

Table 4-6: Results of generalised linear mixed models with the log distance livestock travel to and from markets as the response variable. In the top row of the table the results show the estimated journey distance in kilometres for the baseline group (1 cattle, going to/from primary market on foot having no contacts with local livestock en route), adjusted for Jensen's inequality. In the rows below the coefficients give the estimated multiplicative effect of each variable on journey distance to and from markets with the 95% confidence interval and the. An interaction between species group and market type is fitted in each model and the effects of each group are presented separately in the table (likelihood ratio test P-value <0.001 in journey to model and 0.049 in journey from model)

		Regression coefficient estimate (95% CI)	
		Journey from market (Nobs = 499)	Journey to market (Nobs = 205)
Predicted distance in km (# 1 Cattle, Primary market, On foot, No contact with local livestock on journey)		***22.22 (16.13, 30.61)***	***36.21 (25.00, 52.44)***
Multiplicative effect estimate:	Cattle Secondary market	2.28 (1.01, 5.17) *	0.36 (0.17, 0.76) *
	Small Ruminant Secondary market	0.59 (0.25, 1.39)	0.80 (0.33, 1.93)
	Small Ruminant Primary market	0.76 (0.65, 0.89) ***	0.74 (0.54, 1.00) *
	Log2 Number in batch	1.20 (1.14, 1.27) ***	1.19 (1.11, 1.28) ***
	Transported by Vehicle	2.15 (1.79, 2.58) ***	2.22 (1.69, 2.92) ***
	Contact with local livestock	1.21 (1.04, 1.41) **	-
	Purchased for meat	-	0.69 (0.52, 0.90) **
Markets (N = 22) random effects variance		0.41	0.05
Marginal R squared		0.36	0.29
Conditional R squared		0.67	0.51
Note:			*p<0.05**p<0.01***p<0.001 Market fitted as random effect

4.3.5 Network analysis

4.3.5.1 The observed network

Market movements recorded in this data connect multiple locations outside the study area, including across the international border with Kenya (Figure 4-14).

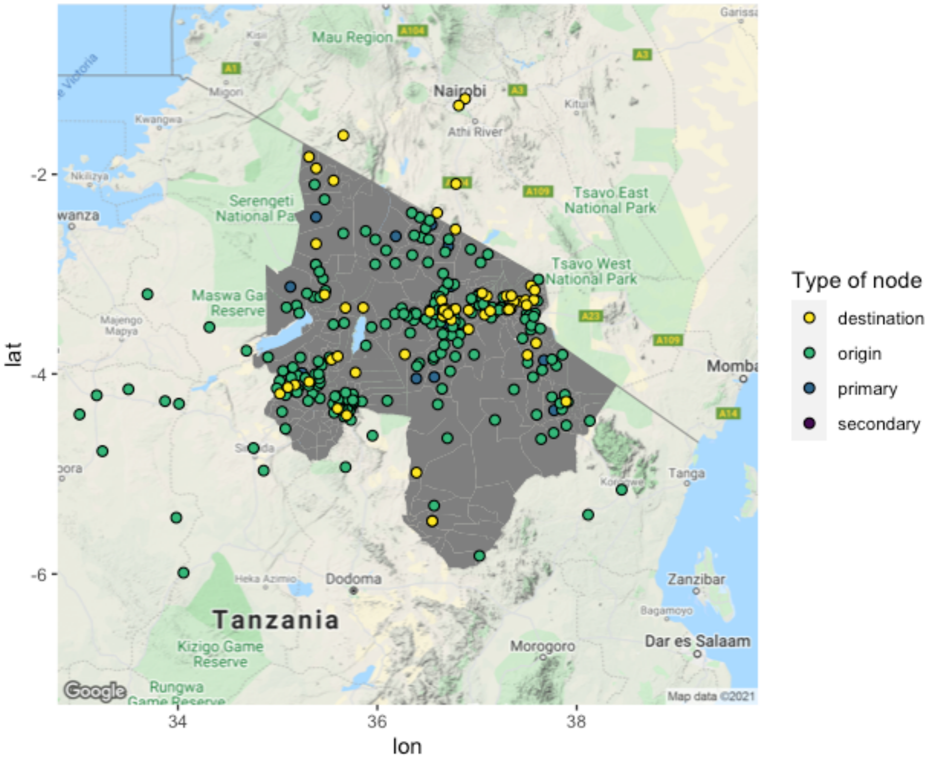
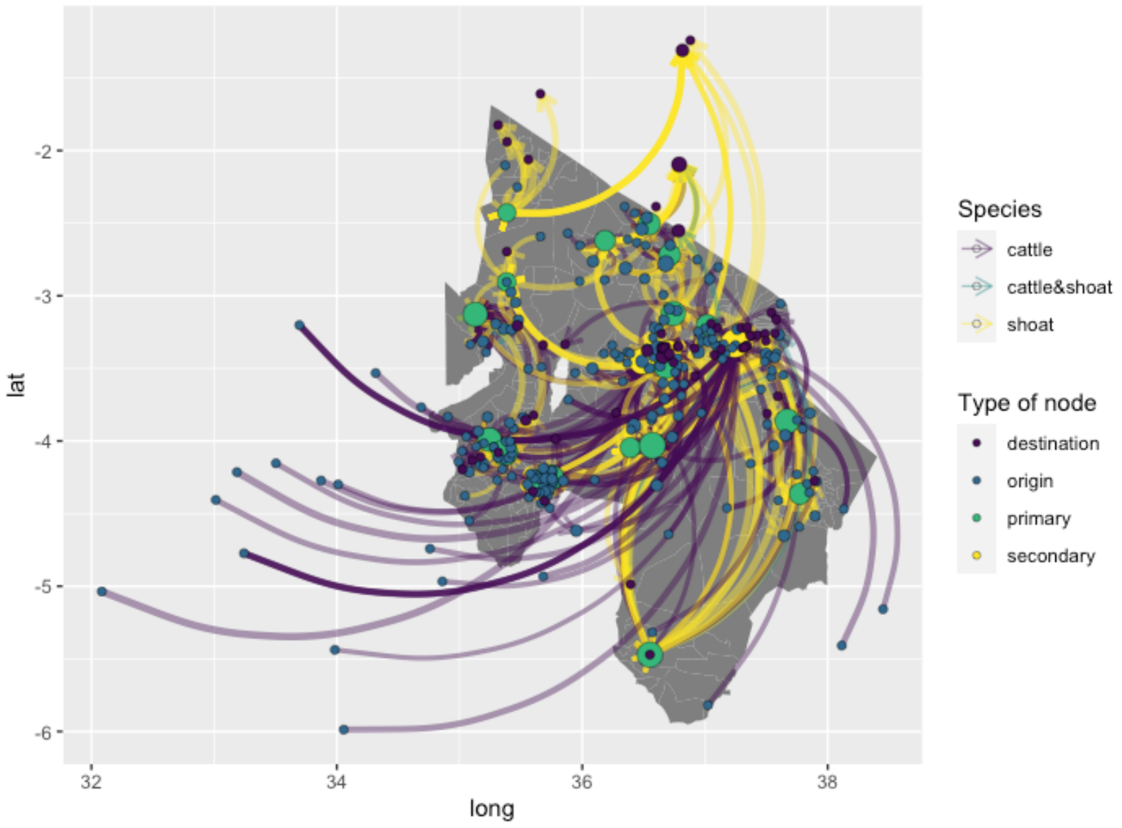


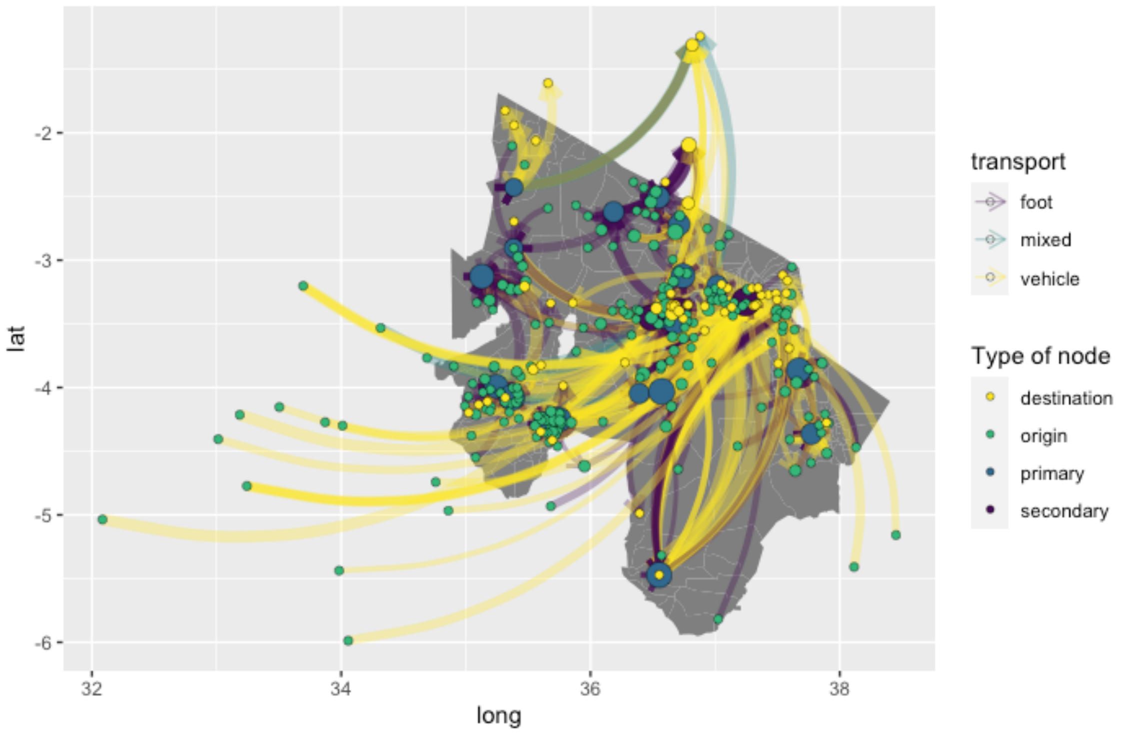
Figure 4-14: Map of the study area highlighted in grey and all identified origins, destinations and markets

Movements into the study area from outside are due to cattle travelling north into the market system and these are all moved by vehicle. Movements out of the study area are all small ruminants moved north by vehicle into Kenya (Figure 4-15 (a,b,c)). Movement plots have links represented by arrows that are coloured by movement type and show the direction of movement. The market movement network constructed using this market survey data is fully connected (assuming links are bidirectional) with 22 market nodes and 280 unique non-sampled origin and destination nodes (61 unique destinations, 184 unique origins and 35 locations listed as both an origin and destination). Of the markets sampled 9 were listed as both origin and destinations, 3 as an origin and 5 as a destination. The nodes were connected by 726 links and 64% (N = 468) of links are unique (only mentioned once as a route livestock travelled in the survey data). 23 origins and 12 destinations could not be matched to known locations thus these 35 locations and the corresponding movements generated from and to them are not plotted in the figures.

a.


-2
-3
-4
-5
-6
lat
32
34
36
38
long
Species
cattle
cattle&shoat
shoat
Type of node
destination
origin
primary
secondary

b.


-2
-3
-4
-5
-6
lat
32
34
36
38
long
transport
foot
mixed
vehicle
Type of node
destination
origin
primary
secondary

c.

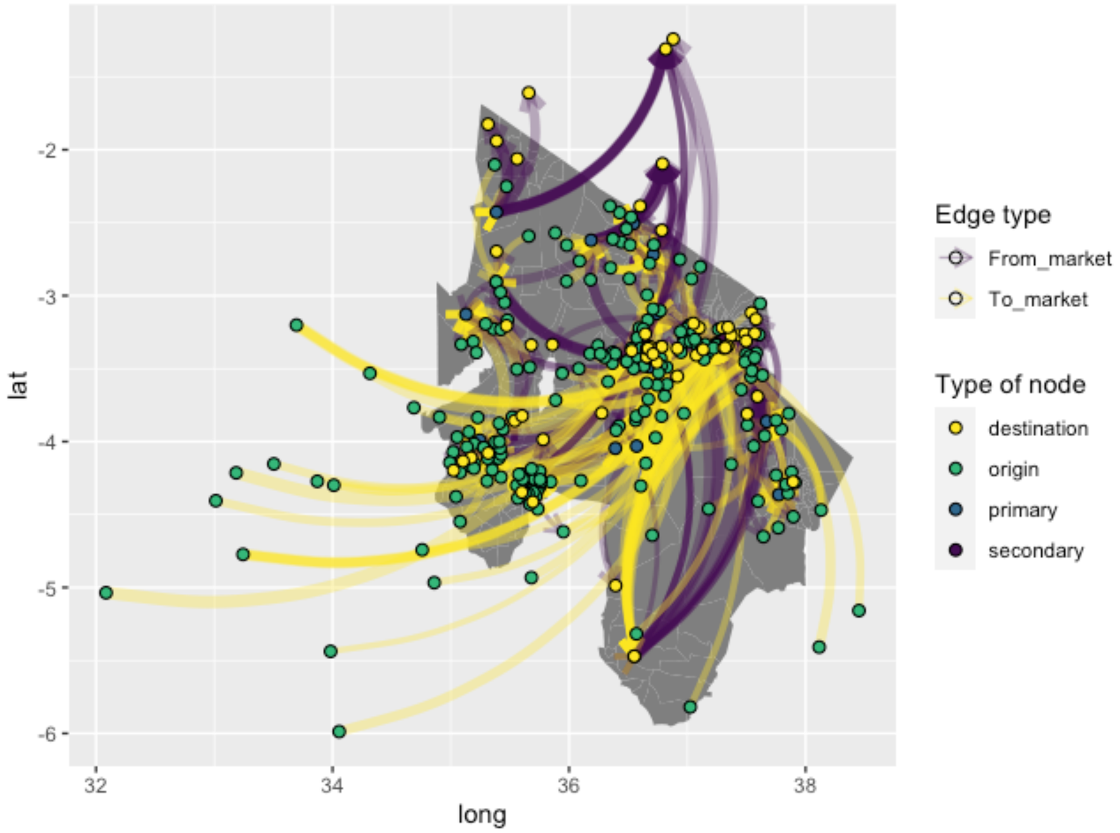


Figure 4-15: Static network of livestock market movements in Arusha, Manyara and Kilimanjaro regions of northern Tanzania. Node size is plotted proportional to its degree centrality (combined in and out degree) and link thickness is proportional to the number of livestock reportedly moved along it on one market day. Links are coloured differently in the three plots by (a) species, (b) transport type, (c) to or from market movements. All market nodes are classed as markets (N=22) regardless of being listed as an origin or destination and locations listed as both origins and destinations (N=35) are classed as origins in the figures.

4.3.5.2 Network properties

The giant weakly connected component (GWCC), calculated by assuming all links are bidirectional, includes all 302 network nodes. This can be used to estimate the maximum size of an epidemic if a pathogen is introduced into the network. Assuming reported links are unidirectional, there were two giant strongly connected components in the network, each made up of 6 nodes. The size of the GSCC can be used to estimate the lower bounds of an epidemic size on the network if a pathogen is introduced. The size of the GSCC should be interpreted with some caution here as we only have a sub-sample of the underlying network. Additionally all nodes involved in the network are not equivalent (some markets, some origin/destination locations) so the pathogen transmission within nodes may vary. The diameter of the network is 10 which shows that a minimum of 10 consecutive livestock movements are needed to move between the two most distant nodes on the network. The average path length on the network is 4 which shows that on average 4 sequential movements of infectious livestock are needed to get between pairs of connected nodes. The reciprocity of the network is 0.11 which shows that 11% of livestock movements captured in this data are bi-directional. Again these results should be interpreted with caution as the density of the underlying network might be higher if more nodes were sampled from, resulting in more links and thus shorter path lengths across the network.

4.3.5.3 Centrality metrics

Market in-degree and geometric mean degree have normal distributions and out-degree and betweeness centrality metrics are skewed to the right Figure 4-16. This shows some markets were connected to many more locations via outgoing movements than others and a small proportion of markets with high betweenness act as bridges in the network connecting many other pairs of nodes.

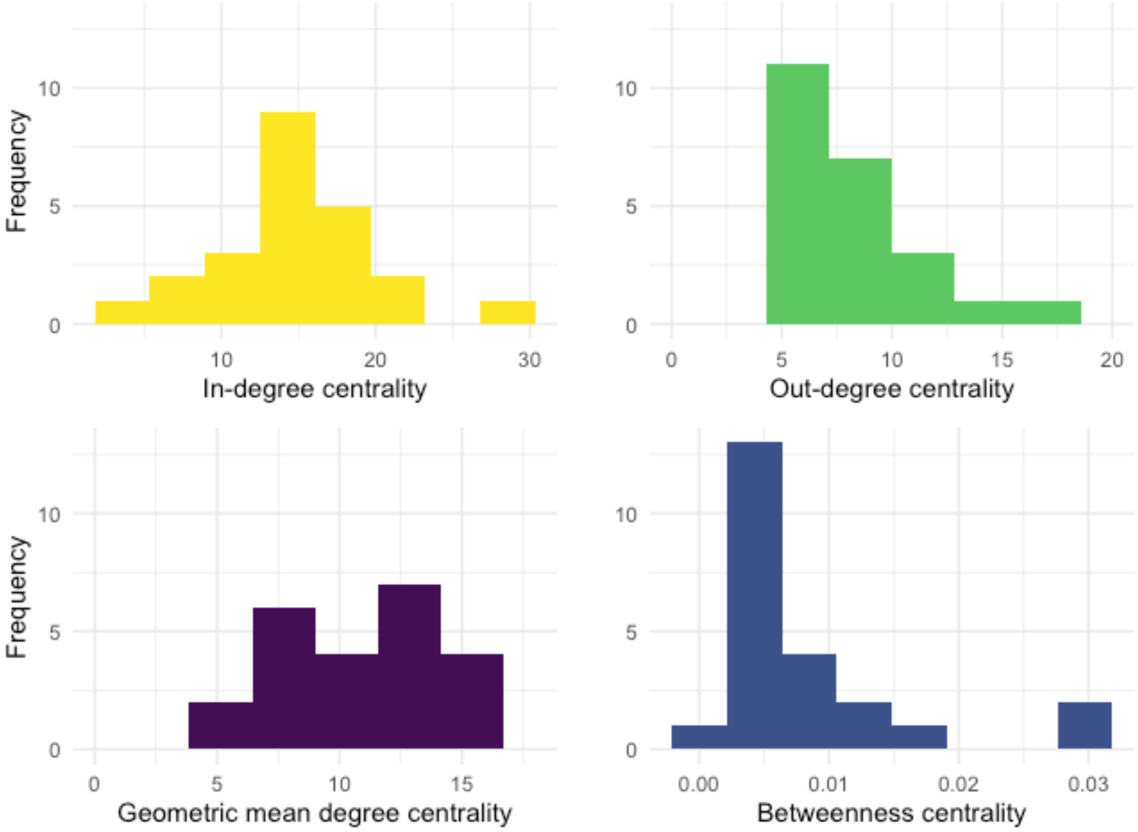


Figure 4-16: Distribution of market node centrality metrics (in-degree, out-degree, geometric mean degree and betweenness centrality) calculated on market movement network

For the non-market nodes the centrality measures were calculated using all links reported in the network. Centrality metric distributions for all non-market nodes identified show strong right skew (Figure 4-17) indicating that through this data we have identified some highly connected non-sampled locations.

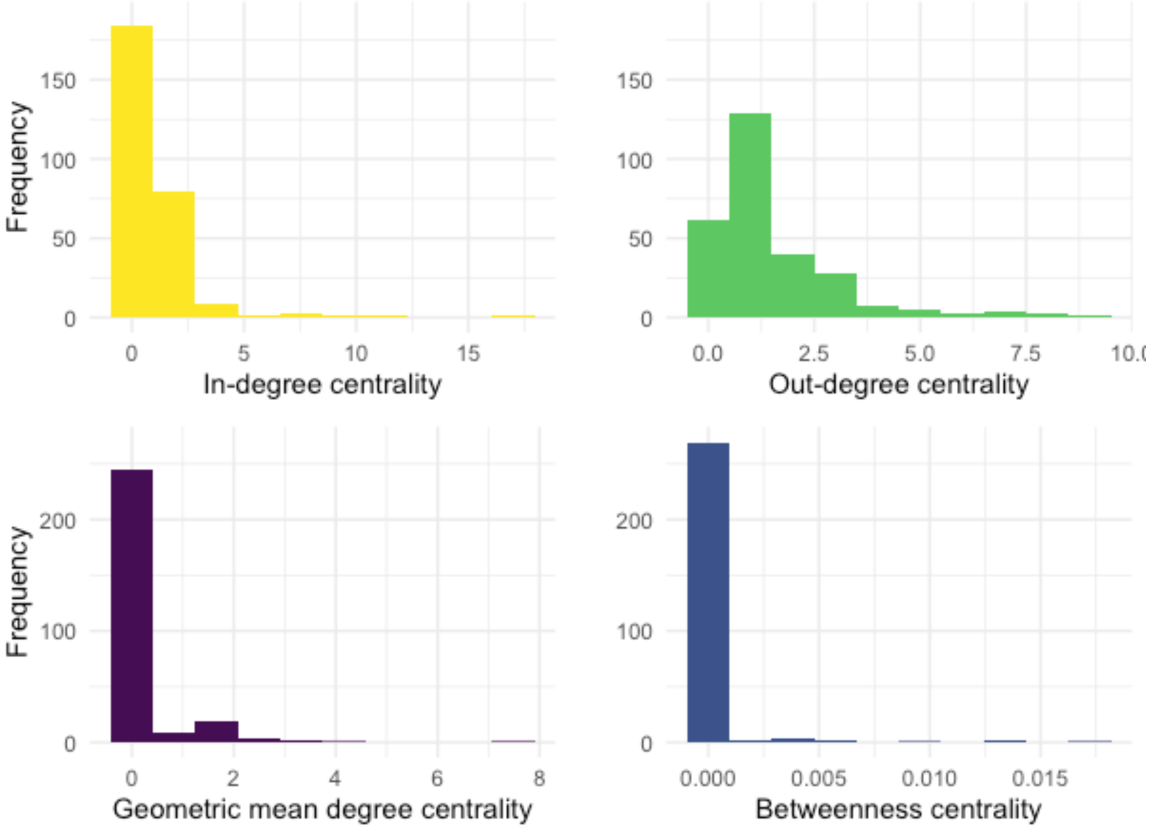


Figure 4-17: Distribution of non-market nodes centrality metrics (in-degree, out-degree, geometric mean degree and betweenness centrality) calculated on market movement network

There was strong positive correlation between market nodes in-degree and both betweenness centrality and geometric mean degree centrality $P < 0.001$ (Figure 4-18). In the non-market nodes there were positive correlations between betweenness centrality and both in-degree ($\rho = 0.29$) and geometric mean degree centrality ($\rho = 0.61$) (Figure 4-19). The positive correlations suggest that a small number of highly connected locations exist, outside the sampled markets, which could be important for pathogen transmission via livestock movements within and beyond northern Tanzania, and these could be identified using this market survey data.

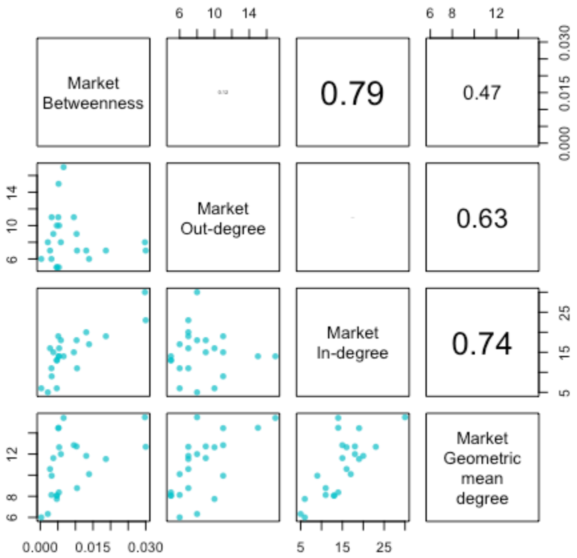


Figure 4-18: Market node centrality metrics correlation plots (bottom left) and corresponding Spearman's rho correlation coefficients (top right) with only significant (AS 86 P-value <0.05) rho values displayed

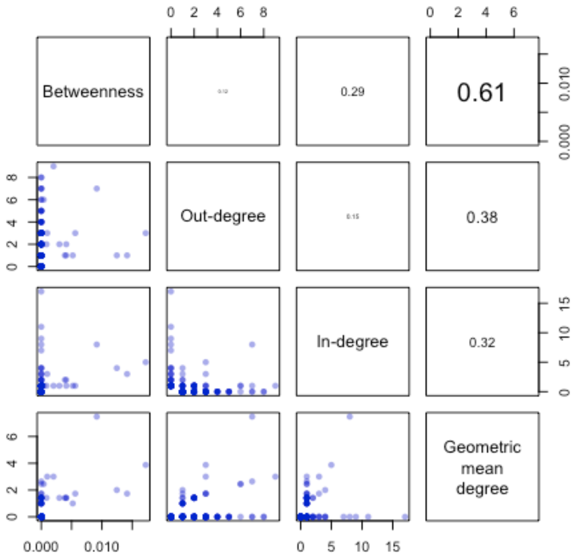


Figure 4-19: Non-market node centrality metrics correlation plots (bottom left) and corresponding Spearman's rho correlation coefficients (top right) with only significant (AS 86 P-value <0.05) rho values displayed

The sample size for comparison between centrality metrics and other market characteristics was small (3 secondary markets and 19 primary markets). Nevertheless there was strong evidence to suggest that secondary markets have higher in-degree (sampling livestock from more unique location) and betweenness centrality (lying on the shortest path between most other pairs of nodes in the network) than the primary markets sampled (P = 0.003 and < 0.001 respectively). Thus, secondary markets sample livestock from a wider range of locations than primary markets and also act as bridges connecting many other locations in the network. There was no evidence to suggest either market centrality metric was associated with the A-E class of the market area (P=value >0.05). Model estimates for in-degree and betweenness centrality metrics for primary and secondary market nodes are given in Table 4-7.

Table 4-7: GLMM Estimated market node's in-degree and betweenness centrality measures

	In-degree	**Betweenness**
Primary markets	13.6 (11.3, 15.8)	0.007 (0.004, 0.009)
Secondary markets	23.7 (17.6, 29.8)***	0.024 (0.017, 0.031)**

4.4 Discussion

4.4.1 Summary

This market survey data shows that the study area in northern Tanzania and beyond, across the international border into Kenya, is epidemiologically connected through livestock market movements. Livestock movements to and from markets covered large distances (over 500km and almost 300km respectively) showing that pathogens can also be transmitted over these distances via a single market movement, if infectious animals are moved (Dean *et al.*, 2013). The network constructed from market movements can be used to identify highly central market and non-market locations that are at high risk of both pathogen introduction and onward transmission. These locations can be made targets for efficient multi-purpose disease control programmes that target multiple pathogens and aim to prevent pathogen introduction and or onward transmission, depending on government priorities (Perry and Grace, 2009; Rich and Perry, 2011; Forman *et al.*, 2012; Sindato, Karimuribo and Mboera, 2012). These results also show key differences between primary and secondary markets and differences in journey characteristics, made by cattle and small ruminants to and from markets that will influence pathogen transmission. The results also provide new information on some of the external pressures that determine livestock flow into and through the market system in northern Tanzania and identify who the key stakeholders are in the market movement network who will be affected by movement restrictions.

4.4.2 Journey contacts and socio-economic factors

Most livestock move to and from markets on foot and livestock moving on foot are more likely to contact local livestock along their journey which provides opportunities for pathogen transmission events with livestock outside the market system. This is despite legislation made in 2000 stating that all livestock market journeys should be made in vehicles (Pica-Ciamarra *et al.*, 2011). The resultant higher rate of local livestock contacts could be risk factors for the spread of highly transmissible pathogens like foot and mouth disease virus (FMDV) and Peste des petits ruminants (PPRV), as well as those transmitted by environmental contamination such as *Leptospira* spp. and *Coxiella burnetii* (Sherman, 2011; Van den Brom *et al.*, 2015; Allan *et al.*, 2018; Herzog *et al.*, 2020). This is an important finding to document as pathogen surveillance programmes targeted solely at markets will miss infections of local livestock that happen along market routes. As has been shown in studies from other countries, both formal and informal trade routes should be considered as additional locations for implementing pathogen surveillance in this setting as well as in markets (Spiegel and Havas, 2019; Ng'asike, Hagmann and Wasonga, 2021). Both local livestock keepers and those moving livestock to markets should actively minimise contact between moving and local livestock to prevent pathogen transmission.

Cattle are more likely to be transported by vehicle, especially for journeys from primary to secondary markets where they achieve a higher price. Livestock transported by vehicle are also more likely to have shorter journey times overall but be mixed together with livestock from other households or purchased batches for the duration of the journey. Journey sharing is a risk factor for pathogen propagation as the chance of at least one animal being infectious increases as the batch size increases and when multiple batches are combined. Vehicle use also provides potential opportunity for fomite pathogen transmission if the vehicle used is contaminated. Vehicle transport might reduce livestock's time in transit, energy required for the journey, stress, and the risk of contact with local livestock but the risk of transmission of infectious pathogens between livestock from different households within the vehicle will be high (Bernini, Bolzoni and Casagrandi, 2019). Longer market journeys of large livestock batches, identified here as driven largely by market economies of scale, will have a greater chance of disseminating pathogens throughout the moving batch and to other livestock along the route if any moving livestock are infectious. The higher rate of journey sharing by livestock transported by vehicle and the increased use of vehicles for cattle journeys and journeys to secondary markets where higher prices are achieved by cattle implies that the cost of vehicle use needs to be offset. This result highlights the restrictive economic component of market movement decisions and is important to document as it shows the cost of vehicle use may be too high for sellers attending primary markets and or poorer sellers are unable to afford vehicle transport. This means that despite multiple initiatives (e.g. the Tanzanian governments Agricultural Sector Development Programme (ASDP) (Michael *et al.*, 2018) and Market infrastructure, value addition and rural finance (MIVARF) funded by the African development bank) to improve livestock keeper's access to markets in Tanzania, access inequality still exists and rural livestock keepers continue to miss out on higher cattle price opportunities available at more distant secondary markets (Pica-Ciamarra *et al.*, 2011).

The primary driver for most livestock sales into the market system is *household need* as the majority of sellers in primary markets were selling their own animals brought from home, primarily to generate money for household purchases. This is an important finding, supported by other studies, and should be used to ensure that the needs of households who rely on livestock sales into the market system are carefully considered in the design and implementation of disease control programmes (Pica-Ciamarra *et al.*, 2011). When people are already living in marginal circumstances it could be detrimental to overall health and wellbeing to remove the primary source of income generation used to meet family needs for a large proportion of the population (Komba *et al.*, 2012).

In secondary markets, located closer to major towns, cattle generate a higher price and are most likely to be sold by traders, as part of commercial business enterprises, who have also travelled further to access the market compared to cattle sold at primary markets. Cattle sold at secondary markets then have relatively short (distance and duration) onward journeys to predominantly final end point slaughter destinations compared to cattle sold at primary markets and small ruminants sold at secondary markets. Similar to findings from other countries this is likely the main reason for secondary markets being located close to major towns: buyers that supply the local meat production units that supply urban areas do not have to travel far to purchase the livestock they need (Martin *et al.*, 2011; Mutua *et* *al.*, 2018; Motta *et al.*, 2019). Urban per-capita consumption of livestock products in Tanzania is twice that of rural households and is largely made up of purchased products, indicating this is what drives the flow of livestock from rural (primary markets) into urban areas (where secondary markets are located) (Covarrubias *et al.*, 2012).

The distinction between market types and the flow of livestock into and out of them is not reflected in the small ruminant data. This could be because small ruminants purchased for meat can be bought and slaughtered locally without entering the market chain because slaughter of smaller animals requires less infrastructure. Alternatively, there may be less demand for small ruminant meat, or it is possible local small ruminant consumers are being effectively priced out of the market by external buyers. Small ruminants in the market system described by this data are commonly being moved on to other markets (48% buyers at primary and 55% at secondary) however these markets are not captured in this data with only a small proportion of surveyed small ruminant sellers selling animals brought from other markets. The mapped survey data helps to explain this misalignment as there were a high number of small ruminant movements out of surveyed markets going north across the international border with Kenya. This finding is supported by other studies that show the international border here does not provide a barrier to livestock movements and thus to pathogens (Mutua *et al.*, 2018; Omondi *et al.*, 2019; Spiegel and Havas, 2019). The long-distance international movements of many small ruminants mixed together in large batches from multiple origins in vehicles, provide potential important routes for rapid regional pathogen dissemination. This is an important finding for policy makers who are currently designing disease surveillance and control interventions, especially for highly transmissible pathogens such as FMDV and PPRV, as it confirms a harmonised regional approach is required (Kivaria, 2003; Motta *et al.*, 2017; Herzog *et al.*, 2019).

4.4.3 Connectivity

If the network links were considered bi-directional all locations reported in this study could reach one-another. The network was built using data on movements reported in one direction. If pathogens can be transferred by fomites, such as the people or vehicles that move the livestock, or by non-sold livestock returning to their origin, it is plausible that for some pathogens the observed links are partially bidirectional, albeit with the two directions weighted differently (Porphyre *et al.*, 2020). As an example the majority of sellers (78%) reported that they would take non-sold animals home. The majority of all sellers surveyed were also selling livestock from home so regardless of sale success the seller is likely to return to their origin and create a bidirectional link.

Secondary markets have higher betweenness and in-degree centrality compared to primary markets which shows they sample livestock from a wider range of locations than primary markets and also act as bridges by connecting many other locations in the network. This makes them good locations to begin targeted pathogen surveillance programmes or in the case of emergency responses to disease outbreaks, where market closures could be targeted to rapidly fragment the network and remove important transmission routes. As we did not have a full sample of all markets in the study area the distinction between market types is an important finding as these results could be generalised from so that all secondary markets in Tanzania become early warning disease surveillance sites.

There was significant right skew of, and positive correlation between, all centrality metrics, in-degree, out-degree, betweenness and geometric mean degree, of the non-market nodes identified by this survey data. This shows that some non-market locations identified in the data could also be targets for efficient disease control interventions such as vaccination or surveillance programmes, depending on the pathogens of primary concern and the type of locations identified. For example, villages with high in-degree and betweenness could be targets for vaccination campaigns, or if locations with high in-degree are slaughter points these could be targets for implementing early warning active surveillance programmes. Other studies have shown efficient network fragmentation can be achieved by removal of the highly connected non-sampled nodes, because these are likely to be some of the most central nodes in the underlying network (Büttner *et al.*, 2013a; Herrera *et al.*, 2016; Colman *et al.*, 2019). If the most central non-sampled locations are other markets, these can become surveillance sites for existing and emerging pathogens. If highly connected non-sampled locations are villages then these might be targets for vaccination campaign starting points so that transmission chains are broken, especially in the face of disease outbreaks such as FMD, RVF and PPR (Bett *et al.*, 2010; Brahmbhatt *et al.*, 2012; Fournié, Waret-Szkuta, Camacho, Laike M Yigezu, *et al.*, 2018)

4.5 Limitations

This data is a small static subsample of the underlying network, we do not control for any adjustments in time or season and we have only sampled from a proportion of markets in the study area. The volume of livestock sales from different areas might change with the wet and dry seasons as household income needs vary through the year (Pica-Ciamarra *et al.*, 2011; Haseeb *et al.*, 2019). The demand for livestock produce might also vary in different areas throughout the year in line with festival seasons, as is seen in other countries (Aklilu *et al.*, 2007; Dean *et al.*, 2013; Nandonde, Gebru and Stapleton, 2017; Mtimet *et al.*, 2021).

Sampling effort was the same in all markets and although we recorded an estimated number of livestock being sold on each market day, with some markets much busier than others, we do not know what proportion of buyers and sellers surveyed at each market. We were unable to measure more specifically how exhaustive or saturated our market sample was on each day as markets are busy with many people and livestock coming and going during the market day and no official recording required. The network, and market node metrics calculated using this data will be skewed left (less heterogeneity identified even if it exists) which we tried to offset by analysing only unique links to calculate market node metrics. The random sampling strategy was followed closely by all interviewers at each market with a total of four people approached unwilling to participate in each case citing shortage of time. This makes the study participation rate look high however market attendees who did not want to participate could have simply avoided the interviewers by moving away if they saw an interviewer close by. There may therefore have been some unknown sampling bias with interviewee participation.

4.6 Conclusions

Northern Tanzania is highly connected via livestock movements to and from markets that provide epidemiological connections and routes for pathogen transmission over long distances between geographically distinct populations. This includes cattle coming into northern Tanzania from the south west of the country and small ruminants moving out of the area northwards across the international border into Kenya. Market journeys predominantly happen on foot with moving livestock at high risk of contacting local livestock and providing opportunities for pathogen dissemination along market routes which should be considered in the design of surveillance programmes (Spiegel and Havas, 2019; Ng'asike, Hagmann and Wasonga, 2021).

There is high price disparity between cattle sold in primary and secondary markets. Cattle sellers in secondary markets are generally selling as part of a livestock trading business and achieve higher prices compared to those in primary markets who are generally selling to generate some income to meet household needs. Secondary markets sample from a higher number of locations and bridge many locations in the observed network which makes them principal targets for implementing efficient pathogen surveillance in northern Tanzania. Efficient surveillance is needed in countries with limited resources and high disease burden with risk of emerging pathogens (Martin *et al.*, 2011; Halliday *et al.*, 2012; Motta *et al.*, 2017).

The findings of this study highlight a disparity in market access and show that market journeys might be responsible for widespread pathogen dissemination across and beyond the study area. Although pathogen control programmes are needed to reduce overall burden in northern Tanzania, the structure of the livestock market system and the key stakeholders in market movements need to be carefully considered in the development of these. Market access needs to be improved so rural livestock keepers can achieve greater parity in prices for their livestock at the same time as reducing the risk of widespread pathogen transmission through market movements.

5 Analysing livestock network data for infectious diseases control: an argument for routine data collection in emerging economies

This chapter is published in Philosophical Transactions B DOI: 10.1098/rstb.2018.0264

Contribution

GC gained access to movement permit data, designed framework for data collection and data processing and carried out data collection and transcription with the help of field staff. GC also contributed to data processing, building of generalised linear mixed model, model fitting and use of model to generate predicted movement data. GC constructed networks from simulated data, carried out network analysis, chose interventions to test in analysis and wrote manuscript.

PCJD contributed to data cleaning and processing and developing GLMM model to predict movements and developed stochastic disease transmission model on network to trial effectiveness of interventions. JC, TD, SM, GR, LCMS and RRK all contributed to review sections 5.3.1, 5.3.2, 5.3.3, 5.4.1, 5.4.2 and 5.4.3 and reviewed manuscript along with PCJD, SC, ES and OMN.

G.L. Chaters[1], P.C.D. Johnson[1], S. Cleaveland[1], J. Crispell[2], W.A. de Glanville[1], T. Doherty[3], L. Matthews[1], S. Mohr[1], O.M. Nyasebwa[6], G. Rossi[3], L.C.M. Salvador[3, 4, 5], E. Swai[6], R.R.Kao[3,2]

[1] Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, UK

[2] School of Veterinary Medicine, University College Dublin, Ireland

[3] Royal (Dick) School of Veterinary Studies and Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian, Scotland, UK

[4] Department of Infectious Diseases, University of Georgia, Athens, Georgia, USA

[5] Institute of Bioinformatics, University of Georgia, Athens, USA

[6] Department of Veterinary Services, Ministry of Livestock and Fisheries, Tanzania

---

5.1 Abstract

Livestock movements are an important mechanism of infectious disease transmission. Where these are well recorded, network analysis tools have been used to successfully identify system properties, highlight vulnerabilities to transmission and inform targeted surveillance and control. Here we highlight the main uses of network properties in understanding livestock disease epidemiology, discuss statistical approaches to infer network characteristics from biased or fragmented datasets. We use a “hurdle model” approach that predicts (i) the probability of movement and (ii) the number of livestock moved to generate synthetic ‘complete’ networks of movements between administrative wards, exploiting routinely collected government movement permit data from northern Tanzania. We demonstrate that this model captures a significant amount of the observed variation. Combining the cattle movement network with a spatial between-ward contact layer we create a multiplex, over which we simulated the spread of ‘fast’ (R_0=3) and ‘slow’ (R_0=1.5) pathogens, and assess the effects of random versus targeted disease control interventions (vaccination and movement ban). The targeted interventions substantially outperform those randomly implemented for both fast and slow pathogens. Our findings provide motivation to encourage routine collection and centralisation of movement data to construct representative networks.

5.2 Introduction

The “static network” concept of a population as a set of “individuals” (nodes) with immutable contacts (links) between them is now well-established in infectious disease modelling. The network representation occurs naturally because the “individual” is typically well-defined (e.g. a person, animal, city, herd, or farm) and the number of potentially infectious contacts per individual is usually few (Watts and Strogatz, 1998; Keeling, 1999; Liljeros *et al.*, 2001; Kao *et al.*, 2006; Robinson, Everett and Christley, 2007). While there are a few studies for human diseases that include comprehensive, explicit network data (Hufnagel, Brockmann and Geisel, 2004), more frequently these are either generated indirectly (for example, using mobile phone data or gravity models to predict commuter flow (Brockmann, Hufnagel and Geisel, 2006; Viboud *et al.*, 2006; Balcan *et al.*, 2009; Wesolowski *et al.*, 2016), or are explicit but at small geographical scales (Meyers, Newman and Pourbohloul, 2006; Gardy *et al.*, 2011). In contrast, in Great Britain (GB) cattle movement data have been recorded for individuals on a daily basis for almost two decades (Green and Kao, 2007). This data richness has presented both challenges and opportunities for the application of network analyses in infectious disease epidemiology (Kao *et al.*, 2006; Robinson, Everett and Christley, 2007). Similar livestock data now exist in many other countries (León *et al.*, 2006; Natale *et al.*, 2009; Bajardi *et al.*, 2011; Nöremark *et al.*, 2011; Konschake *et al.*, 2013; Dutta, Ezanno and Vergu, 2014; VanderWaal *et al.*, 2016). However, they remain rare in emergent economies where disease burden is often high and zoonotic risk is more pronounced due to the high proportion of people who live and work in close contact with livestock (Klous *et al.*, 2016). About one billion of the world’s poorest people (earning < US$2 per day) depend at least partially on livestock for their livelihoods (FAO, 2009), making the trade of livestock and the freedom to move livestock to access natural resources vital in many impoverished communities (Perry and Grace, 2009; Grace *et al.*, 2012; ILRI, 2018). In many regions, such as Sub-Saharan Africa, there are frequent but poorly recorded cross-border movements (Aklilu, 2008; Musemwa *et al.*, 2012; Apolloni *et al.*, 2018) and, when coupled with poor within-country knowledge of livestock movements, this creates risks for international pathogen transmission.

Though network analyses would be greatly aided by systems for comprehensive routine recording of between-farm and market movement, as occurs in GB and elsewhere, in countries with developing infrastructure collecting these data can be onerous and costly and requires well-evidenced justification. Here, we provide an overview of the role of network analysis in epidemiology, paying particular attention to the challenges of exploiting extensive but fragmented data. These insights are used to analyse livestock movements in northern Tanzania, where there is a high burden of livestock disease including zoonoses (Hummel, 1976; Schoonman and Swai, 2010; Crump *et al.*, 2013; Assenga *et al.*, 2015; Sumaye *et al.*, 2015; Wensman *et al.*, 2015; Cash-Goldwasser *et al.*, 2018), no formal livestock traceability system implemented at a national level, and limited resources for disease control. We demonstrate the utility of our network by identifying nodes to target disease control and surveillance interventions, considering both fast and slowly transmitting pathogens, and interrogate their efficacy through simulation, demonstrating substantial potential benefits in reducing disease spread.

5.3 Fundamental network concepts applied to livestock diseases

5.3.1 Centrality measures and transmission patterns

Network centrality measures originated in social science (Wasserman and Faust, 1994), and are used to quantify the importance of nodes and links in a network, with obvious applications to identifying disease risks (Bell, Atkinson and Carlson, 1999; Christley *et al.*, 2005; Natale *et al.*, 2011; VanderWaal *et al.*, 2016; Palisson *et al.*, 2016; Sintayehu *et al.*, 2017). Common measures include degree centrality (the number of links associated with a node[3]), betweenness centrality (the number of times a node or link is traversed by the shortest paths between all other node pairs), and eigenvector centrality (loosely, a measure of how connected a node is to well-connected neighbours)[4]. Network centrality measures have been used to analyse livestock movement data from many countries, with each using different types of data source (Kao *et al.*, 2006; Natale *et al.*, 2009; Büttner *et al.*, 2013a; Palisson *et al.*, 2016; Sintayehu *et al.*, 2017; VanderWaal, Enns, *et al.*, 2017). One example showing the relevance of all three of these centrality measures comes from the analysis of the costly (Haydon, Kao and Kitching, 2004) 2001 foot-and-mouth disease (FMD) epidemic in GB. First, a small number of “cull ewes” were sold and transported

---

[3] For directed networks like livestock movements, where transmission is overwhelmingly in the direction of the movement only, the geometric mean of in- and out-degree can be used.

[4] See the Appendix information for a disease-relevant interpretation.

long distance across GB; these were responsible for seeding virus into many otherwise low risk areas (Gibbens *et al.*, 2001). These seeding movements are a characteristic of “small world” network behaviour (Watts and Strogatz, 1998) with the long-range movements acting as links with high betweenness centrality (Gibbens *et al.*, 2001; M D Shirley and Rushton, 2005). Second, Longtown auction market (the largest in GB) played a dominant role in spreading disease (Kao, 2002), demonstrating the importance of high degree centrality. Third, since the epidemic, prohibition of direct market-to-market livestock movements means that some farms now act as “middlemen” between markets, representing a risk that could be effectively targeted to restrict disease spread (Kao *et al.*, 2006, 2007). This role, linking highly connected nodes, is a well-recognised feature of high eigenvector centrality.

5.3.2 Network Dynamics

In a static network, the infection pressure from a single individual is reduced over time as each daughter infection ‘uses up’ the link it was infected over (Keeling and Grenfell, 2000; Green, Kiss and Kao, 2006b). Further, the components of the network (groups of nodes which can reach each other) are well defined. In dynamic networks, links can shift between individuals over time (rewiring), nodes can appear or disappear and the components of the network can change in size and composition. Rewiring a link away from an infected individual has the potential to expose another susceptible individual, thus increasing the probability of disease persistence (Kao, 2010; Enright and Kao, 2018). Link dynamics also greatly complicate measures of network structure. For example, for an *SIS* infection process on a static network, where susceptible individuals (*S*) can become infected (*I*) and eventually recover to susceptible again, the eigenvector centrality scores of the nodes of the network contact matrix represents the expected proportion of time those nodes are infected over the long term[5]. This is the case so long as the probability of recovery before re-infection is high (e.g. if the density of infected nodes is always low, or the recovery time is substantially shorter than the time between infected generations). However, livestock

---

[5] For an irreducible positive definite matrix (e.g. a contact matrix where all nodes belong to a single strong component), the Perron-Frobenius theorem applies and the matrix is guaranteed to have a unique largest eigenvalue (and positive eigenvector). For directed networks, strong connectivity amongst all nodes is required (all nodes can reach each other reciprocally, i.e. are members of the same strong component). Where this is not the case, eigenvector centrality is not well-defined, and other network measures need to be considered (for example by using singular values).

movements vary daily, seasonally and from year-to-year. Contact patterns between farms and therefore eigenvector centrality measures can change dramatically depending on the season and stochastic progression of the epidemic. This influences epidemic spread (Kao *et al.*, 2006; Green and Kao, 2007; Bajardi *et al.*, 2011; Enright and Kao, 2018), an effect also seen in human diseases (Konschake *et al.*, 2013; Takaguchi, Masuda and Holme, 2013). Individual variability in disease progression and severity will also influence disease generation times and therefore what movements are likely to cause infection spread. Thus, predictions of node importance and targeting can depend strongly both on the dynamic properties of the network and the properties of the underlying disease, making the identification of general principles for the targeting of control more challenging (e.g. (Holme and Masuda, 2015); also Appendix 9.4).

Livestock movements are also an example where the actual contact occurs episodically. Episodic behaviour is a subject of considerable study in the network literature, especially where there are patterns of concentrated bursts ("burstiness") separated by long waiting periods (Barabási, 2005; Oliveira and Barabási, 2005; Vázquez *et al.*, 2006). While an infection may itself cause episodic activity, it is most frequently studied as a property of the underlying network. Episodic activity has been shown to slow an epidemic on simulated (Min, Goh and Vazquez, 2011) and real networks (Iribarren and Moro, 2011) but can also increase epidemic speed, for example, due to observed correlations between the topology of the network and the frequency of episodic contacts (Karsai *et al.*, 2011). Epidemic spread also depends on within-node infection dynamics; in a simulated avian influenza outbreak, patterns of recorded vehicle movement between farms could either slow or accelerate pathogen spread, depending on the disease parameters and detection threshold at the farm level (Nickbakhsh *et al.*, 2013).

Infection events themselves can also change the network structure. If the perceived jeopardy is sufficiently high, rumours of pathogen spread may change contact patterns (Epstein *et al.*, 2008; Funk, Salathé and Jansen, 2010). For livestock, farmers may be inclined to sell infected animals due to their condition, or may be restricted from selling animals until the farm is officially declared disease-free (Green *et al.*, 2008). In human disease, modelling analyses that included changes in the contact process over the course of the recent West African Ebola epidemic were used to inform changes in policy (Drake *et al.*, 2015), highlighting the relevance for detailed datasets on contact patterns and their changes over time, both routinely and in response to an outbreak (Chowell and Nishiura, 2015).

5.3.3 The role of pathogen sequence data for relating transmission networks to livestock networks

Although livestock movements tell us about potentially infectious contacts, the relationship between these contacts and the transmission network of actual infectious contacts is only partially understood. Duration of contact, heterogeneity in immune response, and environmental conditions are some of the factors that could affect which livestock movements transmit infection. The growing availability of high coverage pathogen sequence data provides an unprecedented opportunity to quantify this relationship (Cottam *et al.*, 2008; Kao *et al.*, 2014). A number of tools have been developed to estimate transmission from genetic data (Jombart *et al.*, 2009, 2011; Morelli *et al.*, 2012; Ypma *et al.*, 2012; Ypma, van Ballegooijen and Wallinga, 2013; Numminen *et al.*, 2014; Hall, Woolhouse and Rambaut, 2015; Lau *et al.*, 2015; De Maio, Wu and Wilson, 2016; Worby *et al.*, 2016) and new tools continue to be developed (Lau *et al.*, 2015; Pybus, Tatem and Lemey, 2015; De Maio, Wu and Wilson, 2016). However there remain many challenges (Romero-Severson *et al.*, 2014; Worby, Lipsitch and Hanage, 2014; Biek *et al.*, 2015; Frost *et al.*, 2015; Meehan *et al.*, 2018). A key limitation is that pathogen evolution needs to occur on a similar or faster timescale to the disease generation time in order to infer direction of transmission (Biek *et al.*, 2015). Considering larger epidemiological units (e.g. farms rather than animals) can alleviate this problem, since the generation time will be concomitantly longer (Morelli *et al.*, 2012; Ypma *et al.*, 2012; Lau *et al.*, 2015). Epidemiological information is still required to estimate transmission from genetic data and contact network data is important when trying to identify the most likely transmission events (Di Nardo, Knowles and Paton, 2011; VanderWaal *et al.*, 2014), but there are few tools to formally integrate these (Rasmussen, Volz and Koelle, 2014). Phylodynamic approaches that leverage all available data could provide new insights into pathogen transmission and result in more targeted and improved control interventions, but they must overcome the challenge of appropriate weighting of the often biased and/or fragmented data. Nevertheless, even limited genetic data integrated into transmission models can improve epidemiological insights (Viana *et al.*, 2016) and in situations where other data are fragmented or sparse, sequence data can greatly strengthen the understanding of transmission and inform control.

5.4 Exploiting network properties

5.4.1 Evaluating system resilience

Invasion of a livestock network by an infectious pathogen has the capacity to impair or destroy the function of individual nodes, either by the direct impact on livestock, or by the restrictions resulting from control efforts. The impact on network structure can be considerable, *in extremis* resulting in the destruction of the network as a functioning entity. For infectious diseases, interventions such as movement restrictions, culling or prolonged herd testing are all designed to reduce transmission, but will also have varying degrees of impact on livestock movements and potentially impair the nodes role in the network. Such changes have economic impact (Knight-Jones and Rushton, 2013; Smith *et al.*, 2013) and, if sufficiently harmful, can result in node removal and/or substantial long-term harm to the network. Resilience of a network typically focuses on its ability to recover, retain the same structure, and adapt to maintain system functionality when exposed to disturbances (Carpenter *et al.*, 2001; Holling, 2013). One approach to eliminate disease, such as during the 2001 FMD epidemic, is to disrupt the network by preventing trade for a period (link removal). These movement restrictions, however, can result in excessive livestock welfare issues, welfare culls, and significant long term industry damage (Anderson, 2002). Less disruptively, lasting adjustments (link rewiring) can minimize the impact of highly influential nodes, whilst maintaining overall trade function. An example of this is the implementation of high biosecurity and compartmentalisation in some poultry companies to isolate themselves from disease incursion despite close physical proximity to infected farms, allowing operations to continue in the face of national restrictions (Nickbakhsh *et al.*, 2014).

Minimising the number of affected nodes, or protecting particular ones, may be important for resilience. In dynamic networks, slowing the rate at which contacts occur can slow the rate of pathogen spread and maintain communication between nodes (Kao *et al.*, 2006), improving the networks resilience. Conversely, reducing contact rates can also increase pathogen spread (Nickbakhsh *et al.*, 2013). Additional complications arise when considering multiple layers of a network and multiple diseases that spread on it. Ultimately, targeting control measures that consider the spread of multiple pathogens on a network could be more efficient and robust. Additionally, prior to designing and imposing changes on a network, particularly in economies where livelihoods are heavily dependent on a functioning livestock movement network, the network's resilience to proposed changes should assessed.

