Spatial and Molecular Epidemiological Approaches for Evaluating Dynamics of Foot-and-Mouth Disease Virus in India and Vietnam

A dissertation submitted to the College of Veterinary Medicine of the University of Minnesota

by

Umanga Chathurani Wijewickramalge Gunasekera

In partial fulfillment of the requirements for the degree of Doctor of Philosophy

Advisors: Kimberly VanderWaal, Andres Perez

October, 2021

Acknowledgements

I would like to acknowledge:

School of Veterinary Medicine, University of Minnesota, my advisors, Jonathan Arzt from the United States Department of Agriculture (USDA),

Advisors: Kimberly VanderWaal, Andres Perez,

Committee: Scott Wells, Jeff Bender, Jonathan Arzt,

USDA Agricultural Research Service (ARS), Ministry of Agriculture and Rural development (MARD) and Indian Council of Agriculture Research (ICAR) for providing the data and related information,

All the members in VanderWaal lab and Center for Animal Health and Food Safety (CAHFS) past and present.

Funding:

United States Department of Agriculture (USDA) Agricultural Research Service (ARS) Project 1940-32000-061-00D,

The United States Department of State, Biosecurity Engagement Program through the USDA-ARS Office of International Research Program, and a specific collaborative agreement between USDA-ARS and University of Minnesota.

Dissertation Abstract

Foot and mouth disease virus (FMDV) is still endemic in many regions worldwide. Many widespread FMDV lineages have emerged from South Asia, and subsequently spread to other regions, including Southeast Asia, making these two regions important epicenters for FMDV evolution and transboundary transmission. The progressive control pathway (PCP) proposed by the World Animal Health Organization and the Food and Agriculture Organization (OIE/FAO) provides a framework for countries to reduce the incidence of the disease, with the ultimate goal of achieving zonal or country-wide freedom from disease. The PCP pathway is outcome-oriented and relies on a self-evaluation process before being accredited by the OIE and obtaining the official status for a given stage. Different epidemiological methods are acceptable for evaluation of the progress at each stage, yet OIE/FAO does not provide specific guidelines in regard to methods for epidemiological evaluation. In this dissertation, the objective was to demonstrate how recent and newly developed epidemiological approaches can be applied to support progression through PCP stages in South and Southeast Asia. Two countries in those two regions, India and Vietnam, both countries in stage 3, were used as case studies to examine how commonly available types of epidemiological data, such as reported outbreaks and sequence data, can be analyzed to understand the spatial dynamics of FMDV circulation, assess the effectiveness of vaccination programs, and delineate high-and low-risk areas. Ultimately, this work will serve as a proof-of-concept for novel methods for genomic surveillance, which could be used as a cost-effective means to generate sequence data needed for surveillance and epidemiological analysis and help the countries to move towards stage 4 of the PCP.

The first chapter provides an overview of the PCP with a focus on stages 1-3, epidemiological approaches typically used to support PCP activities, and the FMD situation in India and Vietnam. The second chapter specifically explores FMD situation in India and applies Bayesian space-time regressions to investigate factors underlying the spatial heterogeneity in risk of reported outbreaks, including an assessment of how mass vaccination impacted the spatial and temporal distribution of disease. However, such spatial models account for population connectivity by incorporating spatial autocorrelation amongst contiguous spatial units, which is likely a poor representation of population connectivity for highly mobile hosts such as livestock. The third chapter explores the ways to improve Bayesian space-time regression models using the phylogeography to account for patterns of population connectivity. Finally, monitoring circulating virus strains and rapid detection of novel strains is a necessary component of FMD control as part of the PCP, and sequence data also enables a number of other epidemiological approaches to understand virus circulation in a country. Conventional methods to acquire sequence data in the field are not efficient as a means for routing genomic surveillance, and it may be more effective to

identify sentinel surveillance points to detect emerging outbreaks, such as slaughterhouses. Genomic surveillance at slaughterhouses was explored in the fourth chapter using FMDV sequence data from Vietnam. The final chapter is an overview of demonstrated methods to improve FMD control measures and support progression in PCP stages in FMD endemic countries in endemic regions of South and Southeast Asia.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Foot and Mouth Disease (FMD) is caused by an RNA virus (Foot-and-mouth disease virus, FMDV) in the genus *Apthovirus* and family *Picornaviridae* that affects cattle, goats, pigs, and other cloven-hoofed ruminants. The disease produces blisters in the mouth, tongue, hoof, and udder that cause production loss in the animals (Arzt et al., 2011). Many routes of transmissions have been identified for FMDV. The most common route of transmission from one animal to another is via direct contact and oral route (Alexandersen et al., 2003,Bravo de Rueda et al., 2015). Direct contact occurs via saliva and excretion material from the blisters in the mouth and hooves. There are seven serotypes of the virus, referred to as type O, A, C, Asia 1, SAT1-3, and infection with one serotype does not provide cross-protection for a different serotype. The disease is endemic in much of Asia, Africa, and parts of South America, though not all serotypes occur in all continents.

The FMDV viral genome contains a single open reading frame (ORF) encoding the entire genome, including structural and non-structural proteins. The structural part of the virus is known as the VP1-VP4 protein capsid. Although it represents only 8% of the whole genome, FMDV serotypes and phylogenetic analyses have traditionally been characterized by sequencing the VP1 region (Bachanek-Bankowska et al., 2018,Brito et al., 2017, de Carvalho Ferreira et al., 2017,Brito et al., 2018).

Like most RNA viruses, FMDV is highly susceptible to mutations during replication. As a result, the virus shows high genetic variation and rapid mutation rates that result in different viral phenotypes (Elena and Sanjuán, 2005). From a phylogenetic standpoint, FMDV serotypes are further divided into topotypes, lineages, sub-lineages, and strains. Viral topotypes are defined as groups of viruses that show >85% nucleotide identity (VP1) and typically show a restricted geographic distribution (Knowles et al., 2005). Within a topotype, lineages are differentiated based on a 5% difference in the genetic distance among viral clades. Sub-lineages and strains are nested within the lineages. Because cross-protection amongst strains in the same serotype may only be partial, immune-driven interactions among co-circulating strains at the population level may lead to the replacement of existing strains with new strains (Gupta et al., 1998). New strains can also emerge because the novel strain is more virulent than the existing strains.

The ever-expanding genetic and antigenic diversity of FMDV in Asia presents a challenge to disease control and vaccination programs. Recent work has shown that, in the past ten years (2007-2017), there has been at least 23 sub-lineages originating in South Asia that have spread to other regions such as North Africa, Middle East, East and Southeast Asia considering, including the O/MESA/Ind 2001d lineage that appears to have out-competed existing serotype O

strains that were circulating (Bachanek-Bankowska et al., 2018). In some countries, the Ind-2001d lineage gave rise to new and different strains, such as those identified in the Middle East (O/BAR/2/2015) and Nepal (O/NEP/1/2014) (Bachanek-Bankowska et al., 2018). Thus, many widespread FMD lineages have emerged from Asia, and subsequently spread widely, pose a risk to Europe and other FMD-free regions. This makes the epidemiological dynamics, surveillance, and control of this virus in Asia critical for global control efforts.

The Progressive Control Pathway (PCP), introduced by the World Animal Health Organization (OIE) and Food and Agriculture Organization (FAO), aims to provide a step-by-step framework for the reduction of the risk of FMD in endemic countries. To progress between different stages in the PCP, countries are required to provide epidemiological assessments of the current situation within the country, such as evaluation of the surveillance measures and vaccination program, as well as outline a series of evidence-based plans to control the disease. Thus, analysis of epidemiological data is critical for supporting progress through the PCP, and ultimately controlling FMD more effectively. However, the PCP provides limited guidance on epidemiological analyses, particularly on types of spatiotemporal analyses necessary to identify disease hotspots and understand patterns of circulation, and the available methods are constantly evolving.

In this dissertation, our objective is to demonstrate how recent and newly developed epidemiological approaches can be applied to support progression through PCP stages in endemic countries. Focusing on two countries in Asia, India and Vietnam, both countries in stage 3, we examine how commonly available types of epidemiological data, such as reported outbreaks and sequence data, can be analyzed to understand the spatial dynamics of FMDV circulation, assess the effectiveness of vaccination programs, and delineate high-and low-risk areas. We also demonstrate a proof-of-concept for novel methods for genomic surveillance, which could be used as a cost-effective means to generate sequence data needed for surveillance and epidemiological analysis and help the countries to move towards stage 4 of the PCP.

1.1 FMD in India and Vietnam

FMDV is endemic in India. India is the largest dairy producer and consumer globally, and much of the country is characterized by high densities of bovine (cattle and water buffalo) and small ruminants. There are two regular vaccination programs in India: ASCAD (Assistance to States to Control Animal Disease) and the central government vaccination program (FMDCP). FMDCP is a biannual vaccination program administered to cattle and buffaloes while ASCAD is an annual vaccination program (Hegde et al., 2014). Vaccine production, monitoring, and surveillance are conducted by the government of India. The currently used vaccine for FMDV is a trivalent

(serotypes O, A, and Asia1) inactivated vaccine that produces immunity for up to 6 -12 months duration depending on the adjuvant. Typically, FMDV vaccines reduce clinical signs, protect against clinical infection, and reduce the transmission of the live virus to susceptible animals (Brito et al., 2011). Monitoring of vaccination programs is done by collecting blood samples before and after the vaccination and checking for antibody titers using liquid-phase blocking ELISA (Pattnaik et al., 2012). Such monitoring is a critical component of PCP stage 2 and 3.

Most countries in Southeast Asia (SEA), including Vietnam, are FMDV endemic. Vietnam has established official FMD control program since 2006 which is implemented in a phase wise manner (Lee et al., 2020). Serotypes O and A circulate in the country. Serotype O lineages detected in Vietnam include O/ME-SA/Mya-98, SEA/PanAsia, O/ME-SA/Ind2001d, and O/Cathy. Mya-98 was identified between 1998 and 2014, while PanAsia was first introduced in 2006 and is currently dominant (Le et al., 2016). O/ME-SA/Ind2001d was first detected in the Southern part of the country in 2015 and is still circulating alongside the PanAsia lineage (Vu et al., 2017). The identified Serotype A sequences belong to genotype IX and are closely related to strains identified from Laos and Thailand (Vp et al., 2010). It can be observed that, consecutively, different lineages emerge and decline for both serotypes. To decide on appropriate vaccination, it is important to identify fluctuations in the prevalence/incidence of various FMDVs as early as possible.

Both countries aim at moving towards zonal freedom with vaccination, a requirement of stage 4 where official OIE endorsement is given after evaluation of the progress of each country. The concept of zones within a country usually applies to countries that have set the objectives for zonal freedom and formulated an official control program to eliminate FMD from a distinct geographic area (considering the structure of the livestock industry), including animal movement patterns at national and regional levels, and fulfil the recommendations of the OIE Terrestrial Animal Health Code.

1.2 Progressive Control Pathway

The FAO/OIE's PCP framework for the control of FMD is outcome-oriented and acknowledges that approaches to achieve key outcomes are different from a country to country. It outlines a number of key points that should be fulfilled at each stage. Stages 1-3 provide a global framework that can be applied to control any transboundary animal disease, including FMD, while stages 4 and 5 are oriented towards obtaining OIE-endorsed freedom from FMD. For a country to be in stage 1, they should have completed a risk assessment of FMD. Common activities in stage 1 include having a hypothesis of virus circulation, **identification of circulating strains**, enabling an environment to control the disease including supporting regional corporation, evaluating the

socio-economic impact of the disease, and completing a value chain analysis. A working **hypothesis of virus circulation** includes various building blocks, such as outbreak reporting and outbreak investigations for all the susceptible species, sero-monitoring for FMDV and **identifying FMD risk hot spots**. An enabling environment to control disease includes evaluating the veterinary services of the country as part of a Performance of Veterinary Services (PVS) analysis. Countries are encouraged to participate in regional roadmap meetings with the intention of regional corporation to eliminate FMDV from a given region of the world. Ultimately, the overall aim of PCP stage 1 is to identify the current FMD situation in the country.
Once a country identifies it's FMD situation by completing various steps of stage 1, it can proceed to the stage 2 of the PCP.

By stage 2, a country should have a strategic control plan with an FMD monitoring and evaluation system in place. Ongoing monitoring of FMD risk in different husbandry systems such as cattle, buffalo, goats, and pigs, monitoring of circulating strains, and implementing risk-based control measures in a targeted zone are part of this stage. **There should be evidence that risk of FMD has declined in at least one of the husbandry systems.** Continuing from stage 1, an **evaluation of FMD vaccination campaign by targeted serological surveys should continue in this stage.** There should be a written documentation regarding a risk-based control plan that includes all above activities.