5.4.2 Exploiting network data to improve surveillance

The concepts of network resilience can be used to improve surveillance. Albert *et al.* showed the extent to which different types of complex network can be resilient to breakdown (which makes disease difficult to control) or vulnerable to breakdown (which makes disease easier to control) (Albert, Jeong and Barabási, 2000). Nodes (or links) can be removed from a network randomly or using targeted measures such as removing nodes that are highly ranked by one or more centrality measure. In terms of surveillance, random and targeted node removal can be compared to non-targeted and targeted surveillance (Kao *et al.*, 2006). Network analysis can thus provide an analytical framework to predict which farms to test in targeted surveillance strategies and estimate net gains in performance. While generic network analysis can be valuable (Robinson, Everett and Christley, 2007), it can be made more robust by an understanding of the characteristics of the real system (Rossi *et al.*, 2015) and the dynamics of the considered pathogen (Kao *et al.*, 2007). Network analysis has been used to inform targeted surveillance strategies in many livestock systems (Dubé *et al.*, 2009, 2011; Craft, 2015; Rossi *et al.*, 2015; VanderWaal, Enns, *et al.*, 2017), leading to considerable gains in surveillance efficiency (Bessell *et al.*, 2013; Salvador *et al.*, 2018). Analyses of GB livestock networks have identified highly connected premises with a high risk of both becoming infected with and spreading disease (Christley *et al.*, 2005), and have used simulations to show how targeted surveillance could reduce the size of potential epidemics (Kao *et al.*, 2006). For Swedish cattle and pigs, a bespoke metric was identified to consider the timing and sequence of possible incoming and outgoing infection chains (Nöremark *et al.*, 2011). This metric was subsequently expanded to consider the size of the in- and out- components and then used to analyse the German pig trade movements network to identify high-risk farms (Konschake *et al.*, 2013). Such data are not typically available in low resource settings;

having such network knowledge could enable the use of cost-efficient, network measure-targeted surveillance for disease control, but needs justification for the additional cost and effort required.

5.4.3 Multiplexes, multi-layer networks and multi-host pathogen systems

Complex systems are inherently multi-dimensional, with components linked via a complex set of often directed and weighted interactions, giving rise to diverse and unpredictable behaviours (San Miguel *et al.*, 2012). For infectious diseases, these can arise when spread occurs by more than one mechanism (e.g. animal trade, airborne, fomites, sharing a resource or insect vectors), resulting in a multiplex, or where transmission occurs across more than one species, an example of a multi-layer network. Both can compromise disease control (Webster, Borlase and Rudge, 2017), especially when there are biases in available data or ability to exert control (Godfray *et al.*, 2013). The multiplex representation was first developed in the social sciences to represent different types of inter-personal relationships (Kivelä *et al.*, 2014). It has since been used in a variety of contexts, including ecological systems (Pilosof *et al.*, 2017), air transport (Cardillo *et al.*, 2013), behavioural biology (Barrett, Peter Henzi and Lusseau, 2012), and epidemiology (Brooks-Pollock *et al.*, 2015). In one livestock example, a study of a dairy system in northern Italy explicitly accounted for two independent transmission routes: cattle and veterinarian movements. This study found that at the local scale veterinarian movements explained the spread of *Mycobacterium avium* subspecies *paratuberculosis* better than cattle movements and geographic distance failed to capture the impact of veterinarian visits (Rossi, De Leo, *et al.*, 2017; Rossi, Smith, *et al.*, 2017). This highlights a need to identify the potentially multiple transmission routes beyond discrete livestock movements when collecting data to construct a livestock network that is representative of a transmission network.

Many pathogens are multi-host and therefore the network multi-layer. This complication often has severe implications for humans, livestock and wildlife (Haydon *et al.*, 2002). Unfortunately most analytical frameworks of resilience are unsuitable for multi-dimensional systems (Solé and Montoya, 2001), and network resilience can be influenced by interdependence with other networks (Gao, Barzel and Barabási, 2016). Recent work using percolation theory to study the vulnerability of a system of interdependent networks

(Gao *et al.*, 2013) shows the overlap between network layers can improve network resilience and this makes diseases harder to eradicate (Cellai *et al.*, 2013). By disentangling system dynamics from system structure, network characteristics can be identified that influence resilience (Gao, Barzel and Barabási, 2016). A well-known exemplar is the transmission of *Mycobacterium bovis*, the cause of bovine tuberculosis (bTB), between cows and European badgers (*Meles meles*), where the role of different layers can be quantified by exploiting their spatial patterns (Figure S1) (Green *et al.*, 2008). At finer granularities, radio-collar data were used to quantify inter- and intra-species contacts for cattle and badgers (Böhm, Hutchings and White, 2009); adding a layer of indirect contacts based on badger latrines locations to this network showed better correspondence to badger-to-badger transmission patterns (Silk *et al.*, 2018).

5.5 Movement networks where there is limited resource for explicit traceability

There are many examples where livestock movement data have facilitated the planning of disease control and surveillance (Bigras-Poulin *et al.*, 2006; Kiss, Green and Kao, 2006; Natale *et al.*, 2009; Büttner *et al.*, 2013a; VanderWaal *et al.*, 2016; Motta *et al.*, 2017). Conversely, an absence of movement information can obstruct disease control (Gibbens *et al.*, 2001; Ministry for Primary Industries, 2018). In settings where comprehensive tracing systems are absent, a variety of methods have been used to quantify livestock movement patterns and construct movement networks. These include the use of GPS collar data to describe mobility patterns of pastoral herds and overlaps with wildlife areas (Handcock *et al.*, 2009; Musemwa *et al.*, 2012; Raizman *et al.*, 2013; VanderWaal, Enns, *et al.*, 2017), household and market surveys (Poolkhet *et al.*, 2013), transport vehicle records (Kim *et al.*, 2018), and international movement permits (Lindström *et al.*, 2013; Apolloni *et al.*, 2018).

Movement permits are used in many countries to certify livestock health and/or to regulate movement taxes, and have been used to quantify livestock flow and construct movement networks (Dubé *et al.*, 2008; Lindström *et al.*, 2013). The often ephemeral and patchy nature of these records, due to poor archiving or non-compliance (Poolkhet *et al.*, 2016), can results in substantial non-random “missingness” that is difficult to quantify. In these cases, movement permits have been used in conjunction with household and/or market survey data to estimate the risk of disease introduction and target surveillance and vaccination campaigns, also illustrating the importance of a regional disease control approach (Wongsathapornchai *et al.*, 2008; Selby *et al.*, 2013; Motta *et al.*, 2017; Apolloni *et al.*, 2018). Such analyses have identified traders as key targets for disease control (Poolkhet *et al.*, 2016), demonstrated the effects of cattle movement on regional disease transmission (Dean *et al.*, 2013), identified increased risks of bTB with increased herd introductions (Sintayehu *et al.*, 2017) and, with serology data, identified the role of between-village cattle movements in transmitting Rift Valley fever virus (Nicolas *et al.*, 2013).

Biased network samples can make reconstruction of network characteristics difficult. This was addressed in GB by extrapolating from a small biased network sample via statistical associations between common factors in the network study and a national population survey (Nickbakhsh *et al.*, 2011).

Another approach to network construction is to impose an underlying model on observed population densities. Specifically, if census data (populations and locations) are available or can be estimated, gravity (Xia, Bjørnstad and Grenfell, 2004) and radiation (Simini *et al.*, 2012) models provide two ways of creating network models of population mobility. While there is ongoing research regarding their relative merits (Masucci *et al.*, 2013), they share the property of describing movement in terms of relative population size and a measure of distance. Gravity models, for example, describe the probability of a movement occurring in inverse proportion to spatial distance from each hub.

5.6 Evaluating network-based control strategies for livestock movements in Tanzania.

5.7 Introduction to the study

Tanzania provides an exemplar of a rapidly developing emerging economy. In northern Tanzania there is a heavy reliance upon livestock for food, traction power, income, savings and social status. Movements can be over long distances, often on foot, and occasionally over international boundaries with multiple levels of market activity (Aklilu, 2008; Little, 2009; Di Nardo, Knowles and Paton, 2011; Bouslikhane, 2015). The pathogen burden is often high, and this impacts productivity, creates herd/flock instability and, in the case of zoonoses, directly affects human health (Machang'u, Mgode and Mpanduji, 1997; Karimuribo *et al.*, 2007; Schoonman and Swai, 2010; Biggs *et al.*, 2013; Crump *et al.*, 2013; Assenga *et al.*, 2015; Sumaye *et al.*, 2015). In addition to protecting human health, reducing the burden of endemic livestock pathogens to improve livestock health and productivity is recognised as a route away from poverty and necessary to meet global food demands (Allen, 2003; Steinfeld *et al.*, 2006; Kelly and Marshak, 2007; Randolph *et al.*, 2007; Perry and Grace, 2009; Coker *et al.*, 2011; Muma *et al.*, 2014; Pradère, 2014). Livestock sales are also a major source of income in rural communities (Williams, Spycher and Okike, 2006; Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012). In addition to trade between markets, livestock can be sold privately, borrowed or gifted between households and are regularly moved to access natural resources (Coppolillo, 2000; Sintayehu *et al.*, 2017; VanderWaal, Gilbertson, *et al.*, 2017). A reduction in endemic livestock disease is therefore paramount to improving livelihoods in such emerging economies.

Historically there has been no formal, centralised system for identifying and tracing the movement of individual animals in Tanzania, however a paper movement permit certifying livestock health is officially required whenever animals are traded, recording movements *from* markets, though not movements *to* markets. These data are not digitised and the receipt books are stored at administrative Zonal Veterinary Centres in Tanzania. The aims of this study were to: quantify cattle and small ruminant movements in a large (97000 km^2) area of northern Tanzania (Arusha, Manyara and Kilimanjaro regions) using archived, routinely collected government movement permit data; infer livestock movement networks; and build this information into livestock disease simulations to inform surveillance and control.

5.7.1 Methods

Summary methods are presented here; for full details see the Appendix (Appendix section 9.4).

5.7.1.1 Data source and transcription

Access was granted to archived government movement permit receipt books at the Northern Zonal Veterinary Office, Arusha. Movement permit receipt books were selected for analysis from 2009, 2011, 2013 and 2015. Origin, destination, number of each species (cattle, sheep or goat) moved, and date were manually entered into spreadsheets from 50% of the available permits (30,946 permits), of which 19,438 (63%) permits yielded complete data. Only cattle movements are analysed here.

5.7.1.2 Statistical Modelling

Cattle movements were aggregated temporally by month and spatially at the ward level, because origins and destinations often could not be located at a finer scale. A ward is an administrative unit of mean area 243 km^2 and mean human population of 12,000 across the 398 wards in the study regions (National Bureau of Statistics, 2012). We aimed to infer the inter-ward cattle movement network within the study area; movements to outside the study area and within wards were excluded (local movements from markets are less likely to generate a movement permit due to non-compliance). The resulting data set recorded the movement of 86,195 cattle from 98 origin wards to 239 destination wards over the 4 sampled years.

Due to the large number of non-randomly missing permits, it was not possible to use the movement data directly. Instead, the network was inferred by statistical modelling of the observed movements. First, to distinguish true from artefactual absence of movements (months where an origin ward sent out no cattle) a zero-inflated negative binomial (ZINB) generalised linear model (GLM) was fitted to each origin ward, so that in subsequent modelling steps movements would be imputed in place of false zeroes. Next, inter-ward livestock movement was modelled using a hurdle model. The movement between each pair of wards in a given month is represented by a two-step processes: the binary event of any cattle being moved, modelled by a binomial generalised linear mixed-effects model (GLMM); and the number of animals moved, modelled by a zero-truncated negative binomial (ZTNB) GLMM. Each part of the hurdle model allowed movement to depend multiplicatively on the distance between origin and destination wards and their “masses” (human and cattle population sizes), in addition to other characteristics (Table S1). The combined models can therefore be viewed as a gravity model of the livestock movement network. Unexplained spatial and temporal variation was modelled by fitting random effects for origin and destination ward and for the 48 months.

5.7.1.3 Simulated networks

The fitted model was used to simulate monthly movements amongst the 398 wards for one year, with the number of movements inflated twofold to account for using a 50% subsample of the data.

5.7.1.4 Network measures

The simulated data were used to create an observed year-aggregated, static, directed, weighted cattle movement network. A spatial contact layer, connecting all adjacent wards, was added to the market movements network as a simplified means of accounting for contacts and movements between wards that are not represented by the movement permit data. Network analysis methods were applied to the resulting multiplex network to identify nodes with high in-degree, out-degree, betweenness and eigenvector centrality where disease control interventions could be targeted.

5.7.1.5 Simulating disease outbreaks and control on the network

The spread of a 'fast' (R_0 = 3) and 'slow' (R_0 = 1.5) pathogen was simulated on the multiplex to assess the effects of disease control interventions on the spread of pathogens with varying infectiousness [166]. This was achieved by running a stochastic *SIR* compartmental model within each ward. The total number of cattle in the susceptible (*S*), infectious (*I*) and recovered (*R*) compartments was updated daily, while cattle were moved monthly between wards. The two sources of simulated cattle movement were long distance movements via the market network and short distance movements between adjacent wards to account for unobserved local movements (for a full description see Appendix 9.4; an animation of a simulated fast epidemic is available as a supplementary file). Two types of intervention were trialled: proactive vaccination of 70% of the cattle in a ward before the start of the epidemic, and a reactive ban on cattle movements one month after the start of the epidemic. Vaccine interventions were applied to all wards, or targeted at 20 (5%) of wards that were selected randomly, based on their total cattle population size or based on their network centrality measures. The network centrality measures used for targeting interventions were betweenness centrality, eigenvector centrality, and geometric mean degree. The market movement ban was either implemented in all 111 wards that generated outward cattle movements in the simulations and were therefore assumed to have a market, or were targeted in a subset of 20 of these wards, the same number as in the targeted vaccination interventions, and based on the same selection criteria.

5.7.2 Results

The two parts of the hurdle model explained a substantial proportion (binomial: 40%; ZTNB: 24%) of the spatial and temporal variation in cattle movement, with movement being more probable over shorter distances and into wards containing a secondary market, and the number of animals moved being most strongly associated with the agro-ecological system of the origin wards and the presence of a primary or secondary market in the origin or destination ward (Table S1; Figure S2). All variables were retained in the hurdle model that was used to simulate the monthly cattle market movements.

5.7.2.1 Network and node measures

The multiplex network is fully strongly connected (all wards can be reached by all other wards) and displays ‘small world’ properties. The spatial network layer connects all adjacent wards and the permit-related movements reduce the network diameter (longest path length between two wards) from 18 on the spatial network to 12 (see the Appendix Table S5.2 for cattle market, spatial and multiplex networks summary statistics).

The distributions of the three node centrality measures that were investigated (betweenness, eigenvector, and geometric mean degree) were strongly right-skewed. This indicates that the multiplex may be sensitive to targeted disease control interventions at the highly influential nodes. Figure 5-1 shows the geographical distribution of the top-ranked wards for each centrality measure, showing the potential for substantial differences in the effectiveness of targeting controls based on centrality measures due to their geographical distribution.

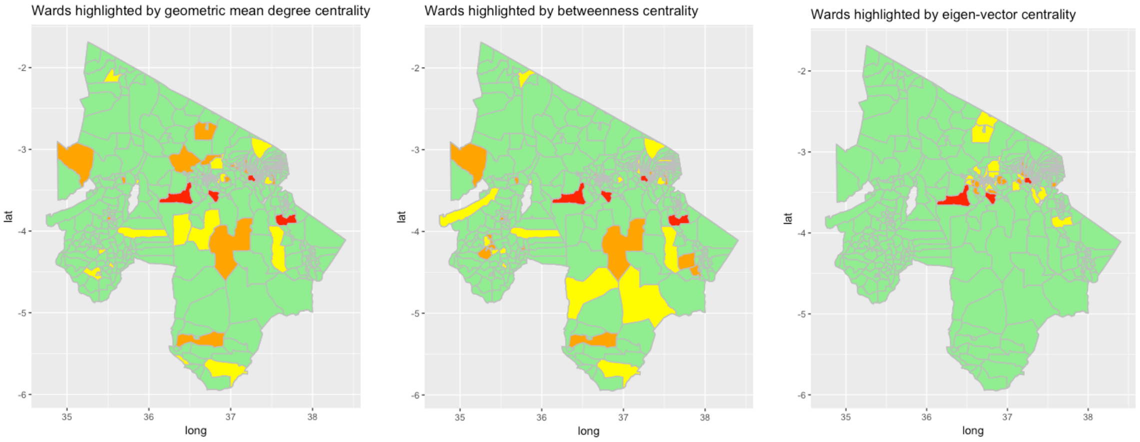


Figure 5-1: Spatial distribution of wards with highest centrality measures in the northern Tanzania livestock movement network, colour shows position in each centrality measure rank, out of 398: red, top 1%; orange, 1-5%; yellow, 5-10%.

5.7.2.2 Simulated movements and pathogen transmission

Mean reductions in population cumulative incidence (PCI) after 1 year for the fast and slow pathogens for each intervention scenario are shown in Figure 5-2. Reductions are relative to PCI reached after 1 year with no intervention (fast: 24%; slow: 1.7%). The higher the reduction in PCI, the more effective the intervention. The list of trialled interventions and associated PCI are given in Tables S3 and S4. All simulated interventions had greater reduction in PCI for the fast pathogen example compared to the slow, although the ranking of intervention efficacy was similar for both fast and slow pathogens. The movement ban implemented in all 111 market wards (high economic and logistical costs) performed only slightly better than when targeted in only 20 wards using network measures, and network-based targeting was more effective than selecting wards using population size or randomly, although there was no substantial difference in performance between the network measures. Vaccination applied to all wards achieved a 100% reduction in PCI for both fast and slow pathogens, while the best-performing targeted intervention, degree centrality, achieved reductions in PCI of 58% (fast) and 31% (slow). The “common sense” intervention of targeting using the total number of cattle performed almost as well as degree centrality, and similarly to the second-best network measure, betweenness, but was much less efficient, requiring 3.5 × more vaccine doses than degree centrality. Targeting vaccination using eigenvalue centrality performed relatively poorly, particularly against the slow disease, where its performance was comparable to selecting wards randomly.

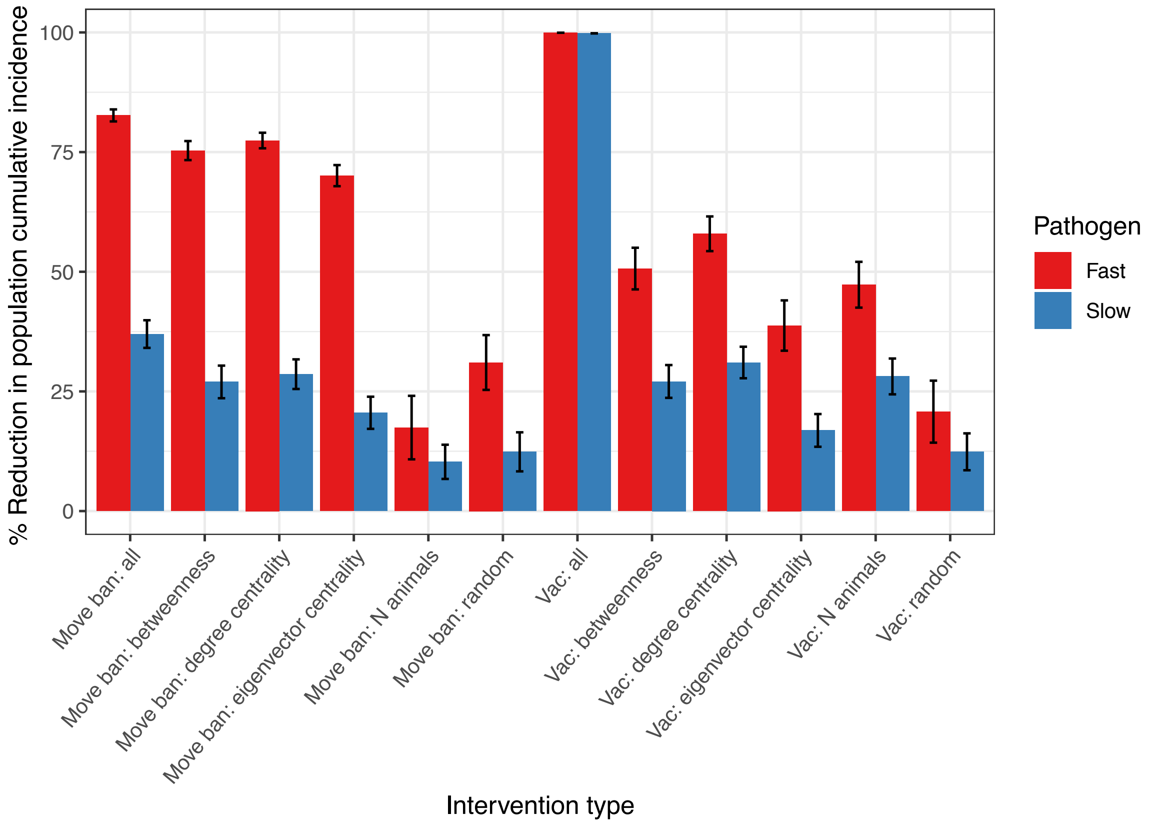


Figure 5-2: Mean (± SE) percentage reduction in population cumulative incidence (PCI) after 1 year for simulated 'fast' and 'slow' transmitting pathogens on the northern Tanzania cattle multiplex network for two types of intervention (market movement ban or vaccination at 70% coverage) applied using six strategies: applied to all wards; targeted to 5% (n = 20) of wards using each of three network centrality measures (betweenness, degree, and eigenvector centrality); targeted to the 5% of wards with the highest cattle population size; and applied to 5% of wards selected randomly. The greater the reduction in PCI, the more effective the intervention is at reducing total number of cases. Mean PCI under each scenario is calculated as the geometric mean of 237 simulated epidemics (full data: Tables S3 and S4).

5.8 Discussion

It is well established that the network analysis of livestock movements can be used to better understand and control diseases of commercial and zoonotic importance in higher income countries where livestock industries tend to be highly structured and movement data are centrally collected and digitised. It is less clear that such approaches are valuable in lower income countries, where movement data are typically unavailable and the cost-benefit ratio less compelling. By exploiting movement permit data collected for health certification and tariff purposes, we have shown that even highly fragmented information about movement patterns can be used to infer network structure. By simulation, we show that the resultant inferred network has the potential to advance strategic understanding. These simulations corroborate that simple network measures can be used to identify good targets for surveillance and disease control that would be appropriate for a range of diseases and reduce the impact of infectious disease at considerably reduced cost and effort. These results could be used to form simple and practical guidelines that could be exploited immediately if, for example, a movement ban was initiated and government needed guidance on where their limited re-enforcement resources should be targeted, although they should not be used for more specific predictions without further data and analysis. They also provide a foundation for deeper research effort, highlighting where the collection of additional empirical data would be useful. For example, the substantial changes in network metrics that result when spatial spread between wards is incorporated highlight the need to augment movement data with more extensive information about local patterns of contact. The homogeneous mixing assumption used at the within-ward level has previously been shown to be useful for developing strategic understanding, even in highly spatially driven scenarios (Keeling *et al.*, 2003), but more detailed recommendations would require modelling of within-ward heterogeneity supported by higher resolution data. This assumption may be less realistic for small urban wards where cattle are tethered, though in larger pastoral and agro-pastoral wards, shared natural resource points might make homogeneous mixing more appropriate (G.L.C., unpublished data and (VanderWaal, Gilbertson, *et al.*, 2017)). Similarly, while the assumption that cattle-to-market movements occur from adjacent wards is consistent with two authors' expert knowledge of livestock management practice (O.M.N. and E.S.), verification with further data collection is an important next step. Finally, simulated movements are dynamically generated based on the random variation generated within the stochastic simulation models. We have not investigated in our dataset evidence of dynamic patterns such as changing network patterns over time because the patchy missingness in our data limits the complexity of the movement model. If more complete data became available for analysis it would be beneficial to assess the evidence for link rewiring throughout the year as this could indicate where control measures should be targeted at specific times. Further potential model deficits include the similar impact of targeting control measures when comparing across centrality measures. This may in part because of the relative crudeness of the disease model; in a more sophisticated model, where the timescales and frequencies of links were considered in greater detail, more substantial differences might be apparent. Similarly, a more explicit model of spatial spread might also prove discriminatory. Finally, the addition of pathogen sequence data where these are available, would provide valuable confirmation of the role of network structure.

5.9 Conclusions

Despite this demonstration of the value of our inferred network approach, we note that data generation was the result of substantial, time consuming effort, and the resultant inferred network, while useful, has limitations as noted above. Mobile broadband technology is becoming increasingly accessible and coupled with the availability of inexpensive scanning devices, the adoption of routine, robust digitised data recording should be achievable. In this paper, we have shown the benefits of having this data to be potentially substantial. This will be particular pertinent in emerging economies such as Tanzania, where changes in industry structure are likely to have unanticipated disease impacts and will require regular monitoring.

5.10 Acknowledgements

We are grateful to Rigobert Tarimo and Sambeke Kiruswa for movement permit data entry, The Ministry of Livestock and Fisheries, Tanzania for access to the movement permit data, Stefan Widgren for assistance with the SimInf package, and three anonymous reviewers whose comments greatly improved this manuscript.

The movement permit study was supported by the UK BBSRC Zoonoses and Emerging Livestock Systems (ZELS) Initiative BB/L018926/1.

Study protocols were approved by the ethical review committees of the Kilimanjaro Christian Medical Centre (KCMC/832) and National Institute of Medical Research (NIMR/2028) in Tanzania and in the UK by the ethics review committee of the College of Medical, Veterinary and Life Sciences, University of Glasgow.

6 Using small ruminant and multi-species movement networks to identify high risk locations for targeted disease control interventions in northern Tanzania

Contribution

As for chapter 5 GC gained access to movement permit data, designed framework for data collection and data processing and carried out data collection and transcription with the help of field staff. GC also contributed to data processing, building of generalised linear mixed model, model fitting and use of model to generate predicted movement data. GC conceptualized research question to be answered by this chapter, combined small ruminant and cattle simulated data, constructed networks and carried out analysis.

6.1 Introduction

People living in the most marginalised communities in northern Tanzania depend on livestock, and increasingly on small ruminants (sheep and goats), for socio-economic and nutritional security (Upton, 2004; Pauw and Thurlow, 2011; Pica-Ciamarra *et al.*, 2011; Covarrubias *et al.*, 2012; Haseeb *et al.*, 2019). Livestock are however susceptible to unpredictable and substantial morbidity and mortality losses due to disease outbreaks (Sindato, Karimuribo and Mboera, 2012; Hughes *et al.*, 2019; Herzog *et al.*, 2020). Minimising livestock disease losses and the uncertainty that accompanies these is needed to ensure people are able to move out of and away from poverty, a key aim of the sustainable development goals (Rich and Perry, 2011; Muma *et al.*, 2014; Pradère, 2014; United Nations, 2015).

Many livestock pathogens can be carried by multiple livestock species and in the case of zoonoses such as Rift Valley Fever virus (RVF) (Wensman *et al.*, 2015; Makuru Nyarobi, 2019), *Brucella* (Assenga *et al.*, 2015) and *Coxiella* (Hummel, 1976), these can spill over to cause disease in people (Crump *et al.*, 2013; Cash-Goldwasser *et al.*, 2018). In northern Tanzania, small ruminant livestock can carry pathogens such as FMDV that may cause mild disease in small ruminants but infect and cause severe disease in cattle, and vice versa, with cattle able to carry the small ruminant virus Peste des petits ruminants virus (PPRV)

(Kivaria, 2003; Herzog *et al.*, 2020). The recent description of PPRV carriage by cattle is an important factor to consider in the development of global PPRV eradication programmes, especially in areas such as northern Tanzania where cattle and small ruminants are frequently kept together (Jones *et al.*, 2016; Fournié, Waret-Szkuta, Camacho, Laike M Yigezu, *et al.*, 2018; Herzog *et al.*, 2019; de Glanville *et al.*, 2020)

The demand for small ruminant livestock products is growing globally, especially in the Middle East, and this growing global demand is increasing small ruminant livestock traffic at an international scale (Sherman, 2011; Meat Livestock Austrailia, 2017). The trade of livestock through markets to meet increasing demands can drive long range movements of livestock between geographically distinct areas and across sometimes multiple international boundaries, providing opportunities for pathogen transmission (Fèvre *et al.*, 2006; Little, 2009; Dean *et al.*, 2013; Nicolas *et al.*, 2018). Where detailed knowledge of livestock movements is available, targeted disease control interventions such as vaccination, active surveillance or temporary movement restrictions can be implemented to reduce pathogen transmission risk (Bessell *et al.*, 2013; Buhnerkempe *et al.*, 2014; VanderWaal, Enns, *et al.*, 2017). In contrast, where data on livestock movements are scarce, disease preparedness and an ability to respond rapidly in the face of a livestock disease outbreak are difficult (Gibbens *et al.*, 2001; Ministry for Primary Industries, 2018).

Livestock movement data is now frequently used in veterinary epidemiology to construct graphs that can be analysed using concepts from network analysis to identify locations or populations for targeted disease control interventions (Vallée *et al.*, 2013; VanderWaal *et al.*, 2016). Graphs are made up of nodes which can represent individuals, populations or locations, connected by links that represent movements or contacts between them (Newman, 2010). For networks that are constructed to assess connectivity for disease control purposes the link between nodes should represent an epidemiological connection along which a pathogen can be transmitted (Craft, 2015).

In addition to some of the NA methods already mentioned in this thesis, community detection is another method which can be used to assess network structure and how homogeneously connected a network is. Communities are groups of nodes in a network with many connections within the group and few between them (Blondel *et al.*, 2008). If highly connected communities exist within networks, their identification can enable groups of nodes within communities to be isolated during disease outbreaks (Guinat *et al.*, 2016). This would be achieved by banning trade or livestock movement between nodes from different communities but allowing movements to continue between nodes that belong to the same community.

When multiple modes of pathogen transmission between individuals or groups of individuals in a population exist, for example via other species or indirectly through environmental contamination, the risk of pathogen persistence and epidemics in a population increases (Craft *et al.*, 2009; Zhao *et al.*, 2014; White, Forester and Craft, 2017; Wilber *et al.*, 2019). Multi-host or multi-species networks are less studied than single species networks, but where they have been used they demonstrate the importance of considering the multiple routes for pathogen transmission (Böhm, Hutchings and White, 2009; Lavelle *et al.*, 2016; Ruget *et al.*, 2021). If high risk individuals or locations can be identified in a multi-species network, targeted disease control measures can be implemented at these locations with multiple species included in the intervention activities. This is an efficient use of disease control resources as the fixed costs are shared and the cost-benefit impacts are maximised across species (Tisdell and Adamson, 2017). Disease control interventions implemented at high-risk nodes on multi-species networks could focus on multi-species pathogens and/or multiple pathogens. This type of integrated and multi-faceted approach to pathogen control is needed for efficient control of endemic diseases that often receive little political attention and limited public funding towards control (Maudlin, Eisler and Welburn, 2009; Carslake *et al.*, 2011; Brooks-Pollock *et al.*, 2015). If single and multi-species networks in a population are highly correlated the multi-species intervention foundations could also be built upon to implement single species pathogen interventions (Böhm, Hutchings and White, 2009). Identification of how correlated single and multi-species networks are in different agro-ecological settings and populations could allow future multi-pathogen disease control programmes to be designed even when data on one or the other species is lacking (Kao *et al.*, 2006; Nöremark *et al.*, 2011; Apolloni *et al.*, 2018; Ruget *et al.*, 2021).

In northern Tanzania there is currently no centralised data available on the movement of small ruminants and thus how these movements contribute to the maintenance or propagation of livestock and zoonotic pathogens in the area. We used government movement permit data, largely representative of livestock movements between and onward from markets in northern Tanzania, to characterize the cattle movement network in chapter 5 (Chaters *et al.*, 2019). The primary objective of this study is to build on the work from chapter 5 (Chaters *et al.*, 2019) to identify how epidemiologically connected northern Tanzania is through the movement of small ruminants alone and through the combination of cattle and small ruminant movements. We will assess if there is a strong community structure in the small ruminants and combined species networks so these communities could be identified and isolated in case of future disease outbreak scenarios.

We will also identify high risk locations for small ruminant and multi-species pathogen introduction and onward transmission and will assess if targeted removal of high-risk locations can successfully fragment the networks' connected components. This will simulate the effectiveness of targeted disease control interventions implemented in nodes ranked based on their risk for small ruminant only or multi-species pathogens. We will compare the effectiveness of targeted node removal at achieving network fragmentation to assess if a multi-species, multi-pathogen intervention would be an efficient use of disease control resources in this setting.

6.2 Methods

6.2.1 Study area and data

This study is focussed on the movement of sheep, goats and cattle in three regions in the northern zone of Tanzania: Arusha, Manyara and Kilimanjaro. The study area and movement data used for this study are described in detail in chapters 2 and 5 (Chaters *et al.*, 2019), respectively. In brief, movement data were collected from movement permit receipt books. Movement permits are required in Tanzania to move livestock from point A to point B. The study area is home to livestock keepers who practice pastoralism (grazing livestock over large areas of pasture) which makes enforcing the use of movement permits difficult. In previous work studying the movement permit data in chapter 4 we found that generally the movement permits are used for trade movements of livestock onward from markets (Chaters *et al.*, 2019). The origin and destination locations, species, number moved, date of movement, mode of transport, any health concerns and price of the movement tax are all recorded on the permit.

In 2016 we were granted access to over 64,000 paper movement permit receipts from Arusha, Manyara and Kilimanjaro regions of northern Tanzania, covering 4 years (2009, 2011, 2013 and 2015). The permit receipt books were ordered by district and year and then photographed and stored as JPEG image files. Photographs of the permits were separated into 20 folders (A:T) in turn so each folder contained a representative sub-sample. Data from 50% of the folders were transcribed (28,421 permits). Data from a further 5,045 permits were entered directly into a spreadsheet so these were also allocated in the same way into 20 batches and subsampled down 50% so the data from these permits were not over represented. This created a final raw data set that comprised 30,906 records. Some permits were unreadable due to fading or illegible writing which resulted in 67% of the permits being successfully transcribed. Through the data cleaning process the to and from locations were assigned to origin and destination locations using the fuzzy text-matching package *stringdist* (van der Loo, 2014), against a database of Tanzanian geographic names compiled from the Geographic Names Database (http://geonames.nga.mil/gns/html/namefiles.html; file dated 10 April 2017) and the National Bureau of Statistics 2012 Population and Housing Census of Tanzania. Central point coordinates were assigned to locations that matched to origins and destinations. Fuzzy matching was used as a guide only; all origins and destinations were checked visually against the JPEG where available, and ambiguous matches were adjudicated with guidance from Tanzanian colleagues with local knowledge. Movements with origins or destinations outside the study area were removed from the analysis dataset. The final cleaned dataset comprised of data from 19,438 movement permits representing movements of 124,491 animals. Further details on data collection and processing are found in chapter 4 (Chaters *et al.*, 2019).

6.2.2 Multi-species husbandry and transmission risk

The proportion of livestock-keeping households that keep both cattle and small ruminants in the study area was estimated using cross-sectional household survey data from the Social, Economic and Environmental Drivers of Zoonoses (SEEDZ) and Bacterial Zoonoses (BacZoo) research projects. Further details on both of these projects are found in chapters

2 and 3 and published studies (Bodenham, Lukambagire, *et al.*, 2020; de Glanville *et al.*, 2020)). The number of cattle, sheep and goats owned for each surveyed household were used to calculate the proportion of surveyed livestock owning households that owned cattle, small ruminants and both. This was done to evaluate the potential for mixing between species groups within households which will provide opportunities for pathogen transmission events in addition to any mixing opportunities and indirect contacts that occur outside the household.

6.2.3 Model fitting

The movement permit data were aggregated to the ward level (398 wards in the study area, each representing an administrative unit of around 12,000 people) for the 48 study months (12 months for each year 2009, 2011, 2013, 2015). Because of the age of some of the movement permit receipt books and the potential for patchy data collection at source due to missing or lost receipt books the data were checked for evidence of zero inflation. Where there was evidence of zero inflation these zeros were taken to represent missing data rather than zero movements.

A mixed effects negative binomial hurdle model with the two response variables *any movement* (binary) and *expected number of animals moved given that a movement occurs* (continuous), was fitted to the data using ward level characteristics as explanatory covariates. The continuous fixed effect covariates were: origin and destination cattle or small ruminant population size, origin and destination human population size, origin and destination geographical area and calendar month (1:12). All continuous covariates were fitted with natural cubic splines with 3 degrees of freedom to allow the model estimates to vary throughout the year and identify up to two peaks or troughs in livestock movement volume. The fixed effect categorical covariates are: year, presence of a primary or secondary market and agro-ecological classification of origin and destination. The model random effects (all categorical) were: from ward, to ward and period (1:48 for all study months). A hurdle model was used so that the strength and effect size of the factors that predict the probability of any movement occurring between two wards in a given month and the expected number of animals moved, given that a movement occurred, could vary for the two separate parts of the model.

6.2.4 Network construction

The cattle and small ruminant negative binomial hurdle models were used to predict the probability of a movement, and the expected number of livestock moved given that a movement occurred, between all potential origin (117 wards as origins in the raw data) and destination wards (398 possible destination wards in the study area) for each month of a year. The model output was boosted by 100% to account for only having transcribed 50% of the available permits. The model prediction outputs were used to create two matrices: a *probability of movement* matrix and an *expected number of animals moved given a movement occurs* matrix, for each month of the year for both cattle and small ruminants. For each month the *probability of a movement* matrix and the *expected number of animals moved given a movement occurs* matrix were multiplied to create 12 monthly *expected number of movements* between all ward pairs for both the small ruminant and cattle data. The 12 monthly *expected number of movements* matrices were then summed to give the *annual expected number of animals moved* between all pairs of wards in Kilimanjaro, Arusha and Manyara regions of northern Tanzania for both cattle and small ruminants. If the expected number of livestock moved between any two wards was less than one animal per year these links were discarded. This was done to avoid creating a network with an unrealistically large well connected component that includes predicted movements between wards that represent almost zero contact. Thus, in the predicted networks some wards appear as inactive, with no movement in and out even though we know this is unlikely as all wards containing some livestock will probably have some movement activity in or out in a year. The cut off decision was made based on the assumption that if recorded movements between wards are so infrequent they are not picked up by the models, discarding these links from the annual networks is unlikely to affect the overall network structure or the identification of central nodes (Wang *et al.*, 2012). The total sum of the discarded small ruminant links (where the expected number moved in a year was <1) is 2,118 (5.0% of the total sum of predicted small ruminant movements) and in the cattle dataset it is 1,847 (2.5% of the total sum of predicted cattle movements). The low percentages of total movements represented by the discarded links (5% small ruminants and 2.5% cattle) shows that the arbitrary cut off of <1 animal per year used here to construct the network doesn’t remove a large proportion of movements. The annual expected number of animal movements matrices were used to construct a small ruminant static annual movement network and a cattle static annual movement network (Chapter

4). In the single species networks, link weights are equal to the expected number of animals moved in a year.

The cattle and small ruminant annual expected number of movements matrices were summed to create a static annual multi-species livestock movement network with link weights equal to the expected total number of cattle and small ruminants moved between node pairs in a year (the *combined equally weighted* network, referred to as combined_E).

Finally, a fourth network was constructed from the combined small ruminant and cattle expected number of annual movements matrices, this time heavily weighting small ruminant movements (multiplying estimated annual number moved by 0.9) and lightly weighting cattle movements (multiplying estimated annual number moved by 0.15) (the *combined small-ruminant-weighted* network, referred to as combined_SR). This was done to simulate a scenario representative of a pathogen circulating at high prevalence (90%) in the small ruminant population and low prevalence (15%) in the cattle population. In this example scenario a multi-species disease control intervention may be needed for pathogen elimination but the relative contribution of each species group to pathogen transmission differs and this may need to be accounted for during the development of interventions.

6.2.5 Network structure descriptive analysis

To assess study area connectivity through the movements of small ruminants alone and through the combination of cattle and small ruminant movements, summary network statistics were calculated for the small ruminant, cattle and combined_E networks. Each network's number of nodes, links, giant strongly connected component (GSCC), giant weakly connected component (GWCC), diameter, average path length and clustering coefficient were calculated (Newman, 2010). Results from analysis of the small ruminant, cattle and combined_E networks were compared to evaluate the effect of adding the cattle data to the small ruminant data on overall study area connectivity.

6.2.6 Community detection

Communities in networks are groups of highly connected nodes with more connections to other nodes within the group compared to nodes outside it. Community detection was done on the small ruminant and combined_E networks to determine whether communities exist in each network, and to count how many individual nodes are isolated and do not belong to a community. In brief, the algorithm used for community detection on these networks assigns each node to its own community then each node is re-assigned to a community with which it achieves the highest contribution to modularity (fraction of nodes that fall within communities minus fraction expected to fall within groups if the network graph were completely random). When no nodes can be reassigned the process is repeated with each community considered a node. The process stops when all nodes are part of one community or modularity cannot be increased (Blondel *et al.*, 2008).

All model fitting, network construction and NA was done in R version 4.0.4 (R Core Team, 2019b) using the *lme4* (Bates *et al.*, 2015), *glmmTMB* (Brooks *et al.*, 2017) and *igraph* packages (Csardi and Nepusz, 2006) and graphs were plotted using *ggplot* (Wickham, 2016). Community detection was done using the *multilevel.community* function in *igraph* which uses the multi-level modularity optimisation algorithm to find community structure (Blondel *et al.*, 2008).

6.2.7 Node centrality metrics

Five node centrality metrics were calculated for each ward in the small ruminant, cattle and combined networks (Newman, 2010): weighted in- and out-degree, eigenvalue centrality (score of a node is proportional to the centrality of its connected neighbours) (Bonacich, 2007), betweenness centrality and geometric mean degree centrality (square root of the product of in- and out-degree). Each of these five centrality measures represent a different aspect of a node’s risk of pathogen acquisition or transmission. Eigenvalue centrality accounts for how well connected a nodes neighbours are and relates directly to a node’s contribution to R_0 on the network (Danon *et al.*, 2011). The five measures assessed here use different methods to quantify a node’s risk and so can be used in combination or alone, depending on the desired outcome of a disease control programme, to select nodes for targeted intervention (De Arruda *et al.*, 2014). The distributions of these five node centrality metrics were described for each network and compared within and between the networks. We assessed if the distributions of node centrality metrics exhibit right skew and compare centrality metric ranks using Spearman’s ρ correlation coefficient, to evaluate if nodes that rank highly for one metric also rank highly for others. This allows us to assess if the different centrality metrics identify the same nodes as high risk for disease acquisition and onward transmission. The P-values presented for the Spearman’s ρ correlation coefficient are calculated using the AS 89 algorithm (Best and Roberts, 1975).

We compared node ranks for three centrality measures (weighted in-degree, betweenness and eigenvalue centrality) between the SR, combined_E and combined_SR networks to evaluate if high risk nodes on the SR network remained high risk when the cattle data were added. These three measures were chosen because they were least correlated within the small ruminant and combined_E networks and represent different risks of pathogen introduction, between group transmission and contribution to transmission on the network through their highly connected neighbours.

The number and proportion of nodes that change rank from zero activity on the small ruminant network to having (i) some activity and (ii) being ranked in the top 20% of most central nodes in the combined_E network was also calculated. This analysis will allow us to assess the extent of the change in nodes’ importance when cattle data are added to the small ruminant data. Understanding the proportion of nodes that change rank when the cattle data are added to the small ruminant data will allow us to assess whether using combined_E node ranks will be a substantial waste of resources when targeting a small ruminant specific pathogen.

6.2.8 Node removal and network fragmentation

Finally, the top 20% and 5% of nodes, ranked based on their weighted in-degree, eigenvalue and betweenness centrality on the combined_E network, were removed from the small ruminant, combined_E and combined_SR networks followed by recalculation of the GSCC and GWCC. Nodes were removed based on their rank in the combined_E network to assess how effective control measures that specifically target locations high risk for multi-host pathogens are at reducing connectivity on the small ruminant only network as well as on the multi-species networks. The top 5% of nodes ranked most highly based on their betweenness centrality on the small-ruminant-only network were also removed from the small ruminant network and the size of the GSCC was recalculated. This was done to assess if targeting using small-ruminant-only network node metrics had a substantially different effect on fragmentation of the small ruminant network compared to removal based on nodes’ ranks in the combined_E network. Complete network fragmentation is achieved when the giant components are reduced to size 1 so that zero nodes remain epidemiologically connected to any others.

6.3 Results

6.3.1 Raw data description

A description of the transcribed raw data from all transcribed movement permits is given in Table 6-1. There was more than double the number of permits recording cattle movements compared to small ruminants. The cattle data identifies more unique origin and destination wards than the small ruminant data and there are some different origins and destinations listed in each dataset as the total numbers of origin and destination locations in the combined dataset are more than in each individual dataset.

The mean number of animals moved on each permit was 10.8 for small ruminants (range 1-190, SD 20.9 and 8.1 for cattle (range 1-337, SD 10.4). The mean movement distance within the study was 90 km for small ruminants (range 0.5 – 850, SD 134) and 79 km for cattle (range 0.5 – 883, SD 115). Distributions of the number of small ruminants and cattle moved per permit (a, b) and distances moved (c, d) are shown in Figure 6-1.

Table 6-1: Number of movement permits transcribed and their characteristics from Arusha, Manyara and Kilimanjaro regions of northern Tanzania

	Number of permits transcribed *Including movement data within study area only					Number of animals moved *Within study area (from 19,438 permits)	Unique links identified in raw data	Number of unique origin and destination wards in raw data	Number of unique Origin wards in raw data	Number of unique Destination wards in raw data
	2009	2011	2013	2015	Total					
Small Ruminant	266	1,199	1,572	2,697	5,734	38,299	2,990	300	95	279
Cattle	1,462	2,755	3,682	5,958	13,857	86,192	5,127	340	101	327
Combined	1,709	3,917	5,222	8,590	19,438	124,491	7,138	389	117	376

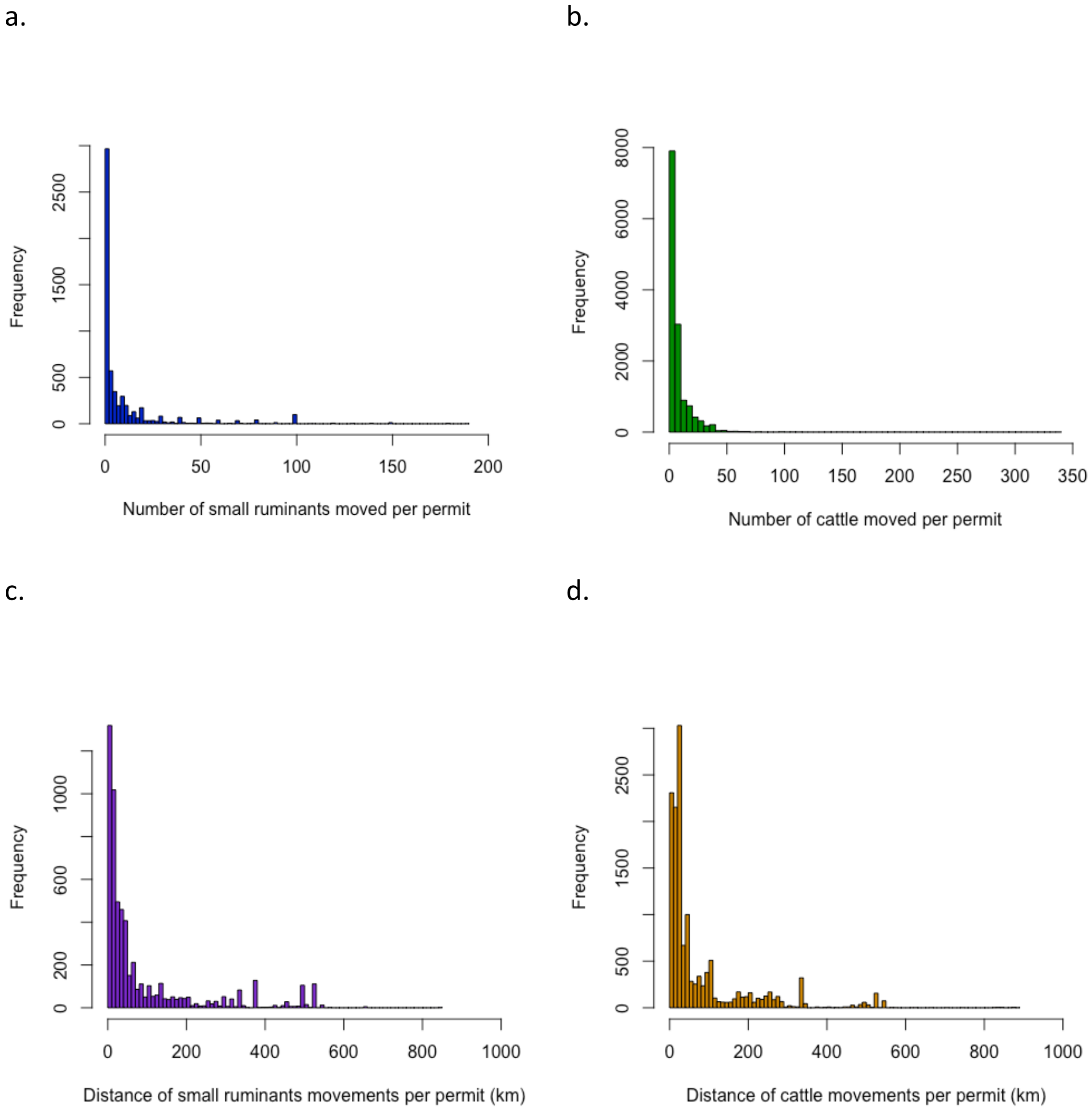


Figure 6-1: Movement permit data distribution of number of number of small ruminants (a) and cattle (b) moved per permit and distance (km) of recorded small ruminant movements (c) and cattle movements (d)

6.3.2 Multi-species husbandry and transmission risk

Out of 524 livestock-keeping households that were surveyed in the study area, 402 (76%) kept both cattle and small ruminants, 61 (12%) kept cattle only, and 61 (12%) kept small ruminants only. The number of small ruminants kept is also on average twice the number of cattle: the mean number of cattle was 42 (range 0-1140, SD 107), while the mean number of small ruminants was 84 (range 0-2600, SD 230).

6.3.3 Hurdle model results

Results from the binomial hurdle models fitted to the transcribed permit data for small ruminants and cattle are given in Table 6-2 and Table 6-3 respectively. The results tables display the effect of each variable on the model R^2 value (effect size of X = 1 – (variance of linear predictor without X / variance of linear predictor of full model)). The size of the effect of each variable, and the significance of the effect (likelihood ratio test P-value) is given for both the binary and count parts of the model. The hurdle models used here are generalised linear mixed models which are fitted by maximising the likelihood of the observed data (different to standard linear regression models which are fitted by least squares). The difference in method model fitting, along with how R^2 is defined, explains why it is possible to get a negative effect on R^2 by removing a variable when this wouldn't be possible with a standard linear regression model.

Similar variables affect the probability of a movement and number of animals moved between wards but the effect sizes vary in the small ruminant and cattle models. Distance has the largest effect on the probability of any movement between wards in both models (small ruminants 51% and cattle 58%, $P < 0.001$) but it has no effect on the number of cattle moved and has a 7% effect on the number of small ruminants moved. Human population size and A-E class of destination, year and presence of a market affect the probability of a small ruminant movement between wards and these variables plus A-E class of origin affect the probability of a cattle movement ($P < 0.05$). In addition to distance, variables that have a significant ($P < 0.05$) effect on the number of small ruminants moved include destination area, presence of a market, A-E class of origin and destination wards and year. In the cattle model variables that affect the number of animals moved include month, year, market presence and origin agro-ecological class ($P < 0.05$).

Table 6-2: Predictors of monthly inter-ward small ruminant movement. Effect: relative reduction in model mean sum of squares when predictor is removed. R2LATENT: proportion of spatial and temporal random effect variation explained by the fixed effects

	Model stage predicting probability of movement (R^2_{LATENT} = 44%)		Model stage predicting number of animals moved (R^2_{LATENT} = 18%)	
Predictor	**Effect**	**P-value**	**Effect**	**P-value**
log$_{10}$(distance/km) [spline]	**51%**	<0.001	**7%**	0.005
log$_{10}$(origin human pop. size) [spline]	3%	0.171	-18%	0.367
log$_{10}$(destination human pop. size) [spline]	**11%**	<0.001	5%	0.256
log$_{10}$(origin cattle pop. size [GWA†]) [spline]	1%	0.122	-1%	0.394
log$_{10}$(destination cattle pop. size [GWA†]) [spline]	0%	0.468	0%	0.503
log$_{10}$(origin area/km^2) [spline]	0%	0.225	0%	0.641
log$_{10}$(destination area/km^2) [spline]	1%	0.157	**7%**	0.028
Calendar month [spline]	0%	<0.001	1%	0.326
Year [categorical]	**8%**	<0.001	**2%**	0.001
1ary/2ary market in origin/destination [categorical]	**20%**	<0.001	**18%**	0.048
Origin agro-ecological system [categorical]	3%	0.419	**12%**	0.002
Destination agro-ecological system [categorical]	**7%**	<0.001	**8%**	0.005

Table 6-3: Predictors of monthly inter-ward cattle movement. Effect: reduction in model mean sum of squares when predictor is removed. R2LATENT: proportion of spatial and temporal random effect variation explained by the fixed effects

	Model stage predicting probability of movement (R^2_{LATENT} = 40%)		Model stage predicting number of animals moved (R^2_{LATENT} = 24%)	
Predictor	**Effect**	**P-value**	**Effect**	**P-value**
log$_{10}$(distance/km) [spline]	**58%**	<0.001	0%	<0.001
log$_{10}$(origin human pop. size) [spline]	4%	0.053	16%	0.125
log$_{10}$(destination human pop. size) [spline]	**14%**	<0.001	0%	0.197
log$_{10}$(origin cattle pop. size [GWA†]) [spline]	0%	0.952	-1%	0.349
log$_{10}$(destination cattle pop. size [GWA†]) [spline]	0%	0.968	4%	0.304
log$_{10}$(origin area/km^2) [spline]	1%	0.274	6%	0.91
log$_{10}$(destination area/km^2) [spline]	2%	0.028	0%	0.28
Calendar month [spline]	0%	0.054	**5%**	<0.001
Year [categorical]	**3%**	<0.001	**2%**	<0.001
1ary/2ary market in origin/destination [categorical]	**27%**	<0.001	**27%**	<0.001
Origin agro-ecological system [categorical]	**6%**	0.036	**29%**	0.039
Destination agro-ecological system [categorical]	**6%**	<0.001	2%	0.759

6.3.4 Network analysis results

Of the total 398 wards in the study area there were 330 active ward nodes in the small ruminant network, 316 in the cattle network and 373 in the combined data network. There are more cattle moved in a year (N = 73,506) than small ruminants (N = 40,529) and more active links on the cattle network (N = 2,926) compared to the small ruminant network (N = 2,653). Descriptive network measures from analysis of the small ruminant, cattle and combined_E networks structures are given in Table 6-4. The combined data network is more connected with more active links, a larger GSCC and GWCC and shorter diameter and average path length compared to both the small ruminant and cattle networks.

The mean small ruminant journey length on the network is 95 km (range 1.7-365, SD 73) and the mean cattle journey length is on average around 30% longer at 118 km (range 2.05 – 435km, SD 85). Both mean distances are similar to the mean movement distances in the raw permit data.

The distributions of link weights (count of number of animals moved along that directed link in a year) for the small ruminant, cattle and combined networks are all skewed to the right. This means that the majority of active links in all of the movement networks represent quiet routes with the movement of few animals and a small proportion of links represent busy livestock traffic corridors with the movement of many animals over the course of a year. For small ruminants the median link weight is 3.2 (mean 15.3) with a range from 1 to 1,275 (SD 58.5). For cattle the median link weight is 4.0 (mean 25.1) with a range from 1 to 2,279 (SD 97.4) and for the combined data network the median link weight is 4.6 (mean 28.0) with a range from 1 to 2,594 (SD 108.1).

Table 6-4: Network measures for small ruminant, cattle and combined annual static movement networks in northern Tanzania where at least one cattle, small ruminant or either, was estimated to be moved between two locations in the year

	Nodes	Active nodes (>1 animal moved in or out per year)	Links	Weighted links (no of animals)	GSCC	GWCC	Diameter	Average path length	Clustering coefficient
Small ruminant	398	330	2,653	40,529	82	330	8	3.31	0.30
Cattle	398	316	2,926	73,506	102	316	8	2.58	0.31
Combined_E	398	373	4,068	114,035	111	373	7	2.6	0.34

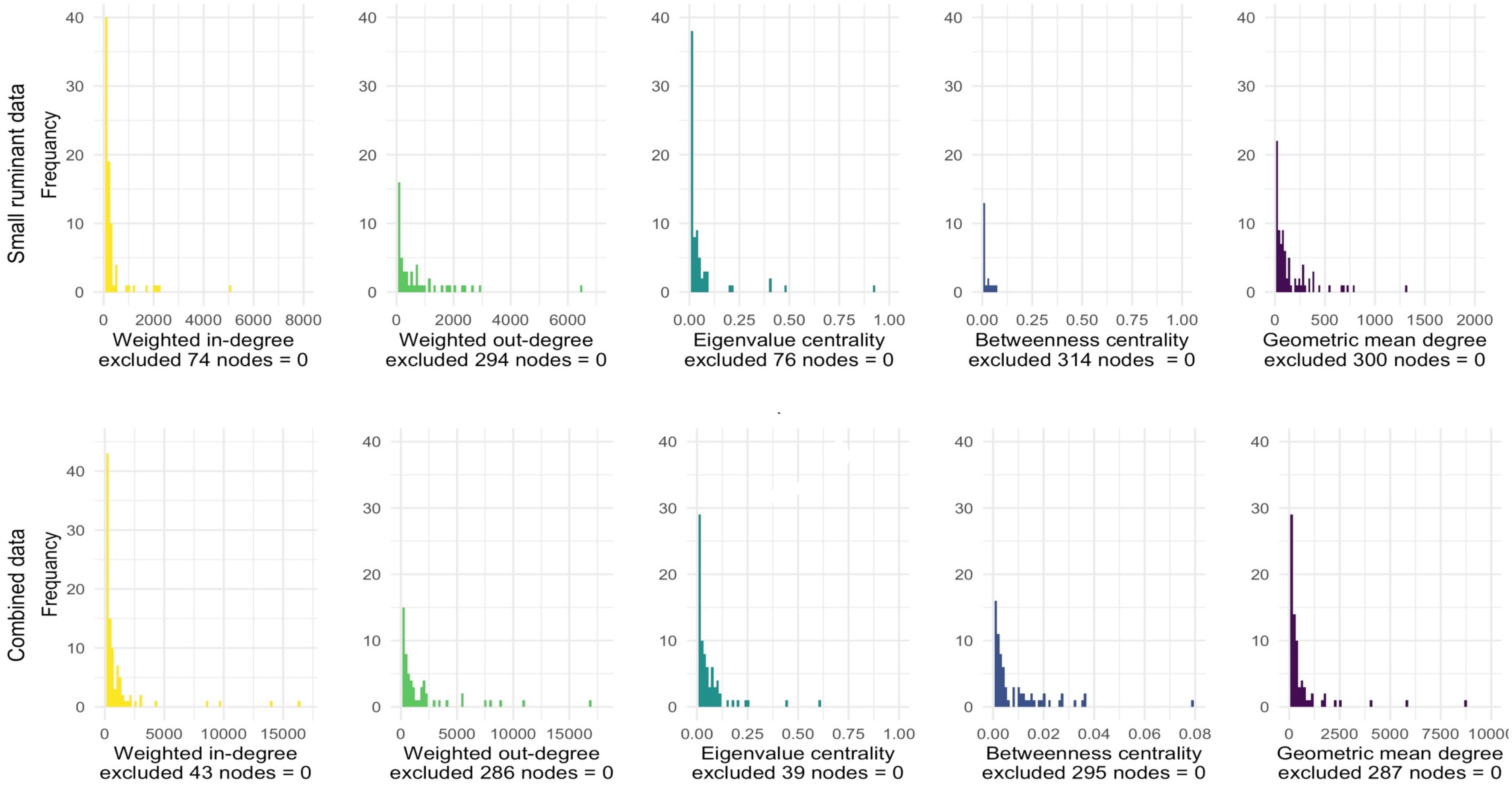
6.3.5 Community detection

The small ruminant network has 3 communities greater than size 1 (69, 125 and 136 nodes) and 68 inactive nodes. Each of the three communities identified in the small ruminant data grow when the cattle data is combined with the small ruminant data. The combined data network also has 3 communities greater than size 1 (77, 134 and 149) with only 38 inactive nodes. Identifying communities in the observed networks means that these could be isolated during disease outbreaks and trade allowed to continue within communities provided there was no contact or movement between nodes from different communities.

6.3.6 Node centrality metrics

6.3.6.1 Centrality metric distributions

There was right skew of all node centrality metrics; weighted in-degree, weighted out-degree, eigenvalue centrality, betweenness centrality and geometric mean degree in the small ruminant, combined_E (cattle and small ruminant links weighted equally) and combined_SR networks (small ruminant links weighted heavily (0.9), cattle links weighted lightly (0.15)). The distributions of all calculated node centrality measures are shown in Figure 6-2 with all zero values excluded from the plots so the distribution of the non-zero values can be observed. The number of excluded zero values is given below the X axis on each plot. Summary data for all node centrality metrics in the small ruminant and combined_E networks can be found in the Appendix Table S5.1.


Small ruminant data
Frequancy
40
30
20
10
0
0
2000
4000
6000
8000
Weighted in-degree
excluded 74 nodes = 0
0
2000
4000
6000
Weighted out-degree
excluded 294 nodes = 0
0.00
0.25
0.50
0.75
1.00
Eigenvalue centrality
excluded 76 nodes = 0
0.00
0.25
0.50
0.75
1.00
Betweenness centrality
excluded 314 nodes = 0
0
500
1000
1500
2000
Geometric mean degree
excluded 300 nodes = 0
Combined data
Frequancy
40
30
20
10
0
0
5000
10000
15000
Weighted in-degree
excluded 43 nodes = 0
0
5000
10000
15000
Weighted out-degree
excluded 286 nodes = 0
0.00
0.25
0.50
0.75
1.00
Eigenvalue centrality
excluded 39 nodes = 0
0.00
0.02
0.04
0.06
0.08
Betweenness centrality
excluded 295 nodes = 0
0
2500
5000
7500
10000
Geometric mean degree
excluded 287 nodes = 0

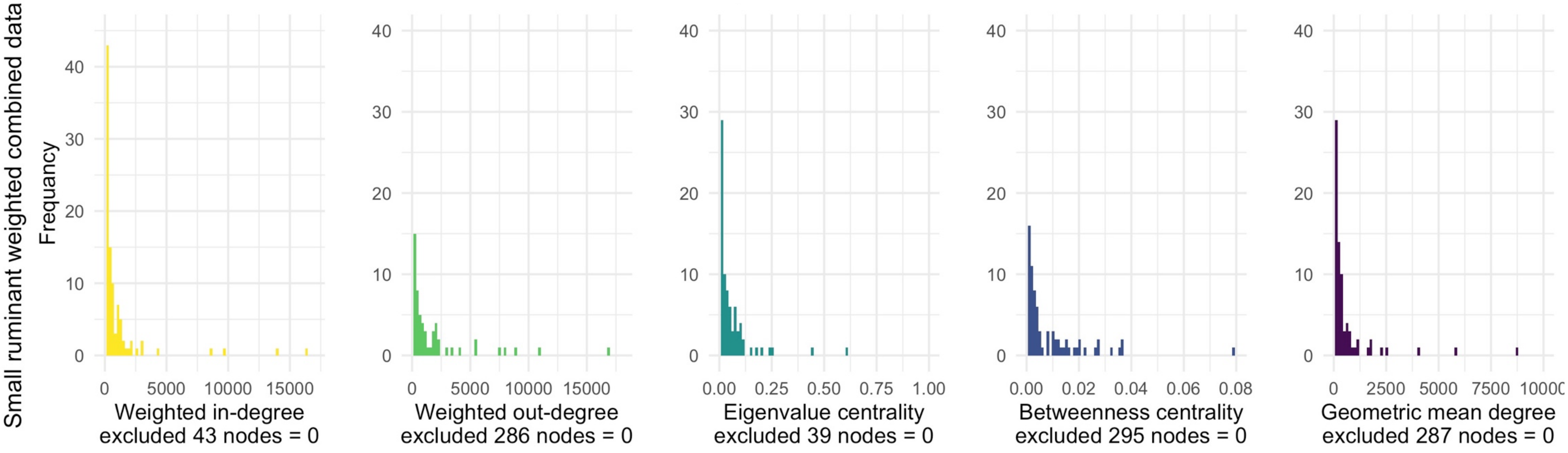


Figure 6-2: Node centrality metric distributions: weighted in-degree, weighted out-degree, eigenvalue centrality, betweenness centrality and geometric mean degree for three networks: small ruminant only, combined cattle and small ruminant (equally weighted) and combined cattle and small ruminant with links weighted heavily for small ruminants (0.9 multiplied by the number of small ruminants moved) and lightly weighted for cattle (0.15 multiplied by the number of cattle moved) to represent an example of the relative impact of the movement of the different species groups on transmission of a pathogen that is shed more by small ruminants that cattle. All zero values are excluded from the plots so the distribution of none-zero values can be visualised. The number of zero values that are excluded for each centrality metric, on each network, is given below each plot.

6.3.6.2 Centrality metric correlations

Node centrality metric ranks were positively correlated on both the small ruminant and the combined_E networks (Figure 6-3). Scatter plots of pairwise comparisons between node centrality metric ranks are in the bottom right panels and Spearman's ρ correlation coefficients are in the top left panels for the small ruminant (Figure 6-3(a)) and combined_E networks (Figure 6-3(b)), all P-values <0.001. Dashed lines on the scatter plots highlight the boundaries for the top five percent (red) and twenty percent (blue) of nodes. Points that fall on the diagonal represent nodes that rank equally high or low for both centrality metrics being compared. Points off the diagonal (above the horizontal lines and to the left of the vertical, or to the right of the vertical line and below the horizontal) represent nodes that are dissimilar in how they rank using the different metrics. The dissimilar nodes rank in the top twenty percent (blue line) or five (red line) percent for one metric but not for the other for each pairwise comparison. Geometric mean degree, out-degree and betweenness centrality were highly positively correlated on both the small ruminant and combined_E networks indicating that they pick up the same information. For the next analysis, comparing how node ranks correlate for different centrality measures on the three networks, the three least correlated centrality measures were used (weighted in-degree, eigenvalue and betweenness) as these metrics contain more diverse information about risk.

Spearman's ρ correlation coefficients (top left) and scatter plots (bottom right) that compare three node centrality metric ranks between the three networks: small ruminant, combined_E and combined_SR networks are shown in Figure 6-4 (weighted in-degree (a), eigenvalue (b) and betweenness (c)). On these scatter plots points above and to the right of the dashed lines are ranked in the top 5% (red) and 20% (blue) of nodes. Points that are off the diagonal (above the horizontal lines and to the left of the vertical and below the horizontal lines and to the right of the vertical) represent nodes with highly discordant ranks. These discordant nodes change rank due to structural changes on the different networks and move into the top twenty percent (top left or bottom right of blue lines) and five percent (top left or bottom right of red lines) high risk groups on the different networks. The scatter plots show a smaller proportion of nodes change rank into the top five and twenty percent of nodes when the cattle data is added to the small ruminant data compared to the proportion of nodes that remain on the diagonal (rank similarly). The

Spearman’s ρ correlation coefficients show that for all pairwise comparisons (using centrality metrics calculated on the different networks), for each centrality metric, they are all strongly positively correlated (Spearman’s ρ 0.82-0.99, all P-values <0.001)).

Table 6-5 shows the number of nodes that change rank from zero activity to (i) active (non-zero centrality metric value) and (ii) ranked in the top twenty percent for that centrality metric, when cattle data is added to the small ruminant data. When the cattle data is added to the small ruminant data, for both methods of link weighting (equal and small ruminant weighted) a few nodes do move from zero activity to active for each centrality measure (range 19-35 nodes (6-47% of zero activity nodes in small ruminant network)). A small number of nodes also move from zero activity into the top 20% of most central (high risk) nodes based on the betweenness centrality measure (4% in combined_E and 5% in the combined_SR network) and eigenvalue centrality measure (4% combined_E). It is important to identify nodes that move into the top 20% of high-risk nodes when cattle data are added to the small ruminant data to ensure these are included in control programmes if the aim is to control a potentially multi-host pathogen.

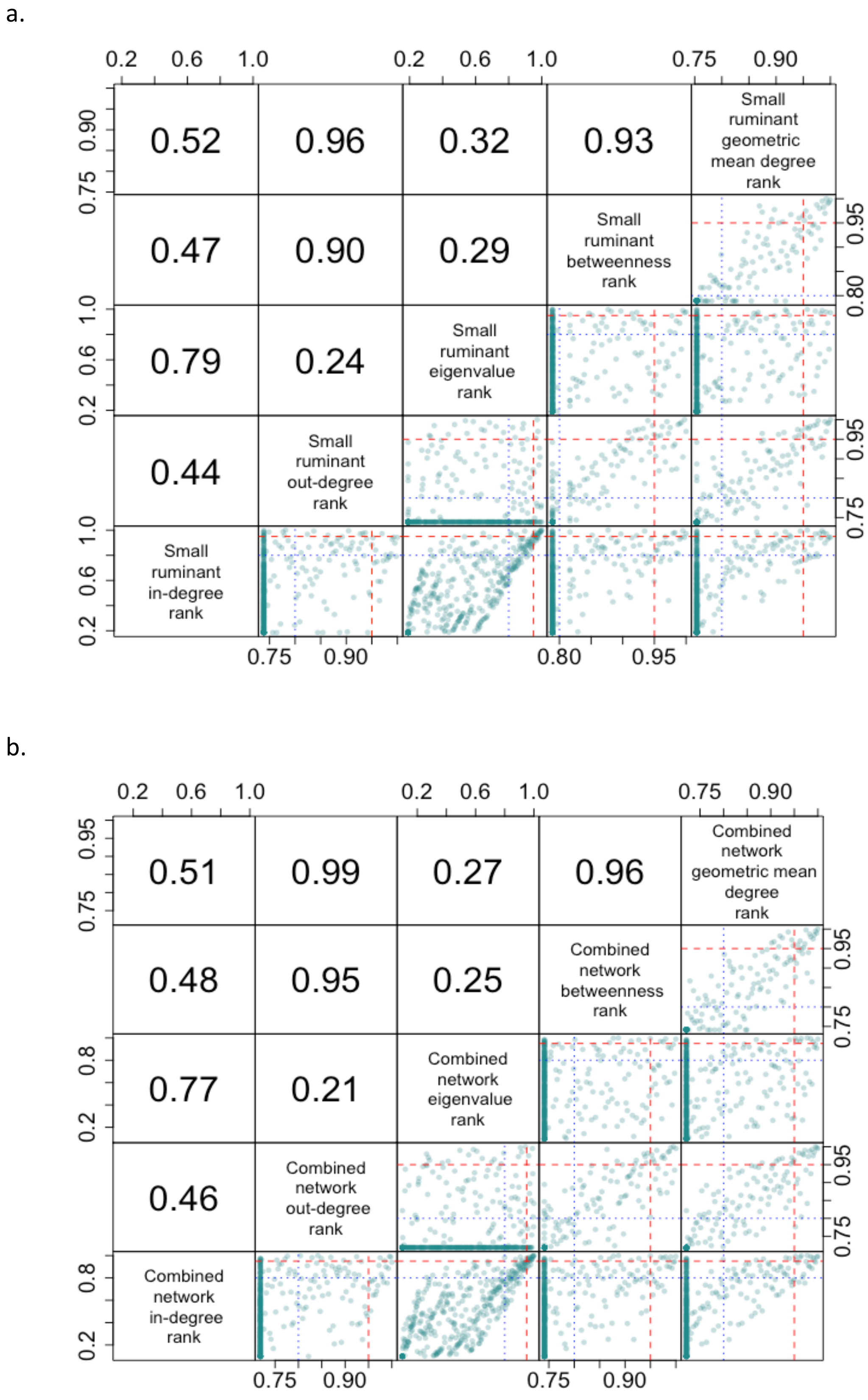
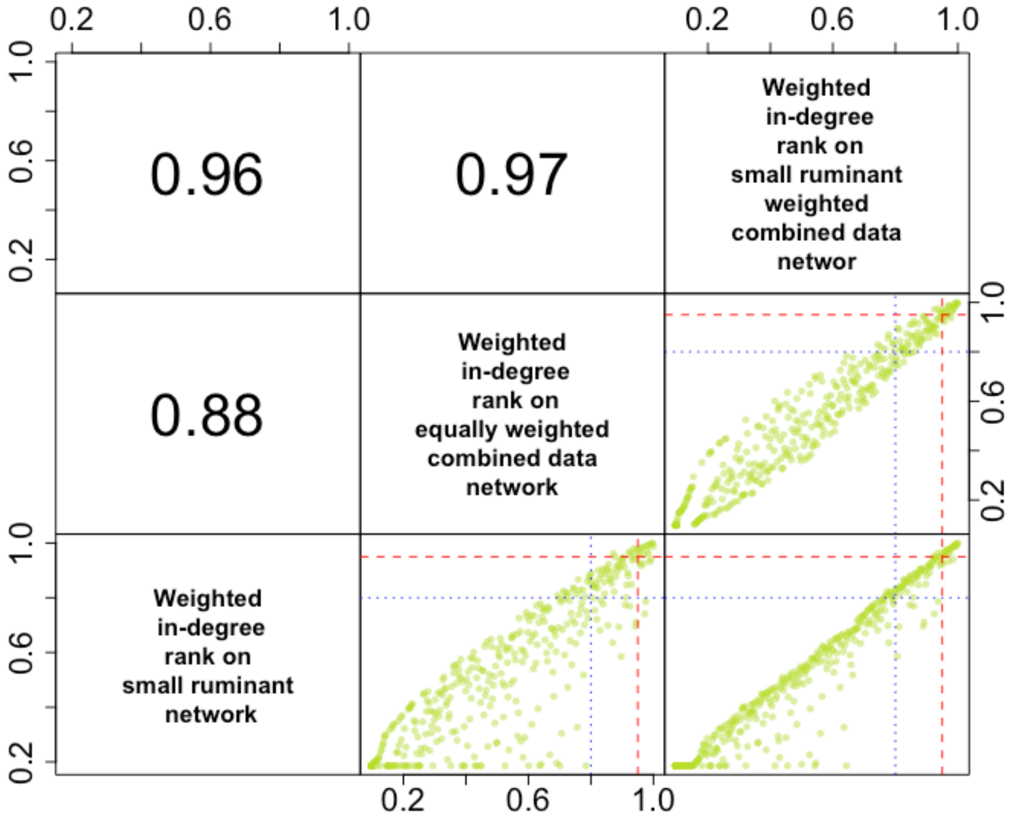
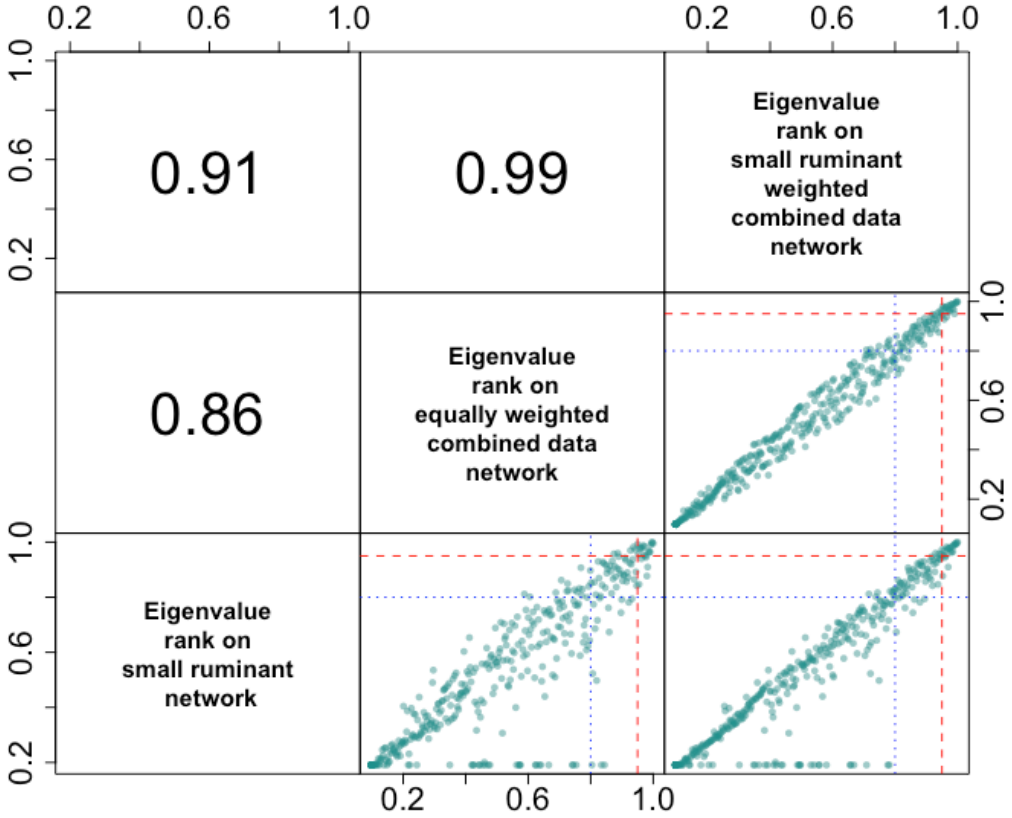


Figure 6-3: Spearman's ρ correlation coefficients (P<0.001) (top left) and correlation plots (bottom right) between five node centrality metric ranks; weighted in-degree, weighted out-degree, eigenvalue, betweeness and geometric mean degree calculated on the (a) small ruminant and (b) the combined cattle and small ruminant networks. Points to the right of the blue and red lines are ranked in the top 20% and 5% of nodes respectively for the centrality metric labelled above the panel. Points above the blue and red lines are ranked in the top 20% and 5% of nodes respectively, for the centrality metric labelled to the left of the panel.

a.


0.2
0.6
1.0
0.96
0.97
Weighted in-degree rank on small ruminant weighted combined data networ
0.88
Weighted in-degree rank on equally weighted combined data network
Weighted in-degree rank on small ruminant network

b.


0.2
0.6
1.0
0.91
0.99
Eigenvalue rank on small ruminant weighted combined data network
0.86
Eigenvalue rank on equally weighted combined data network
Eigenvalue rank on small ruminant network

c.

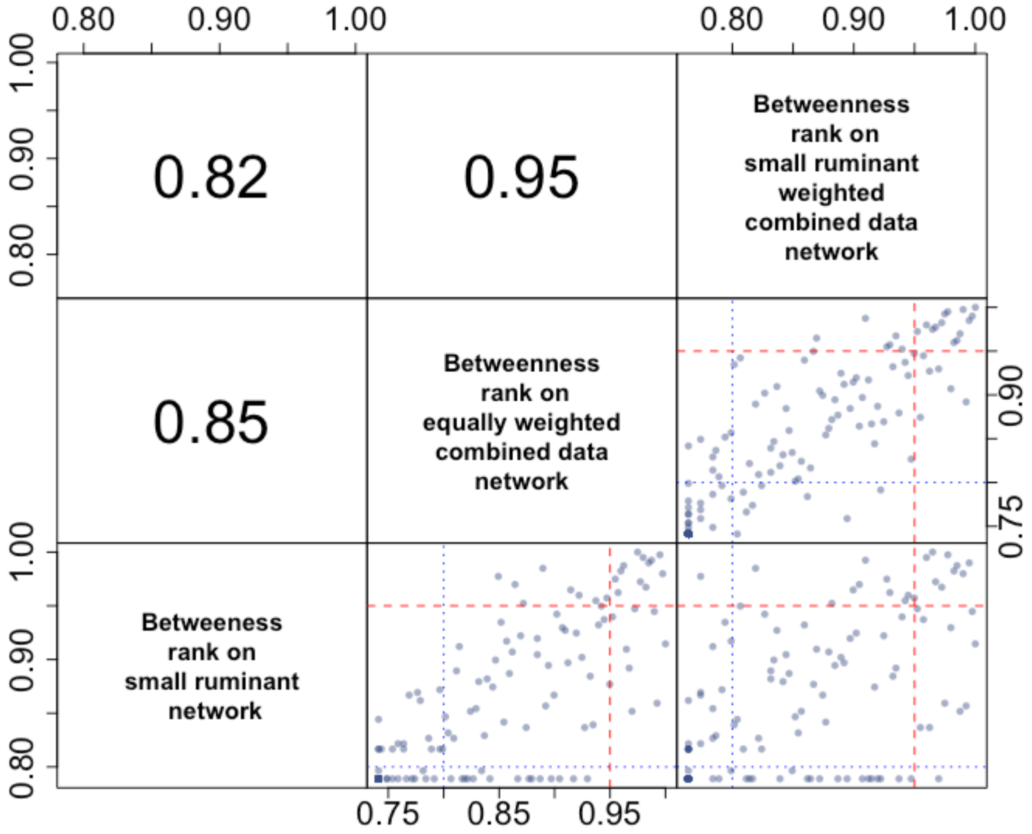


Figure 6-4 Spearman's ρ correlation coefficients (P<0.001) (top left) and correlation plots (bottom right) between three node centrality metrics: weighted in-degree (a), eigenvalue centrality (b) and betweenness centrality (c) on the small ruminant data network, the equally weighted combined cattle and small ruminant data network (combined_E) and the combined data network weighted heavily for small ruminants (0.9) and lightly for cattle (0.15)

Table 6-5: Number of nodes with zero centrality on the small ruminant network that have (i) some activity (non-zero centrality) and (ii) are ranked in the top twenty percent of high risk nodes for: betweenness, weighted in-degree and eigenvalue centrality, when cattle data is added to the small ruminant data on the Combined_E and Combined_SR networks. In the Combined_E network cattle and small ruminant links are equally weighted, in the Combined_SR network small ruminant links are heavily weighted (0.9) and cattle links lightly weighted (0.15).

		Non-zero activity	Top 20% of nodes
Betweenness	Combined _E	23 (7%)	14 (4%)
	Combined_SR	19 (6%)	17 (5%)
Weighted in-degree	Combined_E	35 (47%)	0
	Combined_SR	35 (47%)	0
Eigenvalue centrality	Combined_E	35 (47%)	3 (4%)
	Combined_SR	35 (47%)	0

6.3.7 Fragmentation

Results from targeted node removal analysis, based on nodes’ combined_E weighted indegree, betweenness and eigenvalue rank, to assess if the three networks, small ruminant, combined_E and combined_SR are vulnerable to fragmentation are shown in Table 6-6 (top 20% removed) and Table 6-7 (top 5% removed).

Targeted removal of nodes based on their betweenness centrality rank in the combined_E network resulted in the most efficient fragmentation of all networks. Removal of the top 20% of nodes based on their betweeness centrality resulted in complete fragmentation of all three networks strongly connected components. Removal of the top 5% of nodes ranked based on their betweenness centrality reduced the size of the GSCCs by 81%, 84% and 95% on the combined_SR, combined_E and small ruminant networks respectively.

Removing nodes based on them ranking in the top 5% for betweenness centrality on the small ruminant network did not improve the level of fragmentation achieved on the small ruminant network compared to removing nodes based on their rank on the combined_E network. The GSCC was still reduced to 4 nodes. If node removal is only possible for 5% of nodes then targeting nodes based on their combined_E rank is just as effective at fragmenting the small ruminant network compared to targeting nodes based on their rank in the small ruminant only network.

Table 6-6: Size of giant components and percentage reduction in size of giant components of combined and single species networks when nodes are removed based on belonging to the top 20% of nodes with highest in-degree, eigenvalue centrality and betweenness centrality on the combined data network with equally weighted edges.

	All data		-20% In-degree		-20% Eigenvalue		-20% Betweenness	
	GSCC	GWCC	GSCC (% reduction)	GWCC (% reduction)	GSCC (% reduction)	GWCC (% reduction)	GSCC (% reduction)	GWCC (% reduction)
Small ruminant	82	330	10 (97%)	184 (44%)	25 (92%)	217 (34%)	1 (100%)	21 (94%)
Combined_E	111	360	12 (89%)	224 (38%)	52 (53%)	239 (34%)	1 (100%)	33 (91%)
Combined_SR	111	360	15 (86%)	233 (35%)	49 (56%)	200 (44%)	1 (100%)	38 (89%)

Table 6-7: Size of giant components and percentage reduction in size of giant components of combined and single species networks when nodes are removed based on belonging to the top 5% of nodes with highest in-degree, eigenvalue centrality and betweenness centrality on the combined data network with equally weighted edges.

	All data		-5% In-degree		-5% Eigenvalue		-5% Betweenness	
	GSCC	GWCC	GSCC (% reduction)	GWCC (% reduction)	GSCC (% reduction)	GWCC (% reduction)	GSCC (% reduction)	GWCC (% reduction)
Small ruminant	82	330	51 (38%)	303 (8%)	48 (41%)	287 (13%)	4 (95%)	231 (30%)
Combined_E	111	360	64 (42%)	309 (14%)	83 (25%)	324 (10%)	18 (84%)	241 (33%)
Combined_SR	111	360	65 (41%)	318 (12%)	76 (32%)	319 (11%)	21 (81%)	274 (24%)

6.4 Discussion

The majority (76%) of livestock owning households in the northern Tanzania study area own both cattle and small ruminants. The high prevalence of multi-species ownership means that there is likely to be sufficient opportunity for between-species pathogen transmission events in the homestead, if one species is infectious with a multi-host pathogen. This finding supports our justification for looking at the small ruminant and cattle movement data together in combined networks because it is feasible that multi-host pathogens could be transmitted by either species group in the study area. In addition to homestead contact, extensive management of livestock in northern Tanzania provides opportunity for direct and indirect inter-species contacts between livestock from different households when they move to access grazing and watering points (Coppolillo, 2000; VanderWaal, Gilbertson, *et al.*, 2017; Herzog *et al.*, 2019).

Comparisons between the cattle, small ruminant and combined data (combined_E) movement networks in northern Tanzania show that there are more active wards in the small ruminant network compared to the cattle (SR = 330, C = 316) but the small ruminant network is less well connected with a smaller GSCC (SR = 82, C = 102). When the cattle and small ruminant data are combined the number of active wards increases to 373 and the network's giant connected components increase in size by 35% for the GSCC and 9% for the GWCC. The larger component sizes mean that many more wards are epidemiologically connected through the combination of cattle and small ruminant movements than by movements of a single species groups. This result shows that the mean estimated lower and upper bounds of epidemics caused by multi-host pathogens in this livestock population are significantly larger than what would be estimated for single-host pathogens. The combination of the movement data also increases the number of active links (routes available for pathogen transmission) in the network which shortens the average path length and diameter of the connected components, despite them including more nodes than the single species networks. This means that multi-species pathogens can transmit in fewer consecutive steps across the study area compared to single-species pathogens. The broader reach of multi-host pathogens, which can be achieved in fewer sequential movements, show that disease control interventions for potential multi-host pathogens in northern Tanzania should consider the contribution of both species groups.

All centrality metrics calculated on the small ruminant, combined_E and combined_SR networks exhibit strong right skew which shows that most wards are quiet with little livestock movement in and out and a few are very busy with livestock traffic. The right skew of all centrality distributions on the single and multi-species between ward movement networks indicates that they are going to be vulnerable to fragmentation through targeted node removal by implementing targeted pathogen control interventions.

Pairwise comparisons between five node centrality metrics: weighted in-degree, weighted out-degree, eigenvalue, geometric mean degree and betweenness, on the small ruminant and combined data (combined_E) networks show that they are all positively correlated. The strength of the positive correlations between all five centrality metrics mean that nodes at high risk of pathogen introduction are also those that are high risk for onward pathogen transmission and maintenance on the network. Out-degree, geometric mean degree and betweenness very highly correlated on both networks (Spearman's ρ >0.9, P-value <0.001) which suggests these metric ranks are picking up similar information about nodes risk of pathogen acquisition and onward transmission.

We checked to see how similarly nodes rank for weighted in-degree, eigenvalue and betweenness centrality, when the cattle data was added to the small ruminant data. Pairwise comparisons were made between node ranks on the small ruminant, combined_E (node links were weighted equally) and combined_SR (node links weighted more heavily for small ruminants (0.9) compared to cattle (0.15)) networks. The results show that despite adding cattle data and varying the relative contribution of each species group, node ranks show strong positive correlation (0.82-0.99, P-value <0.001). This shows that high risk nodes for disease introduction and onward transmission on the small ruminant only network remain high risk when the cattle data are added, and vice versa for low risk nodes, regardless of how heavily cattle links are weighted in comparison to small ruminant links. The most correlated metric on the three networks is weighted in-degree which shows that nodes at high risk of small-ruminant-only pathogen introduction are also high risk for multi-species pathogen introduction. The least correlated metric is betweenness which shows that adding the cattle data to the small ruminant data does slightly modify nodes' importance as gate-keepers, connecting other groups of nodes, on the network. The alterations to betweenness rank when the data are combined is important to document as nodes with high betweenness could be targets for movement restrictions or increased surveillance (Molia *et al.*, 2016).

Targeted removal of nodes based on their betweenness centrality was the most efficient way to fragment all networks. Removing the top 20% of nodes based on their betweenness centrality on the combined_E network resulted in complete fragmentation of all networks GSCC. Removing only the top 5% of nodes based on their betweenness centrality, again ranked on the combined_E network, reduced the size of the GSCC's by 81-95%. The level of fragmentation achieved on the small ruminant network was not improved upon by removing nodes that ranked in the top 5% of highest betweenness on the small ruminant network only. This result shows that if resources are limited to only being able to target interventions in 5% of wards then choosing nodes based on their combined betweenness rank is no better than choosing based on their small ruminant only rank, even for small ruminant only pathogens. Control interventions in the top 5% of high-risk wards, chosen based on their betweenness centrality rank can therefore target multi-species pathogens without compromising on efficiency of targeting small ruminant only pathogens. Targeting interventions at nodes that are high risk for multi-species pathogens will be more efficient long term compared to targeting nodes based on their risk rank for single species pathogen. The initial fixed cost of setting up a livestock pathogen intervention will be high but the long-term benefits of a foundation that can be built upon to target multi-host pathogens mean there can be greater public- and livestock-health returns from the investment. The results from this study identify the possibility of a cost-efficient and integrative approach to multi-host, multi-pathogen, livestock disease control interventions in northern Tanzania.

The detection of three highly connected communities on the small ruminant and combined data networks show that isolation of these communities could be achieved by banning movement between them but allowing movement within. For practical feasibility this type of restrictive intervention requires communities to be spatially clustered and also requires increased surveillance within the active communities to ensure pathogens are not introduced and allowed to circulate unchecked. In response to regional disease outbreaks of rapidly transmissible pathogens such as FMDV or RVF, movement restrictions between communities could provide a practical alternative to blanket movement bans. Blanket movement bans can be met with non-compliance due to the economic fragility of livestock trading and the abundant need for households to continue to generate an income (Pica-

Ciamarra *et al.*, 2011; Motta *et al.*, 2018; Haseeb *et al.*, 2019). The policing of movement restrictions between communities would require public financial resources to support a regulation enforcement team but there could be money made available to support this if some legal, taxable, trade continues. Compliance with community isolation can be helped by identifying the communities before a disease outbreak and discussing the practicalities of allowing movements to continue within but not between communities with all community stakeholders. Community isolation could prevent rapid pathogen dissemination across the study area before it is detected or reported and give local vaccination roll-out campaigns an opportunity to get ahead of the epidemic.

This study has limitations. The networks used in this study do not account for movements of less than one animal a year between wards. These are unlikely to change the overall network structure but may provide occasional connections and opportunity for pathogen transmission that remain unaccounted for. Similarly, the networks do not account for unrecorded inter-household movements in the study area and direct and indirect contacts between livestock from different wards. Unrecorded movements and intermittent contacts could provide opportunities for pathogen maintenance in the northern Tanzania livestock population (VanderWaal, Gilbertson, *et al.*, 2017). The extent to which undocumented movements and contacts might undermine targeted pathogen control interventions could be estimated by constructing multi-layer networks that include simulated inter-household movements and intermittent contacts at natural resources.

6.5 Conclusion

We used movement permit data to construct a between-ward contact network for small ruminants, cattle and both species groups combined in northern Tanzania. The networks can be used to identify wards at high risk of pathogen introduction (highest weighted in-degree ranks), or those at high risk of onward pathogen transmission (high betweenness and eigenvalue ranks) for single and multi-host pathogens. The results show that all networks are vulnerable to fragmentation through targeted removal of high-risk wards and this is most efficiently done by removing wards based on their betweenness centrality rank. When only the top 5% of nodes are removed the efficiency of fragmentation of the small ruminant network is not improved upon by selecting wards based on their betweenness rank in the small ruminant network vs in the combined data network. The results from this study show that targeted pathogen control interventions in high-risk wards in northern Tanzania could greatly reduce transmission of single and multi-host pathogens via trade-based livestock movements. Targeted interventions can be developed so that they are multi-host multi-pathogen focussed and reduce overall disease burden in the livestock, and in the case of zoonoses, human populations. Multi-host and multi-pathogen interventions in the highest risk locations can share the initial fixed costs of communication and implementation platforms and result in greater public and livestock health returns from initial investments.

7 Discussion

7.1 General

For this thesis market survey and government movement permit data were gathered, along with household survey data to quantify different types of livestock movements in Arusha, Manyara and Kilimanjaro Regions of northern Tanzania. The results show that inter-household, market-centred and trade-based movements of livestock create networks of epidemiological connectivity between livestock populations across the study area and beyond, including across the international border with Kenya. Movements in and out of Kenya, between households and markets, show that a harmonised, regional approach to disease control is required in the area.

Household survey data show that the majority of movements to and from households, including market movements, are unreported. These movements can cover long distances up to 300 km. This means that undocumented livestock movements (around 75%) have the potential to move pathogens around the study area, acting somewhat as Trojan livestock, undermining disease control programmes should they remain unaccounted for. There is a clear need for improvements to livestock movement reporting and recording so comprehensive data can be accessed in real-time and used to guide effective livestock disease control programmes (Green and Kao, 2007). This platform could be established, for example, by having movements reported using SMS as most households now have access to a mobile ‘phone at some level (Mibei *et al.*, 2017). In addition to ease of access, factors which may be necessary to improve movement reporting compliance include a reduction in movement associated taxes and increased awareness amongst livestock keepers as to why movement data is necessary. Livestock keepers should be made aware of how movement data can be utilised to improve livestock and public health through implementation of efficient disease control interventions.

The inter-household, market-centred and trade networks constructed from movement data all exhibit properties to suggest they are vulnerable to fragmentation through targeted removal of the most central nodes. Fragmentation is the aim of disease control programmes as this effectively breaks up a network into smaller components and removes routes for pathogen transmission between groups of nodes, reducing the extent of spread.

Examples of node removal in this instance include implementing disease control interventions to break pathogen transmission cycles in the most central nodes which are at highest risk of pathogen introduction and onward transmission. Because the majority of movements in the network models are market and trade movements it is likely that money and disease flow similarly on the networks (Xie, Horan and Wolf, 2009; Motta *et al.*, 2018). This factor needs careful consideration in the design of disease control interventions because for these to be successful they must aim to interrupt pathogen transmission but allow commerce to continue wherever possible. This is where vaccination to increase herd immunity prior to any large disease outbreaks can be utilised in high-risk locations as it can reduce the proportion of susceptible animals and reduce the duration of other restrictions such as movement bans.

On all constructed networks (inter-household, market and trade), there were positive correlations between all node centrality metrics. This means that locations identified as high risk for pathogen introduction were also likely to be high risk for onward transmission. On the inter-household movement network there were also positive correlations between weighted and non-weighted node metrics which shows that nodes at high risk of acquiring and transmitting rapidly transmissible pathogens are also high risk for less transmissible pathogens. These positive correlations make targeting disease control interventions more efficient as the fixed costs of logistically setting up intervention platforms are minimised because fewer locations need to be targeted. Interventions can include: vaccination in high risk locations with large resident livestock populations; test and quarantine facilities; and knowledge exchange to improve livestock keeper awareness of disease risks posed to their own animals and others through livestock introductions and movements. On the trade networks constructed from movement permit data, the high risk nodes for single-host pathogens are also high risk for multi-host pathogens. This means that generally, across the study area livestock and zoonotic disease control interventions can be multifaceted, aiming to target multiple pathogens and multi-host pathogens to reduce the risk of both pathogen introduction and propagation. A multifaceted, targeted approach to pathogen control in the northern Tanzanian livestock population is going to be an efficient way to spend limited disease control resources with maximal gains for livestock keepers and livestock dependant societies (Randolph *et al.*, 2007).

7.2 Household movements and disease risk

Through use of household survey data this study shows that largely undocumented household livestock movements create epidemiological connections between geographically close and distant locations, including across the international border with Kenya. The majority of reported movements are to and from markets but gifting and private sale movements create additional epidemiological links between sub-villages and other locations which provide additional routes for pathogen transmission. Similar epidemiological connectivity through inter-household movements has recently been documented in neighbouring Kenya (Omondi *et al.*, 2021). In combination, the inter-household movements identified in this study create a well-connected network despite sampling from only a small proportion of villages and sub-villages (around 2%). If disease control programmes were to only target livestock markets, and other between-household movements such as private sales and gifting continue unchecked, these could provide sufficient opportunities for sustained pathogen transmission (Prentice *et al.*, 2017). There are of course potentially other types of between-household movements such as the shared use of breeding stock or draught animals which could provide further opportunities for pathogen transmission. Increased awareness of the disease risks associated with livestock introductions and contacts amongst livestock keepers (to their own animals, other livestock locally and in the case of zoonoses, themselves), could increase livestock keeper willingness to engage in movement reporting. Increased awareness of the risks associated with introductions and contacts also allows livestock keepers the opportunity to make informed decisions about management factors such as the use of shared grazing and or developing a quarantine area.

This study tests the hypothesis that more household and sub-village introductions will increase pathogen exposure risk for multiple livestock and zoonotic pathogens. This hypothesis embeds an underlying assumption about the risk of introduction and the risk of onward pathogen spread within households and sub-villages following an introduction. The results of the associations between household and sub-village introductions and exposure to production-limiting and zoonotic pathogens tested are complex but summarised in brief schematic form in Figure 7-1. The relationship between introductions and exposure varies depending on the agro-ecological (A-E) setting where livestock are kept and also between pathogens that utilise different transmission routes and have various shedding cycles. This is perhaps not a surprising finding as we know that population structure, mixing and contact dynamics, all of which vary amongst the A-E classes, have significant influence on the R_0 of a pathogen and its ability to both persist and cause epidemics (Keeling and Grenfell, 2000; Kao, 2010). It is also not surprising that within A-E classes the results vary between pathogens because the pathogens tested operate on different time-scales and utilise different transmission routes (Kao *et al.*, 2007), and livestock movements and herd contacts vary widely across different systems (de Glanville *et al.*, 2020). For example, in smallholder settings livestock are tethered and stocking density and human density are high whereas in pastoral settings livestock are extensively grazed and mix with other herds and flocks. Agro-pastoral livestock are managed somewhere between the two with livestock also grazed on croplands following harvest.

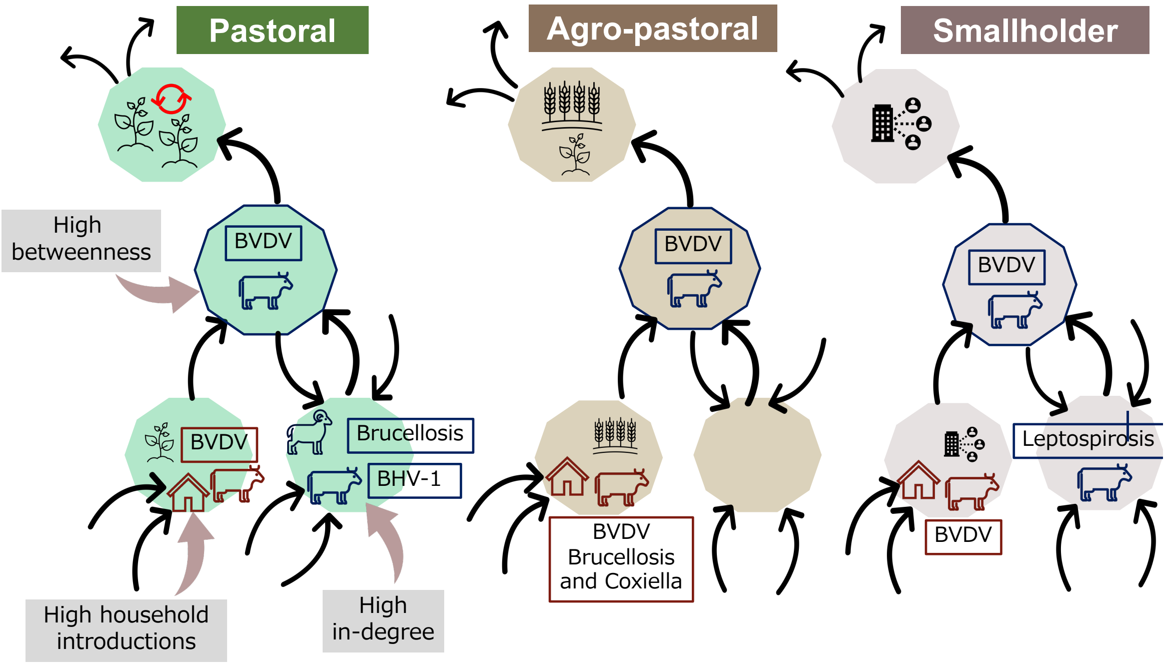


Figure 7-1: Schematic diagram depicting increased risk of pathogen seropositivity (through listing disease in text box alongside graphic of species group affected) and how it is associated with household and sub-village livestock introductions in cattle and small ruminants across three agro-ecological classes: pastoral, agro-pastoral and smallholder. The decagons represent sub-villages with a mock network created for each agro-ecological class (top row depict some characteristics of the A-E class: pastoral have high mixing (red circular arrow) between herds and extensive grazing, agro-pastoral predominantly cropland grazing with some extensive grazing, smallholder have high population density (livestock and people) to depict example sub-villages with high betweenness centrality (middle row – lying on shortest path between most other node pairs), high in-degree (bottom right sub-village for each A-E class) and a household with high numbers of introductions (depicted by a house icon contained within bottom left sub-village for each A-E class). Note high in-degree in agro-pastoral sub-villages is not associated with seroprevalence for any pathogens tested in this work.

Despite the complexities identified using these data, household cattle introductions and sub-village betweenness (calculated on the inter-household movement network from livestock movements reported in and out of sub-village households) are positively associated with exposure to the rapidly transmissible BVDV pathogen in cattle and this association is consistent across A-E classes. Household introductions are also positively associated with exposure to *Brucella* spp. and *C. burnetii* in cattle from agro-pastoral households. Sub-village in-degree is positively associated with *Brucella* spp. exposure in pastoral small ruminants, BHV-1 in pastoral cattle and *Leptospira* serovar Hardjo in smallholder cattle. One example of a factor that may drive the difference in the relationships between pathogen introduction and onward transmission for *Leptospira* serovar Hardjo in smallholder cattle is that the density of other reservoir hosts of *Leptospira* serovar Hardjo could be higher in smallholder areas.

Differences in the way herds are managed and grazed between A-E classes could explain why there is an association between cattle introductions and exposure to the zoonotic pathogens *Brucella* spp. and *C. burnetii* in agro-pastoral households. Agro-pastoral herds are managed as more defined units compared to pastoral herds, and partially grazed on small areas of privately managed cropland following harvesting. This could lead to a higher frequency of within herd contacts vs between herd contacts compared to mixing patterns experienced by smallholder (tethered, not free roaming) and pastoral (extensively grazed on shared pasture) cattle. Additionally, environmental contamination with zoonotic pathogens, by infectious animals from within the herd, of land grazed only by that household herd, could provide further opportunity for within-herd transmission in agro-pastoral cattle. Pastoral herds on the other hand, are managed more extensively with the concept of a herd perhaps more nebulous with more frequent opportunities for between herd contacts and transmission events (VanderWaal, Gilbertson, *et al.*, 2017). Smallholder herds are smaller so will inherently experience smaller within herd outbreaks and because animals are tethered in place there will be less close contact mixing between infectious and susceptible animals. The differences in mixing patterns are likely to make the risks of within herd transmission of pathogens that are shed intermittently or cyclically (after parturition) and through environmental contamination less important for pastoral cattle.

In the case of BVDV, where there is an association between livestock introductions and seropositivity at both the household and sub-village level, interventions that target high-risk households and sub-villages could be implemented. An example of this could be targeting vaccination campaigns in sub-villages with high betweenness, which in this case are most likely to be pastoral sub-villages, to break chains of transmission across the study area. Another example would be provision of test and quarantine facilities so livestock keepers can reduce the risk of introducing pathogens to their own herd and the village herd. At the household level interventions could include those which aim to improve livestock keeper knowledge of rapidly transmissible pathogens such as BVDV to increase risk awareness. The benefit of preventing introduction of pathogens such as BVDV would be felt by the whole sub-village across all A-E classes.

For *Leptospira* serovar Hardjo in smallholder cattle, BHV-1 in pastoral cattle and *Brucella* spp. in pastoral small ruminants the benefits of preventing pathogen introduction are also shared across households throughout the sub-villages. These results support the concept that sub-village level interventions such as vaccination and tests and quarantine facilities should be publicly or collectively funded because they provide a public good, protecting livestock beyond those in the household that engages in the risk behaviour (Rich and Perry, 2011; Maziku, Gebru and Stapleton, 2017). Testing of individual animals introduced into the quarantine facilities could potentially be funded by the individual livestock keeper so they bear some financial responsibility for the cost of mitigating disease risk, based on the protective effect this will have on their herd.

The complexities identified in the relationships between livestock introductions, A-E systems and BHV-1 and the zoonotic pathogens show a disease control programme targeting one pathogen uniformly across the area is not an efficient use of resources. These results show that to reduce livestock and zoonotic pathogen burden for pathogens with variable transmission pathways (different shedding cycles, utilising different transmission routes), in an area with complex between-herd mixing patterns, a more targeted approach to disease control is needed. For zoonotic pathogens, multifaceted disease control interventions which aim to improve livestock keeper awareness, provide opportunities for disease risk mitigation, vaccine access and improved veterinary services are widely advocated (Maudlin, Eisler and Welburn, 2009; Rich and Perry, 2011). The results in this thesis show that multi-pathogen control programmes can be targeted at high-risk locations in northern Tanzania that are identified using the SEEDZ household survey data. For maximal benefits interventions should involve components such as knowledge exchange with livestock keepers as well as improved access to veterinary services including vaccinations and test and quarantine facilities. Knowledge exchange programmes should aim to increase livestock keeper awareness of the risks of pathogen introduction and exposure posed by livestock movements, direct and indirect contacts and introductions. Once armed with this knowledge, livestock keepers can make informed decisions about livestock mixing and introductions and create their own quarantine areas.

The movements and mixing patterns of livestock within the sub-village, and how these are associated with pathogen transmission, have not been described in this thesis. Given the seroprevalence variations identified between A-E classes, it is important to quantify the differences in dynamic intra- and inter-herd and flock mixing patterns within villages belonging to the different A-E classes through further investigation.

A limitation of this study is that the relationship between pathogen exposure and livestock introductions is only investigated at the household and sub-village level. These data could also be explored further to evaluate if the effects of belonging to a village or ward with high betweenness are also associated with high risk of exposure to BVDV so interventions could be rolled out at potentially more practical, larger administrative scales.