By stage 3, countries should have an official control program for FMD control. Once a country establishes an official FMD control program, they request for OIE endorsement to reach stage 4. Ongoing monitoring activities from stage 1 and 2 should continue in the stage 3. **Rapid detection of outbreaks at least in a one zone of the country** is a requirement at this stage. There should also be **evidence of progressive reduction of FMD incidence** in domestic animals in at least one zone in the country, and that the country is moving towards progressive reduction of FMD in the whole country.

In all stages highlighted above, epidemiological analysis is a fundamental component to fulfilling the requirements of each stage and providing evidence for advancing through the PCP.
Many of the activities initiated in stage 1 should continue in later stages, together with optimizing surveillance and risk-based control measures with the available resources. This thesis will focus on how recent and newly developed epidemiological approaches can be applied to support key activities across different stages of the PCP, including identifying FMD risk hotspots, evaluation of the impact of vaccination programs, developing a working hypothesis of virus circulation, and ongoing monitoring of circulating FMDV strains.

1. Identifying FMD risk hotspots

When FMD is endemic in the country, it is not always feasible or cost effective to implement control measures throughout the country. What is economical is to identify high risk areas for FMD which also can be a potential source of FMD and focus on those areas. Several studies have identified spatial-temporal clusters that identify FMD high-risk areas in different countries in Asia, including Thailand (Arjkumpa et al., 2020), China (Chen et al., 2020), Vietnam (Lee et al., 2020), and Bangladesh (Rahman et al., 2020). There are other studies that identify FMD risk hotspots accounting for both outbreak numbers and the risk factors using logistic regression models, such as in Bhutan (Dukpa et al., 2011) Japan (Muroga et al., 2013), Afghanistan (Wajid et al., 2020), and Sri Lanka (Gunasekera et al., 2017).

While spatio-temporal cluster analysis can be utilized to identify areas where more or less outbreaks have been recorded relative to null expectations, regression models are typically used to incorporate risk factors and sometimes to show whether FMD risk has increased or reduced over time. Commonly identified risk factors for FMD in cattle and buffaloes include host density, animal movement-related, and husbandry-related risk factors. For example, transmission occurs when animals share resources, such as grazing in the same field or being fed with grass cut where an infected animal has been grazing. In addition, FMD outbreaks have been shown to be seasonal in some endemic regions (Lee et al., 2020, Arjkumpa et al., 2020). Seasonal, climatic, and environmental factors can correlate with husbandry systems, patterns and timing of animal movements, and survivability of the virus outside the host.

In India, no studies to our knowledge have been carried out to identify risk hotspots and their underlying drivers, which could be important for pursuing zonal freedom. One study conducted in Karnataka identified there is a seasonal pattern for FMD occurrence, livestock density, infected serotype and agroclimatic zone determines the number of outbreaks (Hedge et al.,2014). Several studies have outlined theoretically what should be done to control FMD in India (Pattnaik et al., 2012, Biswal et al., 2015), but a quantitative understanding of spatial dynamics is not available. In Vietnam, high risk FMD clusters have been identified using SatScan analysis, but to our knowledge there are few studies to identify risk factors in a spatial framework. In a farm-based study published in 2017 focusing on north Vietnam, it was identified that FMD risk will change with farm size, age of the animals, and type of animals (beef cattle) (Ferreira et al.,2017).

Drawbacks of many risk factor studies as well as analyses of the impact of vaccination on outbreak numbers are the fact that they are mainly based on the reported outbreak numbers and, importantly here, they do not account for spatial autocorrelation in outbreak occurrence and among risk factors. For example, if a high-risk state/province experiences high numbers of outbreaks, the adjacent states may also document higher numbers of outbreaks based on their proximity to the high-risk area, making it challenging to tease apart the risk factors and identify high-risk areas. Bayesian models provide a better platform to account for autocorrelated data. When epidemiological risk models for FMD are reviewed across the world, however, only 4.1% (2/48) utilized Bayesian models (Souley Kouato et al., 2018).

2. Evaluation of the impact of vaccination programs.

As a part of monitoring the impact of the control measures, it is important to show a reduction in the number of outbreaks over time and serologically confirm the number of FMD positive animals are reducing with time. In addition, pre- and post-vaccine monitoring includes assessing (inferred) protective antibody titers in animals, as measured by LPB ELISA (liquid phase blocking enzyme-linked immunosorbent assay) to check for structural protein-related antibodies. To differentiate vaccinated animals from naturally infected animals, a 3ABC ELISA test is commonly used to detect antibody response against nonstructural proteins (de Carvalho Ferreira et al., 2017, Brito et al., 2017). A highly purified vaccine should not contain the nonstructural proteins. Since the serology monitoring is conducted at the field level, there could be several drawbacks to evaluate the impact of the vaccination program. According to OIE guidelines, it is important to have an animal identification system in place to measure pre- and post-vaccination antibody titers, with sampling conducted in in an age stratified manner as older animals have higher antibody levels. Vaccine monitoring should continue throughout different stages of PCP.

Studies have evaluated the impact of the vaccination program in India as field trials (Sharma et al., 2014) (Mahapatra et al., 2015). Currently used vaccine strain in India against Serotype O, O/IND/R2/75 was evaluated against the circulating serotype O strains in the country using the virus neutralization test and it was identified as having a 79% match (Mahapatra et al., 2015). The efficacy of vaccination was identified comparing districts with the vaccination program to districts without the vaccination program from a longitudinal serological study (Sharma et al 2014). However, a major drawback of these studies, as well as similar studies analyzing the relationship of vaccination to outbreak data through time, is that they often fail to account for spatial and temporal

autocorrelation in the data, which could make it more difficult to quantify the impact of vaccination. As described in (1), Bayesian spatio-temporal models provide a robust framework to disentangle the effect of vaccination programs on outbreak numbers alongside other spatial and temporal factors.

3. A working hypothesis of the virus circulation

Because the PCP pathway is outcome-oriented, different countries will determine patterns of virus circulation in different manners. Understanding FMDV circulation is important to identify where to implement control measures, which is particularly important for building the foundations for transitioning from stage 3 into stage 4 of the PCP. Serological monitoring of the virus across the country, risk factor analysis, and molecular epidemiological analysis for FMDV sequence data isolated from different parts of the country can also be used to identify patterns and test hypotheses regarding FMDV circulation in the country.

Except for few island nations, most countries in Asia are connected by geographical land borders. In addition, three different FMD pools exist in Asia known as pools 1, 2, and pool 3 based on circulating serotypes and topotypes (Paton et al., 2018). Transboundary and trans-pool spread of FMDV has been extensively documented (Bachanek-Bankowska et al., 2018, Brito et al., 2017a). Some of the countries have robust border controls in place for political reasons (India and Pakistan and China) that may act to limit animal movement and segregate these different FMD pools. Otherwise, illegal animal movement is common in Asia (Landes et al., 2020.) and plays a significant role in the spread of infectious diseases such as FMD. It is important to incorporate the potential for transboundary dissemination of FMDV in FMD risk models. Furthermore, tracking both transboundary and domestic movements of animals is a critical component for disease control and preparedness in many countries (though few in Asia). Undocumented animal movement and unavailability of a proper animal identification system in many Asian countries have led to difficulties in quantifying population connectivity and identifying patterns of FMDV circulation. In such instances, proxies may be used for the purpose of modeling, such as road networks and identifying areas where animals are free ranging. Road density may capture aspects of human movement and/or animal trade and animal product movements such as meat and milk. Some FMD risk models use human density as a proxy for areas where there is no livestock and detect human movement as a risk (Chhetri et al., 2010). Even in the absence of a national framework to understand animal movement, farm-based studies often collect data about recent introductions of new animals to the herd, or other animal movement related information. A majority of FMD

risk related studies identified animal movement and animal trade as the most reported risk factor (Souley Kouato et al., 2018).

It is important to identify better methods to build and test hypotheses regarding the virus circulation in the country. Recently, phylogeographic analyses have increasingly been utilized to reconstruct patterns of FMDV dissemination and can be used to uncover the history of both within country spatial and transboundary movement of the virus (Munsey et al., 2021,Muwonge et al., 2021,Bertram et al., 2018). Most likely, viral movement inferred by such analyses reflect underlying patterns of host movement that ultimately drive patterns of population connectivity (Makau et al., 2021, Di Nardo et al., 2011).A new horizon of phylogeographic models is to incorporate phylogeographic outputs on viral movement with other meta-data to explicitly test how different factors influence rates of viral movement (Di Nardo et al., 2021,Munsey 2021). In addition, using outputs of phylogeographic models as a proxy for population connectivity may be useful in Bayesian space-time models by better accounting for the impact of longer distance animal trade networks.

4. Ongoing monitoring of circulating FMDV strains

When a country advances through different stages of PCP, it is important to detect emerging FMD strains to determine appropriate vaccine strains. In addition, generation of geo-referenced sequence data can be instrumental to supporting other aspects of PCP, for example, to quantify patterns of circulation and test hypotheses related to how the virus circulates (as described above). Monitoring of circulating strains can be done via active surveillance, or as part of passive surveillance. More generally, surveillance can be active, passive or sentinel. Active surveillance provides us the magnitude of the problem and a glimpse of population-level virus circulation. This is important related to FMDV, as emerging virulent viral strains are hard to capture by passive surveillance. Passive surveillance is the status-quo in India and Vietnam, as well as in many endemic countries. Typically, once a clinical outbreak occurs in each area, this will be reported to the district level veterinary office and from there, to the central veterinary office. In India, sero-surveillance and serotyping FMDV is carried out in regional laboratories. This includes serotyping of clinical materials from FMD outbreaks. Vaccine matching is performed to circulating field strains and phylogenetic analysis are performed to identify the variation in the virus (Subramaniam et al., 2015). A different study monitored FMDV circulation in goats using serological data (Ranabijuli et al., 2010). Annual reports from the Indian Council of Agriculture Research (ICAR) records sporadic active surveillance of FMDV in different states of India. According to PVS analysis of Vietnam, there is ongoing

active surveillance in Vietnam in selected provinces. As of now, neither of the surveillance systems in either country are risk based. Risk based surveillance is to focus surveillance on identified FMD high risk areas. In PCP stage 3, active surveillance should be conducted for rapid detection of outbreaks in at least in a one zone of the country. Active surveillance of farms can be costly and time consuming, whereas passive surveillance is likely to miss subclinical circulation. However, conducting surveillance at points where animals from many farms congregate (such as markets or slaughterhouses) may provide robust and cost-efficient sentinel surveillance that could both a) identify circulating strains, and b) provide sequence data for molecular epidemiological studies that would support other aspects of the PCP.

1.3 Framework of the research

In summary, this dissertation will demonstrate approaches to maximize the utility of epidemiological data for supporting different stages 1-3 of the PCP pathway and for tailoring control measures in FMD-endemic countries in Asia (Table 1.1).

Table 1.1: Steps of PCP this study target to contribute are highlighted in gray color.