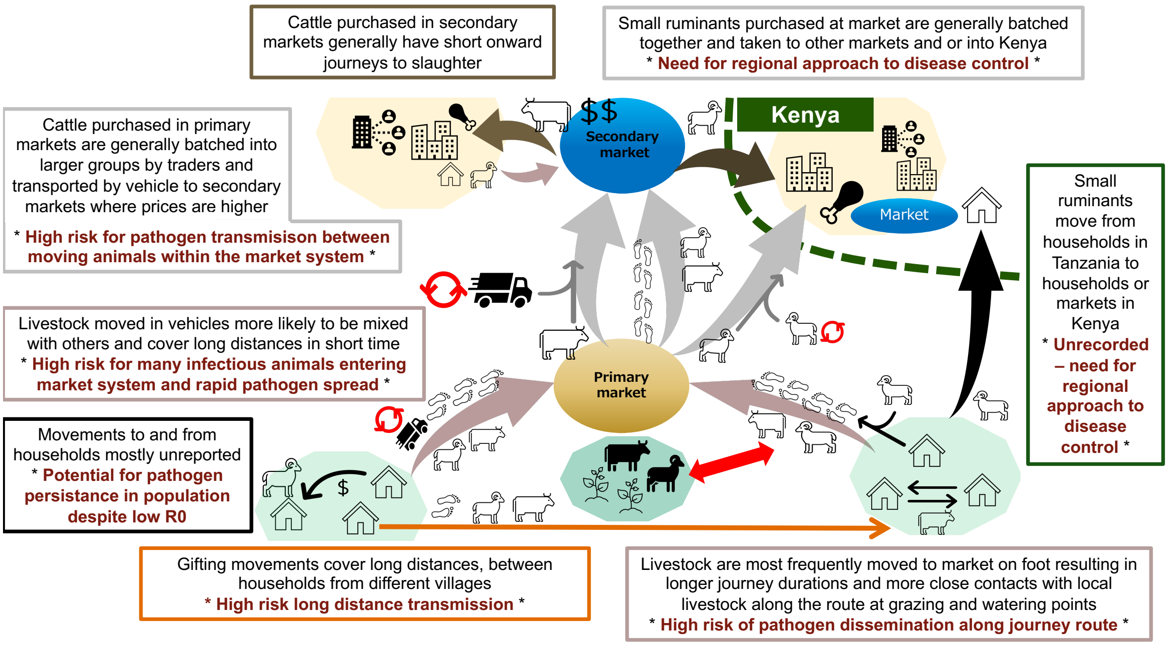
Trialling interventions in sub-villages identified as high risk in this data would be a good place to start to assess if this has any rapid impact on overall livestock stability and productivity. Another important factor to consider here is that the introduction risk factors identified as important for exposure to BVDV may also be similar for other pathogens with similar shedding cycles and transmission routes. The serological samples used in this study could be further tested to evaluate if exposure to pathogens such as Pestes des petits ruminants and foot and mouth disease virus follow similar patterns to BVDV, and thus if vaccination against these pathogens should also be targeted in locations with high betweenness (Kerfua *et al.*, 2018; Herzog *et al.*, 2019).

Disease simulations similar to those done on the trade network were not done on the inter-household and market networks because, despite static network properties being robust to sampling error, epidemic simulations are strongly affected by sub-sampling a network (Génois *et al.*, 2015). The data collected and used in this thesis could be extrapolated from and used to estimate the complete underlying inter-household movement network, which could then be used to run disease simulations and interventions. This was beyond the scope of this thesis but can be done in future work.

7.3 Market movements

Market movements were the most commonly reported type of movement in and out of surveyed households and these are largely un-documented (around 75%). For future traceability and improved disease control this result supports the need for a new system for livestock keepers to easily register livestock movements. Market survey data show that market movements also create long distance epidemiological connections (up to 300 km) across northern Tanzania, including across the international border into Kenya. Consistent with the results from the household survey data this result confirms the need to develop regional disease control programmes.

The majority of market journeys are made on foot, with livestock moved this way spending longer travelling and more likely to come into close contact with local livestock. A schematic diagram showing an overview of livestock market movements and the disease risks associated with these is given in Figure 7-2. Livestock moved on foot are also at greater risk of indirect contacts with local livestock at communal grazing and watering points. Indirect contacts could provide opportunity for transmission of pathogens that can be spread through environmental contamination or through infection of local vectors such as mosquitos and ticks. Large numbers of on-foot market journeys across northern Tanzania are therefore a risk factor for pathogen dissemination throughout the study area. These journeys create additional livestock disease transmission risks beyond those created by livestock moving between two distinct origin and destination locations. Animals that are being moved are also likely to be stressed and thus at increased risk of pathogen shedding (Beach, Murano and Acuff, 2002). Additionally the most social animals within herds are most likely to be the ones to contact livestock from outside the herd, creating opportunities for super-spreading events within and between herds (Böhm, Hutchings and White, 2009; Biswas *et al.*, 2013; Lambert *et al.*, 2018).



Cattle purchased in secondary markets generally have short onward journeys to slaughter

Small ruminants purchased at market are generally batched together and taken to other markets and or into Kenya
* **Need for regional approach to disease control** *

Cattle purchased in primary markets are generally batched into larger groups by traders and transported by vehicle to secondary markets where prices are higher
* **High risk for pathogen transmisison between moving animals within the market system** *

Livestock moved in vehicles more likely to be mixed with others and cover long distances in short time
* **High risk for many infectious animals entering market system and rapid pathogen spread** *

Movements to and from households mostly unreported
* **Potential for pathogen persistance in population despite low R0**

Small ruminants move from households in Tanzania to households or markets in Kenya
* **Unrecorded – need for regional approach to disease control** *

Gifting movements cover long distances, between households from different villages
* **High risk long distance transmission** *

Livestock are most frequently moved to market on foot resulting in longer journey durations and more close contacts with local livestock along the route at grazing and watering points
* **High risk of pathogen dissemination along journey route** *

Figure 7-2: Schematic diagram showing summary of results from market survey data and how these might associate with different factors regarding pathogen transmission and dissemination. Black arrows represent household to household, orange arrow represents gift movement, pink arrows represent household to market movement, grey arrows represent between market movements, dark green arrows represent market to final destination. Red circular arrows depict mixing along journeys, red straight arrow represents contact between livestock moving on foot and local livestock along journey. Dollar signs represent higher cattle price at secondary markets. Green dashed line is international border with Kenya. Pale green decagons are rural villages, pale yellow decagons are urbanised areas

The risk of moving livestock transmitting pathogens such as RVFV and CCHFV to local vectors is particularly important to consider if moving livestock originate from areas with high pathogen prevalence and pass through low risk areas where surveillance and awareness are limited (Aslam *et al.*, 2016; Kim *et al.*, 2021). Markets identified as most central in the study area should be a priority for implementing active surveillance. The positive association identified between secondary markets and the centrality metrics in-degree and betweenness, used to identify locations for targeted interventions, show that secondary markets beyond the study area would also be efficient locations to implement surveillance. If funds and resources were available, additional surveillance sites could be established along well-travelled market routes. This could identify a proportion of infectious animals and associated outbreaks in areas in need of disease control interventions and temporary movement restrictions before mass pathogen dissemination through market mixing and onward movements (Kao *et al.*, 2006). Additionally villages along the busy market routes could be included in vaccination and knowledge exchange disease control campaigns if they are deemed at high risk of pathogen exposure due to contacts with the moving livestock (Dean *et al.*, 2013; Jahel *et al.*, 2020). Stronger enforcement of the 2000 legislation stating that all market journeys should be made in vehicles, while potentially reducing transmission risks, would have the major drawback of further excluding poorer livestock keepers.

Livestock moved in vehicles to and from markets have shorter journey durations but are at a higher risk of being batched together with multiple other livestock. If any transported animals were infectious at the start of the market journey, the increased intensity of mixing within vehicles, stress and associated risk of pathogen shedding, mean that livestock entering the market system following a vehicle journey are more likely to be infected with a pathogen. This might not be a major risk factor for transmission of production-limiting livestock pathogens because most livestock entering the market system are destined for slaughter. A small proportion of livestock purchased in markets are however taken home. Buyers should be made aware of where the livestock they wish to purchase have originated and the intensity of mixing the animals have experienced on their market journey. Access to this type of information could help livestock buyers decide which animals to purchase and/or if a quarantine period in the home village is necessary before introducing purchased animals to their own herd/flock.

The risks of zoonotic disease transmission through intensive mixing of livestock along journeys to people at slaughter points and butcheries is more of a pressing issue. This is especially important for diseases such as Coxiellosis, RVF and CCHF which can infect people directly through contact with livestock blood putting slaughterhouse workers at high risk of exposure (Ergönül, 2006; Aslam *et al.*, 2016; Woldeyohannes *et al.*, 2018; Métras *et al.*, 2020). Market survey and movement permit data show that livestock generally move from rural areas of production, through the market system in multiple journeys along a price gradient, to urban areas of consumption. This finding is in line with other studies in LMICs and warrants the need for transparent traceability of livestock entering the food chain in Tanzania, from point of origin to slaughter to improve disease control and early responses to disease outbreaks (Ntokwane and Dibeela, 2016; Mutua *et al.*, 2018; Napp *et al.*, 2018). This will allow active surveillance in appropriate high risk locations in the face of disease outbreaks as well as early warning surveillance. It will also allow those working in slaughter points and butcheries to be aware of, and act to mitigate, their risks of contracting zoonotic diseases that might usually be associated with infections in distant rural locations.

7.4 Trade movements

The cattle, small ruminant and combined species networks created using movement permit data show a high level of epidemiological connectedness across northern Tanzania through trade-based movements of livestock. Comparison between the networks show that multi-host pathogens have a wider reach in the study area than single host pathogens. Using node centrality metrics, heterogeneity in movement activity is identified between wards on all networks and identify high risk wards that can be made targets for disease control interventions. Through disease simulations on the cattle network, it is shown that implementing interventions in only 5% of nodes identified as high risk based on having high betweeness, eigenvalue and degree centrality, pathogen spread on the network can be reduced by up to 50%. Ward locations identified as high risk for single-host pathogen introduction and onward transmission are also shown to be high risk for multi-host (cattle and small ruminant) pathogens. This finding can be used to encourage policymakers to invest in livestock disease control programmes as these can be initiated in a small proportion of high risk wards and target multiple multi-host pathogens, for efficient use of resources. In the highest risk wards active surveillance can be implemented in markets and used to identify early warning signs of disease outbreaks across the regions. Multi-valent vaccination programmes can also be implemented in the high-risk wards, especially those with large resident livestock populations, as local livestock are at high risk of pathogen exposure due to large volumes of livestock traffic.

Disease simulation models on the cattle trade network showed that movement bans (lasting 11 months) across all wards had the largest effect on reducing pathogen spread, however the socio-economic shocks associated with these restrictions would be large and compliance levels are likely to be low (Sindato, Karimuribo and Mboera, 2012). Movement controls must be carefully matched to epidemiological and economic consequences and well as to livestock keepers perception of what is necessary (Tildesley *et al.*, 2019). To avoid broad scale interventions that are difficult to regulate due to the extensive management of livestock across northern Tanzania, other preventative interventions such as vaccination of livestock in high risk wards can be implemented prior to outbreaks (Coppolillo, 2000; de Glanville *et al.*, 2020). If livestock traceability is improved, for example through development of a quick, cheap and easy to use movement reporting tool (e.g. accessed for free via SMS), surveillance in secondary markets, which sample livestock from the widest range of locations, can be used as early warning for disease outbreaks. Active surveillance could involve thorough clinical examination of livestock and also the use of rapid diagnostic tests either in random samples of traded livestock or in any animal considered potentially unwell. For active surveillance to work efficiently this would require the centralisation of all movement data daily and reports on the health of moving livestock to also be centralised along with movement data so that any deviations from normal can be rapidly recognised. Rapid vaccinations along with temporary movement restrictions can be rolled out in high-risk locations on the network and also in locations where the infected animals originated to stem transmission before pathogens become widespread in the population.

Analysis of the market survey data also shows that the primary driver for livestock movements into the market system is to generate household income to meet household needs. Once within the livestock system animals are traded and moved between markets along a price gradient by livestock traders. The financial needs of livestock keepers and traders will still exist in the face of disease outbreaks and these need to be carefully considered and integrated into the design of any disease control intervention (Sindato, Karimuribo and Mboera, 2012). Blanket movement bans are shown to result in huge economic shock for societies that depend on livestock sales as a source of income (Sindato, Karimuribo and Mboera, 2012). Of course, in some instances, such as regional RVF outbreaks, there may be a need to ban livestock movements but the more that can be done to limit the duration and extent of these bans the better due to their negative socio-economic consequences (Knight-Jones and Rushton, 2013; Tildesley *et al.*, 2019).

If interventions such as vaccination and quarantine can be implemented along with others to reduce livestock mixing this could allow valuable trade movements to continue safely, for longer, while mitigating pathogen transmission risk through the population. Examples of mitigation measures include things already discussed such as vaccination of susceptible livestock in high-risk locations. Large vehicles could also be chartered by buyers and sent to primary markets where livestock can be purchased and taken directly to different slaughter points to avoid long journeys on foot and mixing with other livestock in secondary markets. Infrastructure changes could include development of slaughter points in some rural areas closer to livestock origins with butcheries attached and robust cold chains established. This would mean livestock would have shorter journeys from origin to slaughter and meat could be prepared safely for consumption locally and moved to towns where the demand is highest to supply a market currently met by largely imported meat products (Nandonde, Gebru and Stapleton, 2017). To develop these interventions to reduce livestock pathogen transmission large financial investments are needed. The gains from investments however will be multifactorial with livestock keepers able to access better prices (Pica-Ciamarra *et al.*, 2011; Motta *et al.*, 2019), and societies as a whole experiencing reduced pathogen burden through reduced transmission and dissemination. If livestock keepers were able to access better prices for the animals they sell this might also incentivise them to manage herds in a way that optimized offtake rather than herd size which might be a more sustainable approach in the face of declining access to grazing lands . Livestock keepers will also have more cash capital to spend on disease prevention which will further benefit communities through improved livestock health and productivity and thus nutritional and social security.

The data generated in this thesis show multiple types of livestock movement create epidemiological connections with locations in Kenya. In other studies from LMICs livestock traders were also found to be acting in multiple countries (Dean *et al.*, 2013; Motta *et al.*, 2017). Multiple small ruminant traders were spoken to who openly reported that they were taking livestock across the border into Kenya where better prices were available. Cross border movements are generally not reported as the taxes associated with these are large. From a disease control perspective, given that cross-border movements happen anyway without taxes being paid, these findings show that import and export taxes probably need to be reduced to improve movement reporting compliance and thus pathogen tracing. Epidemiological and economic analysis could be done to address what the potential impacts of movement tax reduction will be on movement reporting compliance, the potential change in numbers of transboundary movements and consequent effects on disease dynamics. The economic analysis could also address what the expected changes in government revenue might be if movement taxes were reduced and the socio-economic gains that could be made through better livestock disease control if movement data were more comprehensive.

7.5 Summary

This is the first detailed descriptive and statistical study of permanent livestock movements in Tanzania. There are multiple layers of epidemiological connectivity across northern

Tanzania created through various types of permanent livestock movements between households, markets and slaughter points. The results presented here provide a clear motivation for digital collection and centralisation of livestock movement data in Tanzania and this could begin with immediate effect. Development of a national digitized, centralised reporting system that is quick, cheap and easy to use is both necessary and achievable. Similar studies in other countries across Africa, where movement data are lacking, have shown similar results but not yet acted to rectify this (Tempia *et al.*, 2010; Dean *et al.*, 2013; Motta *et al.*, 2017). Tanzania now has the opportunity to pave the way for other countries by implementing such a system and then utilising the data generated to move towards improved livestock health and productivity.

Livestock keepers need to be aware of the overarching livestock and societal health benefits that can be achieved as a result of movement data collection, and need to be actively engaged in the data collection design and implementation process. A top down approach is unlikely to be successful in this scenario given the current level of non-compliance with movement reporting through movement permit use. Real-time routine data collection will clarify which locations to target with active surveillance and other strategic disease control interventions such as vaccination or short term movement restrictions (Motta *et al.*, 2017). Real-time movement data will also allow identification of any changes in movement activity throughout the year and changes that occur in response to disease outbreaks which may alter which locations risk level and where control interventions need to be implemented (Funk, Salathé and Jansen, 2010; Chowell and Nishiura, 2015; Enright and Kao, 2018).

The results from my study also clearly show that locations at high risk for single-host pathogens are also high risk for multi-host pathogens. Disease control interventions such as surveillance, test and quarantine and vaccination, which target multiple multi-host pathogens, can therefore be efficiently implemented in high-risk locations. Interventions such as these, which go beyond routine data collection, will require more funding to establish effective platforms. It is however widely recognised that improved veterinary services and livestock health have improved overall public health embedded within them (Upton, 2004; Perry and Grace, 2009; Forman *et al.*, 2012). In Tanzania specifically there remains a profound need for access to transparent and effective livestock health services to allow livestock keepers living in difficult circumstances the opportunity to pursue better livestock health (Virhia, 2019). The results from my study can therefore be considered by Government, NGOs or doners to guide where improved veterinary services will have the largest impact on societal wellbeing by mitigating disease risk and burden. Multifaceted disease control interventions in high-risk sub-villages or wards in general are going to be more beneficial to communities and a more efficient use of resource compared to targeting single pathogens such as Brucella or PPRV (Fournié, Waret-Szkuta, Camacho, Laike M. Yigezu, *et al.*, 2018; GALVmed, 2018). Another relatively tangible intervention would be knowledge dissemination to increase livestock keepers' awareness of pathogen transmission routes and the risk posed by livestock introductions and contacts. Increased knowledge and awareness will not necessarily change livestock keepers' behaviour but it would at least allow them opportunity to make informed decisions about livestock movements that potentially put them, their own livestock and others at risk of disease (Hidano, Gates and Enticott, 2019).

The positive externalities yielded as a result of reduced livestock and zoonotic pathogen burden can be expected to benefit society as a whole. Enhanced livestock health and productivity can lead to improved human health and wellbeing through increased nutritional and socio-economic security and reduced disease burden. The results of this thesis show this is achievable in northern Tanzania through investment and introduction of digitised livestock movement recording and targeted multi-pathogen disease control interventions in a small number of high-risk locations.

8 Bibliography

Aklilu, H. A. *et al.* (2007) ‘Village poultry consumption and marketing in relation to gender, religious festivals and market access’, *Tropical Animal Health and Production*, 39(3), pp. 165–177. doi: 10.1007/s11250-007-9002-8.

Aklilu, Y. (2008) *Livestock Marketing in Kenya and Ethiopia: A Review of Policies and Practice*, *Policy*. Feinstein International Center, Addis Ababa.

Aklilu, Y., Irungu, P. and Reda, A. (2002) *An Audit of the Livestock Marketing Status in Kenya, Ethiopia and Sudan*, *Community-Based Animal Health and Participatory Epidemiology Unit Pan African Programme for the Control of Epizootics Organization of African Unity/Interafrican Bureau for Animal Resources*.

Aktipis, C. A., Cronk, L. and de Aguiar, R. (2011) ‘Risk-Pooling and Herd Survival: An Agent-Based Model of a Maasai Gift-Giving System’, *Human Ecology*. Springer, 39(2), pp. 131–140. doi: 10.1007/s10745-010-9364-9.

Albert, R., Jeong, H. and Barabási, A. L. (2000) ‘Error and attack tolerance of complex networks’, *Nature*. Nature Publishing Group, 406(6794), pp. 378–382. doi: 10.1038/35019019.

Allan, K. J. *et al.* (2015) ‘Epidemiology of Leptospirosis in Africa: A Systematic Review of a Neglected Zoonosis and a Paradigm for “One Health” in Africa’, *PLOS Neglected Tropical Diseases*. Edited by J. Zinsstag. Public Library of Science, 9(9), p. e0003899. doi: 10.1371/journal.pntd.0003899.

Allan, K. J. (2016) *Leptospirosis in northern Tanzania: investigating the role of rodents and ruminant livestock in a neglected public health problem*. University of Glasgow.

Allan, K. J. *et al.* (2018) ‘Assessment of animal hosts of pathogenic Leptospira in northern Tanzania’, *PLOS Neglected Tropical Diseases*. Edited by J. Foley. Public Library of Science, 12(6), p. e0006444. doi: 10.1371/journal.pntd.0006444.

Allen, L. H. (2003) ‘Interventions for Micronutrient Deficiency Control in Developing Countries: Past, Present and Future’, *The Journal of Nutrition*, 133(11), pp. 3875S-3878S. doi: 10.1093/jn/133.11.3875S.

Alonso, S. *et al.* (2016) ‘Prevalence of tuberculosis, brucellosis and trypanosomiasis in cattle in Tanzania: A systematic review and meta-analysis’, *Animal Health Research Reviews*. Cambridge University Press, pp. 16–27. doi: 10.1017/S146625231600013X.

Anderson, I. (2002) *Foot and Mouth Disease 2001: Lessons to be Learned Inquiry*. London.

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

Angelakis, E. and Raoult, D. (2010) ‘Q fever’, *Veterinary Microbiology*, 140(3–4), pp. 297–309. doi: 10.1016/j.vetmic.2009.07.016.

Apolloni, A. *et al.* (2018) ‘Towards the description of livestock mobility in Sahelian Africa: Some results from a survey in Mauritania’, *PLOS ONE*. Edited by I. Sendiña-Nadal. Public Library of Science, 13(1), p. e0191565. doi: 10.1371/journal.pone.0191565.

De Arruda, G. F. *et al.* (2014) ‘Role of centrality for the identification of influential spreaders in complex networks’, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 90(3), p. 032812. doi: 10.1103/PhysRevE.90.032812.

Ashworth, S. W. and Mainland, D. D. (1995) ‘The Economic Impact of BSE on the UK Beef Industry’, *Outlook on Agriculture*. SAGE PublicationsSage UK: London, England, 24(3), pp. 151–154. doi: 10.1177/003072709502400304.

Aslam, S. *et al.* (2016) ‘Crimean-congo hemorrhagic fever: Risk factors and control measures for the infection abatement (Review)’, *Biomedical Reports*. Spandidos Publications, 4(1), pp. 15–20. doi: 10.3892/br.2015.545.

Assenga, J. A. *et al.* (2015) ‘Epidemiology of Brucella infection in the human, livestock and wildlife interface in the Katavi-Rukwa ecosystem, Tanzania’, *BMC Veterinary Research*. BioMed Central, 11(1), p. 189. doi: 10.1186/s12917-015-0504-8.

Ata, A. *et al.* (2006) ‘The effect of subclinical bovine herpesvirus 1 infection on fertility of cows and heifers’, *Acta Veterinaria*, 56(2–3), pp. 267–273. doi: 10.2298/AVB0603267A.

Aune, K. *et al.* (2012) ‘Environmental persistence of Brucella abortus in the Greater Yellowstone Area’, *Journal of Wildlife Management*, 76(2), pp. 253–261. doi: 10.1002/jwmg.274.

Bajardi, P. *et al.* (2011) ‘Dynamical Patterns of Cattle Trade Movements’, *PLoS ONE*. Edited by M. Perc. Public Library of Science, 6(5), p. e19869. doi: 10.1371/journal.pone.0019869.

Balcan, D. *et al.* (2009) ‘Multiscale mobility networks and the spatial spreading of infectious diseases’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(51), pp. 21484–21489. doi: 10.1073/pnas.0906910106.

Bansal, S. *et al.* (2010) ‘The dynamic nature of contact networks in infectious disease epidemiology’, *Journal of Biological Dynamics*. Taylor & Francis , 4(5), pp. 478–489. doi: 10.1080/17513758.2010.503376.

Bansal, S., Grenfell, B. T. and Meyers, L. A. (2007) ‘When individual behaviour matters: Homogeneous and network models in epidemiology’, *Journal of the Royal Society Interface*. The Royal SocietyLondon, 4(16), pp. 879–891. doi: 10.1098/rsif.2007.1100.

Barabási, A. L. (2005) ‘The origin of bursts and heavy tails in human dynamics’, *Nature*. Nature, 435(7039), pp. 207–211. doi: 10.1038/nature03459.

Barnard, R. C. *et al.* (2018) ‘Edge-based compartmental modelling of an SIR epidemic on a dual-layer static-dynamic multiplex network with tunable clustering’, *Bulletin of Mathematical Biology*, 80(10), pp. 2698–2733. doi: DOI: 10.1007/s11538-018-0484-5.

Barragan, V. *et al.* (2017) ‘Critical knowledge gaps in our understanding of environmental cycling and transmission of Leptospira spp.’, *Applied and Environmental Microbiology*. American Society for Microbiology. doi: 10.1128/AEM.01190-17.

Barrett, L., Peter Henzi, S. and Lusseau, D. (2012) ‘Taking sociality seriously: The structure of multi-dimensional social networks as a source of information for individuals’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 367(1599), pp. 2108–2118. doi: 10.1098/rstb.2012.0113.

Bates, D. *et al.* (2015) ‘Fitting linear mixed-effects models using lme4’, *Journal of Statistical Software*. American Statistical Association, 67(1). doi: 10.18637/jss.v067.i01.

Beach, J. C., Murano, E. A. and Acuff, G. R. (2002) ‘Prevalence of Salmonella and Campylobacter in beef cattle from transport to slaughter’, *Journal of Food Protection*. IAMFES, 65(11), pp. 1687–1693. doi: 10.4315/0362-028X-65.11.1687.

Begon, M. *et al.* (2002) ‘A clarification of transmission terms in host-microparasite models : numbers, densities and areas’, *Epidemiol. Infect*. Cambridge University Press, 129, pp. 147–153. doi: 10.1017\S0950268802007148.

Beileh, A. . (2006) ‘Livestock Marketing Project: The United Republic of Tanzania’, *African development fund project completion report, agriculture and rural development department North, East and South Regions (ONAR)*.

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

Bernini, A., Bolzoni, L. and Casagrandi, R. (2019) ‘When resolution does matter: Modelling indirect contacts in dairy farms at different levels of detail’, *PLOS ONE*. Edited by J. Sanchez. Public Library of Science, 14(10), p. e0223652. doi: 10.1371/journal.pone.0223652.

Bessell, P. R. *et al.* (2013) ‘Developing a framework for risk-based surveillance of tuberculosis in cattle: a case study of its application in Scotland.’, *Epidemiology and infection*. Cambridge University Press, 141(2), pp. 314–23. doi: 10.1017/S0950268812000635.

Best, D. J. and Roberts, D. E. (1975) ‘Algorithm AS 89: The Upper Tail Probabilities of Spearman’s Rho’, *Applied Statistics*. JSTOR, 24(3), p. 377. doi: 10.2307/2347111.

Bett, B. *et al.* (2010) ‘Epidemiological Assessment of the Rift Valley Fever Outbreak in Kenya and Tanzania in 2006 and 2007’, *The American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 83(2_Suppl), pp. 65–72. doi: 10.4269/ajtmh.2010.09-0290.

Bharti, A. R. *et al.* (2003) ‘Leptospirosis: A zoonotic disease of global importance’, *Lancet Infectious Diseases*. Lancet Publishing Group, pp. 757–771. doi: 10.1016/S1473-3099(03)00830-2.

Biek, R. *et al.* (2015) ‘Measurably evolving pathogens in the genomic era’, *Trends in Ecology and Evolution*. Elsevier Ltd, pp. 306–313. doi: 10.1016/j.tree.2015.03.009.

Bienstock, E. J. and Bonacich, I. P. (2003) *Balancing Efficiency and Vulnerability in Social Networks | Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*. California: The National Academies Press.

Biggs, H. M. *et al.* (2011) ‘Leptospirosis among hospitalized febrile patients in northern Tanzania’, *American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 85(2), pp. 275–281. doi: 10.4269/ajtmh.2011.11-0176.

Biggs, H. M. *et al.* (2013) ‘Leptospirosis and human immunodeficiency virus co-infection among febrile inpatients in northern Tanzania’, *Vector-Borne and Zoonotic Diseases*. Mary Ann Liebert, Inc., 13(8), pp. 572–580. doi: 10.1089/vbz.2012.1205.

Bigras-Poulin, M. *et al.* (2006) ‘Network analysis of Danish cattle industry trade patterns as an evaluation of risk potential for disease spread’, *Preventive Veterinary Medicine*, 76(1–2), pp. 11–39. doi: 10.1016/j.prevetmed.2006.04.004.

Biswas, S. *et al.* (2013) ‘Bovine herpesvirus-1 (BHV-1) - a re-emerging concern in livestock: A revisit to its biology, epidemiology, diagnosis, and prophylaxis’, *Veterinary Quarterly*. Vet Q, pp. 68–81. doi: 10.1080/01652176.2013.799301.

Blackburn, J. K. *et al.* (2019) ‘Modeling R 0 for pathogens with environmental transmission: Animal movements, pathogen populations, and local infectious zones’, *International Journal of Environmental Research and Public Health*. MDPI AG, 16(6), p. 954. doi: 10.3390/ijerph16060954.

Blondel, V. D. *et al.* (2008) ‘Fast unfolding of communities in large networks’, *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10). doi: 10.1088/1742-5468/2008/10/P10008.

Bodenham, R., Mazeri, S., *et al.* (2020) ‘Latent class evaluation of the performance of serological tests for Brucella spp. in Tanzanian cattle, sheep, and goats’, *in press*.

Bodenham, R., Lukambagire, A., *et al.* (2020) ‘Prevalence and speciation of brucellosis in febrile patients from a pastoralist community of Tanzania’, *Scientific Reports*. Nature Research, 10(1), pp. 1–11. doi: 10.1038/s41598-020-62849-4.

Böhm, M., Hutchings, M. R. and White, P. C. L. (2009) ‘Contact networks in a wildlife-livestock host community: Identifying high-risk individuals in the transmission of bovine TB among badgers and cattle’, *PLoS ONE*. Public Library of Science, 4(4), p. e5016. doi: 10.1371/journal.pone.0005016.

Bonacich, P. (2007) ‘Some unique properties of eigenvector centrality’, *Social Networks*. North-Holland, 29(4), pp. 555–564. doi: 10.1016/j.socnet.2007.04.002.

Borjeson, L., Hodgson, D. L. and Yanda, P. Z. (2008) ‘Northeast Tanzania’s Disappearing Rangelands: Historical Perspectives on Recent Land Use Change on JSTOR’, *International journal of african historical studies*, 41(3), pp. 523–533.

Bouley, A. J. *et al.* (2012) ‘Brucellosis among hospitalized febrile patients in northern Tanzania.’, *The American journal of tropical medicine and hygiene*. The American Society of Tropical Medicine and Hygiene, 87(6), pp. 1105–11. doi: 10.4269/ajtmh.2012.12-0327.

Bouslikhane, M. (2015) ‘Cross border movements of animals and animal products and their relevance to the epidemiology of animal diseases in Arfica’, *Africa – OIE Regional Commission – Bouslikhane*.

Bowling, M. B. *et al.* (2008) ‘REVIEW: Identification and Traceability of Cattle in Selected Countries Outside of North America’, *The Professional Animal Scientist*, 24(4), pp. 287–294. doi: 10.15232/S1080-7446(15)30858-5.

Brackenbury, L. S., Carr, B. V. and Charleston, B. (2003) ‘Aspects of the innate and adaptive immune responses to acute infections with BVDV’, in *Veterinary Microbiology*. Elsevier, pp. 337–344. doi: 10.1016/j.vetmic.2003.09.004.

Brahmbhatt, D. P. *et al.* (2012) ‘Contacts between domestic livestock and wildlife at the Kruger National Park Interface of the Republic of South Africa’, *Preventive Veterinary Medicine*. Elsevier, 103(1), pp. 16–21. doi: 10.1016/J.PREVETMED.2011.08.003.

Brockmann, D., Hufnagel, L. and Geisel, T. (2006) ‘The scaling laws of human travel’, *Nature*. Nature Publishing Group, 439(7075), pp. 462–465. doi: 10.1038/nature04292.

Van den Brom, R. *et al.* (2015) ‘Coxiella burnetii infections in sheep or goats: An opinionated review’, *Veterinary Microbiology*. Elsevier, 181(1–2), pp. 119–129. doi: 10.1016/j.vetmic.2015.07.011.

Brooks-Pollock, E. *et al.* (2015) ‘Eight challenges in modelling infectious livestock diseases’, *Epidemics*. Elsevier, 10, pp. 1–5. doi: 10.1016/j.epidem.2014.08.005.

Brooks, M. E. *et al.* (2017) *glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling*, *The R Journal*. R Foundation for Statistical Computing.

Brownlie, J. *et al.* (1987) ‘Pathogenesis and epidemiology of Bovine virus diarrhoea virus infection of cattle', *Annales de Recherches Vétérinaires, INRA Editions*, 18(2), pp. 157–166.

Bucur, D. and Holme, P. (2020) 'Beyond ranking nodes: Predicting epidemic outbreak sizes by network centralities', *PLOS Computational Biology*. Edited by B. Althouse. Public Library of Science, 16(7), p. e1008052. doi: 10.1371/journal.pcbi.1008052.

Buhnerkempe, M. G. *et al.* (2014) 'The Impact of Movements and Animal Density on Continental Scale Cattle Disease Outbreaks in the United States', *PLoS ONE*. Edited by A. Vespignani. Public Library of Science, 9(3), p. e91724. doi: 10.1371/journal.pone.0091724.

Büttner, K. *et al.* (2013a) 'Efficient Interruption of Infection Chains by Targeted Removal of Central Holdings in an Animal Trade Network', *PLoS ONE*. Edited by Y. Moreno. Public Library of Science, 8(9), p. e74292. doi: 10.1371/journal.pone.0074292.

Büttner, K. *et al.* (2013b) 'Static network analysis of a pork supply chain in Northern Germany—Characterisation of the potential spread of infectious diseases via animal movements', *Preventive Veterinary Medicine*, 110(3–4), pp. 418–428. doi: 10.1016/j.prevetmed.2013.01.008.

Butts, C. T. (2008) 'Social network analysis: A methodological introduction', *Asian Journal of Social Psychology*, 11(1), pp. 13–41. doi: 10.1111/j.1467-839X.2007.00241.x.

Byrne, A. W. *et al.* (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*. Elsevier B.V., 141, pp. 38–47. doi: 10.1016/j.prevetmed.2017.04.007.

Caillaud, D., Craft, M. E. and Meyers, L. A. (2013) 'Epidemiological effects of group size variation in social species.', *Journal of the Royal Society, Interface*. The Royal Society, 10(83), p. 20130206. doi: 10.1098/rsif.2013.0206.

Cardillo, A. *et al.* (2013) 'Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures', *European Physical Journal: Special Topics*. Springer, 215(1), pp. 23–33. doi: 10.1140/epjst/e2013-01712-8.

Carpenter, S. *et al.* (2001) 'From Metaphor to Measurement: Resilience of What to What?', in *Ecosystems*. Springer, pp. 765–781. doi: 10.1007/s10021-001-0045-9.

Carslake, D. *et al.* (2011) 'Endemic cattle diseases: Comparative epidemiology and governance', *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, 366(1573), pp. 1975–1986. doi: 10.1098/rstb.2010.0396.

Cash-Goldwasser, S. *et al.* (2018) 'Risk Factors for Human Brucellosis in Northern Tanzania', *The American Journal of Tropical Medicine and Hygiene*, 98(2), pp. 598–606. doi: 10.4269/ajtmh.17-0125.

Cassius Moreki, J. *et al.* (2012) ‘Cattle Identification and Traceability in Botswana -’, *Journal of Animal Science Advances*, 2(12), pp. 925–933.

Cellai, D. *et al.* (2013) ‘Percolation in Multiplex Networks with Overlap’, *Physical Review E*, 88(5).

Chami, G. F. *et al.* (2017) ‘Social network fragmentation and community health’, *Proceedings of the National Academy of Sciences*, 114(36), pp. 7425–7431. doi: 10.1073/PNAS.1700166114.

Chase, C. C. L. (2013) ‘The impact of BVDV infection on adaptive immunity’, *Biologicals*. Biologicals, 41(1), pp. 52–60. doi: 10.1016/j.biologicals.2012.09.009.

Chaters, G. L. *et al.* (2019) ‘Analysing livestock network data for infectious disease control: An argument for routine data collection in emerging economies’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society Publishing, 374(1776). doi: 10.1098/rstb.2018.0264.

Chen, Y. *et al.* (2007) ‘Percolation theory and fragmentation measures in social networks’, *Physica A: Statistical Mechanics and Its Applications*, 378(1), pp. 11–19. doi: 10.1016/j.physa.2006.11.074.

Chipwaza, B. *et al.* (2014) ‘Community Knowledge and Attitudes and Health Workers’ Practices regarding Non-malaria Febrile Illnesses in Eastern Tanzania’, *PLoS Neglected Tropical Diseases*. Edited by J. O. Gyapong. Public Library of Science, 8(5), p. e2896. doi: 10.1371/journal.pntd.0002896.

Chowell, G. and Nishiura, H. (2015) ‘Characterizing the Transmission Dynamics and Control of Ebola Virus Disease’, *PLoS Biology*. Public Library of Science, 13(1), p. e1002057. doi: 10.1371/journal.pbio.1002057.

Chriél, M., Stryhn, H. and Dauphin, G. (1999) ‘Generalised linear mixed models analysis of risk factors for contamination of Danish broiler flocks with Salmonella typhimurium’, *Preventive Veterinary Medicine*. Elsevier, 40(1), pp. 1–17. doi: 10.1016/S0167-5877(99)00016-1.

Christiaensen, L., Demery, L. and Kuhl, J. (2011) ‘The (evolving) role of agriculture in poverty reduction-An empirical perspective’, *Journal of Development Economics*. North-Holland, 96(2), pp. 239–254. doi: 10.1016/j.jdeveco.2010.10.006.

Christley, R. M. *et al.* (2005) ‘Infection in Social Networks: Using Network Analysis to Identify High-Risk Individuals’, *American Journal of Epidemiology*. Oxford University Press, 162(10), pp. 1024–1031. doi: 10.1093/aje/kwi308.

Christou, L. (2011) ‘The global burden of bacterial and viral zoonotic infections’, *Clinical Microbiology and Infection*. Blackwell Publishing Ltd, 17(3), pp. 326–330. doi: 10.1111/j.1469-0691.2010.03441.x.

Chung, Y. *et al.* (2013) ‘A nondegenerate penalized likelihood estimator for variance parameters in multilevel models’, *Psychometrika*, 78(4), pp. 685–709.

Coker, R. *et al.* (2011) ‘Towards a conceptual framework to support one-health research for policy on emerging zoonoses’, *The Lancet Infectious Diseases*. Elsevier, 11(4), pp. 326–331. doi: 10.1016/S1473-3099(10)70312-1.

Colizza, V. *et al.* (2006) ‘The role of the airline transportation network in the prediction and predictability of global epidemics’, *Proceedings of the National Academy of Sciences*, 103(7), p. 2015:2020. doi: doi.org/10.1073/pnas.0510525103.

Colman, E. *et al.* (2019) ‘Efficient sentinel surveillance strategies for preventing epidemics on networks’. Edited by D. Paolotti. Public Library of Science, 15(11), p. e1007517. doi: 10.1371/journal.pcbi.1007517.

Coppock, D. L., Desta, S. T. P. S. and Gebru, G. (2012) ‘Achieving Development Impact among Pastoral and Agro-Pastoral People: Lessons Learned in Southern Ethiopia, 2000-2009’, *Ethiopian Society for Animal Production*, 65.

Coppolillo, P. B. (2000) ‘The Landscape Ecology of Pastoral Herding: Spatial Analysis of Land Use and Livestock Production in East Africa’, *Human Ecology*. Kluwer Academic Publishers-Plenum Publishers, 28(4), pp. 527–560. doi: 10.1023/A:1026435714109.

Corner, L. A. ., Pfeiffer, D. . and Morris, R. . (2003) ‘Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (Trichosurus vulpecula)’, *Preventive Veterinary Medicine*, 59(3), pp. 147–167. doi: 10.1016/S0167-5877(03)00075-8.

Cottam, E. M. *et al.* (2008) ‘Transmission pathways of foot-and-mouth disease virus in the United Kingdom in 2007’, *PLoS Pathogens*. Public Library of Science, 4(4), p. 1000050. doi: 10.1371/journal.ppat.1000050.

Covarrubias, K. *et al.* (2012) *Livestock and livelihoods in rural Tanzania and A descriptive analysis of the 2009 National Panel Survey*, *The United Republic of Tanzania Ministry of Livestock and Fisharies Development.* World Bank, Washington, DC.

Cowan, R. and Jonard, N. (2004) ‘Network structure and the diffusion of knowledge’, *Journal of Economic Dynamics and Control*. North-Holland, 28(8), pp. 1557–1575. doi: 10.1016/J.JEDC.2003.04.002.

Craft, M. E. *et al.* (2009) ‘Distinguishing epidemic waves from disease spillover in a wildlife population’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 276(1663), pp. 1777–1785. doi: 10.1098/rspb.2008.1636.

Craft, M. E. (2015) ‘Infectious disease transmission and contact networks in wildlife and livestock.’, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 370(1669). doi: 10.1098/rstb.2014.0107.

de Cremoux, R. *et al.* (2012) ‘Coxiella burnetii vaginal shedding and antibody responses in dairy goat herds in a context of clinical Q fever outbreaks’, *FEMS Immunology and Medical Microbiology*. FEMS Immunol Med Microbiol, 64(1), pp. 120–122. doi: 10.1111/j.1574-695X.2011.00893.x.

Croft, S., Chauvenet, A. L. M. and Smith, G. C. (2017) ‘A systematic approach to estimate the distribution and total abundance of British mammals’, *PLOS ONE*. Edited by G. Umapathy. Public Library of Science, 12(6), p. e0176339. doi: 10.1371/journal.pone.0176339.

Cross, P. C. *et al.* (2004) ‘Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo in the Kruger National Park’, *Annales Zoologici Fennici*. Finnish Zoological and Botanical Publishing Board, pp. 879–892. doi: 10.2307/23736148.

Cross, P. C. *et al.* (2005) ‘Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations’, *Ecology Letters*. Blackwell Publishing Ltd, 8(6), pp. 587–595. doi: 10.1111/j.1461-0248.2005.00760.x.

Crump, J. A. *et al.* (2013) ‘Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study’, *PLoS Neglected Tropical Diseases*. Edited by M. Picardeau. Public Library of Science, 7(7), p. e2324. doi: 10.1371/journal.pntd.0002324.

Csardi, G. and Nepusz, T. (2006) ‘The igraph software package for complex network research’, *InterJournal, Complex Systems*, 1695(http://igraph.org), p. 1695.

Danon, L. *et al.* (2011) ‘Networks and the epidemiology of infectious disease.’, *Interdisciplinary perspectives on infectious diseases*. Hindawi Limited, 2011, p. 284909. doi: 10.1155/2011/284909.

Davis, A. and Sharp, J. (2020) ‘Rethinking One Health: Emergent human, animal and environmental assemblages’, *Social Science and Medicine*. Elsevier Ltd, 258, p. 113093. doi: 10.1016/j.socscimed.2020.113093.

Dean, A. S. *et al.* (2013) ‘Potential Risk of Regional Disease Spread in West Africa through Cross-Border Cattle Trade’, *PLoS ONE*. Edited by B. Kaltenboeck. Public Library of Science, 8(10), p. e75570. doi: 10.1371/journal.pone.0075570.

Derah, N. and Mokopasetso (2005) ‘The control of Foot and Mouth Disease in Botswana and Zimbabwe’, *Tropicultura; Special issue*, 23(special), p. 3:7.

Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2001) ‘Giant strongly connected component of directed networks’, *Physical Review E*, 64(2), p. 025101. doi: 10.1103/PhysRevE.64.025101.

Drake, J. M. *et al.* (2015) ‘Ebola Cases and Health System Demand in Liberia’, *PLoS Biology*. Public Library of Science, 13(1), p. 1002056. doi: 10.1371/journal.pbio.1002056.

Dubé, C. *et al.* (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, p. 382:392. doi: 10.1111/j.1865-1682.2008.01053.x.

Dubé, C. *et al.* (2009) ‘A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development’, *Transboundary and Emerging Diseases*. Transbound Emerg Dis, 56(3), pp. 73–85. doi: 10.1111/j.1865-1682.2008.01064.x.

Dubé, C. *et al.* (2011) ‘Introduction to network analysis and its implications for animal disease modelling’, *OIE Revue Scientifique et Technique*. Office International des Epizootes, 30(2), pp. 425–436. doi: 10.20506/rst.30.2.2043.

Dutilly, C. *et al.* (2020) ‘Multi-scale assessment of the livestock sector for policy design in Zambia’, *Journal of Policy Modeling*. Elsevier B.V., 42(2), pp. 401–418. doi: 10.1016/j.jpolmod.2019.07.004.

Dutta, B. L., Ezanno, P. and Vergu, E. (2014) ‘Characteristics of the spatio-temporal network of cattle movements in France over a 5-year period’, *Preventive Veterinary Medicine*, 117(1), pp. 79–94. doi: 10.1016/j.prevetmed.2014.09.005.

Enright, J. A. and O’Hare, A. (2017) ‘Reconstructing disease transmission dynamics from animal movements and test data’, *Stochastic Environmental Research and Risk Assessment*. Springer Berlin Heidelberg, 31(2), pp. 369–377. doi: 10.1007/s00477-016-1354-z.

Enright, J. and Kao, R. R. (2016) ‘A fast algorithm for calculating an expected outbreak size on dynamic contagion networks’, *Epidemics*. Elsevier, 16, pp. 56–62. doi: 10.1016/J.EPIDEM.2016.05.002.

Enright, J. and Kao, R. R. (2018) ‘Epidemics on dynamic networks’, *Epidemics*. Elsevier. doi: 10.1016/J.EPIDEM.2018.04.003.

Epstein, J. M. *et al.* (2008) ‘Coupled contagion dynamics of fear and disease: Mathematical and computational explorations’, *PLoS ONE*. Public Library of Science, 3(12), p. e3955. doi: 10.1371/journal.pone.0003955.

Epstein, P. R. (2002) *Conservation Medicine: Ecological Health in Practice - Biodiversity, climate change and emerging infectios diseases*. Edited by A. Aguirre et al. Oxford University Press.

Ergönül, Ö. (2006) ‘Crimean-Congo haemorrhagic fever’, *Lancet Infectious Diseases*. Elsevier, pp. 203–214. doi: 10.1016/S1473-3099(06)70435-2.

FAO (2009) *The state of food and agriculture. Livestock in the balance |Policy Support and Governance| Food and Agriculture Organization of the United Nations*. Available at:

http://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/1235525/ (Accessed: 20 January 2021).

Farine, D. R. and Whitehead, H. (2015) ‘Constructing, conducting and interpreting animal social network analysis’, *Journal of Animal Ecology*. Edited by S. Altizer. Wiley/Blackwell (10.1111), 84(5), pp. 1144–1163. doi: 10.1111/1365-2656.12418.

Ferdousi, T. *et al.* (2019) ‘Generation of swine movement network and analysis of efficient mitigation strategies for African swine fever virus’, *PLOS ONE*. Edited by W. F. de Boer. Public Library of Science, 14(12), p. e0225785. doi: 10.1371/journal.pone.0225785.

Fèvre, E. M. *et al.* (2006) ‘Animal movements and the spread of infectious diseases’, *Trends in Microbiology*. Elsevier Current Trends, 14(3), pp. 125–131. doi: 10.1016/J.TIM.2006.01.004.

Finn, K. R. *et al.* (2019) ‘The use of multilayer network analysis in animal behaviour’, *Animal Behaviour*. Academic Press, pp. 7–22. doi: 10.1016/j.anbehav.2018.12.016.

Flintan, F. (2012) *Participatory rangeland resource mapping as a valuable tool for village land use planning in Tanzania International Land Coalition, IFAD, CARE Tanzania, KINNAPA Development Programme, Ministry of Livestock and Fisheries Development, local government and CSO partners*. Rome, Italy: International Land Coalition.

Floyd, J. R. *et al.* (2019) ‘Exploring fine-scale human and livestock movement in western Kenya’, *One Health*. Elsevier B.V., 7, p. 100081. doi: 10.1016/j.onehlt.2019.100081.

Lo Fo Wong, D. M. A. *et al.* (2004) ‘Herd-level risk factors for subclinical Salmonella infection in European finishing-pig herds’, *Preventive Veterinary Medicine*, 62(4), pp. 253–266. doi: 10.1016/j.prevetmed.2004.01.001.

Forman, S. *et al.* (2012) ‘Position paper: improving governance for effective veterinary services in developing countries--a priority for donor funding.’, *Revue scientifique et technique (International Office of Epizootics)*, 31(2), pp. 647–60.

Fournie, G. *et al.* (2013) ‘Interventions for avian influenza A (H5N1) risk management in live bird market networks’, *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 110(22), pp. 9177–9182. doi: 10.1073/pnas.1220815110.

Fournié, G. *et al.* (2011) ‘Impact of the implementation of rest days in live bird markets on the dynamics of H5N1 highly pathogenic avian influenza’, *Journal of the Royal Society Interface*. The Royal Society, 8(61), pp. 1079–1089. doi: 10.1098/rsif.2010.0510.

Fournié, G., Waret-Szkuta, A., Camacho, A., Yigezu, Laike M, *et al.* (2018) ‘A dynamic model of transmission and elimination of peste des petits ruminants in Ethiopia.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, p. 201711646. doi: 10.1073/pnas.1711646115.

Fournié, G., Waret-Szkuta, A., Camacho, A., Yigezu, Laike M., *et al.* (2018) ‘A dynamic model of transmission and elimination of peste des petits ruminants in Ethiopia’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 115(33), pp. 8454–8459. doi: 10.1073/pnas.1711646115.

Fournié, G. and Pfeiffer, D. U. (2013) ‘Monitoring and controlling disease spread through live animal market networks’, *The Veterinary Journal*. W.B. Saunders, 195(1), pp. 8–9. doi: 10.1016/J.TVJL.2012.09.014.

Fox, J. and Weisberg, S. (2019) *An R Companion to Applied Regression - John Fox, Sanford Weisberg - Google Books*. Third. Thousand Oaks, CA: Sage Publications.

Freeman, L. C. (1978) ‘Centrality in social networks Conceptual Clarification’, *Social Networks*, 1, pp. 215–239.

Frost, S. D. W. *et al.* (2015) ‘Eight challenges in phylodynamic inference’, *Epidemics*. Elsevier B.V., 10, pp. 88–92. doi: 10.1016/j.epidem.2014.09.001.

Funk, S., Salathé, M. and Jansen, V. A. A. (2010) ‘Modelling the influence of human behaviour on the spread of infectious diseases: A review’, *Journal of the Royal Society Interface*. Royal Society, pp. 1247–1256. doi: 10.1098/rsif.2010.0142.

Galvani, A. P. and May, R. M. (2005) ‘Dimensions of superspreading’, *Nature*, 438(7066), pp. 293–295. doi: 10.1038/438293a.

GALVmed (2018) *brucellosisvaccine*. Available at: https://brucellosisvaccine.org/ (Accessed: 24 May 2021).

Gao, J. *et al.* (2013) ‘Percolation of a general network of networks’, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 88(6), p. 062816. doi: 10.1103/PhysRevE.88.062816.

Gao, J., Barzel, B. and Barabási, A. L. (2016) ‘Universal resilience patterns in complex networks’, *Nature*. Nature Publishing Group, 530(7590), pp. 307–312. doi: 10.1038/nature16948.

Gardy, J. L. *et al.* (2011) ‘Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak’, *New England Journal of Medicine*. Massachusetts Medical Society, 364(8), pp. 730–739. doi: 10.1056/NEJMoa1003176.

Gates, M. C. and Woolhouse, M. E. J. (2015a) ‘Controlling infectious disease through the targeted manipulation of contact network structure’, *Epidemics*, 12, pp. 11–19. doi: 10.1016/j.epidem.2015.02.008.

Gates, M. C. and Woolhouse, M. E. J. (2015b) ‘Controlling infectious disease through the targeted manipulation of contact network structure’, *Epidemics*. Elsevier B.V., 12, pp. 11–19. doi: 10.1016/j.epidem.2015.02.008.

Génois, M. *et al.* (2015) 'Compensating for population sampling in simulations of epidemic spread on temporal contact networks', *Nature Communications*. Nature Publishing Group, 6(1), p. 8860. doi: 10.1038/ncomms9860.

Gibbens, J. C. *et al.* (2001) 'Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months.', *The Veterinary record*, 149(24), pp. 729–43.

Gilbert, M. *et al.* (2005) 'Cattle movements and bovine tuberculosis in Great Britain', *Nature*. Nature Publishing Group, 435(7041), pp. 491–496. doi: 10.1038/nature03548.

de Glanville, W. *et al.* (2020) 'Classification and characterisation of livestock production systems in northern Tanzania', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2020.02.10.941617. doi: 10.1101/2020.02.10.941617.

Godfray, H. C. J. *et al.* (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: Biological Sciences*. Royal Society. doi: 10.1098/rspb.2013.1634.

Goldman, M. J. and Riosmena, F. (2013) 'Adaptive Capacity in Tanzanian Maasailand: Changing strategies to cope with drought in fragmented landscapes.', *Global environmental change : human and policy dimensions*. NIH Public Access, 23(3), pp. 588–597. doi: 10.1016/j.gloenvcha.2013.02.010.

Government of Tanzania (2017) *The United Republic of Tanzania 2016/2017 Annual Agricultural Sample Survey Crop and Livestock Report* . Dar es Salaam, Zanzibar.

Grace, D. *et al.* (2012) 'Mapping of poverty and likely zoonoses hotspots Zoonoses Project 4 Report to Department for International Development, UK'.

Green, D. M. *et al.* (2008) 'Estimates for local and movement-based transmission of bovine tuberculosis in British cattle', *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1638).

Green, D. M. and Kao, R. R. (2007) 'Data quality of the Cattle Tracing System in Great Britain', *Veterinary Record*. British Veterinary Association, 161(13), pp. 439–443. doi: 10.1136/vr.161.13.439.

Green, D. M., Kiss, I. Z. and Kao, R. R. (2006a) 'Modelling the initial spread of foot-and-mouth disease through animal movements.', *Proceedings. Biological sciences*. The Royal Society, 273(1602), pp. 2729–35. doi: 10.1098/rspb.2006.3648.

Green, D. M., Kiss, I. Z. and Kao, R. R. (2006b) 'Parameterization of individual-based models: Comparisons with deterministic mean-field models', *Journal of Theoretical Biology*. J Theor Biol, 239(3), pp. 289–297. doi: 10.1016/j.jtbi.2005.07.018.

Gross, T., D'Lima, C. J. D. and Blasius, B. (2006) 'Epidemic Dynamics on an Adaptive

Network', *Physical Review Letters*. American Physical Society, 96(20), p. 208701. doi: 10.1103/PhysRevLett.96.208701.

Guidoum, K. A. *et al.* (2020) 'Ruminant pestiviruses in North Africa', *Preventive Veterinary Medicine*. Elsevier BV, 184, p. 105156. doi: 10.1016/j.prevetmed.2020.105156.

Guimerà, R. *et al.* (2005) 'The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 102(22), pp. 7794–9. doi: 10.1073/pnas.0407994102.

Guinat, C. *et al.* (2016) 'Exploring pig trade patterns to inform the design of risk-based disease surveillance and control strategies', *Scientific Reports*. Nature Publishing Group, 6(1), p. 28429. doi: 10.1038/srep28429.

Hall, M., Woolhouse, M. and Rambaut, A. (2015) 'Epidemic Reconstruction in a Phylogenetics Framework: Transmission Trees as Partitions of the Node Set', *PLoS Computational Biology*. Public Library of Science, 11(12), p. 1004613. doi: 10.1371/journal.pcbi.1004613.

Halliday, J. *et al.* (2012) 'Bringing together emerging and endemic zoonoses surveillance: shared challenges and a common solution', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1604).