Progressive Control Pathway		**Chapter 2 India**	**Chapter 3 Vietnam**	**Chapter 4 Vietnam**
Stage 1	Implement risk-based approach to reduce FMD			
Activities to understand FMD risk	Hypothesis of how the FMDV circulate in the country including currently circulating strains			
	Value chain analysis			
	Socio economic impact of FMD			
	Evaluation based on OIE-PVS pathway			
	Regional collaboration			
	FMD risk hot spots are identified			
	Identify strategies to control FMD	Risk factors identification		
Stage 2				
	Ongoing monitoring of the circulating strains			
	Implement risk-based zone targeted control measures	High risk area identification		

Risk based strategic plan	Show impact of FMD reduced with control measures	Evaluation of the vaccine program		
	Allocation of sufficient resources			
Stage 3 **Reduction of outbreak incidence and virus circulation in at least on zone of the country**				
	Evaluate the incursion of new serotypes			Enhanced FMD surveillance
	Sustainable veterinary services			
	Legal framework for animal identification			
	Analysis of virological and outbreak data and analysis of serological survey			
	FMD contingency plan			
	Strengthening the veterinary service of the country for own epidemiological investigations			
	Rapid detection of outbreaks at least in a one zone of the country			Improved surveillance
	Endorsement of the official control program by OIE			

Here, we use data available from India and Vietnam as examples of to apply quantitative tools to support PCP. Both India and Vietnam appear to play pivotal roles in FMDV circulation in South and Southeast Asia, respectively; as stage 3 countries that may in the future seek to transition to stage 4, their example can serve as a roadmap for other countries progressing through stages 1-3 of the PCP. The first aim applies Bayesian space-time regressions to explore the FMD situation in India to factors underlying the spatial heterogeneity in risk of reported outbreaks, including an assessment of how mass vaccination impacted the spatial and temporal distribution of disease. However, such spatial models account for population connectivity by incorporating spatial autocorrelation amongst contiguous spatial units, which is likely a poor representation of population connectivity for highly mobile hosts such as livestock. In most endemic countries, databases of livestock transport data are not available. Yet, patterns of viral movement can be inferred from phylogeographic models. Therefore, the second aim focus on Vietnam, where we have a rich dataset of FMDV sequences. Here, we apply the same space-time regression

approach for modeling reported outbreaks but compare and contrast whether inferences on between-province connectivity from phylogeographic models provide a better fit to the outbreak space-time regression than using spatial adjacency as the sole metric of population connectivity. Finally, understanding FMD viral diversity and the risk factors one can identify how important it is to early detection the virus by strengthening the surveillance measures in a country. Monitoring circulating virus strains and rapid detection of outbreaks are a necessity in the stage 3 of PCP. Conventional methods of surveillance will not be cost effective, and it is important to identify sentinel surveillance points to detect emerging outbreaks such as slaughterhouses. This has been explored in the 3rd chapter using data from Vietnam. The final chapter is an overview of proposed methods to improve FMD control measures in FMD endemic countries in Asia following the PCP pathway. Figure 1.1 provides an overview of the framework of this thesis.

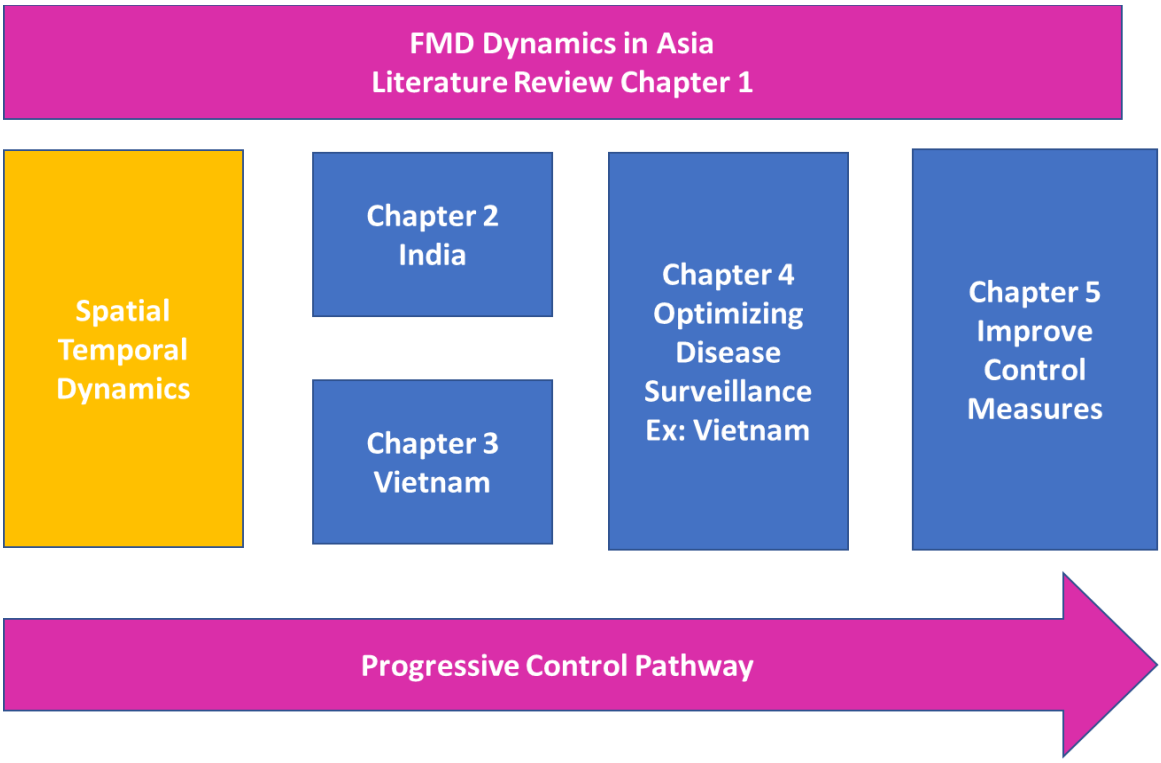


Figure 1.1: Framework of the thesis from the chapter 1 to 5.

Chapter 2: Spatiotemporal dynamics of foot-and and-mouth disease outbreaks in India, 2008-2016

2.1 Introduction

Foot-and-Mouth Disease (FMD) is caused by an *aphthovirus* in the *Picornaviridae* family that affects cattle, buffalo, pigs, and other domestic and wild ungulates. Classical infection produces clinical signs of fever and vesicles in the mouth, tongue, hoof, and udder, affecting production and leading to economic losses (Arzt et al., 2011). However, FMDV is also known to cause various forms of subclinical infection (Stenfeldt and Arzt, 2020) and has been associated with abortion in cattle in India (Ranjan et al., 2016). FMD is endemic across many countries or regions of Asia, South America, and Africa, where estimated economic losses range from $6.5 to 21 billion USD annually (Knight-Jones and Rushton, 2013). Preventing the transboundary spread of FMD into disease-free countries, including many countries in Europe and North America, plays a major role in shaping international trade policies (Shanafelt and Perrings, 2017). In the past two decades, several widespread viral lineages of serotypes O, A and Asia1 have emerged from the Indian sub-continent, suggesting that it is a hotspot for viral evolution and subsequent transboundary spread (Brito et al., 2017b). For example, the O/PanAsia II lineage of FMD emerged in 2003 in the Indian subcontinent, and O/ME-SA/Ind2001d emerged in 2001, re-emerged in 2008 (Knowles et al., 2016,Dahiya et al., 2020, Subramaniam et al., 2015), and has spread to adjacent regions such as North Africa, Middle East, and Southeast Asia through approximately 13 transboundary escape events (Bachanek-Bankowska et al., 2018,Vu et al., 2017). Incidentally some of these lineages are first reported from India, possibly due to having a relatively better surveillance system in-place among all the countries in the South Asia region. Therefore, understanding the epidemiology of FMD in the India is critical and of utmost importance for supporting the regional FMD control initiatives and controlling the disease globally.

India has a population of 302.34 million cattle and buffalo (20th livestock census, 2019).The size of the bovine population in some of the administrative units or states of India is similar to that of individual countries in Africa, Asia and Europe. The country is the largest global producer and consumer of dairy, with dairy products contributing ~70% of total livestock income of India, and the third largest beef exporter (Kumar, 2012, Hemme et al., 2003).

The total farm-level economic loss projected due to FMD in cattle and buffaloes in India was USD 3159 million, USD 270 million and USD 152 million, respectively during the severe, moderate and mild incidence scenarios (Govindaraj et al., 2021).

Presently, a trivalent vaccine, which confers protection against serotypes O, A, and Asia1 has been used in the country. The National FMD Control Program (FMDCP) was started in 54

selected districts in 2003-04, and subsequently expanded in a phase-wise manner. Indian states are enrolled in either of two regular vaccination programs, namely, the central government vaccination program (FMDCP) or the Assistance to States to Control Animal Disease (ASCAD) program, with the programs beginning in different years in different states. FMDCP is a biannual vaccination program in which the trivalent vaccine is administered to cattle and buffaloes, while ASCAD is an annual vaccination program (Hegde et al., 2014). Currently the entire country is covered under National Animal Disease Control Programme in which all the susceptible species are vaccinated against FMD twice a year. For FMDCP, vaccine production, monitoring, and surveillance are conducted by the Indian government. Pre- and post-vaccination monitoring includes the determination of antibody titers by ELISA before and after each round of vaccination (Pattnaik et al., 2012). At the population-level, it is desirable for 80% of animals to have adequate protection (inferred through antibody titers) to minimize the risk of widespread outbreaks (Metwally and Münstermann, 2016).

The World Animal Health Organization (OIE) and the World Food and Agriculture Organization (FAO) have developed a set of outcome-oriented guidelines for FMD endemic countries to reach FMD-free status, which is known as the Progressive Control Pathway (PCP) (OIE/FAO). Most countries in Asia, including India, are in stage 1 to 3 of the PCP. In stage 1, risk from FMD and available control options will be identified. By stage 2, a country is expected to have a risk-based strategic control plan with an FMD monitoring and evaluation system in place. In stage 3, a country should continue to monitor disease risk, analyze passive and active surveillance data to show progressive reductions in FMDV occurrence, and implement its strategic control plan, which may include pursuing FMD-free zones with vaccination within the country. Vaccination plays a major role in achieving this task. To date, very few studies have been carried out to identify risk factors for FMD and the spatial distribution of risk in the country (Hegde et al., 2014, Sharma et al., 2014). In addition, an evaluation of the success of vaccination programs in reducing outbreaks is key to understanding the role of such programs in controlling FMD and achieving FMD zonal freedom with vaccination. Several studies have recognized the importance of optimizing the vaccination program, controlling animal movements, and conducting effective surveillance for FMD control in India (Biswal et al., 2019, Pattnaik et al., 2012). Rigorous spatial epidemiological methods are yet to be applied to understand how vaccination and other factors relate to the spatiotemporal pattern of outbreaks in a changing epidemiological scenario on account of increased population immunity.

The objective of the current study was to model the spatiotemporal dynamics of FMD outbreaks and assess the contribution of mass vaccination campaigns in reducing FMD outbreaks in India. We first assessed vaccination outputs through an evaluation of antibody titer data collected as

part of pre- and post-vaccination monitoring. Using a Bayesian space-time model that accounts for underlying spatial dependencies often present in disease data (Machado et al., 2019, Branscum et al., 2008, Chhetri et al., 2010), we then investigated the impact of mass vaccination programs on the occurrence of reported FMD outbreaks over time alongside other factors that have the potential to influence spread, such as variables related to animal movement, intermingling of animals at grazing areas, proximity to international borders, and environmental factors. Results presented here will ultimately contribute to evaluation of progress of India's mass vaccination campaigns and support country progress in the context of the progressive control pathway for the control of the disease.

2.2 Materials and Methods

2.2.1 Study area and data sources

There are 29 states and seven union territories in India, and each state is further subdivided into administrative districts. The first phase (Phase I) of the FMDCP began in 2003 as a pilot study (Mahapatra et al., 2015). At the beginning of the FMDCP Phase I, nine states and one union territory were part of the mass vaccination program, and not all districts within each state were included. In the second phase of FMDCP (starting between 2010 and 2011, depending on the state), all districts within the participating states were part of the vaccination program except for Uttar Pradesh (an administrative unit in northern India with 50.2 million bovine population). In Uttar Pradesh, only 16 of 75 districts participated in the vaccination program as of 2018. By 2018, all seven union territories and 11 of 29 states were covered by FMDCP without exceptions. During phase 2 of the FMDCP vaccination program, it has been estimated that 38% of the total cattle and buffalo population of India were vaccinated (Mahapatra et al., 2015). The 3PD50 potency trivalent (serotypes O, A, and Asia1) vaccine used in India contains three times the protective dose required to protect 50% of the animal population (Pattnaik et al., 2012).

Data on pre- and post-vaccination FMDV antibody titers (see below) and annual reported number of outbreaks from each state were obtained from the annual summaries of the Directorate of Foot and Mouth Disease of the Indian Council of Agricultural Research (ICAR-DFMD, Ministry of Agriculture), which is the national referral center for FMD diagnosis. Outbreak data were generated through passive surveillance, where an outbreak was defined as a report of clinically FMD-infected animals from the same village/district (OIE) which was further confirmed by laboratory tests conducted on referred clinical samples. The number of infected animals was not available for a given outbreak. Outbreak data were reported at the state and not the district level.

2.2.2 Serological data

As outlined above, mass vaccination of cattle and buffalo was carried out by the Indian government once every six months in the selected states and districts that were part of FMDCP. To determine antibody titers pre-vaccination, sera samples were collected at the time of vaccination for each biannual round of vaccination. Sera samples were also collected at 21 to 30 days post-vaccination. Sampled animals were selected at random and the pre- and post-vaccination samples may or may not come from the same animal. On average, the number of animals per state from which samples were collected ranged from 100 to 1000 animals per sampling round.

Sera samples were tested for reactivity against FMDV using Liquid Phase Blocking (LPB)-ELISA, which was used to infer protective antibody titers against FMDV structural proteins at an inferred protection level of $\log_{10}$ titer of 1.8. Change of log titer values were similar for all three serotypes O, A and Asia 1 as vaccination was conducted with a trivalent vaccine. Since serotype O accounts for more than 80% of the outbreaks (Britto et al., 2017b), only antibody titer change for serotype O are shown in the main text. For this study, antibody titer data was only available from states that were part of FMDCP phase I. Because LPB-ELISA cannot discriminate between antibody responses induced by vaccination verses natural infection, we could not determine whether inferred protection was the sole result of vaccination and not from previous natural exposure to FMDV. Data regarding the percent of animals with inferred protection pre- and post-vaccination were summarized for each six-month round of vaccination.

In contrast to LPB-ELISA, non-structural protein (NSP)-based ELISA differentiates vaccinated animals from naturally infected animals based on elicitation of a response to nonstructural proteins that should be absent in vaccine preparations. There is evidence that vaccination can elicit a transient NSP response in vaccinated animals in India (Mohapatra et al., 2011, Hayer et al., 2018). However, the majority of samples by NSP-based ELISA are expected to be from the previously infected individuals. Annual information on antibody titers of NSP ELISA were available only for certain states/administrative units for some years. Further, the animals sampled for NSP ELISA once in a year are not necessarily the same ones as sampled for LPB-ELISA.

2.2.3 Descriptive analysis

Data from pre- and post-vaccination monitoring obtained through LPB-ELISA from the phase I were analyzed to identify whether there was an increase in percent of animals with inferred protection before and after each individual round of vaccination, as well as to identify trends in inferred protection over time across multiple successive rounds of vaccination.

To evaluate whether there was a correlation between population immunity and FMDV circulation, the association between the percent of animals in each state over a period of eight years (2008-2016) with inferred protection via LPB-ELISA (an indicator of population immunity) and percent of NSP ELISA-positive animals (an indicator of previous natural exposure) was evaluated using a Spearman's correlation test.

Bovine (cattle and buffalo) population data were obtained from the Department of Animal Husbandry and the Dairying census, Government of India, which was available for the years 2008 and 2012. We averaged the two values to represent bovine population size per state. The reported number of outbreaks per state per year along with bovine population size per state was used to calculate a standardized incidence ratio (SIR) at the state level for the years 2008 to 2016. Population size and nationwide outbreak counts were used to calculate the expected number of outbreaks per state per year (e_{it}) if the distribution of outbreaks across space and time was proportional to population size, such that

$$e_{it} = P_{it} \frac{\sum_{it} Y_{it}}{\sum_{it} P_{it}}$$

Where P_{it} is the population of state *i* in year *t*, and Y_{it} is the number of FMD cases in state *i* in year *t.* SIR was defined as the observed to expected ratio (Y_{it}/e_{it}). SIRs were plotted as choropleth maps for all years.

2.2.4 Conceptual framework of outbreak risk

A conceptual diagram was created to represent pathways by which hypothesized risk factors could influence the reported number of outbreaks per state (Figure 2.1). Details on risk factor data are shown in Supplementary Table S2.1. Disease spread is expected to be influenced by contact, which is in turn influenced by host density, environmental factors, animal movement, organized farming practices, and other community activities like animal fairs. In addition to cattle and buffalo, goats and pigs are also affected by FMD and can transmit the virus to cattle and buffalo. If transmission is density dependent, then higher host densities are expected to translate to rapid spread and more outbreaks via increased contact frequency (Hegde et al., 2014). Under Indian socio-ecological conditions, bovine and caprine species are often reared by the same households. Livestock population data of goat density and pig density were therefore included in our model as categorical variables (high/low, split at the mean). The density of livestock was calculated per square km at the state level.

Disease spread is also influenced by transmissibility, which is influenced in part by environmental factors that may affect the survivability of the virus outside the host. From previous studies, it has been identified that droplet nuclei of the virus can occur to a distance of 20 km and can persist in the environment for about a week at a temperature of 20^0C (Alexandersen et al., 2003). FMDV

can survive in temperatures up to 27^0C, but not extremely high temperatures (Donaldson, 1972, Mikkelsen et al., 2003). Temperature may be an important factor not only because of the optimal temperature for the survival of the virus outside the host, but also the efficacy of vaccines in the field depends on storage temperature (OIE Chapter 2.1.5). Wind speeds between 5 and 10 knots have been identified as favorable for FMDV transmission (2.57–5.14 m/s) (Gibbens et al., 2001). Therefore, annual averages for wind speed, rainfall, and temperature were included to capture environmental factors related to outbreak risk (Abatzoglou et al., 2018). Evaluation of intra-annual and seasonal variation was not possible since outbreak reporting was an annual value. However, to quantify areas with more extreme seasonality, annual variance (calculated across 12 months) of each environmental variable was also included. All environmental variables were centered at the mean and standardized. In addition, higher outbreak numbers have been reported in dryer agroclimatic zones in some parts of India (Hegde 2014), potentially due to environmental conditions within those regions or husbandry practices typical of different climatic conditions. For example, communal grazing increases local mixing among livestock herds, and could enhance disease spread. Climatic factors in combination with land cover (waterbody and forest density) may capture variation in husbandry practices common to different agroclimatic zones (Hegde et al., 2014). Thus, these with variables were included in the model (split into high/low categories at the mean, Supplementary Table S2.1).

Animal transport within India (represented by road density as a proxy measure) or across international borders for trading and slaughter may promote disease spread and the occurrence of outbreaks. India is bordered by Pakistan, China, Nepal, Bhutan, Myanmar, and Bangladesh. FAO/OIE has categorized India and the surrounding endemic countries into different “pools” of FMD based on the predominant circulating serotypes and topotypes in each area (Paton et al., 2018). The country and its neighbors are categorized into three different pools. Pool 1 includes Nepal, Bhutan, Myanmar, and China; pool 2 includes India, Sri Lanka, Bangladesh; and pool 3 includes Pakistan. Dummy variables were introduced to the analysis indicating whether each state was bordered by a pool1, pool2 or pool3 country, or if the state did not have any international land borders (Supplementary table S2.1).

For each year, states were categorized into two groups based on whether they were part of the FMDCP. In addition, the presence and efficiency of veterinary services within a state may influence both vaccination as well as outbreak reporting. The coverage of veterinary services in each state was calculated based on the percentage of veterinarians available relative to the number of veterinarians required by that state. This value was obtained from the OIE Performance of Veterinary Services (PVS) analysis for India. There was no substantial correlation between any variables (Supplementary Figure 2.1-2.2).

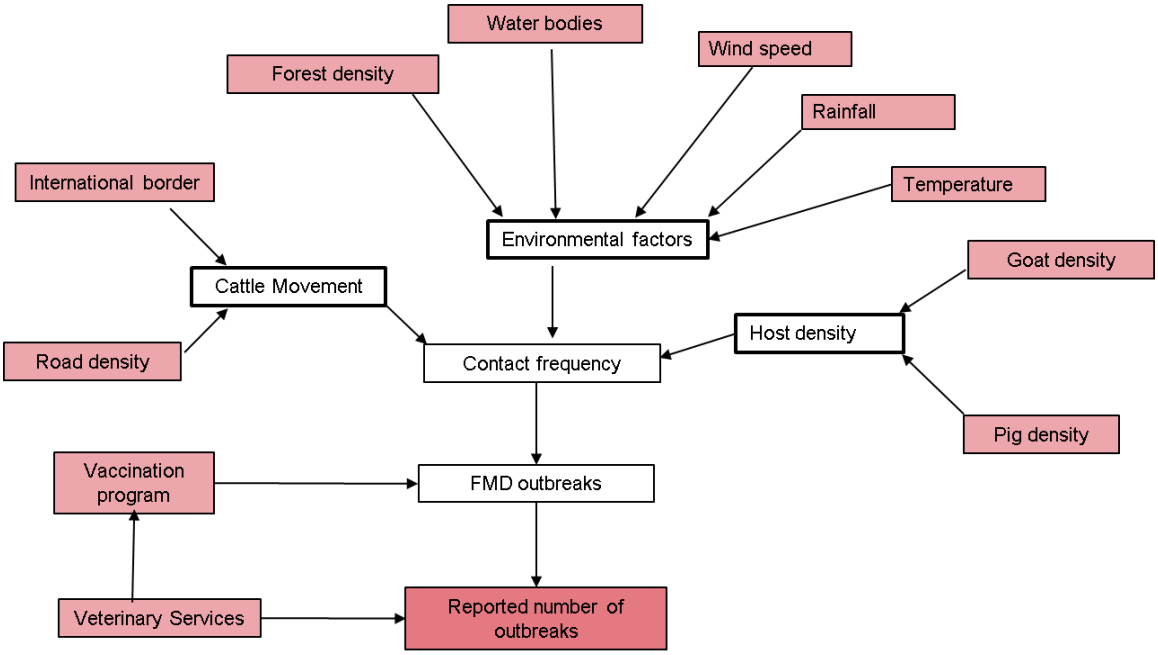


Figure 2.1: Pathways by which hypothesized risk factors may influence the reported number of FMD outbreaks per state per year. The outcome of interest is shown in the dark red box and measured risk factors in the shaded pale red boxes. Measured risk factors were interpreted as proxies for processes, shown in white boxes, that could potentially influence the occurrence of outbreaks, and reporting of outbreaks.

2.2.5 Bayesian space-time hierarchical model

The observed number of outbreaks per state per year was assumed to follow a Poisson distribution $y_{it} \sim Poisson\ (e_{it}, \theta_{it})$, with y_{it} representing the number of FMD outbreaks in state *i* in year *t*; e_{it} representing the expected number of outbreaks defined as above, and θ_{it} is the yearly relative risk for each state. This relative risk incorporated both spatially structured (spatial correlation amongst neighboring states) and unstructured (i.e., random variation) effects, such that:

$$\log(\theta_i) = \alpha + \upsilon_i + \nu_i + \Sigma\beta$$

Where α is the intercept representing the overall level of risk in country, υ_i is the structured spatial effect, and ν_i is the unstructured spatial effect that functions as a random effect for each state. Variables (fixed effects) that modify relative risk are represented by β. This model is known as the BYM2 model (Riebler et al., 2016). Penalized priors were used in the Bayesian analysis following previous studies (Fuglstad et al., 2019).

Spatial-only model

Because data for many of the risk factors were only available for a single point in time, a spatial-only model was initially built to screen important risk factors among potential predictors. Although

we do not expect substantial annual variation in such variables (i.e., cattle population sizes are not expected to change rapidly), the spatial-only model was done so that significance of variables was not inflated due to replication of predictor data across years.

For model selection purposes, univariable analyses were first performed separately for each variable. Backward selection was then performed from a full multivariable model by removing the variables with the widest confidence interval that overlapped zero. From among those different models, the simplest model that was <2 ΔDIC from the model with the lowest DIC value was considered the best-fit spatial-only model (Spiegelhalter et al., 2002). Risk factors in the best-fit spatial-only model were considered as candidate variables in the space-time model alongside temporally variable risk factors (in which yearly data were available from 2008-2016).

Space-time model

To incorporate temporal effects into the model of risk, a BYM2 model was used (Riebler et al., 2016). Several possible model structures exist to incorporate the temporal effect (summarized in Table 2.1): time (year) can be considered as a random effect (ω_t, Equation 1 in Table 2.1), a structured effect (γ_t), in which a random walk is used to account for between-year dependencies ($\omega_t + \gamma_t$, Equation 2) and/or as a random, structured, and space-time interaction ($\omega_t + \gamma_t + \delta_{it}$, Equation 3).The best model structure was selected from amongst these models using DIC. This structure was then used to evaluate the contribution of hypothesized risk factors in shaping relative risk.

Table 2.1: Specification of different model structures, including DIC and posterior predictive p-values. Each model adds additional components to the previous model. * DIC for the spatial-only model is not comparable to space-time model.