Hammouchi, M. *et al.* (2012) 'Experimental infection of alpine goats with a Moroccan strain of peste des petits ruminants virus (PPRV)', *Veterinary Microbiology*. Elsevier, 160(1–2), pp. 240–244. doi: 10.1016/j.vetmic.2012.04.043.

Handcock, R. *et al.* (2009) 'Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing', *Sensors*. Molecular Diversity Preservation International, 9(5), pp. 3586–3603. doi: 10.3390/s90503586.

Hartig, F. (2020) 'DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models', *R package version 0.3.3.0*.

Haseeb, A. *et al.* (2019) 'Economic burden of livestock disease and drought in Northern Tanzania', *Journal of Development and Agricultural Economics*. Academic Journals, 11(6), pp. 140–151. doi: 10.5897/jdae2018.1028.

Haydon, D. T. *et al.* (2002) 'Identifying reservoirs of infection: A conceptual and practical challenge', *Emerging Infectious Diseases*. Centers for Disease Control and Prevention (CDC), pp. 1468–1473. doi: 10.3201/eid0812.010317.

Haydon, D. T., Kao, R. R. and Kitching, R. P. (2004) 'Opinion: The UK foot-and-mouth disease outbreak — the aftermath', *Nature Reviews Microbiology*. Nature Publishing Group, 2(8), pp. 675–681. doi: 10.1038/nrmicro960.

Heffernan, C. and Misturelli, F. (2000) *The Delivery of Veterinary Services to the Poor: Preliminary findings from Kenya*. DFID report, Veterinary Epidemiology and Economics Research Unit, Department of Agriculture, University of Reading, UK.

Herrera, J. L. *et al.* (2016) 'Disease Surveillance on Complex Social Networks', *PLOS Computational Biology*. Edited by M. Salathé. Public Library of Science, 12(7), p. e1004928. doi: 10.1371/journal.pcbi.1004928.

Herzog, C. M. *et al.* (2019) 'Pastoral production is associated with increased peste des petits ruminants seroprevalence in northern Tanzania across sheep, goats and cattle', *Epidemiology and Infection*. Cambridge University Press, 147. doi: 10.1017/S0950268819001262.

Herzog, C. M. *et al.* (2020) 'Identifying age cohorts responsible for peste des petits ruminants virus transmission among sheep, goats, and cattle in northern Tanzania', *Viruses*. MDPI AG, 12(2). doi: 10.3390/v12020186.

Hidano, A. and Gates, M. C. (2019) 'Why sold, not culled? Analysing farm and animal characteristics associated with livestock selling practices', *Preventive Veterinary Medicine*. Elsevier B.V., 166, pp. 65–77. doi: 10.1016/j.prevetmed.2019.03.005.

Hidano, A., Gates, M. C. and Enticott, G. (2019) 'Farmers' Decision Making on Livestock Trading Practices: Cowshed Culture and Behavioral Triggers Amongst New Zealand Dairy Farmers', *Frontiers in Veterinary Science*. Frontiers Media S.A., 6, p. 320. doi: 10.3389/fvets.2019.00320.

Hodgkinson, E. (2017) *Masters degree thesis: Seroprevalence of BVDV in Tanzanian cattle and associated risk factors*. University of Glasgow.

Holling, C. S. (2013) 'Resilience and stability of ecological systems', in *The Future of Nature: Documents of Global Change*. Yale University Press, pp. 245–256. doi: 10.1146/annurev.es.04.110173.000245.

Holme, P. and Litvak, N. (2017) 'Cost-efficient vaccination protocols for network epidemiology', *PLoS Computational Biology*. Edited by M. (Matt) Ferrari. Public Library of Science, 13(9), p. e1005696. doi: 10.1371/journal.pcbi.1005696.

Holme, P. and Masuda, N. (2015) 'The basic reproduction number as a predictor for epidemic outbreaks in temporal networks', *PLoS ONE*. Public Library of Science, 10(3), p. e0120567. doi: 10.1371/journal.pone.0120567.

Hop, G. E. *et al.* (2011) 'Towards an integrated approach to intensify cross-border collaboration in the field of highly contagious livestock diseases: a general framework for decision support.', pp. 15–15.

Hotez, P. J. *et al.* (2009) 'Rescuing the bottom billion through control of neglected tropical diseases', *The Lancet*. Elsevier, 373(9674), pp. 1570–1575. doi: 10.1016/S0140-

6736(09)60233-6.

Hufnagel, L., Brockmann, D. and Geisel, T. (2004) ‘Forecast and control of epidemics in a globalized world’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 101(42), pp. 15124–15129. doi: 10.1073/pnas.0308344101.

Hughes, E. C. *et al.* (2019) ‘Taenia multiceps coenurosis in Tanzania: A major and under-recognised livestock disease problem in pastoral communities’, *Veterinary Record*. British Veterinary Association, 184(6), p. 191. doi: 10.1136/vr.105186.

Hummel, P. H. (1976) ‘Incidence in Tanzania of CF antibody to Coxiella burneti in sera from man, cattle, sheep, goats and game.’, *The Veterinary record*, 98(25), pp. 501–5.

Hyera, J. M. K., Liess, B. and Frey, H. R. (1991) ‘Bovine viral diarrhoea virus infection in cattle, sheep and goats in northern Tanzania’, *Tropical Animal Health and Production*. Kluwer Academic Publishers, 23(2), pp. 83–94. doi: 10.1007/BF02361187.

ILRI (2018) *Why livestock matter | International Livestock Research Institute*. Available at: https://www.ilri.org/whylivestockmatter (Accessed: 11 July 2018).

Iribarren, J. L. and Moro, E. (2011) ‘Branching dynamics of viral information spreading’, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 84(4), p. 046116. doi: 10.1103/PhysRevE.84.046116.

Jahel, C. *et al.* (2020) ‘Mapping livestock movements in Sahelian Africa’, *Scientific Reports*. Nature Research, 10(1), p. 13. doi: 10.1038/s41598-020-65132-8.

James, G. *et al.* (2013) *An introduction to Statistical Learning: with Applications in R*, *Springer Texts in Statistics*. doi: 10.1007/978-1-4614-7138-7.

Johnson, C. K. *et al.* (2020) ‘Global shifts in mammalian population trends reveal key predictors of virus spillover risk’. doi: 10.1098/rspb.2019.2736.

Jombart, T. *et al.* (2009) ‘Spatiotemporal dynamics in the early stages of the 2009 A/H1N1 influenza pandemic’, *PLoS Currents*. PLoS Curr, 1(AUG), pp. 1–11. doi: 10.1371/currents.RRN1026.

Jombart, T. *et al.* (2011) ‘Reconstructing disease outbreaks from genetic data: A graph approach’, *Heredity*. Heredity (Edinb), 106(2), pp. 383–390. doi: 10.1038/hdy.2010.78.

Jones, B. A. *et al.* (2016) ‘The economic impact of eradicating peste des petits ruminants: A benefit-cost analysis’, *PLoS ONE*. Public Library of Science, 11(2), p. 149982. doi: 10.1371/journal.pone.0149982.

Jones, K. E. *et al.* (2008) ‘Global trends in emerging infectious diseases’, *Nature*, 451, pp. 990–994. doi: 10.1038/nature06536.

de Jong, M. C. ., Diekmann, O. and Heesterbeek, H. (1994) ‘How does transmission of infection depend on population size?’, *Publications of the Newton Institute*, 5, pp. 84–94.

Júnior, J. L. S. *et al.* (2017) ‘Livestock markets play an important role in the cattle movement network in Pernambuco, Brazil’, *Braz. J. Vet. Res. Anim. Sci.*, 54(3), p. 225:237. doi: 10.11606/issn.1678-4456.bjvras.2017.124303.

Kaddour, A. *et al.* (2019) ‘Bovine herpesvirus 1 in the northeast of Algiers, Algeria: Seroprevalence and associated risk factors in dairy herd.’, *Journal of advanced veterinary and animal research*. Network for the Veterinarians of Bangladesh, 6(1), pp. 60–65. doi: 10.5455/javar.2019.f312.

Kao, R. R. (2002) ‘The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK’, *Trends in Microbiology*. Elsevier Current Trends, 10(6), pp. 279–286. doi: 10.1016/S0966-842X(02)02371-5.

Kao, R. R. *et al.* (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(16).

Kao, R. R. (2010) ‘Networks and models with heterogeneous population structure in epidemiology’, in *Network Science: Complexity in Nature and Technology*. Springer London, pp. 51–84. doi: 10.1007/978-1-84996-396-1_4.

Kao, R. R. *et al.* (2014) ‘Supersize me: How whole-genome sequencing and big data are transforming epidemiology’, *Trends in Microbiology*. Elsevier Ltd, pp. 282–291. doi: 10.1016/j.tim.2014.02.011.

Kao, R. R. R. *et al.* (2006) ‘Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain.’, *Proceedings. Biological sciences*. The Royal Society, 273(1597), pp. 1999–2007. doi: 10.1098/rspb.2006.3505.

Karimuribo, E. D. *et al.* (2007) ‘Prevalence of brucellosis in crossbred and indigenous cattle in Tanzania’, *Livestock Research for Rural Development.*, 19(10).

Karsai, M. *et al.* (2011) ‘Small but slow world: How network topology and burstiness slow down spreading’, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 83(2), p. 025102. doi: 10.1103/PhysRevE.83.025102.

Keeling, M. J. (1999) ‘The effects of local spatial structure on epidemiological invasions’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 266(1421), pp. 859–867. doi: 10.1098/rspb.1999.0716.

Keeling, M. J. *et al.* (2003) ‘Modelling vaccination strategies against foot-and-mouth disease’, *Nature*, 421(6919), pp. 136–142. doi: 10.1038/nature01343.

Keeling, M. J. and Eames, K. T. D. (2005) ‘Networks and epidemic models.’, *Journal of the*

Royal Society, Interface. The Royal Society, 2(4), pp. 295–307. doi: 10.1098/rsif.2005.0051.

Keeling, M. J. and Grenfell, B. T. (2000) 'Individual-based perspectives on R0', *Journal of Theoretical Biology*. Academic Press, 203(1), pp. 51–61. doi: 10.1006/jtbi.1999.1064.

Keeling, M. J. and Rohani, P. (2002) 'Estimating spatial coupling in epidemiological systems: a mechanistic approach', *Ecology Letters*. Blackwell Science Ltd, 5(1), pp. 20–29. doi: 10.1046/j.1461-0248.2002.00268.x.

Kelly, A. M. and Marshak, R. R. (2007) 'Veterinary medicine, global health', *Journal of the American Veterinary Medical Association*. American Veterinary Medical Association 1931 North Meacham Road - Suite 100, Schaumburg, IL 60173 USA 847-925-8070 847-925-1329 avmajournals@avma.org , 231(12), pp. 1806–1808. doi: 10.2460/javma.231.12.1806.

Kerfua, S. D. *et al.* (2018) 'Spatial and temporal distribution of foot-and-mouth disease in four districts situated along the Uganda-Tanzania border: Implications for cross-border efforts in disease control', *The Onderstepoort journal of veterinary research*. AOSIS, 85(1). doi: 10.4102/ojvr.v85i1.1528.

Van Kerkhove, M. D. *et al.* (2009) 'Poultry movement networks in Cambodia: Implications for surveillance and control of highly pathogenic avian influenza (HPAI/H5N1)', *Vaccine*. Elsevier, 27(45), pp. 6345–6352. doi: 10.1016/J.VACCINE.2009.05.004.

Kersh, G. J. *et al.* (2013) 'Presence and Persistence of Coxiella burnetii in the environments of goat farms associated with a Q fever outbreak', *Applied and Environmental Microbiology*. American Society for Microbiology (ASM), 79(5), pp. 1697–1703. doi: 10.1128/AEM.03472-12.

Kim, Y. *et al.* (2018) 'Livestock trade network: potential for disease transmission and implications for risk-based surveillance on the island of Mayotte', *Scientific Reports*. Nature Publishing Group, 8(1), p. 11550. doi: 10.1038/s41598-018-29999-y.

Kim, Y. *et al.* (2021) 'The role of livestock movements in the spread of rift valley fever virus in animals and humans in Mayotte, 2018–19', *PLoS Neglected Tropical Diseases*. Public Library of Science, 15(3), p. e0009202. doi: 10.1371/journal.pntd.0009202.

Kinsley, A. C. *et al.* (2020) 'Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology', *Frontiers in Veterinary Science*. Frontiers Media S.A., p. 596. doi: 10.3389/fvets.2020.00596.

Kipyego, E. S. *et al.* (2020) 'Sero-prevalence and risk factors of infectious bovine rhinotracheitis virus (type 1) in Meru County, Kenya', *Preventive Veterinary Medicine*. Elsevier B.V., 175, p. 104863. doi: 10.1016/j.prevetmed.2019.104863.

Kiss, I. Z., Green, D. M. and 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(10).

Kivaria, F. M. (2003) 'Foot and mouth disease in Tanzania: An overview of its national status', *Veterinary Quarterly*. Taylor & Francis Group , 25(2), pp. 72–78. doi: 10.1080/01652176.2003.9695147.

Kivelä, M. *et al.* (2014) 'Multilayer networks', *Journal of Complex Networks*. Oxford University Press, 2(3), pp. 203–271. doi: 10.1093/comnet/cnu016.

Klous, G. *et al.* (2016) 'Human–livestock contacts and their relationship to transmission of zoonotic pathogens, a systematic review of literature', *One Health*, 2, pp. 65–76. doi: 10.1016/j.onehlt.2016.03.001.

Knight-Jones, T. J. D. and Rushton, J. (2013) 'The economic impacts of foot and mouth disease - What are they, how big are they and where do they occur?', *Preventive Veterinary Medicine*. Elsevier B.V., pp. 161–173. doi: 10.1016/j.prevetmed.2013.07.013.

Komba, E. V. G. *et al.* (2012) 'Sanitary practices and occurrence of zoonotic conditions in cattle at slaughter in Morogoro Municipality, Tanzania: implications for public health', *Tanzania Journal of Health Research*. National Institute for Medical Research, 14(2). doi: 10.4314/thrb.v14i2.6.

Konschake, M. *et al.* (2013) 'On the Robustness of In- and Out-Components in a Temporal Network', *PLoS ONE*. Public Library of Science, 8(2), p. e55223. doi: 10.1371/journal.pone.0055223.

Kossaibati, M. A. and Esslemont, R. J. (1997) 'The costs of production diseases in dairy herds in England', *The Veterinary Journal*. W.B. Saunders, 154(1), pp. 41–51. doi: 10.1016/S1090-0233(05)80007-3.

Lambert, S. *et al.* (2018) 'High shedding potential and significant individual heterogeneity in naturally-infected Alpine ibex (Capra ibex) with Brucella melitensis', *Frontiers in Microbiology*. Frontiers Media S.A., 9(MAY), p. 1065. doi: 10.3389/fmicb.2018.01065.

Lankester, F. *et al.* (2015) 'The Economic Impact of Malignant Catarrhal Fever on Pastoralist Livelihoods', *PLOS ONE*. Edited by J. P. Stewart. Prentice Hall, 10(1), p. e0116059. doi: 10.1371/journal.pone.0116059.

Lanyon, S. R. and Reichel, M. P. (2013) 'Understanding the Impact and Control of Bovine Viral Diarrhoea in Cattle Populations', *Springer Science Reviews*. Springer Science and Business Media LLC, 1(1–2), pp. 85–93. doi: 10.1007/s40362-013-0007-6.

Latora, V. and Marchiori, M. (2001) 'Efficient Behavior of Small-World Networks', *Physical Review Letters*. American Physical Society, 87(19), p. 198701. doi: 10.1103/PhysRevLett.87.198701.

Lau, M. S. Y. *et al.* (2015) ‘A Systematic Bayesian Integration of Epidemiological and Genetic Data’, *PLoS Computational Biology*. Public Library of Science, 11(11), p. 1004633. doi: 10.1371/journal.pcbi.1004633.

Lavelle, M. J. *et al.* (2016) ‘Evaluating wildlife-cattle contact rates to improve the understanding of dynamics of bovine tuberculosis transmission in Michigan, USA’, *Preventive Veterinary Medicine*, 135, pp. 28–36. doi: 10.1016/j.prevetmed.2016.10.009.

Lee, D. R. and Barrett, C. B. (2001) *Tradeoffs or synergies? Agricultural intensification, economic development, and the environment*, *Wallingford, Oxon, UK; New York, NY*. CABI Pub. doi: 10.1111/j.0002-9092.2005.720_1.x.

Lefcheck, J. S. (2016) ‘piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics’, *Methods in Ecology and Evolution*. Edited by R. Freckleton. British Ecological Society, 7(5), pp. 573–579. doi: 10.1111/2041-210X.12512.

León, E. A. *et al.* (2006) ‘A description of cattle movements in two departments of Buenos Aires province, Argentina’, *Preventive Veterinary Medicine*. Elsevier, 76(1–2), pp. 109–120. doi: 10.1016/J.PREVETMED.2006.04.010.

Leventhal, G. E. *et al.* (2015) ‘Evolution and emergence of infectious diseases in theoretical and real-world networks.’, *Nature communications*. Nature Publishing Group, 6, p. 6101. doi: 10.1038/ncomms7101.

Lichoti, J. K. *et al.* (2016) ‘Social network analysis provides insights into African swine fever epidemiology’, *Preventive Veterinary Medicine*. Elsevier B.V., 126, pp. 1–10. doi: 10.1016/j.prevetmed.2016.01.019.

Liljeros, F. *et al.* (2001) ‘Social networks: The web of human sexual contacts’, *Nature*. Nature, 411(6840), pp. 907–908. doi: 10.1038/35082140.

Lindström, T. *et al.* (2013) ‘A Bayesian Approach for Modeling Cattle Movements in the United States: Scaling up a Partially Observed Network’, *PLoS ONE*. Edited by J. Gomez-Gardenes. Public Library of Science, 8(1), p. e53432. doi: 10.1371/journal.pone.0053432.

Little, P. D. (2009) *Hidden value on the hoof: Cross-border livestock trade in East Africa. Common Market for Eastern and Southern Africa Comprehensive African Agriculture Development Programme.*

Lloyd-Smith, J. O. *et al.* (2005) ‘Superspreading and the effect of individual variation on disease emergence’, *Nature*. Nature Publishing Group, 438(7066), pp. 355–359. doi: 10.1038/nature04153.

van der Loo, M. P. J. (2014) ‘The stringdist package for approximate string matching’, *R Journal*. Technische Universitaet Wien, 6(1), pp. 111–122. doi: 10.32614/rj-2014-011.

Machang’u, R., Mgode, G. and Mpanduji, D. (1997) ‘Leptospirosis in animals and humans in selected areas of Tanzania', *Belgian Journal of Zoology (Belgium)*. Koninklijke Belgische Vereniging voor Dierkunde, 127, p. 97:104.

Mahmoud, M. and Allam, A. (2013) 'Seroprevalence of Bovine Viral Diarrhea Virus (BVDV), Bovine Herpes Virus Type 1 (BHV-1), Parainfluenza Type 3 Virus (PI-3V) and Bovine Respiratory Syncytial Virus (BRSV) among non Vaccinated Cattle', *Global veterinaria*, 10(3), pp. 348–353. doi: 10.5829/idosi.gv.2013.10.3.72119.

De Maio, N., Wu, C. H. and Wilson, D. J. (2016) 'SCOTTI: Efficient Reconstruction of Transmission within Outbreaks with the Structured Coalescent', *PLoS Computational Biology*. Public Library of Science, 12(9), p. e1005130. doi: 10.1371/journal.pcbi.1005130.

Makuru Nyarobi, J. (2019) *The epidemiology of Rift Valley fever in northern Tanzania*. PhD, University of Glasgow.

Mardones, F. *et al.* (2010) 'Parameterization of the duration of infection stages of serotype O foot-and-mouth disease virus: An analytical review and meta-analysis with application to simulation models', *Veterinary Research*. BioMed Central, 41(4), p. 45. doi: 10.1051/vetres/2010017.

Marquetoux, N. *et al.* (2016) 'Using social network analysis to inform disease control interventions', *Preventive Veterinary Medicine*. Elsevier, 126, pp. 94–104.

Martin, V. *et al.* (2011) 'Risk-based surveillance for avian influenza control along poultry market chains in South China: The value of social network analysis', *Preventive Veterinary Medicine*. Elsevier, 102(3), pp. 196–205. doi: 10.1016/J.PREVETMED.2011.07.007.

Martínez-López, Perez, A. M. and Sánchez-Vizcaíno, J. M. (2009a) 'Combined application of social network and cluster detection analyses for temporal-spatial characterization of animal movements in Salamanca, Spain', *Preventive Veterinary Medicine*. Elsevier, 91(1), pp. 29–38. doi: 10.1016/J.PREVETMED.2009.05.007.

Martínez-López, Perez, A. M. and Sánchez-Vizcaíno, J. M. (2009b) 'Social Network Analysis. Review of General Concepts and Use in Preventive Veterinary Medicine', *Transboundary and Emerging Diseases*, 56(4), pp. 109–120. doi: 10.1111/j.1865-1682.2009.01073.x.

Masucci, A. P. *et al.* (2013) 'Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. Phys Rev E Stat Nonlin Soft Matter Phys, 88(2). doi: 10.1103/PhysRevE.88.022812.

Masuda, N. and Holme, P. (2013) 'Predicting and controlling infectious disease epidemics using temporal networks.', *F1000prime reports*, 5(6), p. 6. doi: 10.12703/P5-6.

Mathew, C. *et al.* (2017) 'Reproductive Infections in Cattle in Tanzania – Lessons for Control Priorities', *SOJ Microbiology & Infectious Diseases*. Symbiosis Group, 5(2), pp. 1–9.

doi: 10.15226/sojmid/5/2/00169.

Maudlin, I., Eisler, M. C. and Welburn, S. C. (2009) ‘Neglected and endemic zoonoses.’, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 364(1530), pp. 2777–87. doi: 10.1098/rstb.2009.0067.

Maze, M. J. *et al.* (2018) ‘Risk factors for human acute leptospirosis in northern Tanzania’, *PLOS Neglected Tropical Diseases*. Edited by J. Foley. Public Library of Science, 12(6), p. e0006372. doi: 10.1371/journal.pntd.0006372.

Maziku, M., Gebru, G. and Stapleton, J. I. (2017) *Livestock health priorities in the Tanzania livestock master plan*. CGIR, ILRI, Livestock Master Plan Development Project.

McClowry, A. (2014) ‘A tank of cold: cleantech leapfrog to a more food secure world: Case Study: Tanzania’, *Climate change, Energy, Agriculture and Food*.

McGrann, J. and Wiseman, H. (2001) ‘Animal traceability across national frontiers in the European Union’, *Revue scientifique et technique (International Office of Epizootics)*, 20(2), pp. 406–412.

Mdetele, D. P. and Kassanga, C. (2014) ‘Seroprevalence of foot and mouth disease in Tanzania’, *International Journal of Infectious Diseases*. Elsevier BV, 21, pp. 451–452. doi: 10.1016/j.ijid.2014.03.1352.

Meat Livestock Austrailia (2017) *Market snapshot | Sheep meat | MENA (Middle East & North Africa)*.

Meehan, C. *et al.* (2018) ‘The relationship between transmission time and clustering methods in Mycobacterium tuberculosis epidemiology’, *The relationship between transmission time and clustering methods in Mycobacterium tuberculosis epidemiology*. Cold Spring Harbor Laboratory, p. 302232. doi: 10.1101/302232.

Metcalfe, L. and Metcalfe, L. V. A. (2019) ‘An Update on the Status of BVD Control and Eradication in Europe’, *J Veter Sci Med*, 7(1).

Métras, R. *et al.* (2020) ‘Estimation of Rift Valley fever virus spillover to humans during the Mayotte 2018-2019 epidemic’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, pp. 24567–24574. doi: 10.1073/pnas.2004468117.

Meyers, L. A. *et al.* (2003) ‘Applying Network Theory to Epidemics: Control Measures for *Mycoplasma pneumoniae* Outbreaks’, *Emerging Infectious Diseases*, 9(2), pp. 204–210. doi: 10.3201/eid0902.020188.

Meyers, L. A., Newman, M. E. J. and Pourbohloul, B. (2006) ‘Predicting epidemics on directed contact networks’, *Journal of Theoretical Biology*. J Theor Biol, 240(3), pp. 400–418. doi: 10.1016/j.jtbi.2005.10.004.

Mibei, H. *et al.* (2017) ‘KNOWLEDGE FOR LIFE Mobile Landscape Analysis: Tanzania’. doi: 10.1079/CABICOMM-36-3825.

Michael, S. *et al.* (2018) *Tanzania livestock master plan Developed by the Tanzania Ministry of Livestock and Fisheries (MLF) and the International Livestock Research Institute (ILRI) livestock master plan team*.

Michel, L., Brun, F. and Makowski, D. (2017) ‘A framework based on generalised linear mixed models for analysing pest and disease surveys’, *Crop Protection*. Elsevier Ltd, 94, pp. 1–12. doi: 10.1016/j.cropro.2016.12.013.

Min, B., Goh, K. I. and Vazquez, A. (2011) ‘Spreading dynamics following bursty human activity patterns’, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. Phys Rev E Stat Nonlin Soft Matter Phys, 83(3). doi: 10.1103/PhysRevE.83.036102.

Mineo, T. W. P. *et al.* (2006) ‘Distribution of antibodies against Neospora caninum, BVDV AND BHV-1 among cows in Brazilian dairy herds with reproductive disorders’, *Brazil. J. Vet. Parasitol*, 15(4), pp. 188–192.

Ministry for Primary Industries, N. Z. (2018) *National Animal Identification and Tracing | MPI - Ministry for Primary Industries. A New Zealand Government Department.* Available at: https://www.mpi.govt.nz/growing-and-harvesting/livestock-and-animal-care/national-animal-identification-and-tracing/ (Accessed: 25 June 2018).

Molia, S. *et al.* (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*. Elsevier, 155, pp. 77–88. doi: 10.1016/J.ACTATROPICA.2015.12.003.

Molyneux, D. *et al.* (2011) ‘Zoonoses and marginalised infectious diseases of poverty: Where do we stand?’, *Parasites & Vectors*. BioMed Central, 4(1), p. 106. doi: 10.1186/1756-3305-4-106.

Morelli, M. J. *et al.* (2012) ‘A Bayesian Inference Framework to Reconstruct Transmission Trees Using Epidemiological and Genetic Data’, *PLoS Computational Biology*. Public Library of Science, 8(11), p. 1002768. doi: 10.1371/journal.pcbi.1002768.

Morens, D. M., Folkers, G. K. and Fauci, A. S. (2004) ‘The challenge of emerging and re-emerging infectious diseases’, *Nature*. Nature Publishing Group, 430(6996), pp. 242–249. doi: 10.1038/nature02759.

Motta, P. *et al.* (2017) ‘Implications of the cattle trade network in Cameroon for regional disease prevention and control.’, *Scientific reports*. Nature Publishing Group, 7, p. 43932. doi: 10.1038/srep43932.

Motta, P. *et al.* (2018) ‘Drivers of Live Cattle Price in the Livestock Trading System of Central Cameroon’, *Frontiers in Veterinary Science*. Frontiers, 4, p. 244. doi:

10.3389/fvets.2017.00244.

Motta, P. *et al.* (2019) ‘Characterizing livestock markets, primary diseases, and key management practices along the livestock supply chain in Cameroon’, *Frontiers in Veterinary Science*. Frontiers Media S.A., 6(APR), p. 101. doi: 10.3389/fvets.2019.00101.

Mtimet, N. *et al.* (2021) ‘Zoonotic diseases and the COVID-19 pandemic: Economic impacts on Somaliland’s livestock exports to Saudi Arabia’, *Global Food Security*. Elsevier B.V., 28, p. 100512. doi: 10.1016/j.gfs.2021.100512.

Muma, J. B. *et al.* (2014) ‘The contribution of veterinary medicine to public health and poverty reduction in developing countries’, *Veterinaria italiana*, 50(2), pp. 117–29. doi: 10.12834/VetIt.1405.323.

Munyua, P. *et al.* (2016) ‘Prioritization of zoonotic diseases in Kenya, 2015’, *PLoS ONE*. Public Library of Science, 11(8), p. e0161576. doi: 10.1371/journal.pone.0161576.

Musemwa, L. *et al.* (2012) ‘The Impact of Climate Change on Livestock The Impact of Climate Change on Livestock Production amongst the Resource-Poor Farmers of Third World Countries: A Review’, *Asian Journal of Agriculture and Rural Development*, 2(4), p. 621:631.

Mutua, F. *et al.* (2018) ‘Piloting a livestock identification and traceability system in the northern Tanzania–Narok–Nairobi trade route’, *Tropical Animal Health and Production*. Springer Netherlands, 50(2), pp. 299–308. doi: 10.1007/s11250-017-1431-4.

Myers, S. S. *et al.* (2017) ‘Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition’, *Annual Review of Public Health*. Annual Reviews Inc., pp. 259–277. doi: 10.1146/annurev-publhealth-031816-044356.

Nair, M. K., Augustine, L. F. and Konapur, A. (2016) ‘Food-Based Interventions to Modify Diet Quality and Diversity to Address Multiple Micronutrient Deficiency’, *Frontiers in Public Health*. Frontiers Media SA, 3. doi: 10.3389/fpubh.2015.00277.

Nakagawa, S., Johnson, P. C. D. and Schielzeth, H. (2017) ‘The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded’, *Journal of The Royal Society Interface*. Royal Society Publishing, 14(134), p. 20170213. doi: 10.1098/rsif.2017.0213.

Nakagawa, S. and Schielzeth, H. (2013) ‘A general and simple method for obtaining R2 from generalized linear mixed-effects models’, *Methods in Ecology and Evolution*. John Wiley & Sons, Ltd, 4(2), pp. 133–142. doi: 10.1111/J.2041-210X.2012.00261.X@10.1111/(ISSN)2041-210X.TOPMETHODS.

Nandonde, S., Gebru, G. and Stapleton, J. (2017) ‘Red meat production in the Tanzanian livestock master plan’, *Tanzania LMP Brief*, 8.

Napp, S. *et al.* (2018) 'Understanding the legal trade of cattle and camels and the derived risk of Rift Valley Fever introduction into and transmission within Egypt', *PLoS Neglected Tropical Diseases*. Public Library of Science, 12(1), p. e0006143. doi: 10.1371/journal.pntd.0006143.

Di Nardo, A., Knowles, N. J. and Paton, D. J. (2011) 'Combining livestock trade patterns with phylogenetics to help understand the spread of foot and mouth disease in sub-Saharan Africa, the Middle East and Southeast Asia', *OIE Revue Scientifique et Technique*. Office International des Epizootes, 30(1), pp. 63–85. doi: 10.20506/rst.30.1.2022.

Natale, F. *et al.* (2009) 'Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread', *Preventive Veterinary Medicine*, 92(4), pp. 341–350. doi: 10.1016/j.prevetmed.2009.08.026.

Natale, F. *et al.* (2011) 'Evaluation of risk and vulnerability using a Disease Flow Centrality measure in dynamic cattle trade networks', *Preventive Veterinary Medicine*, 98(2–3), pp. 111–118. doi: 10.1016/j.prevetmed.2010.11.013.

National Bureau of Statistics (2012) *The United Republic of Tanzania 2012 Population and Housing Census, Tanzania in Figures 2012*. Available at: http://nbs.go.tz/nbstz/index.php/english/tanzania-in-figures/229-tanzania-in-figures-2012 (Accessed: 14 September 2018).

National Bureau of Statistics (2013) *The United Republic of Tanzania 2012 POPULATION AND HOUSING CENSUS Population Distribution by Administrative Areas National Bureau of Statistics Ministry of Finance Dar es Salaam*. Dar es Salaam and Zanzibar.

National Bureau of Statistics and Tanzania Government (2016) *Annual agriculture sample survey crop and livestock report 2016/2017*.

Newman, M. E. J. (2002) 'Spread of epidemic disease on networks', *Physical Review E*. American Physical Society, 66(1), p. 016128. doi: 10.1103/PhysRevE.66.016128.

Newman, M. E. J. (2010) *Networks An Introduction*. 1st edn. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199206650.001.0001.

Newman, M. E. J. and Girvan, M. (2004) 'Finding and evaluating community structure in networks', *Physical Review E*. American Physical Society, 69(2), p. 026113. doi: 10.1103/PhysRevE.69.026113.

Ng'asike, O. P., Hagmann, T. and Wasonga, O. V. (2021) 'Brokerage in the borderlands: the political economy of livestock intermediaries in northern Kenya', *Journal of Eastern African Studies*. Routledge, 15(1), pp. 168–188. doi: 10.1080/17531055.2020.1845041.

Nickbakhsh, S. *et al.* (2011) 'Generating social network data using partially described networks: An example informing avian influenza control in the British poultry industry', *BMC Veterinary Research*. BioMed Central, 7(1), pp. 1–16. doi: 10.1186/1746-6148-7-66.

Nickbakhsh, S. *et al.* (2013) ‘Implications of within-farm transmission for network dynamics: Consequences for the spread of avian influenza’, *Epidemics*. Epidemics, 5(2), pp. 67–76. doi: 10.1016/j.epidem.2013.03.001.

Nickbakhsh, S. *et al.* (2014) ‘A metapopulation model for highly pathogenic avian influenza: Implications for compartmentalization as a control measure’, *Epidemiology and Infection*. Cambridge University Press, 142(9), pp. 1813–1825. doi: 10.1017/S0950268813002963.

Nicolas, G. *et al.* (2013) ‘Description and analysis of the cattle trade network in the Madagascar highlands: Potential role in the diffusion of Rift Valley fever virus’, *Acta Tropica*. Elsevier, 126(1), pp. 19–27. doi: 10.1016/J.ACTATROPICA.2012.12.013.

Nicolas, G. *et al.* (2018) ‘Predictive gravity models of livestock mobility in Mauritania: The effects of supply, demand and cultural factors’, *PLOS ONE*. Edited by R. Muneepeerakul. Public Library of Science, 13(7), p. e0199547. doi: 10.1371/journal.pone.0199547.

Njeru, J. *et al.* (2016) ‘Q fever is an old and neglected zoonotic disease in Kenya: a systematic review.’, *BMC public health*. BioMed Central, 16, p. 297. doi: 10.1186/s12889-016-2929-9.

Nöremark, M. *et al.* (2011) ‘Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk based surveillance’, *Preventive Veterinary Medicine*. Elsevier, 99(2–4), pp. 78–90. doi: 10.1016/J.PREVETMED.2010.12.009.

Norris, N. (1940) ‘The Standard Errors of the Geometric and Harmonic Means and Their Application to Index Numbers’, *The Annals of Mathematical Statistics*. Institute of Mathematical Statistics, 11(4), pp. 445–448. doi: 10.1214/aoms/1177731830.

Ntokwane, P. and Dibeela, K. (2016) ‘Utilisation of ICT for improved livestock traceability: a case for Botswana, a developing country perspective’.

Numminen, E. *et al.* (2014) ‘Two-phase importance sampling for inference about transmission trees’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society of London, 281(1794). doi: 10.1098/rspb.2014.1324.

Ogola, J. *et al.* (2018) ‘The topology of between-herd cattle contacts in a mixed farming production system in western Kenya’, *Preventive Veterinary Medicine*. Elsevier B.V., 158, pp. 43–50. doi: 10.1016/j.prevetmed.2018.06.010.

Okumu, T. A. *et al.* (2019) ‘Seroprevalence of antibodies for bovine viral diarrhoea virus, Brucella abortus and Neospora caninum, and their roles in the incidence of abortion/foetal loss in dairy cattle herds in Nakuru District, Kenya’, *BMC Veterinary Research*. BioMed Central Ltd., 15(1), pp. 1–6. doi: 10.1186/s12917-019-1842-8.

Oliveira, J. G. and Barabási, A. L. (2005) ‘Darwin and Einstein correspondence patterns’, *Nature*. Nature Publishing Group, 437(7063), p. 1251. doi: 10.1038/4371251a.

Omondi, G. *et al.* (2019) ‘Phylogeographical and cross-species transmission dynamics of SAT1 and SAT2 foot-and-mouth disease virus in Eastern Africa’, *Molecular Ecology*. Blackwell Publishing Ltd, 28(11), pp. 2903–2916. doi: 10.1111/mec.15125.

Omondi, G. P. *et al.* (2021) ‘Animal movement in a pastoralist population in the Maasai Mara Ecosystem in Kenya and implications for pathogen spread and control’, *Preventive Veterinary Medicine*. Elsevier B.V., 188, p. 105259. doi: 10.1016/j.prevetmed.2021.105259.

Ortiz-Pelaez, A. *et al.* (2006) ‘Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK’, *Preventive Veterinary Medicine*, 76(1–2), pp. 40–55. doi: 10.1016/j.prevetmed.2006.04.007.

Otte, J. and Chilonda, P. (2003) ‘Classification of cattle and small ruminant production systems in Sub-Saharan Africa’, *Outlook on Agriculture*. IP Publishing Ltd, 32(3), pp. 183–190. doi: 10.5367/000000003101294451.

Palisson, A. *et al.* (2016) ‘Role of Cattle Movements in Bovine Tuberculosis Spread in France between 2005 and 2014.’, *PloS one*. Edited by M. V. Cunha. Public Library of Science, 11(3), p. e0152578. doi: 10.1371/journal.pone.0152578.

Patz, J. A. and Khaliq, M. (2002) ‘Global Climate Change and Health: Challenges for Future Practitioners’, *JAMA*. American Medical Association, 287(17), p. 2283. doi: 10.1001/jama.287.17.2283-JMS0501-3-1.

Pauw, K. and Thurlow, J. (2011) ‘Agricultural growth, poverty, and nutrition in Tanzania’, *Food Policy*. Pergamon, 36(6), pp. 795–804. doi: 10.1016/j.foodpol.2011.09.002.

Perry, B. and Grace, D. (2009) ‘The impacts of livestock diseases and their control on growth and development processes that are pro-poor’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1530), pp. 2643–2655. doi: 10.1098/rstb.2009.0097.

Pica-Ciamarra, Ugo *et al.* (2011a) ‘Linking Smallholders to Livestock Markets in Tanzania : Combing Market and Household Survey Data’. World Bank, Washington, DC.

Pica-Ciamarra, Ugo *et al.* (2011b) ‘Linking Smallholders to Livestock Markets in Tanzania : Combing Market and Household Survey Data’.

Pieracci, E. G. *et al.* (2016) ‘Prioritizing zoonotic diseases in Ethiopia using a one health approach’, *One Health*. Elsevier B.V., 2, pp. 131–135. doi: 10.1016/j.onehlt.2016.09.001.

Pilosof, S. *et al.* (2017) ‘The multilayer nature of ecological networks’, *Nature Ecology and Evolution*. Nature Publishing Group, pp. 1–9. doi: 10.1038/s41559-017-0101.

Poester, F. P., Samartino, L. E. and Santos, &r L. (2013) ‘Pathogenesisandpathobiology ofbrucellosisinlivestock’, *Rev.sci.tech.Off.int.Epiz*, 32(1), pp. 105–115. doi: 10.20506/rst.32.1.2193.

Pons, P. and Latapy, M. (2006) ‘Computing Communities in Large Networks Using Random Walks’, *Journal of Graph Algorithms and Applications*, 10(2), pp. 191–218.

Poolkhet, C. *et al.* (2013) ‘Social network analysis used to assess the relationship between the spread of avian influenza and movement patterns of backyard chickens in Ratchaburi, Thailand’, *Research in Veterinary Science*. W.B. Saunders, 95(1), pp. 82–86. doi: 10.1016/J.RVSC.2013.02.016.

Poolkhet, C. *et al.* (2016) ‘Social network analysis of cattle movement in Kampong Cham, Kampong Speu and Takeo, Cambodia’, *Acta Tropica*, 159, pp. 44–49. doi: 10.1016/j.actatropica.2016.03.027.

Porphyre, T. *et al.* (2020) ‘Multilayer network analysis unravels haulage vehicles as a hidden threat to the British swine industry’, *Transboundary and Emerging Diseases*. Blackwell Publishing Ltd, 67(3), pp. 1231–1246. doi: 10.1111/tbed.13459.

Prabhu, M. *et al.* (2011) ‘Q Fever, Spotted Fever Group, and Typhus Group Rickettsioses Among Hospitalized Febrile Patients in Northern Tanzania’, *Clinical Infectious Diseases*. Department of Child and Adolescent Health and Development, Geneva, Switzerland, 53(4), pp. e8–e15. doi: 10.1093/cid/cir411.

Pradère, J.-P. (2014) ‘Improving animal health and livestock productivity to reduce poverty.’, *Revue scientifique et technique (International Office of Epizootics)*, 33(3), pp. 735–44, 723–34.

Prentice, J. C. *et al.* (2017) ‘Complex responses to movement-based disease control: when livestock trading helps.’, *Journal of the Royal Society, Interface*. The Royal Society, 14(126), p. 20160531. doi: 10.1098/rsif.2016.0531.

Pybus, O. G., Tatem, A. J. and Lemey, P. (2015) ‘Virus evolution and transmission in an ever more connected world’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society of London. doi: 10.1098/rspb.2014.2878.

Qi, L. *et al.* (2019) ‘Neighbourhood contacts and trade movements drive the regional spread of bovine viral diarrhoea virus (BVDV)’, *Veterinary Research*. BioMed Central Ltd., 50(1), p. 30. doi: 10.1186/s13567-019-0647-x.

R Core Team (2019a) ‘R: A language and environment for statistical computing. R Foundation for Statistical Computing’. Vienna, Austria.

R Core Team (2019b) ‘R: The R Project for Statistical Computing’. Vienna, Austria.

Raizman, E. A. *et al.* (2013) ‘Feasibility study on the spatial and temporal movement of Samburu’s cattle and wildlife in Kenya using GPS radio-tracking, remote sensing and GIS’, *Preventive Veterinary Medicine*, 111(1–2), pp. 76–80. doi: 10.1016/j.prevetmed.2013.04.007.

Randolph, T. F. *et al.* (2007) ‘Invited Review: Role of livestock in human nutrition and health for poverty reduction in developing countries1,2,3’, *Journal of Animal Science*. Oxford University Press, 85(11), pp. 2788–2800. doi: 10.2527/jas.2007-0467.

Rasamoelina-Andriamanivo, H. *et al.* (2014) ‘Description and analysis of the poultry trading network in the Lake Alaotra region, Madagascar: Implications for the surveillance and control of Newcastle disease’, *Acta Tropica*. Elsevier, 135, pp. 10–18. doi: 10.1016/J.ACTATROPICA.2014.03.008.

Rasmussen, D. A., Volz, E. M. and Koelle, K. (2014) ‘Phylodynamic Inference for Structured Epidemiological Models’, *PLoS Computational Biology*. Public Library of Science, 10(4), p. 1003570. doi: 10.1371/journal.pcbi.1003570.

Rautureau, S., Dufour, B. and Durand, B. (2011) ‘Vulnerability of Animal Trade Networks to The Spread of Infectious Diseases: A Methodological Approach Applied to Evaluation and Emergency Control Strategies in Cattle, France, 2005’, *Transboundary and Emerging Diseases*. Blackwell Publishing Ltd, 58(2), pp. 110–120. doi: 10.1111/j.1865-1682.2010.01187.x.

Rich, K. M. and Perry, B. D. (2011) ‘The economic and poverty impacts of animal diseases in developing countries: New roles, new demands for economics and epidemiology’, *Preventive Veterinary Medicine*. Elsevier, 101(3–4), pp. 133–147. doi: 10.1016/J.PREVETMED.2010.08.002.

Robinson, S. E. and Christley, R. M. (2007) ‘Exploring the role of auction markets in cattle movements within Great Britain’, *Preventive Veterinary Medicine*, 81(1–3), pp. 21–37. doi: 10.1016/j.prevetmed.2007.04.011.

Robinson, S. E., Everett, M. G. and Christley, R. M. (2007) ‘Recent network evolution increases the potential for large epidemics in the British cattle population’, *Journal of the Royal Society Interface*. Royal Society, 4(15), pp. 669–674. doi: 10.1098/rsif.2007.0214.

Roest, Hendrik I J *et al.* (2011) ‘Molecular epidemiology of Coxiella burnetii from ruminants in Q fever outbreak, the Netherlands.’, *Emerging infectious diseases*. Centers for Disease Control and Prevention, 17(4), pp. 668–75. doi: 10.3201/eid1704.101562.

Roest, H. I.J. *et al.* (2011) ‘The Q fever epidemic in the Netherlands: History, onset, response and reflection’, *Epidemiology and Infection*. Cambridge University Press, pp. 1–12. doi: 10.1017/S0950268810002268.

Romero-Severson, E. *et al.* (2014) ‘Timing and order of transmission events is not directly reflected in a pathogen phylogeny’, *Molecular Biology and Evolution*. Oxford University Press, 31(9), pp. 2472–2482. doi: 10.1093/molbev/msu179.

van Roon, A. M. *et al.* (2020) ‘Quantification of risk factors for bovine viral diarrhea virus in cattle herds: A systematic search and meta-analysis of observational studies’, *Journal of Dairy Science*. Elsevier Inc., 103(10), pp. 9446–9463. doi: 10.3168/jds.2020-18193.

Rossi, G. *et al.* (2015) ‘Epidemiological modelling for the assessment of bovine tuberculosis surveillance in the dairy farm network in Emilia-Romagna (Italy)’, *Epidemics*. Elsevier B.V., 11, pp. 62–70. doi: 10.1016/j.epidem.2015.02.007.

Rossi, G., Smith, R. L., *et al.* (2017) ‘Modelling farm-to-farm disease transmission through personnel movements: From visits to contacts, and back’, *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–11. doi: 10.1038/s41598-017-02567-6.

Rossi, G., De Leo, G. A., *et al.* (2017) ‘The Potential Role of Direct and Indirect Contacts on Infection Spread in Dairy Farm Networks’, *PLOS Computational Biology*. Edited by M. (Matt) Ferrari, 13(1), p. e1005301. doi: 10.1371/journal.pcbi.1005301.

Rossi, G. *et al.* (2019) ‘The spread of bovine tuberculosis in Canadian shared pastures: Data, model, and simulations’, *Transboundary and Emerging Diseases*. Blackwell Publishing Ltd, 66(1), pp. 562–577. doi: 10.1111/tbed.13066.

Ruget, A. S. *et al.* (2021) ‘Multi-species temporal network of livestock movements for disease spread’, *Applied Network Science*. Springer Science and Business Media Deutschland GmbH, 6(1), pp. 1–20. doi: 10.1007/s41109-021-00354-x.

Rushmore, J. *et al.* (2014) ‘Network-based vaccination improves prospects for disease control in wild chimpanzees’, *Journal of The Royal Society Interface*. Royal Society, 11(97), p. 20140349. doi: 10.1098/rsif.2014.0349.

Rushton, J. *et al.* (2018) ‘Initiation of Global Burden of Animal Diseases Programme’, *The Lancet*. Lancet Publishing Group, pp. 538–540. doi: 10.1016/S0140-6736(18)31472-7.

Rushton, J., Uggla, A. and Magnusson, U. (2017) *Animal health in development - its role for poverty reduction and human welfare*. Stockholm.

Sah, P., Mann, J. and Bansal, S. (2018) ‘Disease implications of animal social network structure: A synthesis across social systems’, *Journal of Animal Ecology*. Edited by D. Farine. Wiley/Blackwell (10.1111), 87(3), pp. 546–558. doi: 10.1111/1365-2656.12786.

Salathé, M. *et al.* (2010) ‘A high-resolution human contact network for infectious disease transmission.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(51), pp. 22020–5. doi: 10.1073/pnas.1009094108.

Salvador, L. C. M. *et al.* (2018) ‘Risk-based strategies for surveillance of tuberculosis infection in cattle for low-risk areas in England and Scotland’, *Epidemiology and Infection*, 146(01), pp. 107–118. doi: 10.1017/S0950268817001935.

San Miguel, M. *et al.* (2012) ‘Challenges in complex systems science’, *European Physical Journal: Special Topics*. Springer, 214(1), pp. 245–271. doi: 10.1140/epjst/e2012-01694-y.

Schimmer, B. *et al.* (2012) ‘Seroprevalence and Risk Factors for Coxiella burnetii (Q Fever) Seropositivity in Dairy Goat Farmers’ Households in The Netherlands, 2009?2010’, *PLoS ONE*. Edited by D. S. Zamboni. Public Library of Science, 7(7), p. e42364. doi: 10.1371/journal.pone.0042364.

Schoonman, L. and Swai, E. S. (2010) ‘Herd- and animal-level risk factors for bovine leptospirosis in Tanga region of Tanzania’, *Tropical Animal Health and Production*, 42(7), pp. 1565–1572. doi: 10.1007/s11250-010-9607-1.

Segura-Correa, J. C. *et al.* (2016) ‘Seroprevalence and risk factors associated with bovine herpesvirus 1 and bovine viral diarrhea virus in north-eastern Mexico’, *Open Veterinary Journal*. Faculty of Veterinary Medicine, University of Tripoli and National Authority for Scientific Research, 6(2), pp. 143–149. doi: 10.4314/ovj.v6i2.12.

Seitz, R. (2014) ‘Coxiella burnetii - Pathogenic Agent of Q (Query) Fever’, *Transfusion Medicine and Hemotherapy*. S. Karger AG, 41(1), pp. 60–72. doi: 10.1159/000357107.

Selby, R. *et al.* (2013) ‘Cattle movements and trypanosomes: restocking efforts and the spread of Trypanosoma brucei rhodesiense sleeping sickness in post-conflict Uganda’, *Parasites & Vectors*. BioMed Central, 6(1), p. 281. doi: 10.1186/1756-3305-6-281.

Semango, G. *et al.* (2019) ‘The Sero-epidemiology of Neospora caninum in Cattle in Northern Tanzania’, *Frontiers in Veterinary Science*. Frontiers Media S.A., 6(SEP), p. 327. doi: 10.3389/fvets.2019.00327.

Sherman, D. M. (2011) ‘The spread of pathogens through trade in small ruminants and their products’, *Rev. sci. tech. Off. int. Epiz*, 30(1), pp. 207–217.

Shirima, G. M. and Kunda, J. S. (2016) ‘Prevalence of brucellosis in the human, livestock and wildlife interface areas of Serengeti National Park, Tanzania’, *Onderstepoort Journal of Veterinary Research*. Agricultural Research Council, 83(1), p. 4. doi: 10.4102/ojvr.v83i1.1032.

Shirley, Mark D F and Rushton, S. P. (2005) ‘The impacts of network topology on disease spread’. doi: 10.1016/j.ecocom.2005.04.005.

Shirley, M D and Rushton, S. P. (2005) ‘Where diseases and networks collide: lessons to be learnt from a study of the 2001 foot-and-mouth disease epidemic’, *Epidemiology and infection*. Epidemiol Infect, 133(6), pp. 1023–1032. doi: 10.1017/S095026880500453X.

Silk, M. J. *et al.* (2017) ‘The application of statistical network models in disease research’, *Methods in Ecology and Evolution*. Edited by J. Metcalf. Wiley/Blackwell (10.1111), 8(9), pp. 1026–1041. doi: 10.1111/2041-210X.12770.

Silk, M. J. *et al.* (2018) ‘Quantifying direct and indirect contacts for the potential transmission of infection between species using a multilayer contact network’, *Behaviour*. Brill Academic Publishers, 155(7–9), pp. 731–757. doi: 10.1163/1568539X-00003493.

Simini, F. *et al.* (2012) ‘A universal model for mobility and migration patterns’, *Nature*. Nature Publishing Group, 484(7392), pp. 96–100. doi: 10.1038/nature10856.

Sindato, C. *et al.* (2014) ‘Spatial and Temporal Pattern of Rift Valley Fever Outbreaks in Tanzania; 1930 to 2007’, *PLoS ONE*. Edited by T. Ikegami. Public Library of Science, 9(2), p. e88897. doi: 10.1371/journal.pone.0088897.

Sindato, C., Karimuribo, E. and Mboera, L. E. . (2012) ‘The epidemiology and socio-economic impact of rift valley fever in Tanzania: a review’, *Tanzania Journal of Health Research*. National Institute for Medical Research, 13(5). doi: 10.4314/thrb.v13i5.1.

Sinkala, Y. *et al.* (2014) ‘Challenges and economic implications in the control of foot and mouth disease in sub-Saharan Africa: Lessons from the Zambian experience’, *Veterinary Medicine International*. Hindawi Limited. doi: 10.1155/2014/373921.

Sintayehu, D. W. *et al.* (2017) ‘Disease transmission in animal transfer networks’, *Preventive Veterinary Medicine*, 137, pp. 36–42. doi: 10.1016/j.prevetmed.2016.12.017.

Smith, G. C. *et al.* (2008) ‘Post-slaughter traceability’, *Meat Science*. Elsevier, 80(1), pp. 66–74. doi: 10.1016/J.MEATSCI.2008.05.024.

Smith, R. L. *et al.* (2013) ‘Minimization of bovine tuberculosis control costs in US dairy herds’, *Preventive Veterinary Medicine*. Prev Vet Med, 112(3–4), pp. 266–275. doi: 10.1016/j.prevetmed.2013.07.014.

Smith, R. P., Cook, A. J. C. and Christley, R. M. (2013) ‘Descriptive and social network analysis of pig transport data recorded by quality assured pig farms in the UK’, *Preventive Veterinary Medicine*. Elsevier, 108(2–3), pp. 167–177. doi: 10.1016/J.PREVETMED.2012.08.011.

Smith, T. *et al.* (2004) ‘The contribution of small ruminants in alleviating poverty: communicating messages from research: Proceedings of the third DFID Livestock Production Programme Link Project (R7798) workshop for small ruminant keepers’, in. Izaak Walton Inn, Embu, Kenya, 4-7 February 2003. Natural Resources International Ltd, Aylesford, Kent, UK.

Solé, R. V. and Montoya, J. M. (2001) ‘Complexity and fragility in ecological networks’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 268(1480), pp. 2039–2045. doi: 10.1098/rspb.2001.1767.

Spiegel, K. A. and Havas, K. A. (2019) ‘The socioeconomic factors surrounding the initial emergence of peste des petits ruminants in Kenya, Uganda, and Tanzania from 2006 through 2008’, *Transboundary and Emerging Diseases*. Blackwell Publishing Ltd, 66(2), pp.

627–633. doi: 10.1111/tbed.13116.

Steinfeld, H. *et al.* (2006) ‘Livestock’s long shadow, Environmental issues and options’. FAO of the UN.

Stott, A. W. *et al.* (2012) ‘Predicted costs and benefits of eradicating BVDV from Ireland’, *Irish Veterinary Journal*. BioMed Central Ltd., 65(1). doi: 10.1186/2046-0481-65-12.