Model	**Specification**	**DIC**	**P-values (lower, upper)**
Eq. 1. Spatial only model	$\log(\theta_{it}) = \alpha + v_i + \upsilon_t$	224.99*	(15.2,0)
Eq. 2. Space time model (time as an unstructured effect)	$\log(\theta_{it}) = \alpha + v_i + \upsilon_t + \omega_t$	2765.52	(63.7, 26.1)
Eq. 3. Space time model (time as a structured effect)	$\log(\theta_{it}) = \alpha + vi + \upsilon_t + \omega_t + \acute{y}_t$	2768.51	(63.7, 26.1)

Eq. 4. Space time model (space-time interaction)	$\log(\theta_{it}) = \alpha + vi + u_t + \omega_t + \gamma_t + \delta_{it}$	1222.6	(33,0.7)

We also calculated how much variability is explained by each component that made up the final model structure. Once the best model structure was selected, variable selection was performed as described for the spatial-only model, including temporally variable risk factors and spatial-only factors from the spatial-only model as candidate fixed effects. Excess risk (ER) for a given state was calculated as the proportion of the posterior for each fitted θ_{it} that exceeded 0.8.

Prior sensitivity analysis

We used non-informative penalized complexity priors, which are applicable for a large class of hierarchical models (Simpson et al., 2014). Since prior distributions can influence model results, we conducted a sensitivity analysis on the priors. Penalized priors consider that there is a base model and that the complex model that we obtain is a result of deviation from the base model. For Gaussian Random Field distributions, the base model can be given as $\pi(x/\xi)$ where $\xi=0$. The objective of using the penalized priors is to make the model similar as possible to the base model. It has been identified that penalized priors can also account for model overdispersion (Simpson et al., 2014).

The model was refitted with different penalized complexity priors and non-informative priors to evaluate the extent to which our results were sensitive to different prior assumptions (Supplementary figure S2.3).

Model diagnostics

The fit of the final model (selected based on DIC, as described above) was evaluated using posterior predictive p-values, defined as $p(y_i^* \leq y_i|y)$, where y_i^* is the posterior of the predicted distribution from the model. Posterior predictive p-values can be interpreted as an approximation of the proportion of the predicted distribution for y_i that is more extreme than the observed value, and values of $p(y_i^* \leq y_i|y)$ near 0 and 1 indicate poor model fit. If the model is performing well, then a greater portion of the posterior of the predicted values should be >0.1 and <0.9 (Blangiardo and Cameletti, 2015). In addition, the proportion of marginal variance for random effects and each model component was checked in the final model. The explained variability from the covariates was obtained as a percentage of change of standard deviation from the null model to the model with all the selected covariates. We also calculated the correlation between the predicted and observed values (Spearman's correlation).

2.2.6 Software

All analyses were performed in R statistical software. Different packages such as tidyverse 1.2.123 (Wickham 2017), spdep 0.7–425 (Bivand et al.,2015), dplyr, stringr, and ggplot2 were used. For the Bayesian models, INLA 19.09.03 (Rue et al., 2019) was used, and model results were processed with INLAOutputs 19.09.03 (Baquero et al., 2018).

2.3 Results

2.3.1 Descriptive results

A total of 3282 outbreaks were reported over a period of nine years from 2008 to 2016, with substantial heterogeneity in the spatial and temporal occurrence of outbreaks (as shown by SIR values) across states and years (Figure 2.2). An outbreak is defined as occurrence of one or more cases in the studied area with the same likelihood of exposure to FMDV (Terrestrial Code, OIE). During this time period, we summarize the antibody titer data measured in 1,002,437 animals via LBP-ELISA. Antibody titer data were only available for states that were part of FMDCP phase I. This pre- and post-vaccination monitoring demonstrated that the percent of animals with inferred protection for serotype O generally increased after vaccination, but there was high variation between states and through years (Figure 2.3). Similar trends were observed for serotypes A and Asia1 (Supplementary Figure S2.4).

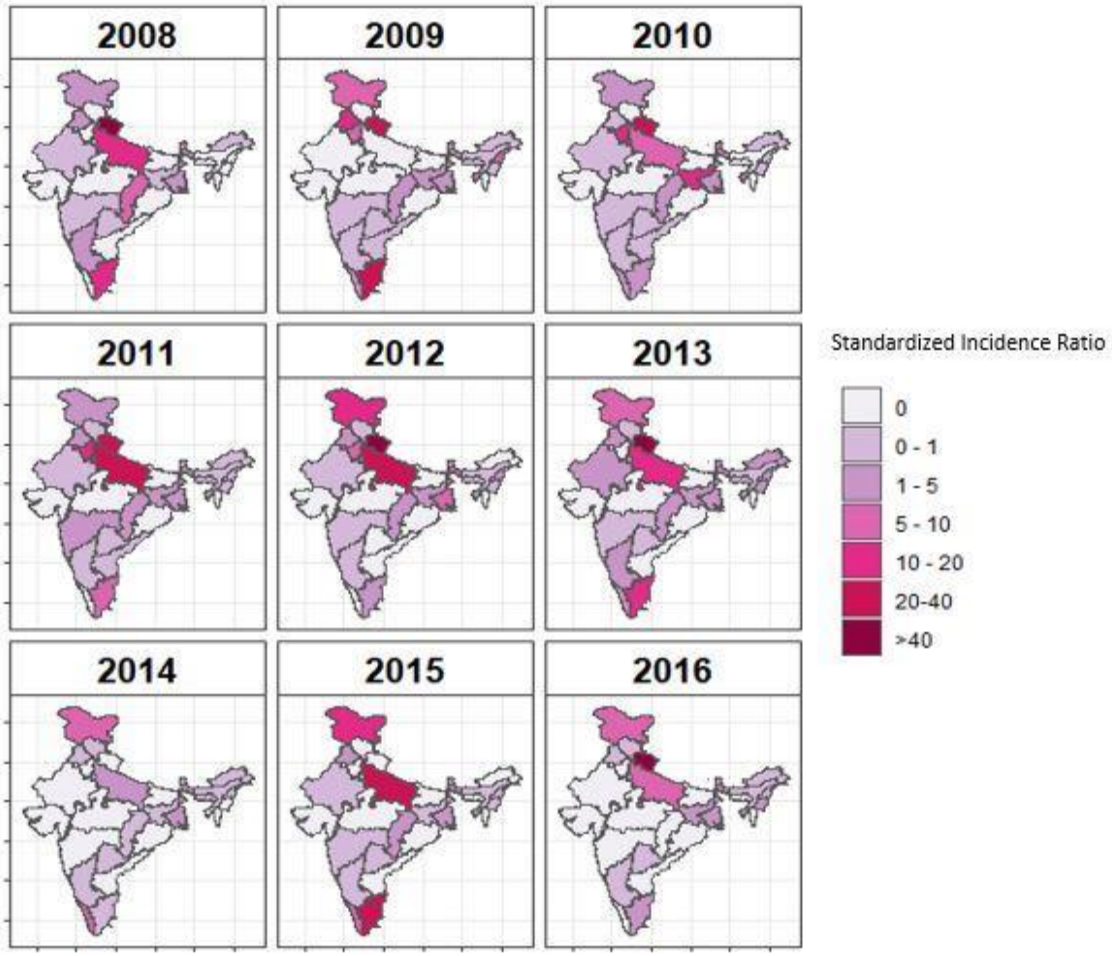


Figure 2.2: Annual standardized incidence ratio (SIR) of reported FMD outbreaks from 2008 to 2016.

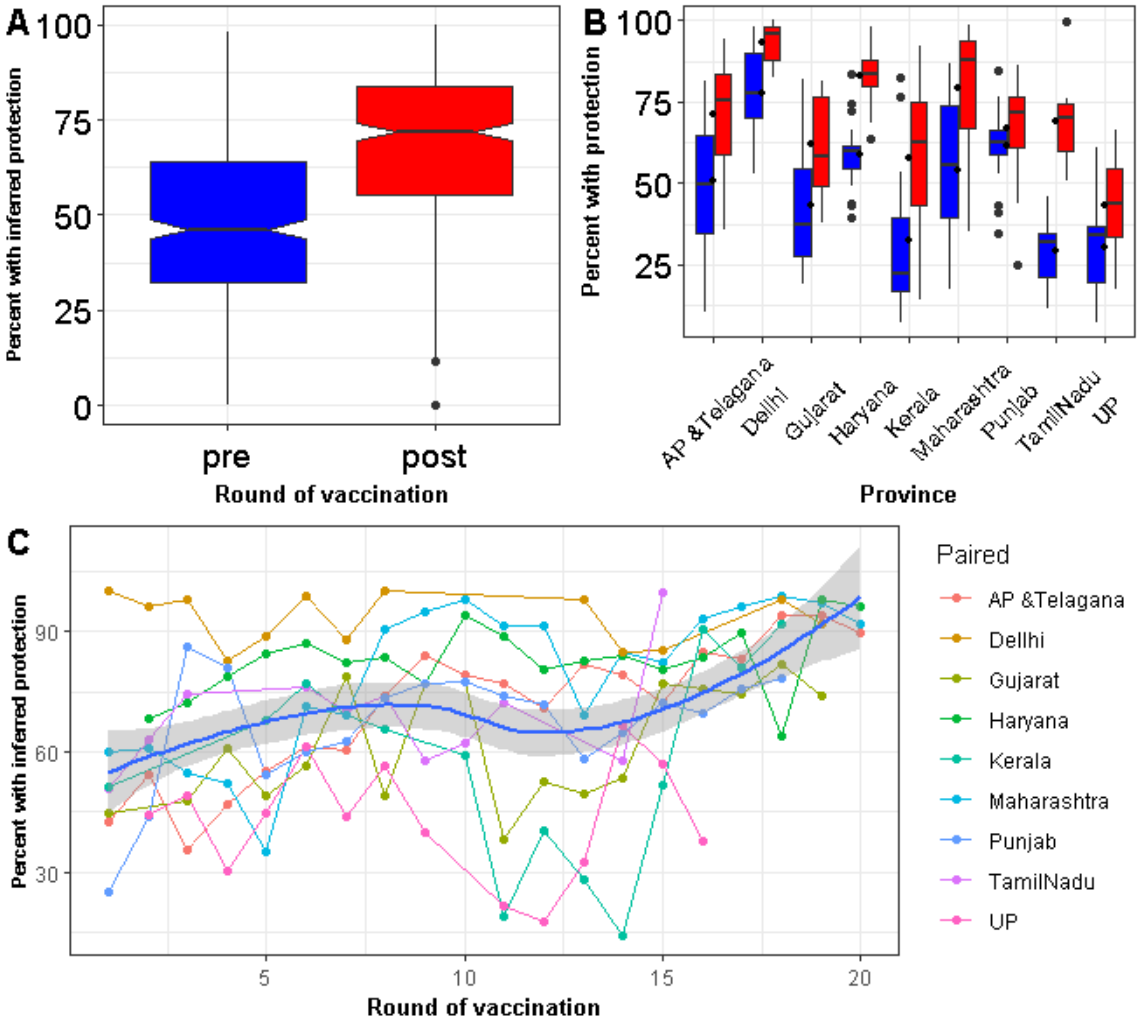


Figure 2.3: Percent of animals with inferred protection based on antibody titers pre- and post-vaccination (A) overall, (B) for each state summarized across all rounds of vaccination, and (C) post-vaccination through time by state. In (A) and (B), pre- and post-vaccination values are shown in blue and red, respectively, for serotype O. In (C), a Loess smoothed line was plotted to visualize an overall increasing trend. States that participated in the FMD PCP phase I are only considered.

For years in which NSP-ELISA data were available (2009-2016), there was a statistically significant negative correlation between the percent of animals positive to LPB-ELISA and NSP-ELISA at the state level (σ = -0.39, $p<0.01$). There was no significant correlation between NSP sero-prevalence and the number of outbreaks per state per year (σ = 0.17, p-value=0.10) or SIR (σ = 0.11, p-value=0.32). However, the percent of animals with inferred protection via LBP-ELISA was negatively correlated with raw outbreak numbers (σ = -0.29, $p<0.001$) and SIR (σ = - 0.25, p-value=0.03).

2.3.2 Bayesian modeling results

Selection of best model structure

Table 2.1 shows the different space-time models that were tested to select the best fitting model structure based on DIC. For the selected model structure that best fit the data (Eq 4, Table 2.1), the unstructured spatial effect accounted for 62% of the variability, whereas the structured spatial effect accounted for only 14.6% of the variability (Supplementary Table S2.2). This means there was relatively little correlation in the occurrence of reported outbreaks across neighboring states through time. The other factor that accounted for substantial variability was the space-time interaction effect.