Sumaye, R. D. *et al.* (2015) ‘Inter-epidemic Acquisition of Rift Valley Fever Virus in Humans in Tanzania’, *PLOS Neglected Tropical Diseases*. Edited by R. C. Sang. Public Library of Science, 9(2), p. e0003536. doi: 10.1371/journal.pntd.0003536.

Swai, E. and Nesella, N. (2010) ‘Using Participatory Epidemiology tools to investigate Contagious Caprine Pleuropneumonia (CCPP) in Maasai flocks, northern Tanzania.’, *International Journal of Animal and Veterinary Advances*, 2(4), pp. 141–147.

Takaguchi, T., Masuda, N. and Holme, P. (2013) ‘Bursty Communication Patterns Facilitate Spreading in a Threshold-Based Epidemic Dynamics’, *PLoS ONE*. Public Library of Science, 8(7), p. 68629. doi: 10.1371/journal.pone.0068629.

Taylor, L. H., Latham, S. M. and Woolhouse, M. E. J. (2001) ‘Risk factors for human disease emergence’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1411), pp. 983–989. doi: 10.1098/rstb.2001.0888.

Tempia, S. *et al.* (2010) ‘Mapping cattle trade routes in southern Somalia: a method for mobile livestock keeping systems.’, *Revue scientifique et technique (International Office of Epizootics)*, 29(3), pp. 485–95.

Thompson, D. *et al.* (2002) ‘Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001’, *Rev. sci. tech. Off. int. Epiz.* , 21(3), pp. 675–687.

Thumbi, S. M. *et al.* (2015) ‘Linking human health and livestock health: A “one-health” platform for integrated analysis of human health, livestock health, and economic welfare in livestock dependent communities’, *PLoS ONE*. Public Library of Science, 10(3). doi: 10.1371/journal.pone.0120761.

Tildesley, M. J. *et al.* (2019) ‘The role of movement restrictions in limiting the economic impact of livestock infections’, *Nature Sustainability*. Nature Publishing Group, 2(9), pp. 834–840. doi: 10.1038/s41893-019-0356-5.

Tisdell, C. A. and Adamson, D. (2017) ‘The importance of fixed costs in animal health systems’, *OIE Revue Scientifique et Technique*. Office International des Epizootes, 36(1), pp. 49–56. doi: 10.20506/rst.36.1.2608.

Tomley, F. M. and Shirley, M. W. (2009) ‘Livestock infectious diseases and zoonoses’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, pp. 2637–2642. doi: 10.1098/rstb.2009.0133.

United Nations (2015) *Sustainable Development Goals | Department of Economic and Social Affairs | Sustainable Development*. Available at: https://sdgs.un.org/goals/goal2 (Accessed: 20 April 2021).

Upton, M. (2004) *A Living from Livestock The Role of Livestock in Economic Development and Poverty Reduction*. doi: 10.22004/AG.ECON.23783.

Valle, P. S. *et al.* (1999) 'Factors associated with being a bovine-virus diarrhoea (BVD) seropositive dairy herd in the More and Romsdal County of Norway', *Preventive Veterinary Medicine*. Elsevier, 40(3–4), pp. 165–177. doi: 10.1016/S0167-5877(99)00030-6.

Vallée, E. *et al.* (2013) 'Analysis of traditional poultry trader networks to improve risk-based surveillance', *The Veterinary Journal*, 195(1), pp. 59–65. doi: 10.1016/j.tvjl.2012.05.017.

Vanderburg, S. *et al.* (2014) 'Epidemiology of Coxiella burnetii Infection in Africa: A OneHealth Systematic Review', *PLoS Neglected Tropical Diseases*. Edited by M. K. Njenga. Public Library of Science, 8(4), p. e2787. doi: 10.1371/journal.pntd.0002787.

Vanderwaal, K. *et al.* (2016) 'Evaluating empirical contact networks as potential transmission pathways for infectious diseases', *J. R. Society Interface*, 13. doi: 10.1098/rsif.2016.0166.

VanderWaal, K., Enns, E. A., *et al.* (2017) 'Optimal surveillance strategies for bovine tuberculosis in a low-prevalence country', *Scientific Reports*. Nature Publishing Group, 7(1), p. 4140. doi: 10.1038/s41598-017-04466-2.

VanderWaal, K., Gilbertson, M., *et al.* (2017) 'Seasonality and pathogen transmission in pastoral cattle contact networks.', *Royal Society open science*. The Royal Society, 4(12), p. 170808. doi: 10.1098/rsos.170808.

VanderWaal, K. L. *et al.* (2014) 'Linking social and pathogen transmission networks using microbial genetics in giraffe (Giraffa camelopardalis)', *Journal of Animal Ecology*. John Wiley & Sons, Ltd, 83(2), pp. 406–414. doi: 10.1111/1365-2656.12137.

VanderWaal, K. L. *et al.* (2016) 'Network analysis of cattle movements in Uruguay: Quantifying heterogeneity for risk-based disease surveillance and control', *Preventive Veterinary Medicine*. Elsevier, 123, pp. 12–22. doi: 10.1016/J.PREVETMED.2015.12.003.

Vázquez, A. *et al.* (2006) 'Modeling bursts and heavy tails in human dynamics', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*. American Physical Society, 73(3), p. 036127. doi: 10.1103/PhysRevE.73.036127.

Vernon, M. C. (2011) 'Demographics of cattle movements in the United Kingdom', *BMC Veterinary Research*. BioMed Central, 7(1), p. 31. doi: 10.1186/1746-6148-7-31.

Vernon, M. C. and Keeling, M. J. (2009) ‘Representing the UK’s cattle herd as static and dynamic networks’, *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1656).

Viana, M. *et al.* (2016) ‘Integrating serological and genetic data to quantify cross-species transmission: Brucellosis as a case study’, *Parasitology*. Cambridge University Press, 143(7), pp. 821–834. doi: 10.1017/S0031182016000044.

Viboud, C. *et al.* (2006) ‘Synchrony, waves, and spatial hierarchies in the spread of influenza’, *Science*. Science, 312(5772), pp. 447–451. doi: 10.1126/science.1125237.

Virhia, J. (2019) *Healthy Animals, Healthy People: Lived Experiences of Zoonotic Febrile Illness in Northern Tanzania*. PhD, University of Glasgow.

Volkova, V. V. *et al.* (2010) ‘Sheep Movement Networks and the Transmission of Infectious Diseases’, *PLoS ONE*. Edited by A. P. Galvani. Public Library of Science, 5(6), p. e11185. doi: 10.1371/journal.pone.0011185.

Wang, D. J. *et al.* (2012) ‘Measurement error in network data: A re-classification’, *Social Networks*. doi: 10.1016/j.socnet.2012.01.003.

Wang, Q. *et al.* (2008) ‘Determination of the selection statistics and best significance level in backward stepwise logistic regression’, *Communications in Statistics: Simulation and Computation*. Taylor & Francis Group , 37(1), pp. 62–72. doi: 10.1080/03610910701723625.

Wasserman, S. and Faust, K. (1994) *Social network analysis : methods and applications. Structural analysis in the social sciences*. Cambridge; New York: Cambridge university press xxxi.

Watts, D. J. (1999) ‘Networks, Dynamics, and the Small-World Phenomenon’, *American Journal of Sociology*, 105(2), pp. 493–527. doi: 10.1086/210318.

Watts, D. J. and Strogatz, S. H. (1998) ‘Collective dynamics of “small-world” networks’, *Nature*. Nature Publishing Group, 393(6684), pp. 440–442. doi: 10.1038/30918.

Webster, J. P., Borlase, A. and Rudge, J. W. (2017) ‘Who acquires infection from whom and how? Disentangling multi-host and multimode transmission dynamics in the “elimination” era’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society. doi: 10.1098/rstb.2016.0091.

Wensman, J. J. *et al.* (2015) ‘A study of Rift Valley fever virus in Morogoro and Arusha regions of Tanzania - serology and farmers’ perceptions.’, *Infection ecology & epidemiology*. Taylor & Francis, 5, p. 30025. doi: 10.3402/IEE.V5.30025.

Wesolowski, A. *et al.* (2016) ‘Connecting mobility to infectious diseases: The promise and limits of mobile phone data’, *Journal of Infectious Diseases*. Oxford University Press, 214(suppl_4), pp. S414–S420. doi: 10.1093/infdis/jiw273.

White, L. A., Forester, J. D. and Craft, M. E. (2017) ‘Using contact networks to explore mechanisms of parasite transmission in wildlife’, *Biological Reviews*. Blackwell Publishing Ltd, 92(1), pp. 389–409. doi: 10.1111/brv.12236.

WHO *et al.* (2005) *The Control of Neglected Zoonotic Diseases A route to poverty alleviation*. WHO Headquarters, Geneva.

Wickham, H. (2016) *ggplot2 - Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Widgren, S. *et al.* (2016) ‘SimInf: An R package for Data-driven Stochastic Disease Spread Simulations’.

Wilber, M. Q. *et al.* (2019) ‘Modelling multi-species and multi-mode contact networks: Implications for persistence of bovine tuberculosis at the wildlife–livestock interface’, *Journal of Applied Ecology*. Blackwell Publishing Ltd, 56(6), pp. 1471–1481. doi: 10.1111/1365-2664.13370.

Williams, T. O., Spycher, B. and Okike, I. (2006) ‘Improving livestock marketing and intra-regional trade in West Africa: determining appropriate economic incentives and policy framework’, *ILRI (International Livestock Research Institute), Nairobi, Kenya*, p. 122.

Wilson, A. J. and Mellor, P. S. (2009) ‘Bluetongue in Europe: past, present and future.’, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 364(1530), pp. 2669–81. doi: 10.1098/rstb.2009.0091.

Wilson, R. T. (2018) ‘The red meat value chain in Tanzania’, *Animal Husbandry, Dairy and Veterinary Science*. Open Access Text Pvt, Ltd., 2(1). doi: 10.15761/ahdvs.1000127.

Wint, G. R. W. and Robinson, T. P. (2007) ‘Gridded livestock of the world 2007’, *FAO Rome*, p. 131.

Wodajo, H. D. *et al.* (2020) ‘Contribution of small ruminants to food security for Ethiopian smallholder farmers’, *Small Ruminant Research*. Elsevier B.V., 184, p. 106064. doi: 10.1016/j.smallrumres.2020.106064.

Woldeyohannes, S. M. *et al.* (2018) ‘Seroprevlance of Coxiella burnetii among abattoir and slaughterhouse workers: A meta-analysis’, *One Health*. Elsevier B.V., 6, pp. 23–28. doi: 10.1016/j.onehlt.2018.09.002.

Wongsathapornchai, K. *et al.* (2008) ‘Assessment of the likelihood of the introduction of foot-and-mouth disease through importation of live animals into the Malaysia-Thailand-Myanmar peninsula’, *American Journal of Veterinary Research*. American Veterinary Medical Association, 69(2), pp. 252–260. doi: 10.2460/ajvr.69.2.252.

Woods, A. (2004) *A Manufactured Plague: The History of Foot-And-mouth Disease in Britain*. London | Sterling VA: Earthscan.

Woolhouse, M. E. J. *et al.* (1997) ‘Heterogeneities in the transmission of infectious agents: Implications for the design of control programs’, *Proceedings of the National Academy of Sciences*, 94(1), pp. 338–342. doi: 10.1073/pnas.94.1.338.

Worby, C. J. *et al.* (2016) ‘Reconstructing transmission trees for communicable diseases using densely sampled genetic data’, *Annals of Applied Statistics*. Institute of Mathematical Statistics, 10(1), pp. 395–417. doi: 10.1214/15-AOAS898.

Worby, C. J., Lipsitch, M. and Hanage, W. P. (2014) ‘Within-Host Bacterial Diversity Hinders Accurate Reconstruction of Transmission Networks from Genomic Distance Data’, *PLoS Computational Biology*. Public Library of Science, 10(3), p. 1003549. doi: 10.1371/journal.pcbi.1003549.

World Bank *et al.* (2011) ‘Numbers for Livelihood Enhancement The Tanzania National Sample Census of Agriculture 2007/2008: A Livestock Perspective’, *Livestock Data Innovation in Africa BRIEF*, (4).

World Bank and TAFS Forum (2011) *World Livestock Disease Atlas : A Quantitative Analysis of Global Animal Health Data (2006-2009)*. World Bank, Washingto, DC and TAFS Forum, Bern.

World Organisation for Animal Health (OIE) (2013) *OIE Tool for the Evaluation of Performance of Veterinary Services Sixth Edition*, *World Organisation for Animal Health, 2013*.

Xia, Y., Bjørnstad, O. N. and Grenfell, B. T. (2004) ‘Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics.’, *The American naturalist*. The University of Chicago Press , 164(2), pp. 267–81. doi: 10.1086/422341.

Xie, F., Horan, R. D. and Wolf, C. A. (2009) ‘A gravity model approach to forecasting tuberculosis transmission in cattle’, *2009 Annual Meeting, July 26-28, 2009, Milwaukee, Wisconsin*. Agricultural and Applied Economics Association.

Ypma, R. J. F. *et al.* (2012) ‘Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data’, *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 279(1728), pp. 444–450. doi: 10.1098/rspb.2011.0913.

Ypma, R. J. F., van Ballegooijen, W. M. and Wallinga, J. (2013) ‘Relating phylogenetic trees to transmission trees of infectious disease outbreaks’, *Genetics*. Genetics, 195(3), pp. 1055–1062. doi: 10.1534/genetics.113.154856.

Zaal, F. *et al.* (2006) ‘9. The Geography of Integration: Cross-Border Livestock Trade in East Africa’, in *Pastoral Livestock Marketing in Eastern Africa*. Rugby, Warwickshire, United Kingdom: Practical Action Publishing, pp. 145–168. doi:

10.3362/9781780440323.009.

Zhang, C. *et al.* (2015) ‘Optimizing hybrid spreading in metapopulations’, *Scientific Reports*. Nature Publishing Group, 5(1), pp. 1–7. doi: 10.1038/srep09924.

Zhang, H. L. *et al.* (2016) ‘Mixed Methods Survey of Zoonotic Disease Awareness and Practice among Animal and Human Healthcare Providers in Moshi, Tanzania’, *PLOS Neglected Tropical Diseases*. Edited by M. Boelaert. Public Library of Science, 10(3), p. e0004476. doi: 10.1371/journal.pntd.0004476.

Zhao, D. *et al.* (2014) ‘Multiple routes transmitted epidemics on multiplex networks’, *Physics Letters, Section A: General, Atomic and Solid State Physics*. Elsevier B.V., 378(10), pp. 770–776. doi: 10.1016/j.physleta.2014.01.014.

Zinsstag, J. *et al.* (2007) ‘Human benefits of animal interventions for zoonosis control.’, *Emerging infectious diseases*. Centers for Disease Control and Prevention, 13(4), pp. 527–31. doi: 10.3201/eid1304.060381.

Zuur, A. F. *et al.* (2009) *Mixed effects models and extensions in ecology with R*. New York, NY: Springer New York (Statistics for Biology and Health). doi: 10.1007/978-0-387-87458-6.

Zuur, A. F., Ieno, E. N. and Elphick, C. S. (2010) ‘A protocol for data exploration to avoid common statistical problems’, *Methods in Ecology and Evolution*. Wiley, 1(1), pp. 3–14. doi: 10.1111/j.2041-210x.2009.00001.x.

9 Appendices

9.1 Chapter 2

9.1.1 Data analysis R scripts summary

1. HH_NW_1

/Users/gemmachaters/Dropbox/PHD/SEEDZ data/making new edge dfs.R

Opens all of SEEDZ household survey data, creates data frame of all livestock movements (number in each batch, destination/origin type and location) for cattle, sheep and goats for all possible movement types which include; sales to households, outside traders, at markets, to slaughter and gifts to; family, clan, friends, relatives and non-ageset persons.

Creates: "**AllEdges**.csv"

2. HH_NW_2

/Users/gemmachaters/Dropbox/PHD/SEEDZ data/1_import seedz hh movements data.R

Using "**AllEdges**.csv"

Explores data, numbers sold/gifted, number of locations and markets people sold at etc

Creates "**NodesUncleaned**.csv"

3. HH_NW_3

/Users/gemmachaters/Dropbox/PHD/SEEDZ data/1cleaningHHdatnodesedgescoordsprodsys.R

Using "**AllEdges**.csv" and "**NodesUncleaned**.csv" this script corrects spellings for edge locations and thus nodes, creates a unique nodes table, assigns production system to nodes based on the ward they are within, condenses node type to be a sub-village, market or village.

Creates "**SEEDZnodes**.csv" and "**SEEDZedges**.csv"

4. HH_NW_4

/Users/gemmachaters/Dropbox/PHD/SEEDZ data/plottingNodesOnMap2.R

Uses "**SEEDZnodes**.csv" and "**SEEDZedges**.csv" plots network and nodes on map of distance and relationship with type of movement, species moved and number moved for use in next chapter.

Creates “**EdgesForPlot**.csv” and “**SeedzDataNodes**.csv”

5. HH_NW_5

/Users/gemmachaters/Dropbox/PHD/SEEDZ data/network analysis on seedz hh data.R

Uses “**EdgesForPlot**.csv” and “**SeedzDataNodes**.csv”, constructs household movement network, calculates social network analysis measures and sensitivity analysis dropping 50% of sampled sub-village and sampled ward data.

Creates “**SeedzNetworkNodes**.csv” and “**SeedzNetworkEdges**.csv”

9.1.2 SEEDZ household questionnaire data dictionary

Data from the following 9 sections of the SEEDZ household survey were used to construct the movement network

9.1.3 Network construction – edge weighting

The unweighted binary network was constructed using unique edge data only (263 edges). All edge weights = 1, regardless of how many times it was reported by households from a sub-village. From the unweighted network an ‘unweighted in-degree’ was calculated by summing the number of in-going edges to each node. The fully weighted network was constructed with edge weights equal to the total number of livestock reported to be moved along that edge. There were 51 ‘NA' responses (6.7%) to the livestock batch size question (number moved), thus the batch size was imputed based on the mean value of the categories ‘species’, ‘in/out movement’, ‘production system moving too’ (12 categories). A fully weighted in-degree was calculated for each sub-village node by summing the weights of all in-going edges on the fully weighted network (this number is equal to the reported total number of livestock introduced into the sub-village). We then assessed the level of correlation between the three values for nodes in-degree by calculating Spearman’s ρ correlation coefficient. Spearman’s ρ can take any value between -1 and 1 with 0 = no correlation.

9.2 Chapter 3

9.2.1 Methods

9.2.1.1 Lab book for data cleaning and references to scripts – files can be made available if requested

6. */Users/gemmachaters/Dropbox/PHD/SEEDZ data/HHaddingdiseasedata_followsNetworkAnalysisScript.R*

Uses

“**allbrucelladata_final_withODvalues_updated1420(1).csv**”,

“**allcoxielladata_final_withODvalues(1).csv**”,

“**gemma_data_BVD_IBR_lepto.csv**”,

“**link**.csv” (a data frame used to link the serum barcodes to the cryovial barcodes. (basically a match between the unique individual animal identification codes on the different sample tubes)

This creates data frame of individual animal disease data using SEEDZ and BacZoo serology data and creates a data frame of aggregated number of each species sampled per household and BVDV herd status for all seed households. If we have link data to link RVF samples to ‘serum_sample_barcodes’ for goats and sheep it would also be possible to include RVF in the final model.

Creates “**individual_dz_dat.**csv**”** and “**HHnumbersampled.**csv”

7. */Users/gemmachaters/Dropbox/PHD/SEEDZ data/SEEDZ merging serology and hh data.R*

Uses:

“AllSEEDZdata.csv” and

“BacZoo_HHQ_DataToGemma_10Apr2019.csv"

This reads in SEEDZ and BacZoo household survey data, creates data frames of livestock numbers owned and introduced (both payed and not payed for), total numbers of livestock owned are calculated by adding adult males, female and juveniles. If this number is less than the reported number owned or the number blood sampled from the household the maximum value is taken to be the true ‘number owned’ for each species.

The column names are cleaned so the datasets match and can be merged. Summary statistics for the data on number owned and number introduced.
Creates SEEDZ "**householdintros**.csv" from seedz data which is used in script, merged to baczoo hh data and creates then "**SEEDZandBacZooHHdata.csv**"

"**SEEDZandBacZooHHdat**.csv

8. /Users/gemmachaters/Dropbox/PHD/SEEDZ data/AddHHdatat.R

Uses "**SEEDZandBacZooHHdata.csv"** and **"HHnumberSampled.csv"**
"hhnumbersampled" data is added to Household data and max value is taken from number sampled, numbers reported owned and calculated (by adding adults and juvenilles) numbers of livestock, to create a 'total_species_number' for each household.

Further cleaning to create mean and estimated number of livestock additions to the sub-village based on households number of introductions, number sampled in the sub-village and total number of livestock owning households in the sub-village.

To create and read back in: **"SDZandBZHHdataTOuse.csv"**

Then uses data on number of households that own livestock from BacZoo
"BZlivestockowninghh.csv" and SEEDZ
"subvillage_numbers%5b1%5d.csv", BacZoo agroecological classifications
"BZagroecoClasses.csv", SEEDZ sub-village 'node' data from network script
"SeedzNetworkNodes.csv".

Creates two datasets of all household data and a refined household data version which will be used to combine with the disease data to create full dataset for use in model.

FinalHHandDZdataForModel.csv (525 observations, 63 variables) full version
HHdataFormodel.csv (525 observations, 20 variables, refined version)

Finally individual disease data **"individual_dz_dat.csv"** for all animals in both studies is read in, the refined household data is merged with the individual animal disease data based on the ".hh_id" variable and the final data set to be used for running the models is created.

Final dataset of individual animal data, household data and sub-village data called **BigTest** in the script and saved as **"FinalHHandDZdataForModel.csv"**

9. /Users/gemmachaters/Dropbox/PHD/SEEDZ data/fitting models to seedz hh sero sv data.R

Uses: "**FinalHHandDZdataForModel.csv"**

Cleans age and sex variables. Creates separate cattle and small ruminant "shoat" data frames. Creates total shoats owned and introduced columns.

"cattle_data.csv"

"shoats_data.csv"

9.2.1.2 Analysis Code

10. "/Users/gemmachaters/Dropbox/PHD/SEEDZ data/UnivarMultivarDZmodels.R"

Uses: "cattle_data.csv"

"shoats_data.csv"

Changes "early seedz" study data to "baczoo"

Assigns household classifications based on most up to date classifications data by de Glanville. Creates descriptive data tables for cattle and small ruminant data.

Finds number to add to log10 betweenness and indegree to avoid adding 1 to zeros and creating huge overdispersion/right skew.

Calculates disease and movement summary statistics (how many livestock belong to disease positive hh's and how many belong to hh's that introduce livestock), calculates disease prevalence's (adjusting for Jensen's inequality because of uneven sampling), creates univariable model for each disease and risk factor variable, and summary tables.

Assesses degree of correlation between risk factor variables and creates multivariable models, checks variance inflation factors and removes variables with VIF >3. Checks for interaction between production system and all other risk factor variables. Keeps interaction terms if likelihood ratio test p value <0.05 for the interaction term. Creates final models with agropastoral production system as baseline and fits models with pastoral and small holder groups as baseline for any models with significant interaction term included. Creates multivariable model tables for each data set (combined and SEEDZ only) for cattle and small ruminants. Calculates marginal and conditional R squared for each multivariable model.

9.2.2 Household survey Questions and data dictionary

9.2.2.1 Entries

cattle_born	How many cattle were born in this compound in the past 12 months [Ng'ombe wangapi wamezaliwa katika boma hili katika miezi 12 iliyopita]?		10
cattle_born_yousure	NOTE TO INTERVIEWER: There are ${cattle_number} cattle in this compound, are you sure none were born in the past 12 months [Kuna ${cattle_number} ng'ombe katika boma hili, unauhakika hakuna aliyezaliwa katika miezi 12 iliyopita]?	yes_no2	10
note_goback_cattle_born	Please go back and change the answer to the question "How many cattle were born in this compound in the past 12 months?" [Tafadhali rudi nyuma na badili jibu la swali "ng'ombe wangapi wamezaliwa katika boma hili katika miezi 12 iliyopita?"]		10
cattle_born_hh	How many cattle were born into this household's herd in the past 12 months [Ng'ombe wangapi wamezaliwa katika kaya hii katika miezi 12 iliyopita]?		10
cattle_intro	Have any cattle been introduced into this COMPOUND in the past 12 months? [Kuna ng'ombe yeyote aliyeliletwa katika kundi lenu kwa kipindi cha miezi 12 iliyopita]	yes_no	10
cattle_intro_comp_pay	How many of the introduced cattle were purchased [Wangapi kati yao walionunuliwa]?		10
cattle_intro_comp_nopay	How many of the introduced cattle were not purchased (e.g. a gift, dowry, exchange etc) [Wangapi hawakununuliwa au waliletwa kama zawadi, mahari, kubadalishana]?		10
cattle_intro_yousure	NOTE TO INTERVIEWER - this is an important question, are you sure no cattle were introduced through purchase or as a gift etc?	yes_no3	10
note_cattle_intro	Please go back and change the answer to the question "Have you cattle been introduced into this compound's herd (walioletwa) in the past 12 months?" [Tafadhali rudi nyuma na badili jibu la swali "Have cattle been introduced into this compound's herd (walioletwa) in the past 12 months?"]		10
cattle_intro_hh	Have cattle been introduced into your HOUSEHOLD'S herd in the past 12 months [Kuna ng'ombe yeyote aliyeliletwa katika kaya yenu kwa kipindi cha miezi 12 iliyopita]?	yes_no	10
cattle_intro_hh_pay	How many of the introduced cattle were purchased [Wangapi kati yao walionunuliwa]?		10

cattle_intro_hh_nopay	How many of the introduced cattle were not purchased (e.g. a gift, dowry etc) [Wangapi hawakununuliwa au waliletwa kama zawadi, mahari, kubadalishana]?		10
cattle_intro_hh_yousure	NOTE TO INTERVIEWER - this is an important question, are you sure no cattle were introduced into the household's herd through purchase or as a gift etc?	yes_no3	10
note_cattle_hh_intro	Please go back and change the answer to the question "Have you cattle been introduced into your household's herd (walioletwa) in the past 12 months?" [Tafadhali rudi nyuma na badili jibu la swali "Have cattle been introduced into your household's herd (walioletwa) in the past 12 months?"]		10
cattle_intro_pay	calculation: if(${hh_number}>1,${cattle_intro_hh_pay}, ${cattle_intro_comp_pay})		10
cattle_intro_pay_origin	Where did you buy the ${cattle_intro_pay} cattle introduced into your household's herd [Kwa hao ng'ombe ${cattle_intro_pay} walionunuliwa walitokea wapi]?	animal_origin	10
cattle_intro_pay_origin_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_intro_source_calc	calculation: count-selected(${cattle_intro_pay_origin})		10
cattle_intro_source_note	NOTE: Only ${cattle_intro_pay} cattle were introduced from ${cattle_intro_source_calc} place(s). This is not correct. Please go back and either change the number of places animals originated from OR the number of animals introduced before proceeding to the next question.		10
cattle_hh	How many of the ${cattle_intro_pay} purchased cattle came direct from this/these other household(s) [Wangapi kati ya ng'ombe ${cattle_intro_pay} walionunuliwa walitoka moja kwa kutoka hii/hizi kaya nyingine]?		10
cattle_hh_number	How many different households did purchased cattle come from [Ni kaya ngapi tofauti walipotoka hao ng'ombe walionunuliwa]?		10
cattle_hh1_origin	Where was the household where most (or all) animals were purchased from [Iko wapi kaya ambayo wanayama wengi zaidi (au wote) walinunuliwa]?	reciprocity_where	10
cattle_hh1_district	District [Wilaya]		10
cattle_hh1_ward	Ward [Kata]		10
cattle_hh1_village	Village [Kijiji]		10
cattle_hh1_relation	What was your relationship with the person/people you bought these animals from [Nini uhusiano wako na mtu/watu uliyeleta wanyama kutoka kwake]?	reciprocity1	10
cattle_hh1_relation_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_hh2_origin	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
cattle_hh2_district	District [Wilaya]		10

cattle_hh2_ward	Ward [Kata]		10
cattle_hh2_village	Village [Kijiji]		10
cattle_hh2_relation	What was your relationship with the person/people you bought these animals from [Nini uhusiano wako na mtu/watu uliyeleta wanyama kutoka kwake]?	reciprocity1	10
cattle_hh2_relation_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_hh3_origin	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
cattle_hh3_district	District [Wilaya]		10
cattle_hh3_ward	Ward [Kata]		10
cattle_hh3_village	Village [Kijiji]		10
cattle_hh3_relation	What was your relationship with the person/people you bought these animals from [Nini uhusiano wako na mtu/watu uliyeleta wanyama kutoka kwake]?	reciprocity1	10
cattle_hh3_relation_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_hh4_origin	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
cattle_hh4_district	District [Wilaya]		10
cattle_hh4_ward	Ward [Kata]		10
cattle_hh4_village	Village [Kijiji]		10
cattle_hh4_relation	What was your relationship with the person/people you bought these animals from [Nini uhusiano wako na mtu/watu uliyeleta wanyama kutoka kwake]?	reciprocity1	10
cattle_hh4_relation_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_hh5_origin	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
cattle_hh5_district	District [Wilaya]		10
cattle_hh5_ward	Ward [Kata]		10
cattle_hh5_village	Village [Kijiji]		10
cattle_hh5_relation	What was your relationship with the person/people you bought these animals from [Nini uhusiano wako na mtu/watu uliyeleta wanyama kutoka kwake]?	reciprocity1	10
cattle_hh5_relation_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_market	How many of the ${cattle_intro_pay} cattle that were purchased came direct from a market [Wangapi kati ya ng'ombe ${cattle_intro_pay} walinunuliwa moja kwa moja kutoka sokoni]?		10
cattle_market_number	How many different markets did cattle come from [Ng'ombe walitoka kwenye masoko mangapi tofauti]?		10
cattle_market1_note	What is the name of the market from which most (or all) cattle came and which district is it in [Jina la soko ambalo ngombe wengi zaidi wametoka (au wote) ni lipi na ni wilaya ipi]?		10
cattle_market1_name	Market name [Jina la soko]		10

cattle_market1_district	District [Wilaya]	10
cattle_market2_note	What is the name of the second market from which cattle came and which district is it in [Jina la soko la pili ambalo ngombe wametoka ni lipi na ni wilaya ipi]?	10
cattle_market2_name	Market name [Jina la soko]	10
cattle_market2_district	District [Wilaya]	10
cattle_market3_note	What is the name of the third market from which cattle came and which district is it in [Jina la soko la tatu ambalo ngombe wametoka ni lipi na ni wilaya ipi]?	10
cattle_market3_name	Market name [Jina la soko]	10
cattle_market3_district	District [Wilaya]	10
cattle_market4_note	What is the name of the fourth market from which cattle came and which district is it in [Jina la soko la nne ambalo ngombe wametoka ni lipi na ni wilaya ipi]?	10
cattle_market4_name	Market name [Jina la soko]	10
cattle_market4_district	District [Wilaya]	10
cattle_market5_note	What is the name of the fifth market from which cattle came and which district is it in [Jina la soko la tano ambalo ngombe wametoka ni lipi na ni wilaya ipi]?	10
cattle_market5_name	Market name [Jina la soko]	10
cattle_market5_district	District [Wilaya]	10
cattle_bought_tradervillage	How many of the ${cattle_intro_pay} cattle that were purchased came from a livestock trader in this village [Wangapi kati ya ng'ombe ${cattle_intro_pay} walionunuliwa kutoka kwa mfanyabiashara wa mifugo katika kijiji hiki]?	10
cattle_bought_trader_othervillage	How many of the ${cattle_intro_pay} purchased cattle came from a livestock trader in another village [Wangapi kati ya ng'ombe ${cattle_intro_pay} walionunuliwa walitoka kwa mfanyabiashara wa mifugo kwenye kijiji kingine]?	10
cattle_bought_trader_othervillage_number	How many different villages were these livestock traders in [Ni vijiji vingapi tofauti walikwepo wafanyabishara wa mifugo ambao uliwauzia hii ng'ombe]?	10
note_cattle_bought_outsidetrader1	Which village did most (or all) cattle come from [Ni kijiji gani ng'ombe mingi (au yote) ilitoka]?	10
cattle_bought_outsidetrader1_village	Village [Kijiji]	10
cattle_bought_outsidetrader1_ward	Ward [Kata]	10
cattle_bought_outsidetrader1_district	District [Wilaya]	10
note_cattle_bought_outsidetrader2	Where was the second village [Kijiji cha pili kilikuwa wapi]?	10
cattle_bought_outsidetrader2_village	Village [Kijiji]	10
cattle_bought_outsidetrader2_ward	Ward [Kata]	10

cattle_bought_outsidetrader2_district	District [Wilaya]		10
note_cattle_bought_outsidetrader3	Where was the third village [Kijiji cha tatu kilikuwa wapi]?		10
cattle_bought_outsidetrader3_village	Village [Kijiji]		10
cattle_bought_outsidetrader3_ward	Ward [Kata]		10
cattle_bought_outsidetrader3_district	District [Wilaya]		10
note_cattle_bought_outsidetrader4	Where was the fourth village [Kijiji cha nne kilikuwa wapi]?		10
cattle_bought_outsidetrader4_village	Village [Kijiji]		10
cattle_bought_outsidetrader4_ward	Ward [Kata]		10
cattle_bought_outsidetrader4_district	District [Wilaya]		10
note_cattle_bought_outsidetrader5	Where was the fifth village [Kijiji cha tano kilikuwa wapi]?		10
cattle_bought_outsidetrader5_village	Village [Kijiji]		10
cattle_bought_outsidetrader5_ward	Ward [Kata]		10
cattle_bought_outsidetrader5_district	District [Wilaya]		10
cattle_bought_other	How many of the ${cattle_intro_pay} purchased cattle came direct from this other source (${cattle_intro_pay_origin_other}) [Ni wangapi kati ya ng'ombe ${cattle_intro_pay} walionunuliwa moja kwa moja kutoka chanzo hiki)]?		10
cattle_bought_other_place	Where was this other source (please give village and district, if appropriate) [Chanzo hiki kilikuwa wapi (tafadhali nipe jina la kijiji na wilaya, kama ni sahihi)]?		10
cattle_intro_nopay	calculation: if(${hh_number}>1,${cattle_intro_hh_nopay}, ${cattle_intro_comp_nopay})		10
cattle_intro_nopay_origin	Where did the ${cattle_intro_nopay} cattle that were introduced but not purchased come from [Ng'ombe${cattle_intro_nopay} walioletwa walitokea wapi]?	reciprocity1	10
cattle_intro_nopay_origin_other	Specify other [Vinginevyo (ainisha)]:		10
cattle_intro_nopay_source_calc	calculation: count-selected(${cattle_intro_nopay_origin})		10
cattle_intro_nopay_source_note	NOTE: Only ${cattle_intro_nopay} cattle were introduced from ${cattle_intro_nopay_source_calc} place(s). This is not correct. Please go back and either change the number of places animals originated from OR the number of animals introduced before proceeding to the next question.		10
cattle_intro_compound_number	How many of the ${cattle_intro_nopay} cattle came from family members in this compound [Wangapi ng'ombe walitoka kwa wanafamilia katika boma hii]?		10

cattle_intro_family_number	How many of the ${cattle_intro_nopay} cattle came from family members outside this compound [Wangapi ng'ombe walitoka kwa wanafamilia]?		10
cattle_intro_family_hh_no	How many different households of family members did cattle come from [Je walitoka kwa wanafamilia wanaoishi nje au mabali ya boma hili]?		10
cattle_intro_family_village	Where was the household of family members where most (or all) animals came from [Iko wapi kaya ya wanafamilia ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromfamily_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfamily_district1	District [Wilaya]		10
cattlefromfamily_ward1	Ward [Kata]		10
cattlefromfamily_village1	Village [Kijiji]		10
cattlefromfamily_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfamily_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfamily_district2	District [Wilaya]		10
cattlefromfamily_ward2	Ward [Kata]		10
cattlefromfamily_village2	Village [Kijiji]		10
cattlefromfamily_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfamily_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfamily_district3	District [Wilaya]		10
cattlefromfamily_ward3	Ward [Kata]		10
cattlefromfamily_village3	Village [Kijiji]		10
cattlefromfamily_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfamily_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfamily_district4	District [Wilaya]		10
cattlefromfamily_ward4	Ward [Kata]		10
cattlefromfamily_village4	Village [Kijiji]		10
cattlefromfamily_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfamily_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfamily_district5	District [Wilaya]		10
cattlefromfamily_ward5	Ward [Kata]		10
cattlefromfamily_village5	Village [Kijiji]		10
cattle_intro_relatives_number	How many of the ${cattle_intro_nopay} cattle came from relatives through marriage [Wangapi ng'ombe walitoka kwa ndugu kupitia kuoa]?		10
cattle_intro_relatives_hh_no	How many different households of relatives through marriage did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za ndugu au jamaa]?		10

cattle_intro_relatives_village	Where was the household of relatives through marriage where most (or all) animals came from [Iko wapi kaya ya ndugu au jamaa ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromrelatives_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromrelatives_district1	District [Wilaya]		10
cattlefromrelatives_ward1	Ward [Kata]		10
cattlefromrelatives_village1	Village [Kijiji]		10
cattlefromrelatives_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromrelatives_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromrelatives_district2	District [Wilaya]		10
cattlefromrelatives_ward2	Ward [Kata]		10
cattlefromrelatives_village2	Village [Kijiji]		10
cattlefromrelatives_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromrelatives_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromrelatives_district3	District [Wilaya]		10
cattlefromrelatives_ward3	Ward [Kata]		10
cattlefromrelatives_village3	Village [Kijiji]		10
cattlefromrelatives_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromrelatives_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromrelatives_district4	District [Wilaya]		10
cattlefromrelatives_ward4	Ward [Kata]		10
cattlefromrelatives_village4	Village [Kijiji]		10
cattlefromrelatives_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromrelatives_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromrelatives_district5	District [Wilaya]		10
cattlefromrelatives_ward5	Ward [Kata]		10
cattlefromrelatives_village5	Village [Kijiji]		10
cattle_intro_friends_number	How many of the ${cattle_intro_nopay} cattle came from friends/age mates [Wangapi ng'ombe walitoka kwa marafiki/watu wa rika moja]?		10

cattle_intro_friends_hh_no	How many different households of friends/age mates did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za marafiki/watu wa rika moja]		10
cattle_intro_friends_village	Where was the household of friends/age mates where most (or all) animals came from [Iko wapi kaya ya marafiki/watu wa rika moja ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromfriends_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfriends_district1	District [Wilaya]		10
cattlefromfriends_ward1	Ward [Kata]		10
cattlefromfriends_village1	Village [Kijiji]		10
cattlefromfriends_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfriends_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfriends_district2	District [Wilaya]		10
cattlefromfriends_ward2	Ward [Kata]		10
cattlefromfriends_village2	Village [Kijiji]		10
cattlefromfriends_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfriends_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfriends_district3	District [Wilaya]		10
cattlefromfriends_ward3	Ward [Kata]		10
cattlefromfriends_village3	Village [Kijiji]		10
cattlefromfriends_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfriends_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfriends_district4	District [Wilaya]		10
cattlefromfriends_ward4	Ward [Kata]		10
cattlefromfriends_village4	Village [Kijiji]		10
cattlefromfriends_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromfriends_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromfriends_district5	District [Wilaya]		10
cattlefromfriends_ward5	Ward [Kata]		10
cattlefromfriends_village5	Village [Kijiji]		10
cattle_intro_non_ageset_number	How many of the ${cattle_intro_nopay} cattle came from people outside your age set [Wangapi ng'ombe walitoka kwa watu nje ya rika lako]?		10

cattle_intro_non_ageset_hh_no	How many different households of people outside your ageset did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za watu nje ya rika lako]?		10
cattle_intro_non_ageset_village	Where was the household of people outside your ageset where most (or all) animals came from [Iko wapi kaya ya watu nje ya rika lako ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromnon_ageset_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_ageset_district1	District [Wilaya]		10
cattlefromnon_ageset_ward1	Ward [Kata]		10
cattlefromnon_ageset_village1	Village [Kijiji]		10
cattlefromnon_ageset_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_ageset_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_ageset_district2	District [Wilaya]		10
cattlefromnon_ageset_ward2	Ward [Kata]		10
cattlefromnon_ageset_village2	Village [Kijiji]		10
cattlefromnon_ageset_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_ageset_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_ageset_district3	District [Wilaya]		10
cattlefromnon_ageset_ward3	Ward [Kata]		10
cattlefromnon_ageset_village3	Village [Kijiji]		10
cattlefromnon_ageset_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_ageset_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_ageset_district4	District [Wilaya]		10
cattlefromnon_ageset_ward4	Ward [Kata]		10
cattlefromnon_ageset_village4	Village [Kijiji]		10
cattlefromnon_ageset_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_ageset_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_ageset_district5	District [Wilaya]		10
cattlefromnon_ageset_ward5	Ward [Kata]		10
cattlefromnon_ageset_village5	Village [Kijiji]		10
cattle_intro_clan_number	How many of the ${cattle_intro_nopay} cattle came from people in your clan		10

	[Wangapi ng'ombe walitoka kwa watu wa ukoo wako]?		
cattle_intro_clan_hh_no	How many different households of people in your clan did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za watu wa ukoo wako]?		10
cattle_intro_clan_village	Where was the household of people in your clan where most (or all) animals came from [Iko wapi kaya ya watu wa ukoo wako ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromclan_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromclan_district1	District [Wilaya]		10
cattlefromclan_ward1	Ward [Kata]		10
cattlefromclan_village1	Village [Kijiji]		10
cattlefromclan_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromclan_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromclan_district2	District [Wilaya]		10
cattlefromclan_ward2	Ward [Kata]		10
cattlefromclan_village2	Village [Kijiji]		10
cattlefromclan_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromclan_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromclan_district3	District [Wilaya]		10
cattlefromclan_ward3	Ward [Kata]		10
cattlefromclan_village3	Village [Kijiji]		10
cattlefromclan_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromclan_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromclan_district4	District [Wilaya]		10
cattlefromclan_ward4	Ward [Kata]		10
cattlefromclan_village4	Village [Kijiji]		10
cattlefromclan_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromclan_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromclan_district5	District [Wilaya]		10
cattlefromclan_ward5	Ward [Kata]		10
cattlefromclan_village5	Village [Kijiji]		10
cattle_intro_non_clan_number	How many of the ${cattle_intro_nopay} cattle came from people outside your clan [Wangapi ng'ombe walitoka kwa watu nje ya ukoo wako]?		10
cattle_intro_non_clan_hh_no	How many different households of people outside your clan did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za watu nje ya ukoo wako]?		10

cattle_intro_non_clan_village	Where was the household of people outside your clan where most (or all) animals came from [Iko wapi kaya ya watu nje ya ukoo wako ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromnon_clan_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_clan_district1	District [Wilaya]		10
cattlefromnon_clan_ward1	Ward [Kata]		10
cattlefromnon_clan_village1	Village [Kijiji]		10
cattlefromnon_clan_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_clan_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_clan_district2	District [Wilaya]		10
cattlefromnon_clan_ward2	Ward [Kata]		10
cattlefromnon_clan_village2	Village [Kijiji]		10
cattlefromnon_clan_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_clan_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_clan_district3	District [Wilaya]		10
cattlefromnon_clan_ward3	Ward [Kata]		10
cattlefromnon_clan_village3	Village [Kijiji]		10
cattlefromnon_clan_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_clan_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_clan_district4	District [Wilaya]		10
cattlefromnon_clan_ward4	Ward [Kata]		10
cattlefromnon_clan_village4	Village [Kijiji]		10
cattlefromnon_clan_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromnon_clan_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromnon_clan_district5	District [Wilaya]		10
cattlefromnon_clan_ward5	Ward [Kata]		10
cattlefromnon_clan_village5	Village [Kijiji]		10
cattle_intro_other_number	How many of the ${cattle_intro_nopay} cattle came from these other people (${cattle_intro_nopay_origin_other}) [Wangapi ng'ombe walitoka kwa hawa watu wengine]?		10

cattle_intro_other_hh_no	How many different households of these other people did cattle come from [Ng'ombe walioletwa walitoka katika kaya ngapi tofauti za watu watu wengine]?		10
cattle_intro_other_village	Where was the household of these other people where most (or all) animals came from [Iko wapi kaya ya watu wengine ambayo wanayama wengi zaidi (au wote) walitoka]?	reciprocity_where	10
note_cattlefromother_where1	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromother_district1	District [Wilaya]		10
cattlefromother_ward1	Ward [Kata]		10
cattlefromother_village1	Village [Kijiji]		10
cattlefromother_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	10
note_cattlefromother_where2	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromother_district2	District [Wilaya]		10
cattlefromother_ward2	Ward [Kata]		10
cattlefromother_village2	Village [Kijiji]		10
cattlefromother_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	10
note_cattlefromother_where3	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromother_district3	District [Wilaya]		10
cattlefromother_ward3	Ward [Kata]		10
cattlefromother_village3	Village [Kijiji]		10
cattlefromother_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	10
note_cattlefromother_where4	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromother_district4	District [Wilaya]		10
cattlefromother_ward4	Ward [Kata]		10
cattlefromother_village4	Village [Kijiji]		10
cattlefromother_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	10
note_cattlefromother_where5	Where is this village [Hiki kijiji kipo wapi]?		10
cattlefromother_district5	District [Wilaya]		10
cattlefromother_ward5	Ward [Kata]		10
cattlefromother_village5	Village [Kijiji]		10

9.2.2.2 Exits

cattle_death	Have any cattle kept by your household died in the past 12 months (not through slaughter) [Kuna ng'ombe yoyote anayefugwa kwenye boma lako amekufa katika miezi 12 iliyopita (siyo kwa kuchinjwa)]?	yes_no	11

cattle_death_type	What did these animals die of [Hawa wanyama walikufa na nini]?	deaths	11
cattle_death_type_other	Specify other [Vinginevyo (ainisha)]:		11
cattle_death_drought	How many cattle died from drought [Ng'ombe wangapi walikufa kutokana na ukame]?		11
cattle_death_predation	How many died from predation [Wangapi walikufa kwa kuliwa na wanyama]?		11
cattle_death_disease	How many died from disease [Wangapi walikufa kutokana na magonjwa]?		11
cattle_death_trauma	How many died from trauma [Wangapi walikufa kwa kuumia]?		11
cattle_death_other	How many died from ${cattle_death_type_other} [Wangapi walikufa kutokana ${cattle_death_type_other}]?		11
cattle_sold	Have you sold any cattle out of your own herd in the past 12 months [Je, umeuza ng'ombe yoyote nje ya kundi lako katika miezi 12 iliyopita]?	yes_no	11
cattle_sold_number	How many cattle have been sold [Ng'ombe wangapi wameuzwa]?		11
cattle_sold_dest	Who were they sold to/how were they sold [Waliuzwa kwa nani/waliuzwaje]?	animal_dest	11
cattle_sold_dest_other	Specify other [Vinginevyo (ainisha)]:		11
cattle_sold_dest_calc	calculation: count-selected(${cattle_sold_dest})		11
cattle_sold_dest_note	NOTE: Only ${cattle_sold_number} cattle were sold to ${cattle_sold_dest_calc} place(s). This is not correct. Please go back and either change the number of places animals were sold to from OR the number of animals sold before proceeding to the next question.		11
cattle_sold_hh	How many of the ${cattle_sold_number} cattle were sold directly to this/these household(s) [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa moja kwa moja kwenye kaya]?		11
cattle_sold_hh_number	How many different households were cattle sold to [Nyumba ngapi tofauti ng'ombe waliuzwa]?		11

note_cattle_sold_hh1	Where was the household where most (or all) cattle were sold to [Iko wapi kaya ambayo ng'ombe wengi zaidi (au wote) waliouzwa]?		11
cattle_sold_hh1_district	District [Wilaya]		11
cattle_sold_hh1_ward	Ward [Kata]		11
cattle_sold_hh1_village	Village [Kijiji]		11
cattle_sold_hh1_relationship	What is your relationship with this person/household [Uhusiano wako na huyu mtu/kaya ukoje]?	reciprocity1	11
cattle_sold_hh1_relationship_other	Specify other [Vinginevyo (ainisha)]:		11
note_cattle_sold_hh2	Where was the second household [Kaya ya pili ilikuwa wapi]?		11
cattle_sold_hh2_district	District [Wilaya]		11
cattle_sold_hh2_ward	Ward [Kata]		11
cattle_sold_hh2_village	Village [Kijiji]		11
cattle_sold_hh2_relationship	What is your relationship with this person/household [Uhusiano wako na huyu mtu/kaya ukoje]?	reciprocity1	11
cattle_sold_hh2_relationship_other	Specify other [Vinginevyo (ainisha)]:		11
note_cattle_sold_hh3	Where was the third household [Kaya ya tatu ilikuwa wapi]?		11
cattle_sold_hh3_district	District [Wilaya]		11
cattle_sold_hh3_ward	Ward [Kata]		11
cattle_sold_hh3_village	Village [Kijiji]		11
cattle_sold_hh3_relationship	What is your relationship with this person/household [Uhusiano wako na huyu mtu/kaya ukoje]?	reciprocity1	11
cattle_sold_hh3_relationship_other	Specify other [Vinginevyo (ainisha)]:		11
note_cattle_sold_hh4	Where was the fourth household [Kaya ya nne ilikuwa wapi]?		11
cattle_sold_hh4_district	District [Wilaya]		11
cattle_sold_hh4_ward	Ward [Kata]		11
cattle_sold_hh4_village	Village [Kijiji]		11

cattle_sold_hh4_relationship	What is your relationship with this person/household [Uhusiano wako na huyu mtu/kaya ukoje]?	reciprocity1	11
cattle_sold_hh4_relationship_other	Specify other [Vinginevyo (ainisha)]:		11
note_cattle_sold_hh5	Where was the fifth household [Kaya ya tano ilikuwa wapi]?		11
cattle_sold_hh5_district	District [Wilaya]		11
cattle_sold_hh5_ward	Ward [Kata]		11
cattle_sold_hh5_village	Village [Kijiji]		11
cattle_sold_hh5_relationship	What is your relationship with this person/household [Uhusiano wako na huyu mtu/kaya ukoje]?	reciprocity1	11
cattle_sold_hh5_relationship_other	Specify other [Vinginevyo (ainisha)]:		11
cattle_sold_market	How many of the ${cattle_sold_number} cattle were sold directly at this/these market(s) [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa moja kwa moja hili/haya masoko]?		11
cattle_sold_market_number	How many different markets were cattle sold at [Ng'ombe waliuzwa kwenye masoko mangapi tofauti]?		11
note_cattle_sold_market1	What was the name of the market where most (or all) cattle went and what district was it in [Jina la soko ambalo ngombe wengi zaidi wameenda (au wote) ni lipi na ni wilaya ipi]?		11
cattle_sold_market1	Market name [Jina la soko]		11
cattle_sold_market1_district	District [Wilaya]		11
note_cattle_sold_market2	What was the name of the second market and what district was it in [Nini jina la soko la pili na lipo ndani ya wilaya gani]?		11
cattle_sold_market2	Market name [Jina la soko]		11
cattle_sold_market2_district	District [Wilaya]		11
note_cattle_sold_market3	What was the name of the third market and what district was it in [Nini jina la soko la tatu na lipo ndani ya wilaya gani]?		11
cattle_sold_market3	Market name [Jina la soko]		11

cattle_sold_market3_district	District [Wilaya]	11
note_cattle_sold_market4	What was the name of the fourth market and what district was it in [Nini jina la soko la nne na lipo ndani ya wilaya gani]?	11
cattle_sold_market4	Market name [Jina la soko]	11
cattle_sold_market4_district	District [Wilaya]	11
note_cattle_sold_market5	What was the name of the fifth market and what district was it in [Nini jina la soko la tano na lipo ndani ya wilaya gani]?	11
cattle_sold_market5	Market name [Jina la soko]	11
cattle_sold_market5_district	District [Wilaya]	11
cattle_sold_tradervillage	How many of the ${cattle_sold_number} cattle were sold directly to a livestock trader(s) in this village [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa moja kwa moja kwa mfanyabiashara wa mifugo wa kijiji hiki]?	11
cattle_sold_trader_othervillage	How many of the ${cattle_sold_number} cattle were sold directly to livestock traders in other villages [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa moja kwa moja kwa mfanyabiashara wa mifugo wa kijiji kingine]?	11
cattle_sold_trader_othervillage_number	How many different villages were the livestock traders to which you sold these cattle in [Ni vijiji vingapi tofauti walikwepo wafanyabishara wa mifugo ambao uliwauzia hii ng'ombe]?	11
note_cattle_sold_outsidetrader1	Where was the village where most cattle were sold [Ni kijiji gani mifugo mingi zaidi iliuzwa]?	11
cattle_sold_outsidetrader1_district	District [Wilaya]	11
cattle_sold_outsidetrader1_ward	Ward [Kata]	11
cattle_sold_outsidetrader1_village	Village [Kijiji]	11
note_cattle_sold_outsidetrader2	Where was the second village [Kijiji cha pili kilikuwa wapi]?	11
cattle_sold_outsidetrader2_district	District [Wilaya]	11

cattle_sold_outsidetrader2_ward	Ward [Kata]	11
cattle_sold_outsidetrader2_village	Village [Kijiji]	11
note_cattle_sold_outsidetrader3	Where was the third village [Kijiji cha tatu kilikuwa wapi]?	11
cattle_sold_outsidetrader3_district	District [Wilaya]	11
cattle_sold_outsidetrader3_ward	Ward [Kata]	11
cattle_sold_outsidetrader3_village	Village [Kijiji]	11
note_cattle_sold_outsidetrader4	Where was the fourth village [Kijiji cha nne kilikuwa wapi]?	11
cattle_sold_outsidetrader4_district	District [Wilaya]	11
cattle_sold_outsidetrader4_ward	Ward [Kata]	11
cattle_sold_outsidetrader4_village	Village [Kijiji]	11
note_cattle_sold_outsidetrader5	Where was the fifth village [Kijiji cha tano kilikuwa wapi]?	11
cattle_sold_outsidetrader5_district	District [Wilaya]	11
cattle_sold_outsidetrader5_ward	Ward [Kata]	11
cattle_sold_outsidetrader5_village	Village [Kijiji]	11
cattle_sold_slaughter	How many of the ${cattle_sold_number} cattle were sold directly to slaughter place(s) [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa moja kwa moja sehemu ya machinjio]?	11
cattle_sold_slaughter_number	How many different slaughter places were cattle sold to [Ni sehemu ngapi tofauti za machinjio ngombe waliuzwa]?	11
note_cattle_sold_slaughter1	Where was the slaughter place where most animals went [Machinjio ambayo walienda wanayama wengi iko wapi]?	11
cattle_sold_slaughter_district1	District [Wilaya]	11
cattle_sold_slaughter_village1	Village/Town [Kijiji/Mji]	11
note_cattle_sold_slaughter2	Where was the second slaughter place [Sehemu ya machinjio ya pili ilikuwa wapi]?	11