Univariable analyses of potential risk factors

Variables for which data were only available for one year were first screened in univariable spatial-only models, whereas time-varying variables were screened in univariable spatiotemporal models. Variables that were associated with reported outbreaks (credible interval of odds ratio does not overlap one) are shown in the Table 2.2. The complete list of variables is included in the supplementary materials (Supplementary table S2.3 and S2.4). In the spatial-only models, bordering a country of pool 1, having higher pig density and higher forest coverage was associated with increased relative risk of outbreaks, whereas states with no international border had reduced numbers of reported outbreaks. In the space-time univariable analysis, the risk of reported outbreaks decreased if a given state was in the vaccination program. All univariable models included the underlying terms that accounted for spatially structured and unstructured effects, as well as temporal effects if applicable.

Table 2.2: Results of the univariable analysis for the a) spatial-only model and b) space-time model (coefficients and credible intervals are exponentiated to be on the odds scale).

Univariable model	**DIC**	**Coefficient (Credible Interval)**
A) **Spatial-only models**		
No international border	228.35	0.52 (0.06,0.74)
Bordered by country of FMD pool 1	229.33	6.11 (1.72, 21.93)
Pig density (reference: low)	229.81	2.23 (1.06, 4.72)
Forest coverage density (reference: low)	230.25	3.81 (2.07, 7.06)
B) Space time model		
Participation in the vaccination program (reference: No)	1224.01	0.41 (0.22,0.78)

Best-fit multivariable model

Fixed effects identified from the best-fit multivariable spatial-only model were bordering a pool 1 country, not having an international border, waterbody density, road density, pig density, forest coverage, and veterinary service fulfillment percentage at the state level (Supplementary Table S2.5).These variables were included as candidates in the multivariable space-time model. Two predictors were retained in the final space-time model: No international border and participation in the vaccination program (Table 2.3). Fitted relative risk values calculated from the best-fit model for each state are shown in Figure 2.4. Most border areas show continuous high risk throughout the years. In addition, relative risk increased in some areas while decreased in others. Excess risk peaked in many states between 2011 and 2013 (Supplementary figure S2.5).

The sensitivity analysis of model priors demonstrated that similar DIC and p-values were produced regardless of choice of priors. Fitted and observed values have a Spearman's correlation of 0.93.

Table 2.3: Results from the final Bayesian space-time model. Coefficients and credible intervals have been exponentiated to be on the odds scale.

Fixed effect	**Coefficient 95% Credible Interval**
Intercept	1.64 (0.72, 3.68)
No international border	0.27 (0.08, 0.99)
Participation in the vaccination program (reference: No)	0.45 (0.24, 0.84)

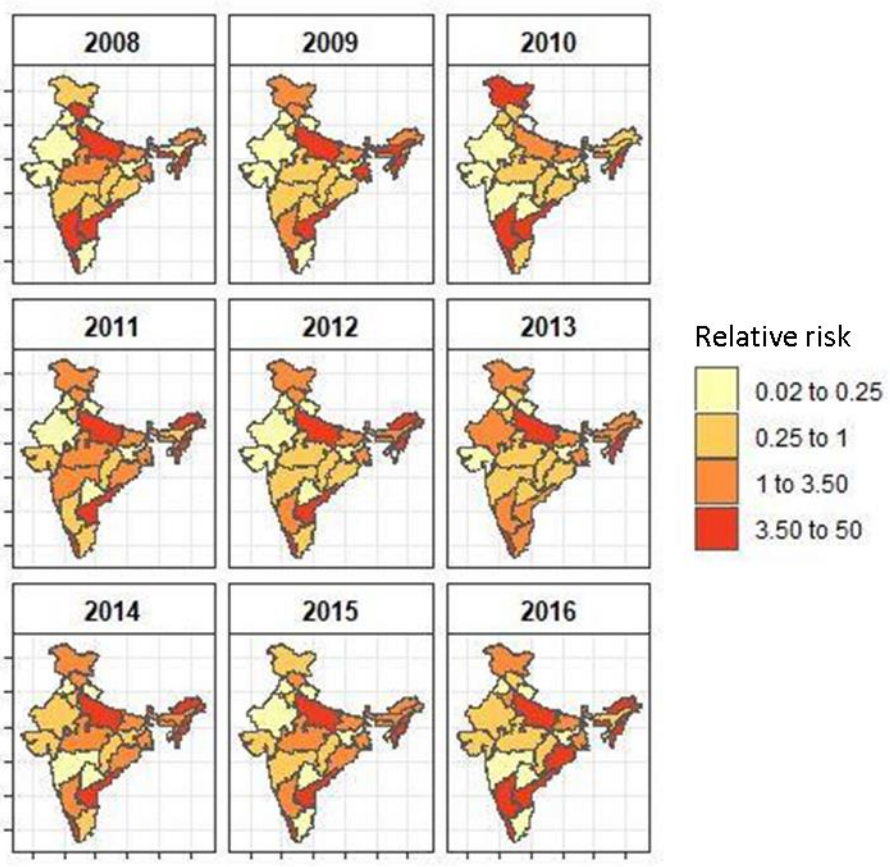


Figure 2.4: Fitted relative risk of outbreaks for each state from the best-fit multivariable model.

2.4 Discussion

In this study, we first conducted a descriptive analysis of epidemiological outcomes of governmental FMDV vaccination programs in India. This analysis showed that the states that were in the vaccination program had fluctuations in the percent of animals with inferred protective antibody titers through each round of vaccination across years and states, but there was a general increase in the percent of animals with protective antibody titers after each round of vaccination and across time. We then analyzed the distribution of reported FMD outbreaks by using a Bayesian space-time model to map high-risk areas and identify factors that influence risk in order to inform risk-based control strategies. This model demonstrated that states that were included in the vaccination program and did not have an international border experienced reduced risk of FMD outbreaks. This result warrants more stringent vaccination and sero-monitoring and movement restrictions at international borders.

India has used FMDV vaccination as a control measure since the 1980s. The trivalent vaccine produced in India has a protective effect against the circulating outbreak strains of serotype O, A, and Asia1, and studies have been carried out on vaccine safety and efficacy (Mahapatra et al., 2015, Mohanty et al., 2015, Subramaniam et al., 2015). Approximately one third of India's cattle and buffalo population have been vaccinated through the government's FMDCP program (Pattnaik 2012). At the beginning of the program, only some districts in each state were included in the program, but the number of districts and states included has increased through time. Following PCP guidelines for stage 3, India has conducted pre- and post-vaccination sero-monitoring according to OIE guidelines to monitor the population-level immunity since 2008. For a given round of vaccination, as expected, antibody titers were higher post-vaccination compared to pre-vaccination (Figure 2.3A, Supplementary Figure S2.4). However, there was substantial heterogeneity across states and between years (Figure 2.3B and C), and the percentage of animals with inferred protection often fell below OIE's recommendation of >80% coverage (Metwally and Münstermann, 2016).

Population demographics, turnover, and waning immunity all may contribute to periodic dips below the 80% threshold (Knight-Jones et al., 2016). According to a study conducted in Turkey, biannual mass vaccination can leave gaps in population-level immunity. Among other factors, young animals may have received insufficient vaccine doses to attain long-lasting immunity, and also animals in late pregnancy are sometimes not vaccinated, resulting in declines in population-level immunity just prior to the subsequent round of vaccination (Knight-Jones et al., 2016). Due to herd demographics and semi-intensive management practices, it was concluded that vaccination without biosecurity may not be able to control FMD in Turkey (Knight-Jones et al., 2016). Similar dynamics may also occur in India, as shown by the high spatial and temporal variation in the percent of animals with inferred protection, and these spatial and temporal gaps in herd immunity may allow for the persistence and spread of FMDV in the country. The proportion of animals with inferred protection could also have been influenced by inconsistent vaccine administration, delay in re-vaccination, lack of booster doses in the primo-vaccinated calves, transboundary introduction of naïve animals, and transport conditions, which could have contributed to variable antibody titers.

We also investigated the relationship between the occurrence of FMDV within states and vaccination data (i.e., participation in FMDCP or the percent of animals with inferred protection via LPB-ELISA). We used two imperfect measures to quantify the extent of FMD circulation: standardized incidence ratios (SIR, based on reported outbreaks) and NSP-based sero-prevalence. Outbreak reporting can be inconsistent and likely provides an incomplete picture of FMDV incidence. In contrast, the NSP-ELISA data captured the percentage of animals with an

anti-NSP response, which is indicative of natural infection. A naturally infected animal is also expected to be positive on LPB-ELISA, thus the percent of animals with inferred protection (based on LPB-ELISA) cannot discriminate between immunity due to vaccination or natural infection. However, during the period under study, the percent of animals positive on LBP-ELISA and NSP-ELISA was negatively correlated. These results suggest that a) LPB-ELISA data can be interpreted as an indicator of vaccine coverage rather than natural virus circulation, and b) areas with higher vaccine coverage experienced reduced circulation of FMDV (as shown by low NSP sero-prevalence and fewer reported outbreaks). These results are in agreement with a study conducted in 2014, which identified that the states in the biannual vaccination program performed better in terms of reporting lower disease incidence in herds by a series of cross-sectional studies (Sharma et al., 2014). In our study, though fewer outbreaks and lower SIRs were reported in states with higher LPB seroprevalence. Once we accounted for spatiotemporal dynamics in the space-time model, participation in the FMDCP reduced outbreaks by ~55%.

To better understand heterogeneities in outbreak occurrence within India, we developed a Bayesian space-time model that allowed us to examine risk factors associated with outbreak risk alongside model components that accounted for the spatial interdependency of risk across states. Although reported outbreak numbers are likely an underestimate of the true number of outbreaks, analyzing patterns of reported outbreak occurrence does advance our understanding about the factors that cause outbreaks. An examination of the variance explained by each component making up the model's structural backbone (Supplementary table S2.2) revealed initial insights into processes shaping outbreak risk. First, the unstructured spatial effect (which essentially operates as a random effect for each state) contributed the most to explaining variability in the outcome, which suggests that there were unaccounted for variables at the state level that were important in structuring outbreak risk. These could include animal movement for grazing and trading, and human movement related to biosecurity of farms, among others. In contrast, the structured spatial effect explained relatively little variation, indicating that the outbreak risk in one state was not closely correlated with the occurrence of outbreaks in neighboring states. Likely, this pattern may be because states in India are large, and a smaller spatial scale would better capture the local spatial dynamics of outbreak propagation. Also, this result suggests that outbreaks or control programs in one state would not have large impacts on the adjacent state.

From our model, it is evident that relative risk of outbreaks changes through time and space, though there are some states that were more consistently at higher risk (Figure 2.4). The two variables retained in the final model were participation in the FMDCP vaccination program and not having an international border. The relative risk of outbreaks in states that were part of the FMDCP during 2008 to 2016 was about one half that of states that were not part of the program.

This is consistent with our descriptive analysis on the importance of vaccination. It is also notable that we observe a benefit of the FMDCP in the number of reported outbreaks state-wide, despite the observed variability in percent of animals with inferred protection and that the FMDCP did not always extend to all districts within the state.

The other important risk factor identified by the model was a ~70% reduction in the relative risk of outbreaks in states with no international border. Thus, international borders increased the relative risk of outbreaks. In a longitudinal study conducted in 2014 to determine serological herd immunity, it was identified that for that year, border states such as Assam, Rajasthan, Jammu and Kashmir, West Bengal, and Uttar Pradesh were at high risk due to low population immunity. There were also instances where high incidence of FMD was observed in border states even where herd immunity was high (Sharma et al., 2014). We observed the same pattern of border states having greater excess risk (Supplementary Figure S2.5), including the states of Meghalaya, Assam, Arunachal, and Jammu & Kashmir.

The potential for transboundary introductions of novel FMDVs into India from neighboring countries may in part explain the risks associated with international borders. This may occur through movement of subclinically infected animals or fomites (Stenfeldt and Arzt, 2020). Alternatively, transboundary value chains may result in high risk in certain border states if animals are transported to border states from elsewhere in India prior to exportation. Legal and illegal animal movement occurs between neighboring countries, but the extent of such transboundary movements depends on the countries involved (Landes et al., 2017). Previous studies have identified that cattle and buffaloes are transported from India to Malaysia through Myanmar and Vietnam (all pool 1 countries), which may lead to the dissemination of FMDV (Rweyemamu et al., 2008, Smith et al., 2016), and OIE has identified that the virus can spread extensively to the Southeast Asia region due to intensive livestock trade (Bartels et al., 2017). Bayesian phylogeographic reconstruction has previously effectively demonstrated transboundary and within-country movements of lineages of FMDV O/ME-SA/PanAsia (Brito et al., 2017). Interestingly, our spatial-only model suggests that bordering a country in pool 1 carried a higher risk, which would be consistent with the idea that transboundary movements with pool 1 countries shapes FMD risk within India, although this variable was not retained in the space-time model.

Interestingly, excess risk peaked between 2009, 2011 and 2013 in almost all the states (Supplementary Figure S2.5). During 2013, widespread FMD outbreaks occurred in India, caused by the strain O/ME-SA/Ind2001d within serotype O. This strain also spread to other countries in the Middle East and Southeast Asia at this time (Subramaniam et al., 2015, Brito et al., 2017a),

suggesting that periods of excess risk in India may also translate to heightened frequencies of transboundary transmission.

India is in the stage 3 of the Progressive Control Pathway (PCP) for FMD. Countries within this stage should engage in ongoing monitoring of risk and implementation of risk-based strategic to define a pathway to obtain freedom from FMD (with vaccination) in at least one geographic zone, including analysis of passive/active surveillance data to document epidemiological evidence of reductions in FMD incidence. Related to this, our results suggest that a feasible strategy may be to continue trying to decrease prevalence in identified high risk areas to mitigate the impact of the disease with special focus on states that are part of international borders. Alternatively, low risk areas identified from this spatial analysis could help delineate areas in which zonal freedom may be more readily attainable.

There are several caveats to the interpretation of the serological data that present limitations to this study. First, NSP data were available only from 24 states out of 29 states for six years. In addition, transient increases of NSP titers can occur within 21 days of vaccination in up to 15% percent of previously uninfected animals, which complicates the interpretation of NSP results particularly if vaccination history is not available (Hayer et al., 2018, Mohapatra 2011). Second, animals can be positive on an LPB-ELISA from either vaccination or natural infection. The negative correlation between NSP-ELISA and LPB-ELISA data suggests that a) rates of LPB-ELISA positivity likely represented vaccination rather than natural infection, and b) rates of NSP-ELISA positivity were not coupled with vaccination. However, serial testing and monitoring for clinical signs is necessary to identify the changes in antibody titers in infected and vaccinated animals to determine whether animals have acquired antibodies due to infection or vaccination (Mohanty et al., 2015). Related to this, when pre- and post-vaccination antibody titers were compared at the state level, samples were not coming from the same animal which limits the conclusions we can draw from this comparison.

Another limitation of this study related to the space-time modeling is that outbreak data come from passive surveillance, and there may be substantial under-reporting. If there are spatial biases in the extent to which outbreaks are under-reported, then this could introduce spatial biases to the SIR data and the data used for the space-time model. We attempted to partially address this by including veterinary service coverage as a potential predictor, though it is unclear whether veterinary service coverage is a useful proxy for variable reporting and this factor was not retained in the final model. These types of potential bias are common in observational epidemiological studies that rely on passive surveillance; however, we believe there is still value in describing large-scale patterns of FMD incidence. In addition, we have no information about

the number of animals infected in each outbreak, which means that small and large outbreaks receive equal weight in our analysis. Finally, no environmental or climatic factors were retained in our best-fit model. This may be an artefact and limitation of the state-level spatial and yearly temporal scale of our analysis, which did not allow us to capture finer-scale spatial variation or seasonal effects. For example, it is suspected viral spread and FMD incidence increases with monsoon heavy rains in November through January, and more outbreaks are reported after the rains (ICAR reports). Future analyses could overcome the limitations imposed by the spatial and temporal resolution of our outbreak data by tabulating outbreak data on a finer spatiotemporal scale, thus enabling a better evaluation of the importance of environmental risk factors. The data for the year beyond 2016 has not been included in the present study because the country started using Solid Phase Competitive ELISA (SPCE) assessment of herd immunity, and this could create confusion in the pattern of protection at population level.

Conclusion

In this study, we have shown that the standardized incidence of FMD outbreaks has reduced over time with the implementation of mass vaccination, though the percent of animals with inferred protection was highly variable through space and time and often fell below the desired threshold of >80%. Over the same time period, the percentage of animals with inferred protection was negatively correlated with the number of reported outbreaks in a state. Through implementing a Bayesian space-time model, we have demonstrated that states that were part of the FMDCP experienced a ~50% reduction in the risk of reported outbreaks. Our results also demonstrate a substantial risk of outbreaks associated with international borders, suggesting a role of transboundary movements of animals or fomites in shaping FMD incidence. For India to proceed with a risk-based strategic control plan, it is important to reinforce surveillance activities and animal movement control at states with international borders. This study advances understanding of risk factors associated with high risk areas, which will contribute to a better understanding of viral circulation and contribute towards efforts to reduce disease prevalence.

Table S2.1. All the variables used in the analysis.

Variable	**Description and units**	**mean (1^{st} & 3^{rd} quantile)**	**Resolution/ calculation**	**Source**
Cattle density		Mean 95.039 /Km^2	Total cattle population in the state/ total Sq Km area of the state	India Department of Animal Husbandry and the Dairying census, India (http://www.dahd.nic.in/about-us/divisions/statistics)
Pig density	2	Mean 7.60 (0, 51.44)	Total pig population in	India Department of Animal Husbandry and the Dairying census, India

	(reference: low)	High 7.6 to 51.44/km^2 Low 0 to 7.6/km^2	the state/ total Sq Km area of the state	(http://www.dahd.nic.in/about-us/divisions/statistics)
Goat density	2 (reference: low)	Mean 36.33 High: 36.32-129.64 /km^2 Low: 0.05-36.32 / km^2	Total goat population in the state/ total Sq Km area of the state	India Department of Animal Husbandry and the Dairying census, India http://www.dahd.nic.in/about-us/divisions/statistics
Road density	2 (reference: low)	Mean 0.52 High 0.52-2.94/km^2 Low 0.004 to 0.52/km^2	Total road Sq Km in the state/ total Sq Km area of the state	India Department of Animal Husbandry and the Dairying census, India http://www.dahd.nic.in/about-us/divisions/statistics
Waterbody density	2 (reference: low)	Mean 0.0365 High 3.65E-02 to 0.156/Km2 Low 1.12E-05 to 3.65E-02/Km^2	Total water body area Sq Km in the state/ total Sq Km area of the state	India Department of Animal Husbandry and the Dairying census, India http://www.dahd.nic.in/about-us/divisions/statistics
Forest coverage	2 (reference: low)	0.33(0.035, 0.862) High 0.33 to 0.86/Km^2 Low 0.035 to 0.33/Km^2	Total forest coverage area Sq Km in the state/ total Sq Km area of the state	India Department of Animal Husbandry and the Dairying census, India http://www.dahd.nic.in/about-us/divisions/statistics
Veterinary service	Percentage		Available number of veterinarians/ Required number of Veterinarians	OIE PVS analysis https://www.oie.int/solidarity/pvs-gap-analysis/pvs-gap-analysis-reports/

Bordered by a pool1 country	1 or 0		Paton et al., 2018
Bordered by a pool2 country	1 or 0		Paton et al., 2018
Bordered by a pool3 country	1 or 0		Paton et al., 2018
No international border	1 or 0		Paton et al., 2018
Vaccination program	1 or 0		ICAR annual report
Annual windspeed average	Standardized based on the mean value	Monthly windspeed average 10m wind speed from raster (high-spatial resolution (1/24°, ~4-Km))	Terra Climate data (Abatzoglou et al., 2018)
Annual windspeed variance	Standardized based on a mean value		Terra Climate data (Abatzoglou et al., 2018)
Annual temperature average	Standardized based on mean value	Monthly average maximum and minimum temperature raster (high resolution)	Terra Climate data (Abatzoglou et al., 2018)
Annual temperature variance	Standardized based on mean value	Cumulative variance of each monthly average temperature	Terra Climate data (Abatzoglou et al., 2018)
Annual rainfall average	Standardized based on mean value		India meteorological department http://www.imd.gov.in/Welcome%20To%20IMD/Welcome.php

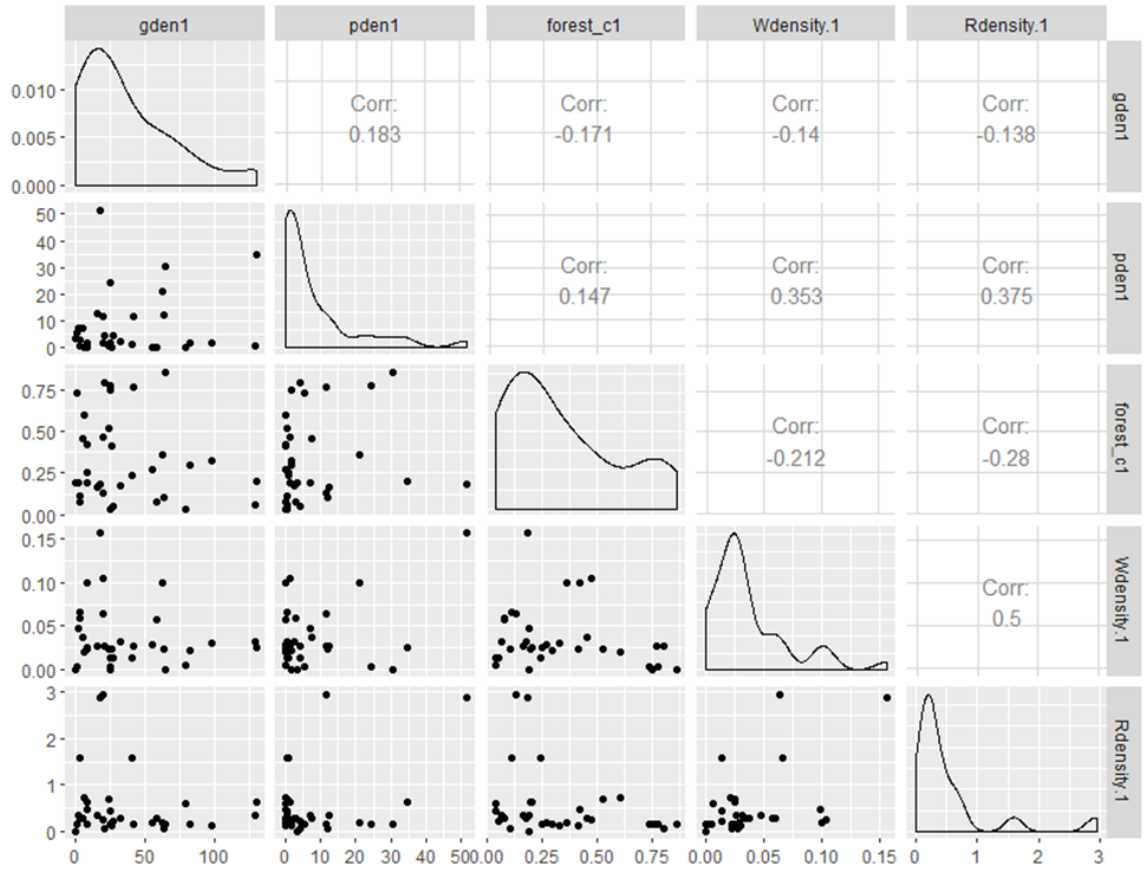


Figure S2.1: Correlation plot for variables used in the spatial only model.