cattle_sold_slaughter_district2	District [Wilaya]		11
cattle_sold_slaughter_village2	Village/Town [Kijiji/Mji]		11
note_cattle_sold_slaughter3	Where was the third slaughter place [Sehemu ya machinjio ya tatu ilikuwa wapi]?		11
cattle_sold_slaughter_district3	District [Wilaya]		11
cattle_sold_slaughter_village3	Village/Town [Kijiji/Mji]		11
note_cattle_sold_slaughter4	Where was the fourth place [Sehemu ya machinjio ya nne ilikuwa wapi]?		11
cattle_sold_slaughter_district4	District [Wilaya]		11
cattle_sold_slaughter_village4	Village/Town [Kijiji/Mji]		11
note_cattle_sold_slaughter5	Where was the fifth slaughter place [Sehemu ya machinjio ya tano ilikuwa wapi]?		11
cattle_sold_slaughter_district5	District [Wilaya]		11
cattle_sold_slaughter_village5	Village/Town [Kijiji/Mji]		11
cattle_sold_other	How many of the ${cattle_sold_number} cattle were sold to this other source (${cattle_sold_dest_other}) [Wangapi kati ya ng'ombe ${cattle_sold_number} waliuzwa kwenye chanzo hiki kingine]?		11
cattle_sold_other_place	Where is this other place (please give village and district, if appropriate) [Hii sehemu nyingine ipo wapi]?		11
cattle_out_nopay	Have you given away any cattle from your herd over the past 12 months [Je, umepeana ng'ombe kutoka kundi lako zaidi ya miezi 12 iliyopita]?	yes_no	11
cattle_out_nopay_number	How many cattle have been given away [Ng'ombe wangapi umepeana]?		11
cattle_out_nopay_dest	Who did you give these ${cattle_out_nopay_number} cattle to [Umepeana hawa ng'ombe na nani]?	reciprocity1	11
cattle_out_nopay_dest_other	Specify other [Vinginevyo (ainisha)]:		11
cattle_nopay_dest_calc	calculation: count-selected(${cattle_out_nopay_dest})		11

cattle_nopay_dest_note	NOTE: Only ${cattle_out_nopay_number} cattle were given to ${cattle_nopay_dest_calc} place(s). This is not correct. Please go back and either change the number of places animals were given to OR the number of animals given away before proceeding to the next question.		11
cattle_out_compound_number	How many of the ${cattle_out_nopay_number} cattle went to family members in this compound [Ng'ombe wangapi walienda kwa familia katika boma hili]?		11
cattle_out_family_number	How many of the ${cattle_out_nopay_number} cattle went to family members outside this compound [Ng'ombe wangapi walienda kwa wanakaya]?		11
cattle_out_family_hh_no	How many different households of family members did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_family_village	Where was the household of family members where most (or all) animals went [Iko wapi kaya ya wanafamalia ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11
note_cattletofamily_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofamily_district1	District [Wilaya]		11
cattletofamily_ward1	Ward [Kata]		11
cattletofamily_village1	Village [Kijiji]		11
cattletofamily_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	11
note_cattletofamily_where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofamily_district2	District [Wilaya]		11
cattletofamily_ward2	Ward [Kata]		11
cattletofamily_village2	Village [Kijiji]		11
cattletofamily_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	11
note_cattletofamily_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofamily_district3	District [Wilaya]		11

cattletofamily_ward3	Ward [Kata]		11
cattletofamily_village3	Village [Kijiji]		11
cattletofamily_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	11
note_cattletofamily_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofamily_district4	District [Wilaya]		11
cattletofamily_ward4	Ward [Kata]		11
cattletofamily_village4	Village [Kijiji]		11
cattletofamily_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	11
note_cattletofamily_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofamily_district5	District [Wilaya]		11
cattletofamily_ward5	Ward [Kata]		11
cattletofamily_village5	Village [Kijiji]		11
cattle_out_relatives_number	How many of the ${cattle_out_nopay_number} cattle went to relatives through marriage [Wangapi ng'ombe walienda kwa ndugu kupitia kuoa]?		11
cattle_out_relatives_hh_no	How many different households of relatives through marriage did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_relatives_village	Where was the household of relatives through marriage where most (or all) animals went to [Iko wapi kaya ya ndugu kupitia kuoa ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11
note_cattletorelatives_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletorelatives_district1	District [Wilaya]		11
cattletorelatives_ward1	Ward [Kata]		11
cattletorelatives_village1	Village [Kijiji]		11
cattletorelatives_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	11
note_cattletorelatives_where2	Where is this village [Hiki kijiji kipo wapi]?		11

cattletorelatives_district2	District [Wilaya]		11
cattletorelatives_ward2	Ward [Kata]		11
cattletorelatives_village2	Village [Kijiji]		11
cattletorelatives_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	11
note_cattletorelatives_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletorelatives_district3	District [Wilaya]		11
cattletorelatives_ward3	Ward [Kata]		11
cattletorelatives_village3	Village [Kijiji]		11
cattletorelatives_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	11
note_cattletorelatives_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletorelatives_district4	District [Wilaya]		11
cattletorelatives_ward4	Ward [Kata]		11
cattletorelatives_village4	Village [Kijiji]		11
cattletorelatives_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	11
note_cattletorelatives_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletorelatives_district5	District [Wilaya]		11
cattletorelatives_ward5	Ward [Kata]		11
cattletorelatives_village5	Village [Kijiji]		11
cattle_out_friends_number	How many of the ${cattle_out_nopay_number} cattle went to friends/age mates [Wangapi ng'ombe walienda kwa marafiki/watu wa rika moja]?		11
cattle_out_friends_hh_no	How many different households of friends/age mates did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_friends_village	Where was the household of friends/age mates where most (or all) animals went to [Iko wapi kaya ya marafiki/watu wa rika moja ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11

note_cattletofriends_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofriends_district1	District [Wilaya]		11
cattletofriends_ward1	Ward [Kata]		11
cattletofriends_village1	Village [Kijiji]		11
cattletofriends_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	11
note_cattletofriends_where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofriends_district2	District [Wilaya]		11
cattletofriends_ward2	Ward [Kata]		11
cattletofriends_village2	Village [Kijiji]		11
cattletofriends_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	11
note_cattletofriends_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofriends_district3	District [Wilaya]		11
cattletofriends_ward3	Ward [Kata]		11
cattletofriends_village3	Village [Kijiji]		11
cattletofriends_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	11
note_cattletofriends_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofriends_district4	District [Wilaya]		11
cattletofriends_ward4	Ward [Kata]		11
cattletofriends_village4	Village [Kijiji]		11
cattletofriends_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	11
note_cattletofriends_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletofriends_district5	District [Wilaya]		11
cattletofriends_ward5	Ward [Kata]		11
cattletofriends_village5	Village [Kijiji]		11

cattle_out_non_ageset_number	How many of the ${cattle_out_nopay_number} cattle went to people outside your age set [Wangapi ng'ombe walienda kwa watu nje ya rika lako]?		11
cattle_out_non_ageset_hh_no	How many different households of people outside your age set did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_non_ageset_village	Where was the household of people outside your age set where most (or all) animals went to [Iko wapi kaya ya watu nje ya rika lako ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11
note_cattletonon_ageset_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_ageset_district1	District [Wilaya]		11
cattletonon_ageset_ward1	Ward [Kata]		11
cattletonon_ageset_village1	Village [Kijiji]		11
cattletonon_ageset_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_ageset_where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_ageset_district2	District [Wilaya]		11
cattletonon_ageset_ward2	Ward [Kata]		11
cattletonon_ageset_village2	Village [Kijiji]		11
cattletonon_ageset_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_ageset_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_ageset_district3	District [Wilaya]		11
cattletonon_ageset_ward3	Ward [Kata]		11
cattletonon_ageset_village3	Village [Kijiji]		11
cattletonon_ageset_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_ageset_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_ageset_district4	District [Wilaya]		11

cattletonon_ageset_ward4	Ward [Kata]		11
cattletonon_ageset_village4	Village [Kijiji]		11
cattletonon_ageset_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_ageset_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_ageset_district5	District [Wilaya]		11
cattletonon_ageset_ward5	Ward [Kata]		11
cattletonon_ageset_village5	Village [Kijiji]		11
cattle_out_clan_number	How many of the ${cattle_out_nopay_number} cattle went to people in your clan [Wangapi ng'ombe walienda kwa watu wa ukoo wako]?		11
cattle_out_clan_hh_no	How many different households of people in your clan did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_clan_village	Where was the household of people in your clan where most (or all) animals went to [Iko wapi kaya ya watu wa ukoo wako ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11
note_cattletoclan_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoclan_district1	District [Wilaya]		11
cattletoclan_ward1	Ward [Kata]		11
cattletoclan_village1	Village [Kijiji]		11
cattletoclan_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_where	11
note_cattletoclan_where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoclan_district2	District [Wilaya]		11
cattletoclan_ward2	Ward [Kata]		11
cattletoclan_village2	Village [Kijiji]		11
cattletoclan_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_where	11
note_cattletoclan_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoclan_district3	District [Wilaya]		11
cattletoclan_ward3	Ward [Kata]		11
cattletoclan_village3	Village [Kijiji]		11

cattletoclan_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_ where	11
note_cattletoclan_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoclan_district4	District [Wilaya]		11
cattletoclan_ward4	Ward [Kata]		11
cattletoclan_village4	Village [Kijiji]		11
cattletoclan_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_ where	11
note_cattletoclan_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoclan_district5	District [Wilaya]		11
cattletoclan_ward5	Ward [Kata]		11
cattletoclan_village5	Village [Kijiji]		11
cattle_out_non_clan _number	How many of the ${cattle_out_nopay_number} cattle went to people outside your clan [Wangapi ng'ombe walienda kwa watu nje ya ukoo wako]?		11
cattle_out_non_clan _hh_no	How many different households of people outside your clan did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_non_clan _village	Where was the household of people outside your clan where most (or all) animals went to [Iko wapi kaya ya watu nje ya ukoo wako ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_ where	11
note_cattletonon_clan_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_clan_district1	District [Wilaya]		11
cattletonon_clan_ward1	Ward [Kata]		11
cattletonon_clan_village1	Village [Kijiji]		11
cattletonon_clan_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_ where	11
note_cattletonon_clan_where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_clan_district2	District [Wilaya]		11
cattletonon_clan_ward2	Ward [Kata]		11
cattletonon_clan_village2	Village [Kijiji]		11
cattletonon_clan_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_ where	11

note_cattletonon_clan_where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_clan_district3	District [Wilaya]		11
cattletonon_clan_ward3	Ward [Kata]		11
cattletonon_clan_village3	Village [Kijiji]		11
cattletonon_clan_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_clan_where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_clan_district4	District [Wilaya]		11
cattletonon_clan_ward4	Ward [Kata]		11
cattletonon_clan_village4	Village [Kijiji]		11
cattletonon_clan_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_where	11
note_cattletonon_clan_where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletonon_clan_district5	District [Wilaya]		11
cattletonon_clan_ward5	Ward [Kata]		11
cattletonon_clan_village5	Village [Kijiji]		11
cattle_out_other_number	How many of the ${cattle_out_nopay_number} cattle went to these other people (${cattle_out_nopay_dest_other}) [Wangapi ng'ombe walienda kwa hawa watu wengine]?		11
cattle_out_other_hh_no	How many different households of these other people did cattle go to [Ni kaya ngapi tofauti walienda]?		11
cattle_out_other_village	Where was the household of these other people where most (or all) animals went to [Iko wapi kaya ya watu wengine ambayo wanayama wengi zaidi (au wote) walienda]?	reciprocity_where	11
note_cattletoother_where1	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoother_district1	District [Wilaya]		11
cattletoother_ward1	Ward [Kata]		11
cattletoother_village1	Village [Kijiji]		11

cattletoother_other1	Where was the second household [Kaya ya pili ilikuwa wapi]?	reciprocity_ where	11
note_cattletoother_ where2	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoother_district 2	District [Wilaya]		11
cattletoother_ward2	Ward [Kata]		11
cattletoother_village 2	Village [Kijiji]		11
cattletoother_other2	Where was the third household [Kaya ya tatu ilikuwa wapi]?	reciprocity_ where	11
note_cattletoother_ where3	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoother_district 3	District [Wilaya]		11
cattletoother_ward3	Ward [Kata]		11
cattletoother_village 3	Village [Kijiji]		11
cattletoother_other3	Where was the fourth household [Kaya ya nne ilikuwa wapi]?	reciprocity_ where	11
note_cattletoother_ where4	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoother_district 4	District [Wilaya]		11
cattletoother_ward4	Ward [Kata]		11
cattletoother_village 4	Village [Kijiji]		11
cattletoother_other4	Where was the fifth household [Kaya ya tano ilikuwa wapi]?	reciprocity_ where	11
note_cattletoother_ where5	Where is this village [Hiki kijiji kipo wapi]?		11
cattletoother_district 5	District [Wilaya]		11
cattletoother_ward5	Ward [Kata]		11
cattletoother_village 5	Village [Kijiji]		11

9.2.3 Model variables description

Table S1. Risk factor and potential confounding variables for disease risk models, description and motivation for inclusion in model.

Number of household introductions	We hypothesised that higher numbers of livestock introductions will result in higher risk of pathogen

	introduction and thus higher prevalence of seropositivity in animals from these households. For the SEEDZ data the number of household introductions was calculated for each species by adding the numbers of animals purchased from markets, purchased from other households, and received as gifts in the 12 months preceding the study. For the BacZoo data the number of reported introductions of each species was used directly because the original survey question was not separated into introduction type.
Estimated number of sub-village introductions	We hypothesised that belonging to a sub-village with high numbers of livestock introductions will increase an individual's risk of pathogen exposure. The standardised number of sub-village introductions was calculated to use as a risk factor in the model to assess whether being a household in a 'busy' sub-village increased the risk of disease exposure. This was calculated by adding the number of reported introductions from all surveyed households in the sub-village, dividing the total by the number of surveyed households and multiplying this 'mean' number of introductions by the total number of livestock owning households in the sub-village.
Sub-village in-degree	In-degree is the sum of the incoming edges into a node (edges are weighted by the number of times they are reported between the sub-village and location X). In-degree is used to test the hypothesis that livestock belonging to a sub-village that has connected more frequently, to more locations, via batches of livestock introductions, are at greater risk of pathogen exposure. Node metrics from the partially weighted network are used for the analysis rather than from the unweighted network, based on the

	assumption that multiple introductions from location X will increase the risk of pathogen introduction compared to one/few introduction(s).
Sub-village betweenness	This metric is again ascertained from the constructed household movement network and is a measure of the number of times a sub-village lies on the shortest path between two other locations in the observed network. Betweeness will be used to test the hypothesis that belonging to a sub-village that is relatively ‘busy’ with many livestock passing through it on journeys between other locations, increases an individual’s risk of exposure to infectious pathogens. It should be noted here that because this network is a sub-sample of the underlying network, betweenness measures may be falsely inflated because locations with the truest high betweenness may not be included in the sample – this should be considered when interpreting the model results.
Herd BVDV status	A household herd was considered BVDV positive if one or more of the sampled cattle were seropositive for BVDV antibodies. Herds were thus assigned a value 0 (negative) or 1 (positive). This is considered as a potential risk factor for exposure to other infectious pathogens as, if circulating, it could cause immunosuppression of animals in the herd and also potentially increase the risk of bacterial pathogen shedding.
Household herd/flock size	For each household the ‘total number owned’ of each species (cattle and small ruminants (goats and sheep) was calculated by adding the reported number of ‘juveniles’, ‘adult females’ and ‘adult males’ owned. Where breakdown numbers were unreported (N households = 2) then the

	‘number reported owned’ was imputed as the ‘total number owned’. For a number of households in the SEEDZ data set more livestock were blood sampled than the ‘total number owned’ (Cattle = 43 households, Goats = 52 households, Sheep = 51 households). For these households the maximum number from ‘number sampled’ or ‘number reported owned’ was imputed in place of the ‘total number owned’.
Sub-village herd/flock size	Number of livestock owning households and the size of these herds/flocks, in a sub-village was considered a likely risk factor for disease exposure at the household level. More livestock owning households with larger herds/flocks could potentially mean there are inherently more livestock introductions and/or higher opportunity for contact and transmission events between herds/flocks. Sub-village herd/flock size was calculated by taking the mean herd/flock size(s) from the surveyed households and multiplying by the total number of livestock owning households in the sub-village. Participating households were randomly selected from a list of households that attended a meeting at the sub-village central point thus they are assumed to be representative of others in the same sub-village.
Household production system classification	Production system classification data for each household was taken from the household classification model developed by de Glanville (de Glanville *et al.*, 2020). Agro-ecological classification is assigned based on a number of production system metrics of households. For the SEEDZ cattle owning households 49% are agro-pastoral, 39% are pastoral and 12% small holder, in contrast to the BacZoo data set which is comprised of predominantly small holder households (63%), two agro-pastoral (3%) and the remaining pastoral (34%). For the SEEDZ small ruminant

owning households 48% are agropastoral, 52% pastoral, 0 smallholder and for the BacZoo small ruminant households 1% are agropastoral, 26% pastoral and 73% small holder.

9.2.4 Zero cases of brucella in small ruminant data

Using BGLMER

There were zero animals with *Brucella* spp. exposure in sampled smallholder small ruminants, resulting in quasi-complete separation of the data in the univariable model investigating household agro-ecological classification as a risk factor for Brucella seropositivity. In consequence, this model could not be fitted by maximum likelihood, because in such cases the maximum likelihood prevalence is zero, which cannot be represented on the log odds scale. Thus this univariable GLMM was fitted using a Bayesian maximum *a posteriori* approach implemented in the *bglmer* function from the *blme* package in R (Chung *et al.*, 2013). A weak normally distributed prior with mean of zero and variance of nine was placed on each fixed effect, effectively penalising prevalence's very close to zero (e.g. less than ~10^{-6}). For the small ruminant *Brucella* spp. Bayesian univariate GLMM there was no evidence of a difference in seroprevalence between small holder and agro-pastoralist small ruminant disease risk. These classes were therefore combined into 'agro-past/smallholder' for the SEEDZ *Brucella spp*. multivariable analysis only, so that the model could be fitted using maximum likelihood, giving comparable results with the other models in this analysis.

Further details found in R script *UnivarMultivarDZmodels.R*

9.2.5 Supplementary results

9.2.5.1 Serology

For cattle and small ruminants the *Brucella* seroprevalence was low at 4.0% (95% Confidence Interval (CI) 2.5 - 5.8%) and 2.6% (95% CI 1.1-4.5%) respectively, with 90 (20.5%) and 69 (16.8%) cattle- and small ruminant-owning households owning positive animals.

Similarly, C. *burnetii* seroprevalence was low in cattle 4.4% (95% CI 3.1-6.0%), with 106 (16.8%) cattle owning households owning seropositive cattle. In small ruminants the *C. burnetii* seroprevalence was higher at 20.2% (95% CI 16.3-24.3%) with 283 (68.9%) small ruminant owning households owning at least one seropositive small ruminant.

The individual seroprevalence of L .*hardjo* in cattle is high, compared to the other two bacterial zoonoses at 32.6% and these animals belong to 281 households (77.8%) cattle owning households.

Overall, ‘any zoonoses’ seroprevalence in cattle using SEEDZ only data (*Brucella* and/or C. *burnetii* and/or L .*hardjo*) is 36.1% (95% CI 30.9-41.4%) and in small ruminants (Brucella and/or Coxiella using combined SEEDZ and BacZoo data) is 21.5% (95% CI 17.3-25.8%). The household prevalence of ‘any zoonoses’ within the SEEDZ cattle owning households is 82.2% and in the SEEDZ and BacZoo small ruminant owning households it is 70 %.

For the SEEDZ cattle the sero-prevalence of the viral pathogens BHV-1 and BVDV were 65.8% (95% CI 61.2-70.2) and 39.5% (95% CI 32.5-46.5%) respectively with herd prevalence of 95.8% and 80.3%.

Analysis was done in R version 3.6.0 using packages *devtools and lme4.* More details can be found in supplementary material section 1.1 and 1.2 with links to the analysis code.

9.2.5.2 Univariable results

In the univariable cattle models increasing household herd cattle introductions is positively associated with an increased probability of exposure to BVDV (OR 1.54, CI 1.09 – 2.18). For all other pathogens tested, the estimated associations with introductions were positive (odds ratios range between 1.17 and 1.27), although all confidence intervals cross 1, meaning that there is insufficient evidence to support these associations.

Belonging to a sub-village with increasing numbers of cattle introductions is positively associated with the unadjusted odds of BVDV sero-positivity (OR 1.98, CI 1.23 – 3.18) and in the other pathogen models the effect estimates are again positive (OR’s ranging from 1.12 – 1.35) but with confidence limits crossing 1 again suggesting insufficient evidence.

Using the SEEDZ only data the univariable models for the network measures risk factors for cattle pathogen seropositivity show that increasing sub-village in-degree is associated with an increased odds of BVDV seropositivity (OR 4.55, CI 1.38 – 14.95). Increasing betweeness is associated with an increased odds of seropositivity for *Brucella sp.*, (OR 2.05, CI 1.42 – 2.94), *L. hardjo* (OR 1.55 , CI 1.15 – 2.09), ‘any zoonoses’ (OR 1.51, CI 1.16 – 1.98) and BVDV (OR 2.25, CI 1.57 – 3.21). There is insufficient evidence to support a significant association between increasing in-degree and seropositivity for all other investigated pathogens. The odds ratio effect estimates are all positive (range from 1.39 – 3.15) but the confidence limits all cross 1. The effect estimates for increasing betweeness and the odds of *C.burnetii* and IBR seropositivity are also positive (OR 1.19 and 1.25 respectively) but again the confidence limits cross 1 suggesting more evidence is needed to assess the effect of this association.

For the second hypothesis; belonging to a BVDV positive herd will be associated with an increased risk of exposure to other infectious pathogens; the univariable models show strong evidence to suggest a positive association between herd seropositivity and the unadjusted odds of *Brucella sp.* (OR 4.15, CI 1.15 – 14.94) and *L.hardjo* 1.63 (CI 1.04 – 2.55) seropositivity.

Table S2.2. Results of univariable models for potential risk factors for Brucella sp., C.burnetii, L.hardjo, 'Any zoonoses', BHV-1 and BVDV seropositivity in cattle

	Dependent variable: Unadjusted odds ratio (95% confidence limits)					
	Brucella sp.	*C. burnetii*	*L.hardjo*	Any Zoonoses	BHV-1	BVDV
Age	1.28 (1.20, 1.38)***	1.10 (1.03, 1.17) **	1.18 (1.14, 1.21) ***	1.20 (1.16, 1.23) ***	1.62 (1.56, 1.70) ***	1.09 (1.05, 1.12) ***
Sex: Male	0.29 (0.17, 0.51) ***	0.59 (0.39, 0.90)*	0.77 (0.63, 0.94)*	0.70 (0.58, 0.85) ***	0.58 (0.48, 0.70) ***	0.86 (0.70, 1.05)

Household herd size (log 10)	1.93 (1.20, 3.10) **	1.15 (0.76, 1.74)	1.58 (1.19, 2.09) **	1.60 (1.22, 2.10) ***	1.96 (1.52, 2.54) ***	1.25 (0.95, 1.64)
Household cattle introductions (log 10)	1.17 (0.60, 2.27)	1.22 (0.65, 2.27)	1.26 (0.87, 1.84)	1.27 (0.88, 1.84)	1.27 (0.86, 1.88)	1.54 (1.09, 2.18) *
Sub-village herd size (log 10)	1.60 (0.87, 2.92)	0.85 (0.52, 1.39)	2.19 (1.22, 3.92)**	2.16 (1.26, 3.68) **	1.96 (1.24, 3.09)**	7.65 (4.18, 14.01) ***
Sub-village cattle introductions (log 10)	1.26 (0.81, 1.97)	0.96 (0.68, 1.37)	1.31 (0.90, 1.91)	1.35 (0.96, 1.90)	1.12 (0.83, 1.52)	1.98 (1.23, 3.18) **
Sub-village in-degree (log 10)	3.15 (0.95, 10.51)	1.39 (0.50, 3.85)	1.72 (0.68, 4.39)	1.79 (0.76, 4.21)	1.85 (0.91, 3.78)	4.55 (1.38, 14.95)*
Sub-village betweenness (log 10)	2.05 (1.42, 2.94) ***	1.19 (0.85, 1.66)	1.55 (1.15, 2.09)**	1.51 (1.16, 1.98) **	1.25 (0.98, 1.60)	2.25 (1.57, 3.21) ***
herd BVDV status	4.15 (1.15, 14.94)*	0.73 (0.38, 1.39)	1.63 (1.04, 2.55)*	1.36 (0.90, 2.06)	1.43 (0.99, 2.07)	na
Agropastoral production system	0.0078 (0.41, 1.51)	0.025 (0.015, 0.042)	0.325 (0.239, 0.442)	0.396 (0.301, 0.521)	1.76 (1.35, 2.30)	0.369 (0.250, 0.547)
Pastoral production system	5.19 (2.69, 10.02) ***	1.30 (0.74, 2.29)	1.77 (1.20, 2.62)***	1.86 (1.30, 2.65) ***	1.63 (1.15, 2.30)*	2.17 (1.42, 3.32)***
Small holder production system	1.85 (0.70, 4.91)	0.96 (0.42, 2.18)	0.60 (0.33, 1.06)	0.51 (0.29, 0.89)**	0.86 (0.53, 1.39)	1.77 (1.07, 2.93)**
Observations	3,357#	3,330##	2,848	2,853	2,848	2,848

Note: # SEEDZ only variables N observations = 2,846, ## SEEDZ only variables N observations = 2,845

Household ID and sub-village fitted as random effects

$^{*}p<0.05^{**}p<0.01^{***}p<0.001$

The univariable small ruminant models using both SEEDZ and BacZoo data, show a significant association between increasing household and sub-village introductions and the odds of *Brucella sp.* seropositivity (OR 2.61, CI 1.09 – 6.29 and OR 2.12, CI 1.20 – 3.74 respectively). The odds ratio effect estimates for increasing household and sub-village introductions and *C. burnetii* and 'Any Zoonoses' exposure are also positive however the confidence limits cross 1 suggesting insufficient evidence to support this association.

The univariable models using SEEDZ data only to assess the effect of the network measure risk factor variables on pathogen exposure in small ruminants show there is good evidence to suggest increasing sub-village 'betweenness' is positively associated with *C. burnetii* and 'any zoonoses' exposure (OR 1.62, CI 1.10 – 2.39 and OR 1.65, CI 1.10 – 2.46 respectively). Increasing in-degree and betweeness appear to be positively associated with the odds of *Brucella sp.* seropositivity (OR 3.10 and 1.61 respectively) but the confidence limits for these estimates cross 1 which means that there is insufficient evidence to support the association. The univariable model assessing the association between in-degree and *C. burnetii* exposure shows no evidence of an effect as the OR estimate is close to 1 with a wide confidence interval.

Table S2.3. Results of univariate models for potential risk factors for Brucella, Coxiella and 'Any zoonoses' seropositivity in Sheep and Goats

	Dependent variable:		
	Unadjusted odds ratio (95% confidence interval)		
	Brucella sp.	*C.burnetii*	Any Zoonoses
Age	1.27 (1.15, 1.39) ***	1.33 (1.28, 1.38) ***	1.33 (1.29, 1.38) ***
Sex	0.45 (0.28, 0.73) **	0.40 (0.33, 0.48) ***	0.40 (0.33, 0.48) ***
Household herd size (log 10)	4.94 (2.24, 10.86) ***	1.75 (1.36, 2.26) ***	1.88 (1.46, 2.43) ***

Household small ruminant introductions (log 10)	2.61	1.12	1.12
	(1.09, 6.29)*	(0.88, 1.42)	(0.88, 1.44)
Sub-village herd size (log 10)	4.53	2.68	2.89
	(1.63, 12.61)**	(1.77, 4.07) ***	(1.88, 4.44) ***
Sub-village small ruminant introductions (log 10)	2.12	1.31	1.34
	(1.20, 3.74)**	(0.95, 1.79)	(0.96, 1.87)
Sub-village in-degree	3.10	1.05	1.00
	(0.55, 17.55)	(0.40, 2.71)	(0.37, 2.70)
Sub-village betweenness	1.61	1.62	1.65
	(0.78, 3.33)	(1.10, 2.39)*	(1.10, 2.46)*
Pastoral production system #	4.83	3.42	3.64
	(1.72, 13.59)**	(2.13, 5.51) ***	(2.24, 5.93) ***
Small holder production system #	0.10	0.48	0.44
	(0.00, 2.89)	(0.23, 1.00)*	(0.21, 0.93)*
Observations	5,524	5,500	5,524
SEEDZ data only N obs.	4,318	4,316	4,318

Note: # model uses bglmer. See analysis methods.

$^{*}p<0.05^{**}p<0.01^{***}p<0.001$

Household ID and sub-village fitted as random effects

9.2.5.3 Checking for multicoliniarity between risk factor and confounding variables for inclusion in the disease risk models.

In-degree and sub-village introductions appear to be correlated in the small ruminant dataset (correlation 0.66, Kendall’s tau P value <0.001) . All variables were included in the initial multivariable models and the variable with the highest ‘variance inflation factor’ (VIF) was removed in a backwards step-wise process until all model variables had a VIF below 3.0.

Table S2.4. Estimates of level of correlation between livestock disease risk factor variables using Kendall’s tau statistic – assessed prior to inclusion in multivariable models

	Correlation	Kendall’s tau P value
Household number of cattle and number of cattle introductions	0.21 (0.17 – 0.24)	<0.01
Household cattle herd size and sub-village cattle herd size	0.43 (0.41 - 0.46)	<0.01
Household number of cattle introductions and sub-village estimated number of cattle introductions	0.27 (0.24 - 0.30)	<0.01

Sub-village herd size and sub-village estimated number of cattle introductions	0.44 (0.41 – 0.46)	<0.01
Sub-village number of cattle introductions and in-degree (SEEDZ only data)	0.47 (0.44 – 0.50)	<0.01
Cattle data sub-village degree and in-degree	0.75 (0.73 – 0.76)	<0.01
Cattle data sub-village degree and betweenness	0.64 (0.62 – 0.66)	<0.01
Cattle data sub-village in-degree and betweenness	0.52 (0.50 – 0.54)	<0.01
Household number of shoats and number of shoat introductions	0.21 (0.18-0.23)	<0.01
Household shoat herd size and sub-village shoat herd size	0.49 (0.47-0.51)	<0.01
Household number of shoat introductions and sub-village estimated number of shoat introductions	0.32 (0.3-0.35)	<0.01
Sub-village shoat herd size and sub-village estimated number of shoat introductions	0.41 (0.39-0.43)	<0.01
Sub-village number of shoat introductions and in-degree (SEEDZ only data)	0.66 (0.64-0.67)	<0.01
Shoat data sub-village degree and in-degree	0.75 (0.73-0.76)	<0.01
Shoat data sub-village degree and betweenness	0.64 (0.62-0.66)	<0.01
Shoat data sub-village in-degree and betweenness	0.52 (0.5-0.54)	<0.01

9.2.5.4 Multivariable results

Small holder data were excluded from the SEEDZ only data analysis for the *Brucella spp.* and *C.Burnetii* models because there were so few events in these groups to model (N =2) their inclusion resulted in model instability and a failure to converge. In the combined dataset there were N=12 events for each *Brucella spp.* and *C.Burnetii* thus the models were able to be fitted with the small holder data included.

9.3 Chapter 4

9.3.1 Tax collection at markets

Primary markets: Under local government control, so there is a local government tax collector who collects the 'market fee'
Secondary markets are under central government control so there should only be someone from the central government present collecting the market fee. (However, there are also local government revenue collectors at the secondary markets collecting a fee and people do not know that this is not actually necessary and they central government collectors and people who run the market do not want to cause conflict with the local government people collecting extra revenue, this is specifically at Themi and Mesarani markets, I did not see any local government revenue collectors at Weruweru)

It is the buyers who pay the revenue market tax, per head of cattle = 5,000 TzSh, per head of small ruminant = 1,500 TzSh. The tax is collected, by the market local government revenue collector, whether the buyer takes a movement permit or not. The tax receipt should match the movement permit for the number of animals being purchased and moved and these should be checked together as the buyers leave the market with their animals. At primary markets the local government revenue collector uses a receipt machine to issue a receipt of payment and a smart phone app to register the tax/sale. This is supposed to be done at all markets but we didn't see it happening very often. We do not know where the data is stored and exactly what is collected (origin/destination information for example). At secondary markets the market fee is collected by the accountant or similar and an exchequer receipt issued to prove the market tax has been paid.

9.3.2 Supplementary details about livestock movement permits and taxes

By law an animal moving from point A to point B should have a movement permit. If animals are being moved for grazing and then back to a boma they are not required to have a movement permit. Village to village movements do require a permit and the process to get this permit is complex; the animals must be inspected by the LFO to ensure they are healthy (I am not sure how this is done - I assume they have to not be exhibiting any obvious signs of clinical disease, I do not know if they check for ticks, body condition etc.) and there must be 1.a letter from the receiving village or district to to show that there is communal land or that the animal owner has private land in the receiving village that is fine to be grazed. Thus, due to the logistical complexities involved with getting the permit, and the fee involved, it is very rare that they are ever used. My source who spent years checking the permit books at Arusha office estimated that <1% of permits we have will represent village to village movements.

Movement permits are issued at the end of the market day after all of the sales are completed, as people leave the market. Generally animals owned by different people who are going to the same destination are grouped together and one permit is written for the truck or tracker taking the animals onward.

In Mbulu there is one LFO issuing all of the permits at the different markets on different days so he said he could issue one permit to cover the buyer buying animals at more than one market. Movements of animals from markets, within the district, are less likely to take movement permits. They are supposed to, however they are unlikely to get stopped if they are making short journeys so they may try to not take permits. This will be slightly different at different markets depending on the LFO's and peoples attitudes towards the permits and revenue collection and it will be quite difficult to quantify the missing 'within district' movements not represented, we should be able to see from the market Q the proportion of movements that are 'within' and 'out with' the district but in total we interviewed fewer buyers that sellers.

The movement permit fee is different to the market tax revenue and should be paid to the central government. (This is possibly why the revenue collection which stays with local government could be a higher priority than issuing the movement permit and collecting the fee for the permits at primary markets?)
Cattle within district = 1000 TzSh
Cattle outwith district, within region = 1500 TzSh
Cattle outwith region = 2500 TzSh
Shoats within district = 500 TzSh
Shoats outwith district, within region = 1000 TzSh
Shoats outwith region = 1500 TzSh
The DVO is responsible for issuing the movement permits for his district however there are many markets within districts thus he delegates to LFO's associated with local

markets. An LFO can be responsible for attending more than one market and can also delegate some of his role to LEO's. The market master is responsible for overseeing the revenue collection by the local government revenue collectors. The market masters can be other members of local government such as the ward executive officers. The district executive officer is responsible for appointing the market master/supervisor and he can chose anyone he see's fit.

One person is responsible for issuing the movement permit at Meserani and they work with a cashier who collects the money for the permit. This should match the numbers on the permit and is attached to the movement permit. At Mesarani there are 2 livestock field officers (LFO's) and 2 livestock extension officers (LEO's) and an accountant collecting revenue tax, writing the movement permits, collecting the movement permit fee and checking the permits and tax receipts on animals exit from the market. They are still very busy and feel like the job could be done more thoroughly with more man power.

9.3.3 Supplementary notes about livestock markets

Prices at secondary markets appear to be much higher than at primary markets (Pica-Ciamarra *et al.*, 2011).
There is illegal movement of animals from the Northern Zone of Tanzania markets across the border into Kenya which is going to be difficult to capture; the demand and prices in Kenya are higher and there are larger meat processing plants. People moving animals north of the border are generally buying high numbers and trucking them together with others to Longido, or areas close to the border that are accessible by road, then the animals are walked by Massai across the border and picked up by trucks at the other side.

The 2009 Tanzanian National Panel Survey data showed that 63% of livestock keepers sold livestock in their own village, 46% had sold in a neighbouring village, 6% sold in the same region but a different district within the region and 2% sold out with their region (Covarrubias *et al.*, 2012).

Efforts have been made to standardise sale prices of livestock by introducing a standard price per kilo and weigh scales at some markets however there is a reluctance to use this system from both sellers and buyers. Each party thinks this system will allow the other to achieve a more desirable price thus traditional bartering remain the way most livestock prices are agreed. Price is not standardised across districts, regions or times of year.

Within markets, animals in better body condition were found to fetch a higher price (Pica-Ciamarra *et al.*, 2011) but despite this only 20% of surveyed households reported purchasing extra fodder for livestock that would improve their body condition before sale (Covarrubias *et al.*, 2012). The full effects of limiting or redirecting livestock movements and markets are important to consider when developing a pathogen control programme as pathogen flow will not be the only thing that is hindered. Money and information transfer will also be affected.

9.3.4 Market visits and notes on the specific markets:

Participant selection

Middle men were often present at the market and work by both buying and selling livestock in the market on that day but not moving the livestock themselves to or from the market. Although the middle-men are involved in the value chain and run a business as a result of livestock sales they do not physically move livestock between different points and often do not know the details of the origin of the animals they are selling thus they were not asked to participate in the survey .

Market specific details

25/04/17 Weruweru, Hai, Kilimanjaro

Secondary market, close to Moshi, large and busy. Many middle men as people bring large numbers of animals then sell to middle men and they sell on from the smaller groups. for example, a man sells 20 cattle to 4 middle men and they each sell 5 to buyers. Mixed species, animals arriving and leaving by truck. See evidence of people using movement permits to exit and some permits are being collected on arrival.

26/04/17 Makanya, Same, Kilimanjaro, -4.359809, 37.821463

Primary market, shoats only. People predominantly arriving on foot. Some people buying to take home to their own flocks, some buying to truck together with others specifically for meat, some buying for local meat consumption, leaving on pikipiki or daladala. Today was a small market but every other week there is a larger market with cattle here.

28/04/17 Mgagao, Same, Kilimanjaro
Primary market, large and mixed stock. Animals all seem in fair condition. Three large trucks left with shoats. Animals appeared to be arriving on foot. Some middle men for the cattle make finding actual sellers and buyers who are taking away from the markets .

29/04/17 Naibili, Siha, Kilimanjaro
Primary market, mixed livestock, poor to fair condition, animals arriving and leaving on foot, poor access by road, some trading before official market started, some middle men, once market started (market master arrived and everyone inside fenced area) everyone leaving needed a pink slip.

Purchases of 1-2 shoats for local consumption, selling to get money to meet needs. West Kilimanjaro area, the weather here has been good recently with lots of rain.

01/05/17 Arash, Ngorongoro, Arusha

Primary market, shoats only, very rural, all Massai, 2 trucks buying and taking the animals north, they say they are going to Wasso but we suspect they are going further north to Kenya, some do name the Kenyan market. Animals here are in fair condition, there has been some good weather here recently and sheep and goats are a little fitter and ready for sale in comparison to the cattle. People are selling animals to get money to meet the family needs and shoats specifically because they are fit for sale but the cattle are still recovering from the drought and are in poorer condition (so comparatively price for shoats is better at the minute). People tend to be selling small numbers 1-5, but many shoats arrive in large groups of 10-20, being walked together by a number of owners from their own villages. Quite a lot of ticks here and some sheep with raised respiratory rates. One buyer was buying 20 sheep to expand his own flock, most groups being bought were for meat.

02/05/17 Endulen, Ngorongoro, Arusha
Primary market, mixed , animals appear in fair to good condition, prices for cattle here are very low. there are many more sellers than buyers, some people have been coming here repeatedly for a number of weeks and taking their cattle home again due to lack of

buyers and very low prices. There is a noticeably high proportion of smaller cattle intended for sale for breeding and there appears to be a mix of stock being sold, in that people are selling some of their own cattle and some cattle that they bought at other markets and fattened for a few months and are now trying to sell again here. This has not worked out very well though it seems as prices being offered are lower than the original purchase price (Avg. 300,000 TZsh) which is why people have been taking the animals home again in recent weeks.

There are also two women selling here, they have been given money by the NCA to buy/rear and sell cattle as a business. There is a thought here that this is when people are selling large cattle suitable for slaughter due to money needs and that they should want to replace them with new breeding stock.

05/05/17 Themi, Monduli, Arusha

Secondary market, mixed livestock. Many cattle are walked here from Simanjiro. The only market i have noticed so far where the LFO is collecting permits from treckers as they arrive into the market area. The majority of buyers are butchers. Some people are buying to take to Mesarani market, Weruweru (both other secondary markets) or to Arusha meats. A number of buyers will take the cattle to Kenya. The animals here are in fair-good condition. The shoats are gathered and sold in a smaller walled enclosure separate to the cattle and have to pass the LFO to leave the walled area.

06/05/17 Oldonyosambu, Arumeru, Arusha

Primary market, mixed livestock, just off the main road outside Arusha on the Longido road, about 30 minutes drive from Arusha, fairly close to Sakina abattoir (Arusha meats). Animals are arriving here on foot, generally from not far away (1-3 hours walk?) and the animals seem to be in fair condition. Many shoats are leaving in trucks. There are buyers here buying 20-40 shoats and transporting them with others north to Kenya though they are not willing to talk to us about where they are going officially. Animals appear clean and healthy, there are also some calves for sale here which is different, I haven't noticed so many calves for sale at other markets.

08/05/17 Mesarani

Secondary market, only cattle on a Monday, cattle arriving by truck and leaving on trucks. Many walking from Singida to Karatu (6 days) then taking a truck to Mesarani (in groups of 12-30) with other groups of cattle. Cattle are in good condition and going for high average prices, around 750-800,000. Cattle are going to Sakina Abattoir in Arusha or to Longido and then north into Kenya. The people taking cattle north were not very willing to give that answer as a destination as it is illegal.

09/05/17 Njia Panda, Moshi

Primary market, small, 1 hour east of Moshi on the Dar road, very easy road access. Mixed livestock, cattle in particularly are in poor to fair condition. Animals generally arriving on foot in small groups and have traveled for 0-3 hours. The price of cattle here is quite low in comparison to Mesarani/Weruweru. Buyers are using this market as there are animals here that meet their needs, small butchers are buying 1 or 2 because it is there day to slaughter. The slaughter slabs in a village take turns to slaughter (village timetable) and they use the market that fits best with the day they are allocated to slaughter. The prices vary between the local markets but they say they just use whichever is on their allocated day to buy because they don't have the spare capital to buy a cow too far in advance. (But the difference in price could be 200,000 tzsh so I need to look into this, it is interesting that they wouldn't just buy at the closest/cheapest market and hold the animal at there house/butchery till their slaughter day. I wonder how often slaughter days are? Something to ask at village level.)

10/05/17 Mateves, Arusha
Primary market, shoats only, 10km outside Arusha. Some people buying for local abattoirs and slaughter slabs in Arusha, buyers have orders from processors or something similar. There are also a couple of buyers taking many (30-40) shoats to go north to Longido (most likely then into Kenya). These people say they are too busy to answer questions but are more than happy to let us sample the animals they have bought as they are for meat. Animals tend to be arriving on foot and have not travelled far (0-2 hours) and are being sold for family needs. The body condition of sheep is worse than goats and they are dirtier. There are some noticeably ill animals here (n=3). Ormilo? Heart water?

13/05/17 Gilai Bomba, Longido, Arusha
Medium sized, mixed livestock Massai market, all animals arriving on foot, no good road access. Noticed a couple of cattle with cutaneous nodules (lumpy skin disease? cutaneous TB? I need to ask someone about these cattle with lumps all over them, they usually look in poorer condition than other cattle being sold at the same market/in same group).

Some people mentioned that they were buying animals here because they were good prices and then they would keep for a couple of months to fatten and sell - although there isn't a lot of grass here and the rainy season is through so I am not sure how this will work? But it seems like a sensible economic decision, if the animals do grow/fatten. Note, the mapping of grass land and feed back to villages is happening here so people may feel like they have the ability to successfully graze and fatten animals here where previously they may have not.

17/05/17 Gendi, Babati Town council, Manyara

Primary mixed market, by the road side, around 5k out of town, the only market in this town district, this is a relatively new administrative district.
There was slaughter of livestock here and they were cooked on the fires at the market to make nyama chime which is one of the main reasons people like to visit the market. This is an agri pastoral area, the crops are looking good, there has been a lot of rain over the past few weeks. People are selling small livestock animals to get money to buy food, some for school fees. Yesterday there was a large market at Galapo and 200-300 shoats were sold so today there are fewer sellers and animals here and thus the prices are high. Some of the buyers here today are people who we met selling at Oldonyosambu market close to Arusha last week.

18/05/17 Riroda, Babati district council, Manyara
Primary market, fairly small, mixed livestock, 30 minutes drive from Babati, Irac tribe area. People are much less willing to talk and felt like more of a hostile environment. The livestock, particularly the cattle, were in poorer condition, I noticed a couple of thinner animals with SC lumps on the skin over the ribs and shoulder area. There is a notably high proportion of younger animals here. People are buying young stock for growing and fattening.
The price at this market varies, from December to May the price is low. During this time people do not have food so they sell animals to get money to buy food. Then when the crops are harvested they have food and can sell surplus crops for money and then use this money to buy animals to replace the ones they were forced to sell during times of hardship. During this time (June/July onwards) after the rains cattle may be in better condition and the numbers supplied at the market are much lower (there is not the need to sell) so the prices are much higher. (they quoted 250,000 now (mid May), 700,000+

later in the year).

People do not seem concerned with disease and losing the animals that they invest in, they appear to be happy to continue investing their money in this way. (I need to ask more about this at village level to understand how assets and basic needs and luxuries are prioritized).

19/05/17 Dongobesh, Mbulu, Manyara

Primary market, large and mixed stock and separate holding pens for larger (intended for meat market) and smaller cattle (intended for breeding) estimated 200 cattle being sold here today. Shoats are being sold in a separate area. This market happens twice a month and we were there on the busier day. People were buying here and buy at other markets locally then take cattle to Mesarani, the cattle seemed in fairly good condition, the cattle intended for breeding or keeping to fatten up for draught power use were noticeably smaller and thinner and going for much lower prices in the separate area. Animals here appeared to be predominantly arriving and leaving on foot. There were many sellers but few buyers in comparison so the sellers complained that the price was very low in comparison to what they wanted as there was less competition between buyers they could wait and force the price down (market saturation).

20/05/17 Dirim, Mbulu, Manyara

Primary mixed market, also only happens twice a month, crazy cattle (around 300+ today), brought here by different tribes. Cattle mostly in very good condition, Zebu breed. All arriving on foot and leaving on foot, animals here are coming from the rift valley, mainly Tatanga tribe, where people don't cultivate the land they simply graze it so the animals are looking fit and healthy, in good condition. I noticed one cow here with S/C lumps on its back legs. A buyer here was going to sell to a butcher, but said the butcher varies the price he is willing to pay so a good price purchase today doesn't necessarily equate to good profit margins for him.

22/05/17 Engaranaibor, Longido (Mairowa village), Longido

Primary market, mixed small and large ruminants, separate walled pens for both the cattle and the small ruminants. Cattle here were in good condition, mostly Boran breed cattle, around 50 being sold today, there appeared to be a lot of discussion going on, a good sellers:buyers ratio. There were a lot of smaller cattle that were being sold and the owners thought they would be good for breeding but the buyers were mainly there looking for meat/purchasing cattle to sell on either in Kenya or at Oldonyosambu or Mesarani for markets. Again people were bringing 1-3 cattle for sale and buyers were looking to buy around 5. One man was looking for 10-20 depending on price, to supply a small abattoir in Namanga at the border, he will only buy this number if the price is good. This market is close to the border and people often buy here and take to Kenya.

Recently (over the past 1-2 months) people have been finding Ruby at the local mine and the influx of cash is being spent by local people here on livestock, people are buying cattle and goats to expand their existing herds and flocks (possibly why there were more people trying to sell younger good breeding animals), meaning at the minute more buyers here are keeping the animals in Longido rather than exporting them to Kenya so perhaps more livestock are staying in the area in comparison to during normal financial times when the demand for livestock meat from Kenya is higher than it is locally thus driving the flow of livestock into Kenya from this area. There is a higher demand for goats than cattle and sheep here, from people who are buying to expand their herds/ flocks, as the climate and grazing conditions more recently (over the past few years) are thought to suit goats the best (there is less grass available).

4 cattle being sold were tagged to say they were vaccinated against ECF.
This again is in the grazing management satellite imaging district, I am meeting with these people on Monday.

24/05/17 Sukuro, Simanjiro district, Manyara
Primary market, medium sized mixed livestock, cattle and shoats being sold at different sides of the large market. Estimate around 150 cattle being sold here today, it looked like there were some pink slips being given out to people who were buying cattle and there was something happening with the shoats being placed in a pen after purchase, before they left the market they likely needed to get a permit too and it was being enforced.
A buyer told me that this is a good market to buy at as there are few buyers thus competition is low and generally the sellers need to sell. This lady was going to take some animals to her own butchers,

25/05/17 Terrat, Simanjiro district, Manyara
Primary, small mixed local town market, estimated 40 cattle being sold today, perhaps

around 100 shoats, trucks parked close by on the main road which is a good dirt road that goes to Themi market, people are here buying who were also buying at Sukuro market yesterday.

26/05/17 Emboret, Simanjiro district, Manyara
Primary market, small ruminant only, supposed to be a pastoral area but there are some fields of crops here now. 27km from Terrat, there are lots of wildebeest and cattle on the plains around this area. Estimated 150 shoats being sold today, people are selling 1-6 animals, arriving on foot some individually and some as parts of large groups. People interviewed tend to use this market, Sukuru or Terrat for selling their stock. One person is here because he failed to finish seeing all of his goats at Terrat yesterday. There are a number of buyers at this market not willing to talk to us, they are taking the animals to sell at Kenya, Keryobogy market was the only market in Kenya mentioned.

30/05/17 Weruweru, Hai district, Kilimanjaro

Secondary market, re visit to market to collect more questionnaire responses as this was the first market visited. In comparison to the last time we were here, 4 weeks ago, there are more animals here this week, the market appears less hectic because buyers are standing back waiting for the sellers to lower there prices

The market master was collecting permits from people as they arrived with batches of cattle. The permits would be for around 40 animals but those animals are owned by a number of people. All people that I interviewed here were bringing animals here that they had bought at other markets in different areas. At this time (Mid May) a lot of the animals were coming from Tabora region and they were in very good condition, this is within the lake zone where they usually get the rains earlier (November to December) so the cattle are fit for sale earlier than those in other regions where they get the rains later. The price is lower in Tabora from January to March as there are many cattle getting fat and ready for sale, however the price in Tabora goes up from April - May as there are many people going there to buy from Dar Es Salaam, Sinray (a border market) and Nairobi due to the quality of the cattle.

31/05/17 Longido, Longido, Arusha (-2.727363, 36.688137)

Very small informal market, the official market takes place on a Saturday. There were only 3 buyers, all buying small ruminants to take to Kenya. The sales were just being carried out beside a water point. There were mentions of no money changing hands between some people, the sellers would get there money when the buyers returned from Kenya having sold the animals on. I estimated that there were 35 goats and 3 sheep sold here today. The sellers were not hanging around, it was a simple arrive with animals, take money and leave situation. Perhaps because there were so few buyers there was no space for price negotiations or competition. There were no permits and no one collecting revenue.

03/06/17 Longido, Longido, Arusha (-2.727363, 36.688137)

Primary market, official market functioning on a Saturday, mixed small and large ruminants. Animals were held within spikes bush fenced areas, cattle separate from shoats. Estimated 200 shoats (predominantly goats), 60-70 cattle, all buyers interviewed were taking purchased animals to Kenya. Prices were thought to be good at this market. Generally the shoats were in good condition.

9.3.5 Market survey

type	Name	Label	Hint	required	relevant
start	start				
integer	q_number	Questionnaire number			
text	market	Market name		yes	
text	district	District		yes	
text	region	Region		yes	
date	date	Date			
text	origin	Where have you travelled from (Ulisafiri kutoka wapi)		yes	
text	origin_district	Which district did you travel from (Ulisafiri kutoka wilaya gani)			
integer	journey_to_time	How long did it take you to get here (Ilikuchukua muda gani kufika hapa)	Answer must be number of hours	yes	
select_one transport	transport_to_m	What mode of transport did you use to get here (Ulitumia usafiri gani kufika hapa)		yes	
select_one origin_type	origin_type	What type of place did you travel from (Ni sehemu gani ulizotembelea)		yes	
select_one yes_no	selling	Are you selling animals here today (Je unauza wanyama hapa leo)		yes	

select_one species	species_selling	What species are you selling today (Je ni aina gani ya wenyama unauza leo)	yes	selected(${selling},'yes')
integer	nof_selling	How many animals are you selling today (Unauza wangama wangapi leo)	yes	selected(${selling},'yes')
select_one age	age_selling	How old are the animals you are selling (Je ni wanyama wa umri gani)		selected(${selling},'yes')
select_one body_cond_score	avg_bcs_selling	What is the animals (average) Body Condition Score (Je wana afya nzuri)	yes	selected(${selling},'yes')
select_multiple disease_evident	dz_evident_selling	Is there any evidence of disease in this animal(s) (Je kuna dalili yeyote ya ugonjwa)	yes	selected(${selling},'yes')
text	dz_selling_other	What other disease is evident		selected(${dz_evident_selling},'other')
select_one own_animals	own_animals	Are the animals you are selling your own (Je wanyama unaouza ni wako?)	yes	selected(${selling},'yes')
select_one nof_vil_from	nof_villages_from	How many villages are the animals from (Wanyama hawa wanatoka vijiji vinapi?)		selected(${selling},'yes')
select_multiple journey_contacts	contacts_journey_to	Did you contact other herds or flocks on your journey here (Ulikutana na makundi mengine ulipoku unakuja hapa?)	yes	selected(${selling},'yes')
select_multiple reason_species_sold	reason_sold_cattle	Why are you selling Cattle (Kwanini unauza ng'ombe)		selected(${species_selling},'cattle')
select_multiple reason_species_sold	reason_sold_sheep	Why are you selling Sheep (Kwanini unauza kondoo)		selected(${species_selling},'sheep')

select_multiple reason_species_sold	reason_sold_goat	Why are you selling Goats (Kwanini unauza mbuzi)		selected(${species_selling},'goat')
select_multiple reason_species_sold	reason_sold_sheepgoat	Why are you selling Sheep and Goats (Kwanini unauza kondoo na mbuzi)		selected(${species_selling},'shoat')
select_multiple reason_species_sold	reason_sold_cattlesheepgoat	Why are you selling Cattle and Sheep/Goats (Kwanini unauza ng'ombe na kondoo na mbuzi)		selected(${species_selling},'cattle&shoat')
select_multiple reason_animal_sold	reason_sold_animal	Why are you selling these specific animals (Kwanini unauza wanyama hawa?)	yes	selected(${selling},'yes')
text	reason_sold_animal_other	Enter other reason for selling these specific animals		selected(${reason_sold_animal},'other')
select_multiple money_use	money_use	What is the money for (Je pesa unatumia kufanya nini?)	yes	selected(${selling},'yes')
text	money_use_other	Enter other reason the money is for		selected(${money_use},'other')
select_one aim_sell_to	who_selling_to	Who do you wish to sell these animals to (Je ni nani ungependa kumuuzia wanyama hawa?)	yes	selected(${selling},'yes')
text	who_selling_to_other	Enter other person you wish to sell to		selected(${who_selling_to},'other')
select_one no_sale_plan	no_sale_plan	If you do not sell the animals here today what will you do (Je usipouza wanyama hawa leo utafanya nini?)	yes	selected(${selling},'yes')
text	no_sale_plan_other	Enter other no sale plan		selected(${no_sale_plan},'other')

integer	min_price	What is the minimum price per animal you wish to take (Mi kiasi gani ungependa kuuza kwa kila mnyama?)		selected(${selling},'yes')
integer	price_achieved	What price per animal did you get (Je unapata bei gani kwa kila mnyama)		selected(${selling},'yes')
select_one yes_no	buying	Are you buying animals today (Je unanunua wanyama leo)	yes	
select_one species	species_purchased	What species are you buying (Unanunua wanyama gani)	yes	selected(${buying},'yes')
integer	nof_purchased	How many animals are you buying (Unanunua wanyama wangapi)	yes	selected(${buying},'yes')
select_one age	age_buying	how old are the animals you are buying today (Wangama unaonunua wana umri gani)	yes	selected(${buying},'yes')
select_one body_cond_score	avg_bcs_purchased	What is the animals (average) Body Condition Score (Je wana afya nzuri)	yes	selected(${buying},'yes')
select_multiple disease_evident	dz_evident_purchased	Is there any evidence of disease in this animal(s) (Je kuna dalili yeyote ya ugonjwa)	yes	selected(${buying},'yes')
text	dz_evident_purchased_other	Enter other disease evident		selected(${dz_evident_purchased},'other')
select_multiple reason_species_purchased	reason_cattle_purch	What is the reason for purchasing cattle (Je ni kwanini unanunua ng'ombe)		selected(${species_purchased},'cattle')
select_multiple reason_species_purchased	reason_goat_purch	What is the reason for purchasing goat (Je ni kwanini unanunua kondoo)		selected(${species_purchased},'sheep')
select_multiple reason_species_purchased	reason_sheep_purch	What is the reason for purchasing sheep (Je ni kwanini unanunua mbuzi)		selected(${species_purchased},'goat')

select_multiple reason_species_purchased	reason_sheepgoat_purch	What is the reason for purchasing sheep and goats (Je ni kwanini unanunua kondoo na mbuzi)			selected(${species_purchased},'shoat')
select_multiple reason_species_purchased	reason_cattlesheepgoat_purch	What is the reason for purchasing cattle, sheep, goats (Je ni kwanini unanunua ng'ombe, kondoo / mbuzi)			selected(${species_purchased},'cattle&shoat')
select_multiple reason_animal_purchased	reason_animals_purch	What is the reason for purchasing these animals (Je wanyama hawa utawapeleka wapi?)			selected(${buying},'yes')
text	reason_purch_other	Enter reason these animals were purchased			selected(${reason_animals_purch},'other')
text	destination_name	What is the name of the place you take these animals (Je wanyama hawa utawapeleka wapi)		yes	selected(${buying},'yes')
select_one destination_type	destination_type	What sort of place is your destination (Je ni maeneo gani utawapeleka wanyama hawa)		yes	selected(${buying},'yes')
text	destination_type_other	Enter type of destination			selected(${destination_type},'other')
integer	time_destination	How long will journey take (Je safari yako itachukua muda gani)	Enter number of hours journey will take	yes	selected(${buying},'yes')
select_one transport	transport_from_market	How will you travel to the next destination (Je utatumia usafiri gani kuwafikisha wanyama hawa)		yes	selected(${buying},'yes')
select_multiple why_this_market	why_market	Why did you choose to use this market (Kwanini ulichagua kutumia soko hili)		yes	selected(${buying},'yes')

select_one yes_no	more_q_time	Do you have time to answer a couple more questions (Je una muda wa kujibu maswali mengine zaidi)		yes	
select_one yes_no	other_markets	Do you use other markets (Je unatumia masoko mengine)			selected(${more_q_time},'yes')
text	other_markets_name	What other markets do you use (Ni masoko gani mengine unatumia)			selected(${other_markets},'yes')
text	other_markets_name2	What other markets do you use 2 (Ni masoko gani mengine unatumia)			selected(${other_markets},'yes')
select_multiple why_other_markets	why_other_market	Why do you use other markets (Kwanini unatumia masoko mengine)			selected(${other_markets},'yes')
select_multiple others_sale_reasons	others_reason_market	What are the main reasons that other people sell animals at Market (Ni sababu gani inawafanya watu kuuza wanyama sokoni)			selected(${more_q_time},'yes')
select_one yes_no	vag_swab	swab collected (Je mnakusanya ufusi)		yes	
integer	nof_swabs	How many swab samples were taken (Je mnakusanya fusi ngapi)	If more than 1 animal in batch		
select_one yes_no	faeces_sample	faeces collected (Je mlikusanya kinyesi)		yes	
integer	nof_faeces_samples	How many faeces samples were taken (Je mlichukua aina ngapi za kinyesi)	If more than 1 animal in batch		
end	end				

Response Choices

list name	**name**	**label::eng**
transport	a	On foot - alone (kwa miguu - pekeyako)
transport	b	On foot - with other people (kwa miguu - na watu wengine)
transport	c	vehicle - alone (kwa gari - pekeyako)
transport	d	vehicle - with other herders (kwa gari - na watu wengine)
origin_type	homevillage	Home - village (nyumbani - kijijini)
origin_type	hometown	Home - town (nyumbani - mjini)
origin_type	othervillage	Village - other (kijijini - na watu wengine)
origin_type	market	Market (sokoni)
origin_type	slaughterpoint	Slaughter place (sehemu za machinjio)
origin_type	other	Other
yes_no	yes	Yes (ndio)
yes_no	no	No (hapana)
species	cattle	Cattle (ng'ombe)
species	sheep	Sheep (kondoo)
species	goat	Goat (mbuzi)
species	shoat	Sheep & Goat (kondoo na mbuzi)
species	cattle&shoat	Cattle & Sheep/Goat (ng'ombe, kondoo / mbuzi)
species	other	Other
age	<1	less than 1 year (chini ya mwaka mmoja)
age	1-3	1 - 3 years (mwaka mmoja hadi miaka mitatu)
age	4<	4 years or over (miaka minne zaidi)
age	mixed	mixed (miaka mcha nganyiko)
body_cond_score	<1.5	emaciated (amekonda sana)
body_cond_score	1.5-2	thin (amekonda)
body_cond_score	2.5-3	normal (kawaida)
body_cond_score	3.5<	fat (mnene)
disease_evident	skincond	Skin Condition
disease_evident	metritis	Metritis or Endometritis
disease_evident	RFM	Retained Foetal Membranes
disease_evident	lameness	Lameness
disease_evident	Diarrhoea	Diarrhoea
disease_evident	NoneEvident	None Evident

disease_evident	other	Other (mengineyo)
own_animals	all	all (wote)
own_animals	most	most (wengi wao)
own_animals	some	some (baadhi)
own_animals	none	none (hakuna)
nof_vil_from	OneVillage	One village (kijiji kimojo)
nof_vil_from	TwoVillages	Two Villages (vijiji viwili)
nof_vil_from	ThreeOrMoreVillage	Three or more villages (vijiji vitatu zaidi)
journey_contacts	No	No (hapana)
journey_contacts	YesTravelledWith	Yes - made journey with other herds (ndio - ninasafiri na makundi mengine)
journey_contacts	YesWater	Yes - at watering points on route (ndio - katika kituo cha kunywea maji)
journey_contacts	YesGrazing	Yes - at grazing points on route (ndio - katika kituo cha kulisha majani)
reason_species_sold	moneyworth	Worth more money (kupata pesa zaidi)
reason_species_sold	moreavailable	I have more of these available (nina wanyama wengi zaidi)
reason_species_sold	lessimportant	They are less important to me (hawana umuhimu kwangu)
reason_species_sold	highdemand	There is a high demand for these (wanahitajika zaidi)
reason_species_sold	allavail	They are all I have to sell (nahitaji kuuza wote)
reason_species_sold	other	Other (mengineyo)
reason_animal_sold	dzathome	High disease risk at home (hatari kubwa magonjwa nyumbani)
reason_animal_sold	animaldz	animals have disease (wangama wanamagonjwa)
reason_animal_sold	slaughter	For slaughter (kwaajili yakuchinja)
reason_animal_sold	nofoodavail	Not enough food (hakuna chakula chakutosha)
reason_animal_sold	nowateravail	Not enough water (hakuna maji yakutosha)
reason_animal_sold	environmentcatast	Environment problem (matatizo ya mazingira)
reason_animal_sold	moneyotherlivestock	Need money to buy other stock (nahitaji pesa kununua wanyama zaidi)

reason_animal_sold	moneyotherthing	Need money for other things (nahitaji pesa kwa matumizi mengine)
reason_animal_sold	nospace	Not enough housing space (hakuna nafasi ya kuweka mifugo nyumbani)
reason_animal_sold	geneticsbreeding	Genetic reasons / for breeding (kwasababu ya kuzaliana/maumbile)
reason_animal_sold	timeofyear	Time of year you always sell (muda wa kufanya mauzo kwa mwaka)
reason_animal_sold	abortion	Abortion (kutoka kwa ujauzito)
reason_animal_sold	other	Other (mengineyo)
money_use	schoolfee	School fees (ada ya schule)
money_use	celebration	Celebration (sherehe)
money_use	illness	Illness / emergency (ugonjwa / dharura)
money_use	householdpurch	household purchase (kununua vita vya nyumbani)
money_use	goodpriceop	Opportunity for a good price (bahati ya kupata bei nzuri)
money_use	buyothers	to buy other animals (kununua wanyama wengine)
money_use	buisness	buisness (biashara)
money_use	other	Other (mengineyo)
aim_sell_to	meatco	Meat company (kiwanda cha nyama)
aim_sell_to	butcher	Butcher (kwaajili ya kushinja)
aim_sell_to	farmer	Farmer (wakulima)
aim_sell_to	livestockkeeper	Livestock keeper (wafugaji)
aim_sell_to	samevilperson	Same village (mtu wa kijiji kimoja)
aim_sell_to	samedistperson	Same district (mtu wa wilaya mojo)
aim_sell_to	other	Other (mengineyo)
no_sale_plan	leave	Leave them here till next sale (unawaacha hapa)
no_sale_plan	takehome	Take home (utawapeleka nyumbani)
no_sale_plan	othermarket	Take to other market (utawapeleka katika masoko mengine)
no_sale_plan	mustsell	Must sell (lazima uuze)
no_sale_plan	other	Other (menginyo)
reason_species_purchased	replace	to replace lost/dead animal (kurejesha wanyama waliopotea/kufa
reason_species_purchased	expand	to expand herd/flock (kuongeza idadi/akiba)
reason_species_purchased	savedtobuy	have saved money to buy (ulihifadhi fedha ya kununua)

reason_species_purchased	moremeatcarcas	more meat on carcas (nyama zaid kwenye mizoga)
reason_species_purchased	moreavail	more animals available (kuna upatikanaji wa wanyama zaidi)
reason_species_purchased	goodprice	Good price (bei nzuri)
reason_species_purchased	affordmore	can afford to buy more (unaweza kununua zaidi)
reason_species_purchased	meatdemand	they are what people want to eat (watu wanahitaji kula zaidi)
reason_species_purchased	milksale	milk production - buisness/to sell (uzalishaji wa maziwa - kuuza)
reason_species_purchased	milkhome	milk production - home consumption (uzalishaji wa maziwa - matumizi ya nyumbani)
reason_species_purchased	other	other (menginyo)
reason_animal_purchased	health	Looks healthy (wana afya nzuri)
reason_animal_purchased	milkproduction	Milk production (uzalishaji wa maziwa)
reason_animal_purchased	pricemake	Will sell next for a good price (nitowaza wakati mwingine bei nzuri)
reason_animal_purchased	qualitymeat	Good quality meat (nyama nzuri)
reason_animal_purchased	knowownwer	I know the owner (ninafahamu mmiliki)
reason_animal_purchased	qualityanimal	The quality of animal I want in my herd (nilihitaji kuongeza kundi langu)
reason_animal_purchased	pricepurch	Good price (bei nzuri)
reason_animal_purchased	other	Other (menginyo)
destination_type	a10	Other Market (masoko mengine)
destination_type	b10	village (kijijini)
destination_type	c10	Slaughter slab (machinjioni)
destination_type	d10	Abattoir (machinjio makubwa)
destination_type	other	Other (mangineyo)
why_this_market	distance_home	It is closest to my home (karibu na nyumba yangu)
why_this_market	prices	The prices are good (bei ni nzuri)
why_this_market	proximity_town	It is closest to large town / buisness (karibu na mji mkubwa/biashara)
why_this_market	proximity_road	It is easy access by roads (ni rahisi kufika kwa barabara)
why_this_market	good_animals	The animals are good quality for my needs (wanyama niwazuri kwa mahitaji yangu)
why_this_market	tradition	tradition (ni utamaduni wa kutumia soko hili)

why_this_market	social	It is a socially good market to see friends (ni soko jumuia kuona ndungu na marafiki)
why_this_market	type_animal	There is the type of animal here I need to sell or buy (kuna aina wanyama nahitaji kuona/kununua
why_this_market	other	Other (mengineyo)
why_other_markets	price	Prices (bei)
why_other_markets	season	Season (msimu)
why_other_markets	sell_more	Selling more (kuuza zaidi)
why_other_markets	sell_less	Selling less(kuuza kidogo)
why_other_markets	distance	Distance to travel (umbali kusafiri)
why_other_markets	weather	Weather change (mabodiliko ya hali ya hewa)
why_other_markets	buy_more	Buying more (kununua zaid)
why_other_markets	buy_less	Buying less (kununua kudogo)
why_other_markets	social	Social factors (sababu za kijamii)
why_other_markets	other	Other (mengineyo)
others_sale_reasons	others_money	To get Money (kupata pesa)
others_sale_reasons	others_tomany	thave Too many animals (wana wanyama wengi sana)
others_sale_reasons	others_dz	Disease (magonjwa)
others_sale_reasons	others_goodprices	Good prices (bei nzuri)
others_sale_reasons	others_other	other (mengineyo)

9.3.6 Table S3.1

Table S3.1. Summary of seller and buyers market survey data variables stratified by market type and species group

	Primary		Secondary	
Variable	**Small ruminant** **N (%) /** **Mean (min, max)**	**Cattle** **N (%) /** **Mean (min, max)**	**Small ruminant** **N (%) /** **Mean (min, max)**	**Cattle** **N (%) /** **Mean (min, max)**
Seller surveys	**283**	**137**	**16**	**66**
Number in batch being sold	3.4 (1,30)	2.8 (1, 35)	5.7 (1, 38)	11.6 (1,46)
Vehicle to market	52 (18%)	7 (5%)	2 (13%)	39 (59%)

Foot to market	230 (82%)	130 (95%)	14 (88%)	25 (38%)
Mixed transport to market	1 (<1%)	0	0	2 (3%)
Shared journey to market	150 (53%)	82 (60%)	9 (56%)	48 (73%)
Contact with local livestock on journey	49 (17%)	35 (26%)	3 (19%)	11 (17%)
Business sale	34 (12%)	20 (15%)	6 (38%)	32 (48%)
Own animals	265 (94%)	126 (92%)	14 (88%)	40 (61%)
Livestock originate from more than 1 village	15	8	4	29
Livestock originate from another market	8	10	0	25
Livestock originate from another village	36	20	2	8
Livestock originate from another home	238	107	14	31
Journey to market duration (hours)	3.4 (0.25, 96)	10.3 (0.3, 120)	13.6 (1, 48)	30.1 (1, 148)
Journey to market distance (km)	30.1 (0.9, 227.9)	28.6 (1.2, 299.9)	19.1 (1.1, 85.8)	177.0 (4.6, 510.8)
Buyer surveys	**132**	**46**	**9**	**22**

Number in batch purchased	**17.4 (1, 500)**	**5.6 (1, 25)**	**7.9 (1, 40)**	**9.6 (1, 100)**
Vehicle From market	60 (45%)	10 (22%)	1 (11%)	13 (59%)
Foot from market	70 (53%)	34 (74%)	8 (89%)	9 (41%)
Mixed transport from market	2 (2%)	2 (4%)	0	0
Shared journey from market	77 (58%)	29 (63%)	3 (33%)	17 (77%)
Destination type home	35 (27%)	10 (22%)	1 (11%)	1 (5%)
Destination type slaughter	26 (20%)	12 (26%)	2 (22%)	15 (68%)
Destination type market	61 (46%)	20 (43%)	5 (55%)	3 (14%)
Destination type other/village	8 (6%)	2 (4%)	1 (11%)	3 (14%)
Purchased for meat	27 (20%)	13 (28%)	1 (11%)	12 (55%)
Purchased for price	73 (55%)	28 (61%)	7 (77%)	9 (41%)
Purchased for health	46 (35%)	20 (43%)	0	11 (50%)
Journey from market duration (hours)	24.1 (0.25, 192)	20.9 (0.25, 96.0)	3.3 (1,0, 6.0)	5.4 (1.0, 72.0)
Journey from market distance (kms)	73.4 (1.6, 296.6)	57.8 (7.5, 226.7)	23.4 (7.3, 66.9)	25.9 (4.8, 154.0)

9.3.7 Distance imputed where origins and destinations were not located

23 listed origins and 12 destinations were not assigned location coordinates thus were not plotted on the network map and distances for these journeys were imputed from a GLMM that estimated distance for livestock sold and purchased based on market type, species, number in batch and mode of transport used.

In total 20% of sellers used vehicle to get there livestock to market. 82% of these shared the journey with others. Stratified by market type only 14% of sellers at primary markets used vehicle to get their livestock there and again 80% of these journeys were shared with others. At secondary markets 50% of sellers used vehicle to get there livestock to the market and 85% of journeys were shared with others. For livestock being taken to market on foot around 50% of journeys were made with others (livestock mixed but less intensely compared to sharing a journey in a vehicle).

BLME was used to fit model to shared journey data due to small number of small ruminants (N=2) transported by vehicle to secondary markets using a prior with a very low estimate.

Code for models for differences in modes of transport used in 310:360 R script Market_data_analysis.R

Reason for sale

Reason sold animal – this is important for us to understand as it is the reasons for sale that drive the movement of livestock which in turn drives the spread of pathogen across the study area

Table S3.2. Summary of reported reasons for livestock sale

Reported reason for sale	Number of respondents listing this reason
Business	96
Value of animal/healthy/ good price at this market	76
Demand for meat/slaughter	13
For breeding	4
To generate money to buy other livestock	10
To generate money to buy other things	199
We need money for a problem at home	41
We have no food available	18
There is no water available	7
We have surplus stock/no space	9
Environmental problem	2
There is disease at home	2
Other	214

Money use

Table S3.3. Summary of what the money generated from livestock sales will be used for

Reported money use	Number of respondents listing this use
Business	96
Buying other livestock	19
Celebration	20
Household purchase	223
To treat Illness	125
For savings	6
School fees	154

Good price opportunity	1
Buy food	68
Buy drugs	13
Building work	13
Buy clothes	3

Who are you selling to

Table S3.4. Summary of the people sellers are hoping to sell their livestock to	
Anyone	**353**
Butcher/meat company	**57**
Livestock keeper	**33**
Other	**65**

No sale plan

Table S3.5. Summary of what people will do if they don't sell the livestock at this market	
Leave animals here	42
No option – they must be sold	12
Take to other market	47
Take home	392
Graze/take home then return here	8
Graze/take home then other market	2
Don't know	1

9.3.8 Price

Price wanted

Table S3.6. summary of minimum price wanted for livestock being sold at primary and secondary markets

	Minimum	**25th quintile**	**Median**	**Mean**	**75th quintile**	**Max**	**NA**
Cattle primary	**0**	**230,000**	**300,000**	**320,636**	**400,000**	**800,000**	**27**

Cattle secondary	**30,000**	**400,000**	**500,000**	**524,830**	**600,000**	**1,200,000**	**7**
Small ruminant primary	**0**	**45,000**	**55,000**	**57,588**	**70,000**	**200,000**	**23**
Small ruminant secondary	**0**	**38,750**	**55,000**	**49,375**	**60,000**	**95,000**	

Price achieved

Table S3.7. summary of price achieved for livestock being sold at primary and secondary markets

	Minimum	25th quintile	Median	Mean	75th quintile	Max	NA
Cattle primary	**0**	**250,000**	**300,000**	**350,233**	**450,000**	**1,200,000**	**51**
Cattle secondary	**250,000**	**400,000**	**460,000**	**489,130**	**550,000**	**1,000,000**	**20**
Small ruminant primary	**0**	**40,000**	**50,000**	**51,357**	**60,000**	**160,000**	**53**
Small ruminant secondary	**30,000**	**31,250**	**47,500**	**46,071**	**50,000**	**60,000**	**2**

9.3.9 Origins

Own animals

Table S3.8. Number of sellers that are selling their own animals or others for trade

	Primary market cattle	Secondary market cattle	Primary market small ruminant	Secondary market small ruminant	Primary market cattle and small ruminants
All own	126 (92%)	36 (55%)	265 (94%)	14 (88%)	3 (60%)
Some own	6 (4%)	7 (10%)	8 (3%)		2 (40%)

None own	5 (4%)	23 (35%)	10 (3%)	2 (12%)

The majority of sellers at primary markets are selling their own animals (92% cattle 94% small ruminants), at secondary markets a higher proportion of sellers are selling livestock for trade that are not their own (45% cattle (10% selling some of their own 35% selling none of their own), and 12% of small ruminant sellers).

Table S3.9. Descriptive data binomial and gaussian generalised linear mixed models summary statistics.

Model	**Model type**	**No. of obs.**	**No. of market random effects**	**Variance**	**Marginal R^2**	**Conditional R^2**
Sellers Vehicle transport	*Binomial GLMM*	*499*	*22*	*1.74*	*0.05*	*0.25*
Buyers Vehicle transport	*Binomial GLMM*	*205*	*22*	*2.98*	*0.06*	*0.46*
Sellers share journey	Binomial GLMM	499	22	0.06	0.07	0.08
Buyers share journey	Binomial GLMM	205	22	0.55	0.17	0.27
Sellers contacts at grazing and water points	Binomial GLMM	499	22	0.17	0.02	0.04
Sellers reason sale business	Binomial GLMM	499	22	0.92	0.07	0.18
Sellers Cattle price wanted	Gaussian GLMM	167	17	7.2×10^9	0.21	0.39
Sellers Cattle price achieved	Gaussian GLMM	131	16	9.7×10^9	0.12	0.37
Sellers Small ruminants price wanted	Gaussian GLMM	269	21	4.8×10^7	0.004	0.10
Sellers Small ruminants price achieved	Gaussian GLMM	235	21	8.2×10^7	0.01	0.19
Sellers from home	Binomial GLMM	499	22	0.24	0.05	0.09
Destination market	Binomial GLMM	205	22	0.81	0.09	0.24
Destination meat	Binomial GLMM	205	22	0.34	0.08	0.14
Destination Home	Binomial GLMM	205	22	0.11	0.06	0.08
Batch size sold	Gaussian GLMM (log transformed outcome)	502	22	0.12	0.19	0.33
Batch sizes purchased	Gaussian GLMM (log transformed outcome)	209	22	0.20	0.04	0.18
Journey duration To markets	Gaussian GLMM (log transformed outcome)	499	205	0.03	0.35	0.37

Journey duration From markets	Gaussian GLMM (log transformed outcome)	205	205	0.05	0.32	0.44
Journey distance To markets	Gaussian GLMM (log transformed outcome)	499	205	0.41	0.31	0.67
Journey distance From markets	Gaussian GLMM (log transformed outcome)	205	205	0.05	0.29	0.51

9.4 Chapter 5

Supplementary information for : Analysing livestock network data for infectious diseases control: an argument for routine data collection in emerging economies

G.L. Chaters[1,a], P.C.D. Johnson[1,a], S. Cleaveland[1], J. Crispell[2], W.A. de Glanville[1], T. Doherty[3], L. Matthews[1], S. Mohr[1], O.M. Nyasebwa[4], G. Rossi[3], L.C.M. Salvador[3, 4, 5], E. Swai[6], R.R.Kao[3,b]

[1] Boyd Orr Centre for Population and Ecosystem Health, Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, UK
[2] School of Veterinary Medicine, University College Dublin, Ireland
[3] Royal (Dick) School of Veterinary Studies and Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian, Scotland, UK
[4] Department of Infectious Diseases, University of Georgia, Athens, Georgia, USA
[5] Institute of Bioinformatics, University of Georgia, Athens, USA
[6] Department of Veterinary Services, Ministry of Livestock and Fisheries, Tanzania;

[a] These authors contributed equally to this work
[b] Corresponding author

9.4.1 Network dynamics

Eigenvector centrality and dynamic infection processes.

If the contact network between nodes in a network is represented as an unweighted adjacency or contact matrix *M* (where a link between two individuals *i* and *j* is represented by a 1 at position m_{ij} or 0 otherwise), then for some initial disease vector v_0, and some probability of infection *p,* the product $p.M.v_0$ estimates the distribution of infection probability across all nodes in the first generation so long as the clustering coefficient (i.e. the proportion of triplet nodes where A is connected to B is connected to C and A and C are also connected) and/or *p* are sufficiently small. For an *SIS* infection process where susceptible individuals when exposed become immediately infectious and upon recovery, become completely susceptible again and the probability of recovery before re-infection is high (e.g. if the density of infected locations is always low, or the recovery time is shorter than the intergeneration time), then $p^n.M^n.v_0$ is an estimate of the distribution of infection probability in the n^{th} generation. We shall assume here that the matrix is square and irreducible (this is only the case if all nodes are part of the same strong component), and note that it is also positive definite (i.e. real and non-negative); in this case, the Perron-Frobenius theorem applies and the lead eigenvalue of the matrix is guaranteed to be real and positive. By definition, for an eigenvector e_i of M (i = 1 to n_p, where n_p is the dimension of the matrix),

$$\begin{aligned}\lim_{n\to\infty}\sqrt[n]{(p.M)^n}\,.e &= p.\lim_{n\to\infty}\sqrt[n]{M^n}\,.e\\ &= \lambda.e\end{aligned}$$

where λ is an eigenvalue of M, and $e_i=\sum_{j=1}^{n_p}\alpha_j u_j$ where α_j are coefficients associated with the unit vectors u_j. As any vector v can be written as a sum of eigenvectors, provided v includes the lead eigenvector, then not only is α_j the eigenvector centrality value associated with node j, it also represents the relative proportion of time that the node is infected over the long-term evolution of the epidemic. This is analogous to the next generation matrix (NGM) definition of the basic reproduction number R_0 except that the contact matrix considers individuals, while the NGM definition considers populations and thus issues of node clustering and the requirement that infectious nodes quickly become susceptible again are not issues in the NGM definition.

9.4.2 Mulitplexes, multi-layer networks and multi-host pathogen systems

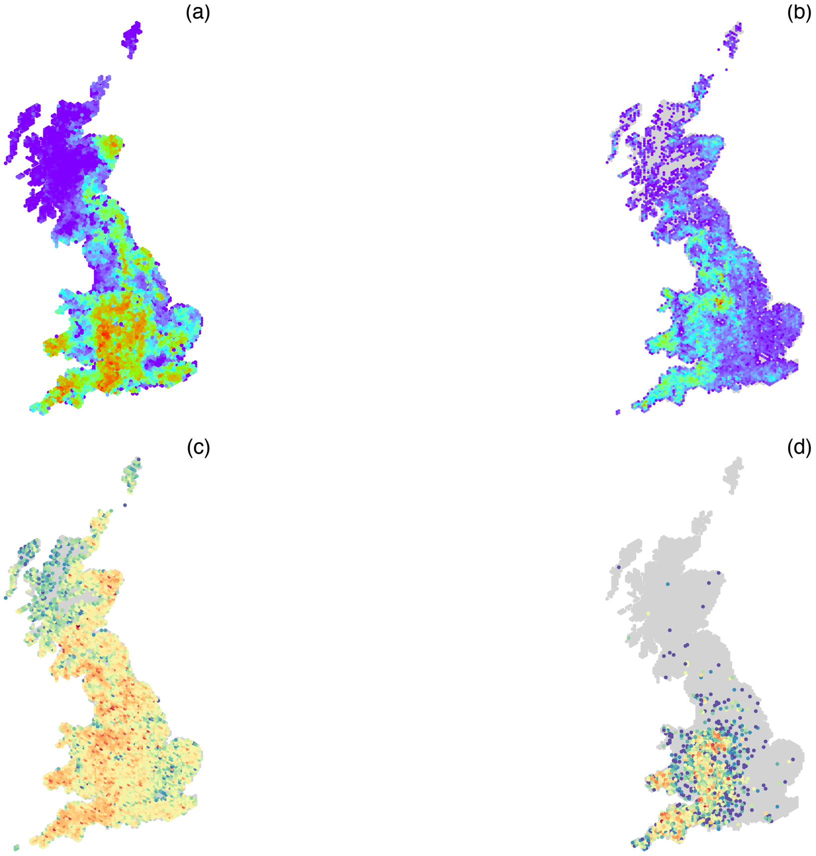


Figure S1. Maps of Great Britain, showing density (per 100 km^2 hexagonal tile) of activity associated with three layers contributing to incidence of cattle with positive bovine Tb test results (reactors). Colour scale reflects proportional reduction compared to the largest calculated value (red highest, blue lowest). (a) Estimated density of badger main setts across GB, based on (Croft, Chauvenet and Smith, 2017) (b) Cattle numbers as of 01/01/2013. (c) Geometric mean of total number of inward cattle movements multiplied by total number of outward cattle movements. (d) Number of test positive cattle.

Livestock movement permit data analysis methods

Data source

Access was granted to all archived government movement permit receipt books from the study regions (Arusha, Manyara and Kilimanjaro) at the Northern Zonal Veterinary Office, Arusha. Movement permit receipt books were selected for analysis from 2009, 2011, 2013 and 2015. A permit receipt book consists of 50 consecutively numbered permit receipts (referred to here as permits).

Data processing

Data from all of the available receipt books from 2009, 2011, 2013 and 2015 were entered directly into a spreadsheet (n = 5,045) or photographed and stored as JPEG files (n = 56,849) between September 2016 and March 2017. The data recorded were date (year, month and day), origin, destination, number of animals of each species (cattle, sheep and goat) moved, and permit number. Owing to the considerable effort required to enter data from thousands of permits, data was entered from only 50% of the permit JPEGs, as selected follows. Prior to data entry, permit JPEGs were ordered by district, then by year within each district. Consecutive permit JPEGs were allocated to each of twenty batches in turn so that each batch contained a representative subsample. Data from ten of the twenty batches of JPEGs were entered into spreadsheets, resulting in a database from 28,421 (50%) photographed permits. The 5,045 directly entered permits were allocated to batches and subsampled down to 2525 (50%) records in the same way as the JPEGs, to avoid these permits being over-represented in the final raw data set, which contained 30,946 records.

Data cleaning

Data were cleaned using an R program to detect anomalous data, correcting it where possible using the stored JPEGs, and deleting records that could not be corrected. First, removal of records where the data entry technicians had indicated that the permit was blank or unreadable reduced the database to 26,855 permits. The origin and destinations recorded on the permits were matched, using a fuzzy text-matching program written in R, against a database of Tanzanian geographic names compiled from the Geographic Names Database (http://geonames.nga.mil/gns/html/namefiles.html; file dated 10 April 2017) and the National Bureau of Statistics 2012 Population and Housing Census of Tanzania. Central point coordinates were assigned to locations that matched to origins and destinations. Fuzzy matching was used as a guide only; all origins and destinations were checked visually against the JPEG, where available, and ambiguous matches were adjudicated with guidance from Tanzanian colleagues with local knowledge. Where possible, missing year was imputed from the preceding and subsequent permits, ordered by permit number, where the two permit numbers differed by less than 100 and bore the same year. Month was imputed similarly, except that where the subsequent permit bore a later month than the preceding permit, the mean month was imputed. Following imputation of missing dates, 22,538 records with all of the following data were retained: year, month, origin, destination, and number of each species moved. Duplicate records were identified either by having the same permit number or highly similar data, verified by comparing JPEGs, then removed, leaving 21,316 records. Finally, permits from years outside the four target years were removed, leaving 19,438 complete permit records, recording 112,531 cattle movements (mean 8.1 cattle per permit, range 1-337), 11,900 sheep movements (mean 6.0 sheep per permit, range 1-85) and 47,201 goat movements (mean 10.4 goats per permit, range 1-180).

For the present analysis, only cattle movements were analysed. Data were aggregated temporally within 48 (12 months × 4 years) calendar months and spatially within the 398 wards (administrative units of around 12,000 people across the study regions), resulting in a database recording the number of cattle moved in each of the 48 months between each pair of wards. Local (within-ward) movements were not analysed because of suspected non-compliance with the permit system for short movements, and movements to outside the three study regions were also omitted.

Detection of missing data

A major obstacle to inferring the movement network from the permit data was the large number of non-randomly missing permits. Frequently permits were missing from locations and time periods that were known to be active from local knowledge and trade volume data (Livestock Information Network Knowledge System; http://www.lmistz.net). To distinguish true from artefactual absence of movement (months where an origin ward sent out no cattle) a zero-inflated negative binomial (ZINB) generalised linear model (GLM) was fitted to each origin ward. Significant zero-inflation ($P < 0.05$) was detected at 16 of the 112 origin wards (wards with at least one outward movement of any species), and in each case a clear majority of zeroes were predicted to be false (range 89-100%), therefore all zero months for these wards were assumed to be due to missing data, and were removed to allow them to be imputed in the subsequent statistical modelling. The final data set recorded the movement of 86,195 cattle from 98 origin wards to 239 destination wards over the 4 sampled years.

Statistical model of inter-ward cattle movement

Inter-ward livestock movement was modelled using a hurdle model, which represented movement between each pair of wards in a given month as two processes: the binary event of any cattle being moved, modelled as a binomial generalised linear mixed-effects model (GLMM); and the number of animals moved, given that at least one animal was moved, modelled as a zero-truncated negative binomial (ZTNB) GLMM. Both parts of the hurdle model allowed movement to depend on the distance between origin and destination wards and their masses (human and cattle population sizes), in addition to other characteristics (Table S1), so the combined model components can be viewed as a gravity model of the livestock movement network. Unexplained spatial and temporal variation was modelled by fitting normal random effects for origin and destination ward and for time period (48 months). The same fixed and random effects were fitted in both parts of the model. Characteristics of origin and destination wards that were included as fixed effects were: $\log_{10}$ human population size, $\log_{10}$ cattle population size, $\log_{10}$ area in km^2, all of which were continuous, and presence of a primary (N=81) or secondary market (N=3), and production system classification (agropastoral N=159; pastoral N=55; smallholder N=150; urban N=34), which were categorical. In addition, log distance in km between ward centroids (continuous), calendar month (continuous), and year (categorical) were fitted as fixed effects. To allow their relationships with movement to deviate from linearity, all continuous variables were fitted as natural cubic splines with three degrees of freedom. Models were fitted using the glmmTMB (Brooks *et al.*, 2017) package for R version 3.5.0 (R Core Team, 2019a). Distributional assumptions were checked by inspecting plots of residuals against fitted values. The predictive performance of each stage of the hurdle model was assessed by calculating a modified R^2 statistic which we term R^2_{LATENT} because it focusses on the variance components on the latent scale (i.e. the transformed scale where the model is linear). R^2_{LATENT} is the fixed effects variance as a proportion of the total linear predictor variance (composed of the fixed effects plus the three random effects), and can be interpreted as gauging the predictive power of the fixed effects to explain spatial and temporal variation in cattle movements.

Predictor variables for the hurdle model

- **Ward area** shapefile with ward boundaries from 2012 Tanzania national census data (World Bank *et al.*, 2011).

- **Human population size** from 2012 Tanzania national census data (World Bank *et al.*, 2011).
- **Cattle population size** Estimates of cattle population numbers in 313 of the 398 wards were provided by local District Veterinary Officers via the Directorate of Veterinary Services (DVS) of Tanzania. Cattle numbers for the remaining 85 wards were imputed by linear regression (R^2 = 42%) of the log-transformed 313 DVS estimates onto log-transformed estimates from Gridded Livestock of the World (Wint and Robinson, 2007) (map: "Predicted global cattle density (2005), corrected for unsuitability, adjusted to match observed totals)", downloaded from *http://www.fao.org/ag/againfo/resources/en/glw/ GLW_dens.html* on 2017-08-11).
- **Ward classification** created using recently updated Tanzania northern zone village classification data developed by W.A.d.G. The village classification model assigns a classification based on the highest probability of it being 'agripastoral', 'pastoral', or 'smallholder' generated by the model. Ward classification was determined by the most common classification among constituent villages.
- **Primary or Secondary market presence**; A list of all active or recently active markets in the study regions was taken from The Zonal Veterinary Centre in Arusha and two binary variables were created for each ward; 'presence of primary market' and 'presence of secondary market'.

Simulation of livestock movement

Three quantities (fitted probability of any movement between each pair of wards; fitted rate of cattle movement given any movement; and the estimated dispersion parameter of the ZTNB distribution) were used to simulate livestock movements out of markets among the 398 wards in the study area for each of the 12 months of the year 2015, conditioning on both the fixed and random effects. Before simulating the movements, the probabilities and rates of movement were amplified twofold to account for the 50% subsampling of the permit data. This was achieved by first multiplying the probability of movement using the expression $1 – (1 – p)^2$, which approximately doubles low probabilities ($p < 0.1$) but has less effect on larger probabilities, then scaling the rate of movement so that the overall increase in expected number of movements was exactly twofold.

We simulated five kinds of market-related movement: from herd to primary market; from primary market to secondary market; from primary market to ward; from secondary market to secondary market; and from secondary market to ward. The three wards containing secondary markets (Bwawani, Machame Kusini and Meserani) were treated as if the ward was a secondary market, so that these wards had to be empty of cattle after each round of movements. Where the model predicted imbalances in cattle flows through secondary markets (for example, typically inflow to a secondary market was less than outflow, probably because permits stored at large secondary markets are less likely to be lost than those stored at primary markets), this was dealt with by boosting the deficient cattle flow rate so that inflow and outflow were balanced. Because permits are only generated by markets, wards with no recorded outflow of cattle over the four years studied were assumed not to contain markets. Such wards were assumed to export cattle only locally, supplying the nearest market. The assumption that most cattle sold at primary markets come from local herds is based on the experiences of E.S. and O.M.N., who have worked in the livestock industry and for the government veterinary services for several years, and G.C., who has conducted an unpublished survey investigating the origins of livestock at 24 primary livestock markets. These wards with no markets (n = 287) were therefore linked as feeder wards to the nearest active ward containing a primary market (n = 108), and the outflow from this active primary market ward was balanced by creating inflow divided evenly among the nearest feeder wards. Any remaining imbalances were evened out by creating births and deaths. By following this scheme, the cattle population remained stable at 3.7 million, with approximately 30,000 movements, 1,200 births and 1,200 deaths occurring each month.

Network measures for targeting interventions

The simulated livestock movement data were used to calculate three measures of network centrality (betweenness centrality, degree centrality and eigenvector centrality) for each ward with the aim of targeting the two types of intervention (market movement ban and vaccination at 70% coverage) to influential wards that are potentially important for disease transmission. The package *igraph* (Csardi and Nepusz, 2006) for R (R Core Team, 2019a) to derive a year-aggregated, static, directed, weighted movement network for cattle. A spatial contact layer was added to the market movements network as a simplified means of accounting for contacts that happen at waterholes and grazing points and via the transfer of animals between households as gifts or financial support. For this example, each ward was connected to all spatially adjacent wards via a single link with a probability one. Betweenness centrality, degree centrality and eigenvector centrality were calculated for each ward from the resulting multiplex network.

Simulating disease outbreaks

To conduct an example simulation of pathogen transmission on the network and investigate the effects of targeted interventions of epidemic spread, some simplifying assumptions were made. Each ward was assumed to have a homogeneously mixing population of cattle, with the population size estimated as described above. To avoid underflow in the simulations, wards with fewer than 1000 cattle (n = 38) were assumed to have 1000 cattle, except for the three secondary market wards which were assumed to hold zero cattle. One cattle per month was moved between adjacent wards (along the spatial network) to capture short-distance non-market movements across wards boundaries. The choice of one animal moved per month between adjacent wards was taken from a separate unpublished model of Rift Valley fever (RVF) spread among cattle across the same regions and using the same movement data. In the RVF model, we calibrated the number of cattle moved across ward boundaries so that when combined with longer-range market movements we observed a realistic rate of spread across the study area when compared with published data (Sindato *et al.*, 2014).

Random introduction and transmission of both fast and slow transmitting pathogens was simulated on the cattle movement multiplex network using the SimInf (Widgren *et al.*, 2016) package for R. A stochastic SIR model with frequency-dependent transmission was simulated within each ward. R_0 was 3 for the fast disease and 1.5 for the slow disease, and the mean infectious period was 7 days, corresponding to transmission rates (β) of 0.429 (fast) and 0.214 (slow) and a recovery rate (γ) of 0.143. In each simulated epidemic, 10 infected cattle were introduced into each of five “seed” wards, and the epidemic was simulated over 12 months. To allow the average intervention effects to be estimated, each epidemic scenario was run 79 times, each time starting from a different set of five seed wards, after which all the disease had been seeded into all 395 wards with cattle populations (i.e. not including the three secondary markets), allowing the effect of seed wards to be balanced between scenarios. To further reduce the effect of sampling variation when comparing interventions, each set of 79 simulations was repeated 3 times, so that each scenario was run a total of 237 times. For each type of intervention, market movement ban and vaccination at 70% coverage, seven intervention scenarios were simulated, including worst- and best-case scenarios, three interventions targeted using network measures, and two non-network-targeted controls:

- No intervention (worst case scenario)
- Intervention applied to all wards (best case scenario)
- Targeting of the intervention in 5% all 398 wards (n = 20), selected for:
    - Highest betweenness centrality
    - Highest degree centrality (geometric mean of indegree and outdegree)
    - Highest eigenvector centrality
    - Highest number of cattle ("common sense" network-free intervention scenario)
    - Random selection (non-targeted negative control to gauge the effect of reducing effort from all wards to 5% of wards)

The effect of each intervention was estimated as the percentage reduction in population cumulative incidence (PCI) at 1 year relative to the no-intervention scenario. This was calculated as one minus the geometric mean across the 237 simulations of the intervention PCI divided by the no-intervention PCI, expressed as a percentage. The standard error of the geometric mean was calculated according to (Norris, 1940).

Results

Table S5.1. **Predictors of monthly inter-ward cattle movement fitted in the two stages of the hurdle model of cattle movement. Continuous variables were fitted as natural cubic splines with three degrees of freedom. Effect size: reduction in model mean sum of squares when predictor is removed. R^2_{LATENT}: proportion of spatial and temporal variation explained by the fixed effects. LRT: likelihood ratio test.**

	Model stage predicting probability of movement (R^2_{LATENT} = 40%)		Model stage predicting no of animals moved (R^2_{LATENT} = 24%)	
Predictor	**Effect size**	**LRT P-value**	**Effect size**	**LRT P-value**
$\log_{10}$(distance/km) [spline]	58%	<0.001	0%	<0.001
$\log_{10}$(origin human pop. size) [spline]	4%	0.053	16%	0.125
$\log_{10}$(destination human pop. size) [spline]	14%	<0.001	0%	0.197
$\log_{10}$(origin cattle pop. size) [spline]	0%	0.952	-1%	0.349
$\log_{10}$(destination cattle pop. size) [spline]	0%	0.968	4%	0.304
$\log_{10}$(origin area/km^2) [spline]	1%	0.274	6%	0.910
$\log_{10}$(destination area/km^2) [spline]	2%	0.028	0%	0.280
Calendar month [spline]	0%	0.054	5%	<0.001
Year [categorical]	3%	<0.001	2%	<0.001
1ary/2ary market in origin/destination [categorical]	27%	<0.001	27%	<0.001
Origin production system [categorical]	6%	0.036	29%	0.039
Destination production system [categorical]	6%	<0.001	2%	0.759

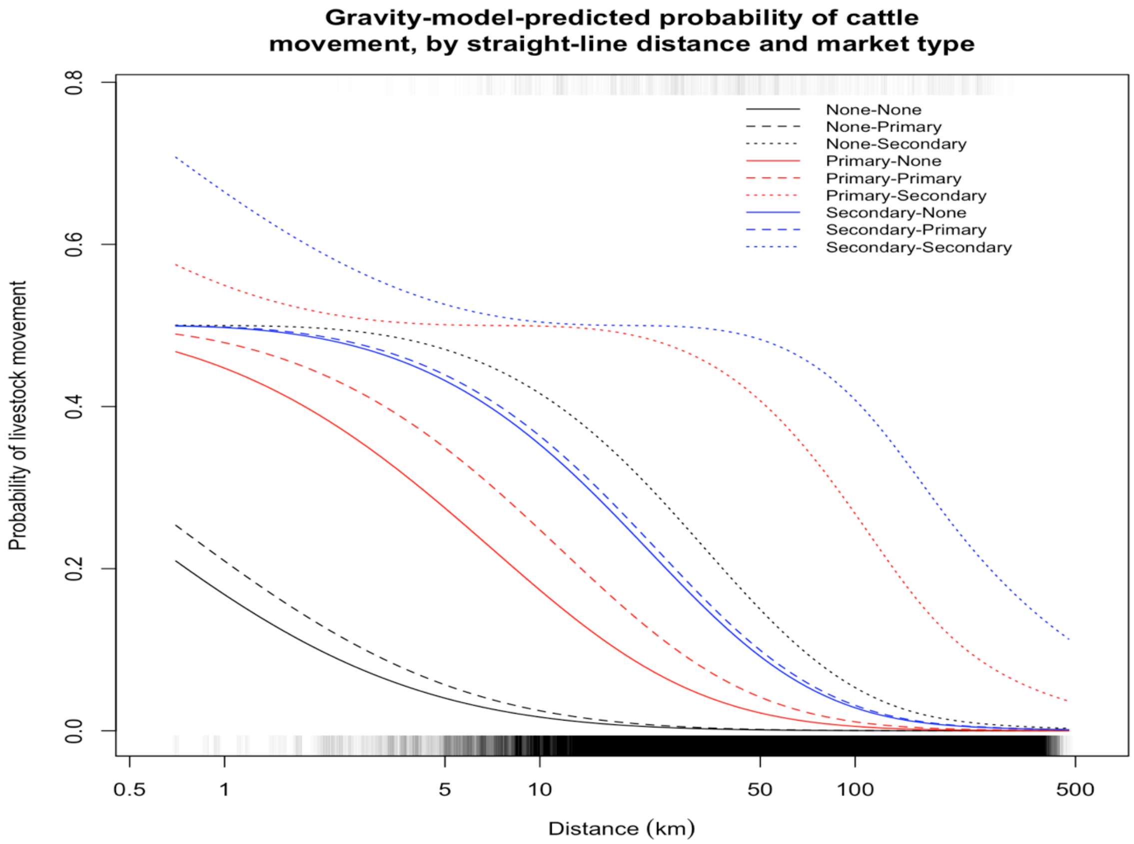


Figure S2. Probability of any cattle movement between wards in a given month, by straight line distance and market type in the origin and destination wards. Predictions are conditional on all continuous variables except month being set at their geometric means. Month is set to January 2015, while production system in both origin and destination wards is agropastoral.

Network measures

Table S5.2. Network measures calculate from the cattle market movement network, the spatial contact network, and the multiplex network

Network measure	Cattle	Spatial	Multiplex
Giant weakly connected component	344	398	398
Giant strongly connected component	143	398	398
Diameter	8	18	12
Transitivity	0.17	0.39	0.21
Reciprocity	0.19	1.00	0.66
Number of edges	1760	2222	3792
Edge weight	89,229	2,222	91,451
Edge density	0.011	0.014	0.024

Table S5.3. Mean percentage population cumulative incidence (PCI) of the fast-transmitting pathogen ($R_0 = 3$) after one year, with no intervention and under six strategies for targeting interventions to wards. The two targeted interventions were a ban on cattle movements through markets, and vaccination of 70% of the cattle in a ward. Mean (SE) absolute reduction in cumulative incidence relative to the no-intervention scenario is also given. The simulated scenarios are: no intervention; application of the intervention to all wards; and targeting of the intervention to 20 wards (5% of the total 398 wards) selected either for network centrality (betweenness, degree and eigenvector), the total number of cattle, or randomly. The total cattle population size in each simulation was 3,707,830.

Targeting method	**Movement ban**		**Vaccination**	
	Mean PCI (%)	**Mean % reduction in PCI (SE)**	**Mean PCI (%)**	**Mean % reduction in PCI (SE)**
No intervention	23.9	-	23.5	-
All wards	3.8	82.7 (1.3)	0.0	99.9 (0.0)
Betweenness centrality	6.3	75.3 (2.0)	12.5	50.7 (4.4)
Degree centrality	5.4	77.4 (1.6)	11.0	57.9 (3.6)
Eigenvector centrality	7.8	70.1 (2.2)	16.4	38.8 (5.3)
Number of cattle	19.7	17.4 (6.6)	14.4	47.3 (4.8)
Random	17.9	31.0 (5.7)	19.5	20.8 (6.5)

Table S5.4. Mean percentage population cumulative incidence (PCI) of the slow-transmitting pathogen ($R_0 = 1.5$) after one year, with no intervention and under six strategies for targeting interventions to wards. The two targeted interventions were a ban on cattle movements through markets, and vaccination of 70% of the cattle in a ward. Mean (SE) absolute reduction in cumulative incidence relative to the no-intervention scenario is also given. The simulated scenarios are: no intervention; application of the intervention to all wards; and targeting of the intervention to 20 wards (5% of the total 398 wards) selected either for network centrality (betweenness, degree and eigenvector), the total number of cattle, or randomly. The total cattle population size in each simulation was 3,707,830.

Targeting method	**Movement ban**		**Vaccination**	
	Mean PCI (%)	**Mean % reduction in PCI (SE)**	**Mean PCI (%)**	**Mean % reduction in PCI (SE)**
No intervention	1.7	-	1.7	-
All wards	0.9	37.0 (2.9)	0.0	99.8 (0.0)
Betweenness centrality	1.1	27.0 (3.4)	1.1	27.1 (3.4)
Degree centrality	1.1	28.6 (3.1)	1.0	31.0 (3.3)
Eigenvector centrality	1.2	20.5 (3.4)	1.3	16.9 (3.4)
Number of cattle	1.5	10.3 (3.6)	1.1	28.1 (3.7)
Random	1.5	12.4 (4.1)	1.5	12.4 (3.9)

9.5 Chapter 6

Table S5.1. Node centrality measure distributions, for nodes with any activity (annual degree > 1) on the small ruminant, cattle and combined networks

	Small ruminants (Nodes = 330) Mean [Range] (SD)	Cattle (Nodes = 344) Mean [Range] (SD)	Combined (Nodes = 373) Mean [Range] (SD)
In degree	8.0 [0 - 91] (12.5)	*9.3* *[0 - 98]* *(14.6)*	*11.3* *[0 - 100]* *(16.6)*
Out degree	8.0 [0 - 143] (23.1)	*9.3* *[0 - 162]* *(25.2)*	*11.3* *[0 - 170]* *(29.5)*
Degree	16.1 [1 - 177] (27.8)	*18.5* *[1 - 247]* *(32.9)*	*22.6* *[1 - 249]* *(37.5)*
Weighted In-degree	*122.8* *[0.0 - 8,904]* *(610.1)*	*232.6* *[0.0 - 14,767]* *(1,068)*	*316.8* *[0.0 - 16,460]* *(1,379)*
Weighted Out-degree	*122.8* *[0.0 – 6,426]* *(1,354)*	*232.6* *[0.0 – 10,647]* *(1,111)*	*316.8* *[0.0 – 16,963]* *(1,420)*
Betweenness	*0.002* *[0.0 - 0.071]* *(0.007)*	*0.002* *[0.0 - 0.070]* *(0.007)*	*0.002* *[0.0 - 0.079]* *(0.007)*
Eigenvalue	*0.017* *[0.0 - 1.0]* *(0.087)*	*0.017* *[0.0 - 1.0]* *(0.070)*	*0.017* *[0.0 - 1.0]* *(0.072)*
Geometric mean degree	*40.9* *[0.0 -1,304]* *(128.7)*	*109.6* *[0.0 - 8,303]* *(605.3)*	*141.1* *[0.0 - 8,738]* *(648.7)*

Table 5.2. Spearmans rank rho value measureing how correlated ward node centrality measures are on the small ruminant, cattle, combined, combined small ruminant high transmissibility and combined cattle high transmissibility networks constructed using livestock movement permit data.

	In-degree Betweenness	In-degree Geometric mean degree	In-degree Eigenvalue	Betweenness Geometric mean degree	Betweenness Eigen value	Geometric mean degree Eigen value
Small ruminant	0.47	0.52	0.79	0.93	0.28	0.32
Cattle	0.46	0.49	0.86	0.90	0.27	0.31
Combined	0.48	0.51	0.77	0.96	0.25	0.27
Comb. small ruminant high trans	0.46	0.51	0.74	0.91	0.24	0.27
Comb. cattle high trans.	0.46	0.50	0.80	0.91	0.26	0.27