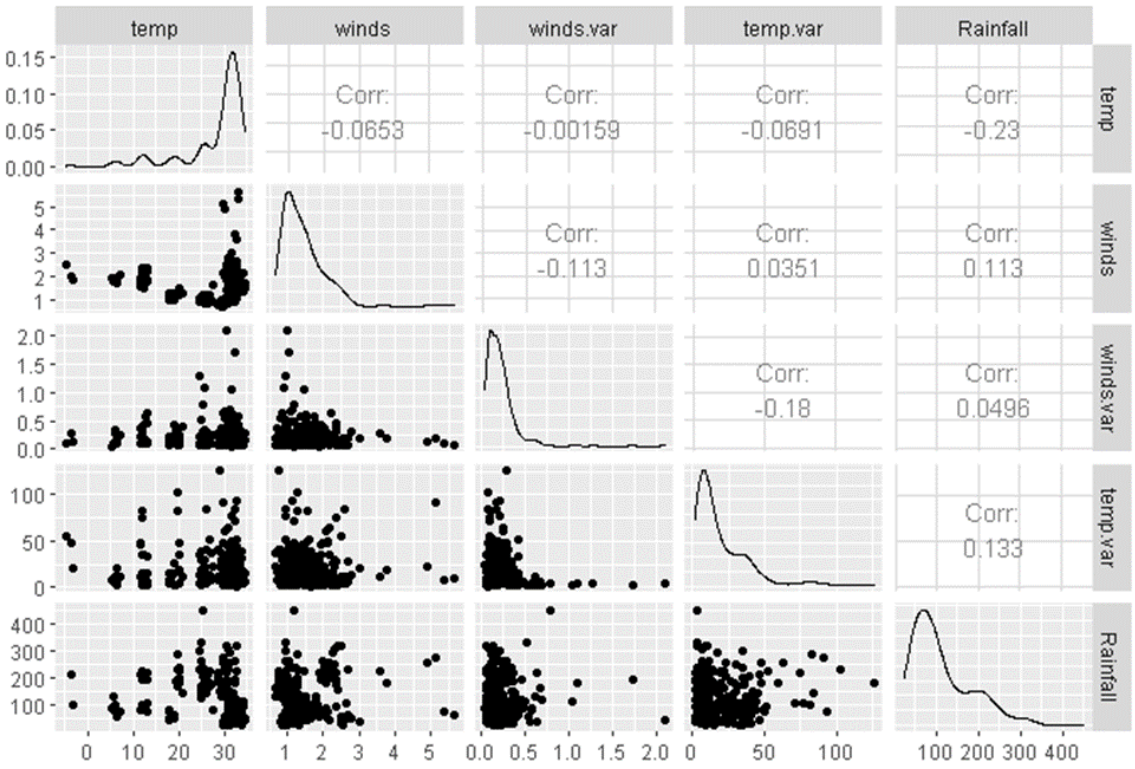
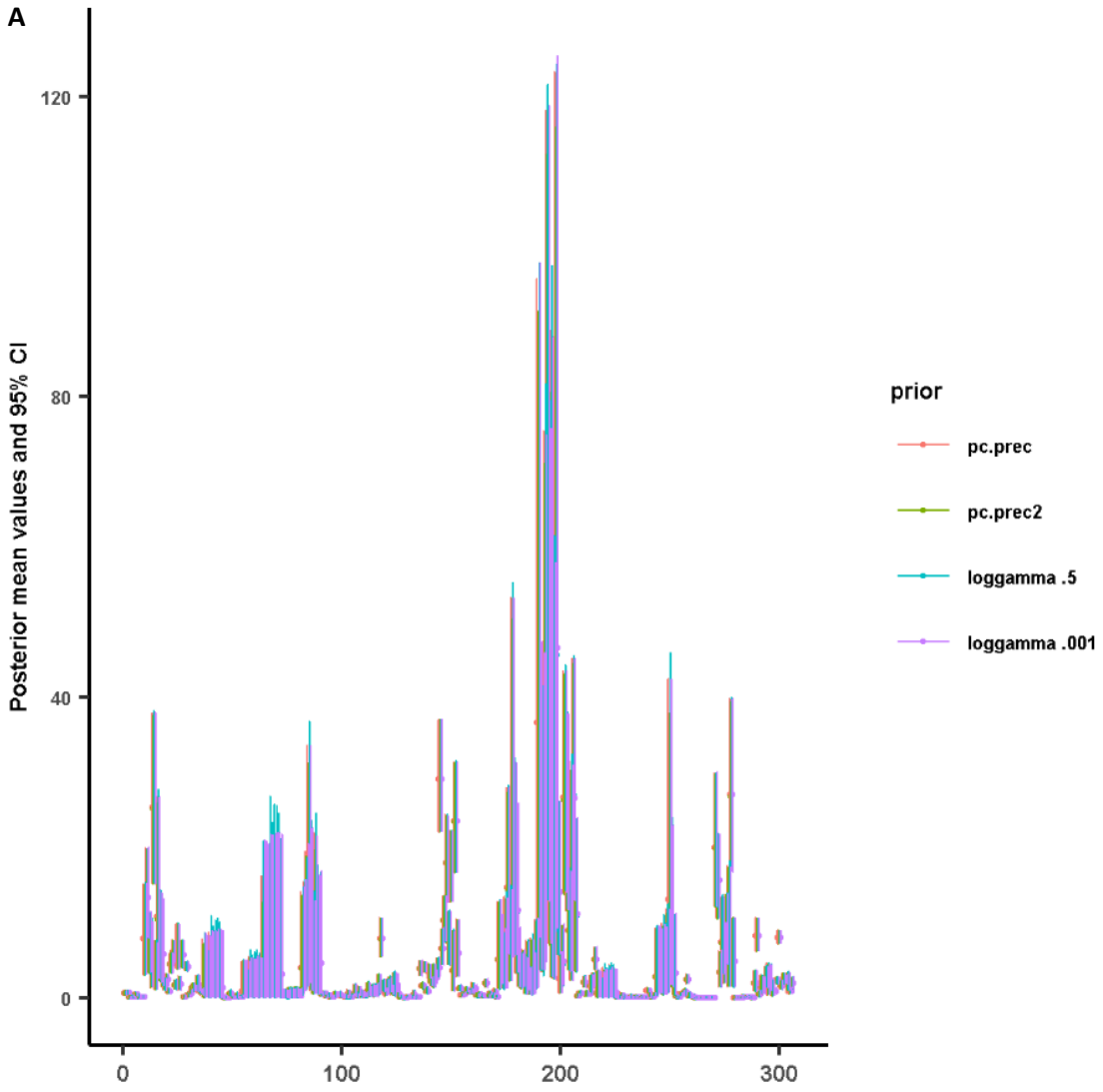
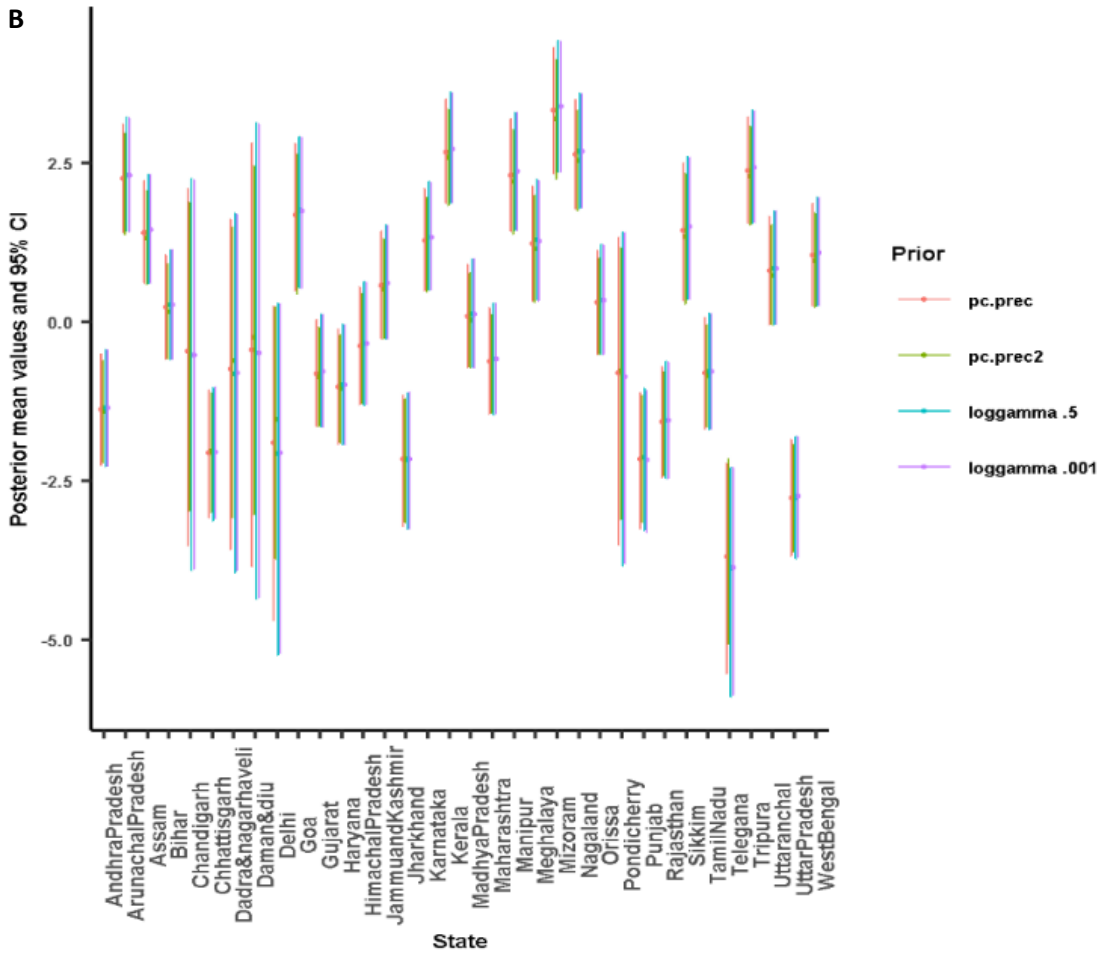


Figure S2.2: Correlation plot for variables used in the space time only model.


A
Posterior mean values and 95% CI
120
80
40
0
0
100
200
300
prior
pc.prec
pc.prec2
loggamma .5
loggamma .001


B
Posterior mean values and 95% CI
2.5
0.0
-2.5
-5.0
AndhraPradesh
ArunachalPradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra&nagarhaveli
Daman&diu
Delhi
Goa
Gujarat
Haryana
HimachalPradesh
JammuandKashmir
Jharkhand
Karnataka
Kerala
MadhyaPradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Orissa
Pondicherry
Punjab
Rajasthan
Sikkim
TamilNadu
Telegana
Tripura
Uttaranchal
UttarPradesh
WestBengal
State
Prior
pc.prec
pc.prec2
loggamma .5
loggamma .001

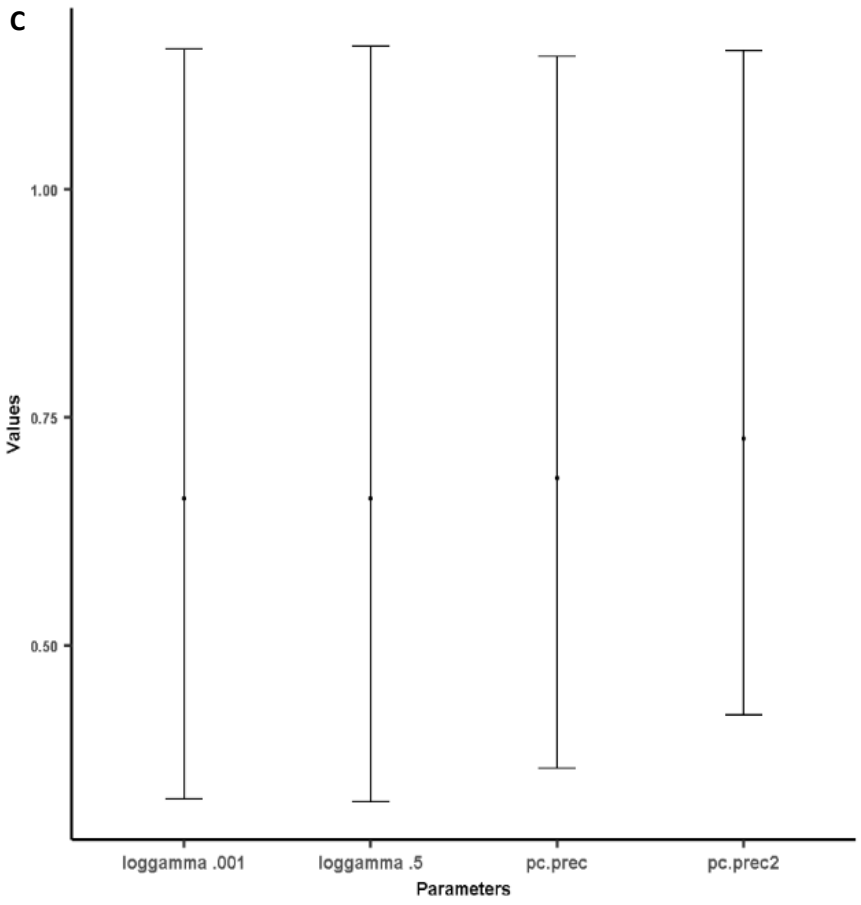


Figure S2.3: Results from the prior sensitivity analysis showing that different prior combinations produce similar results. A) Mean, B) Random effect and C) Fixed effects.

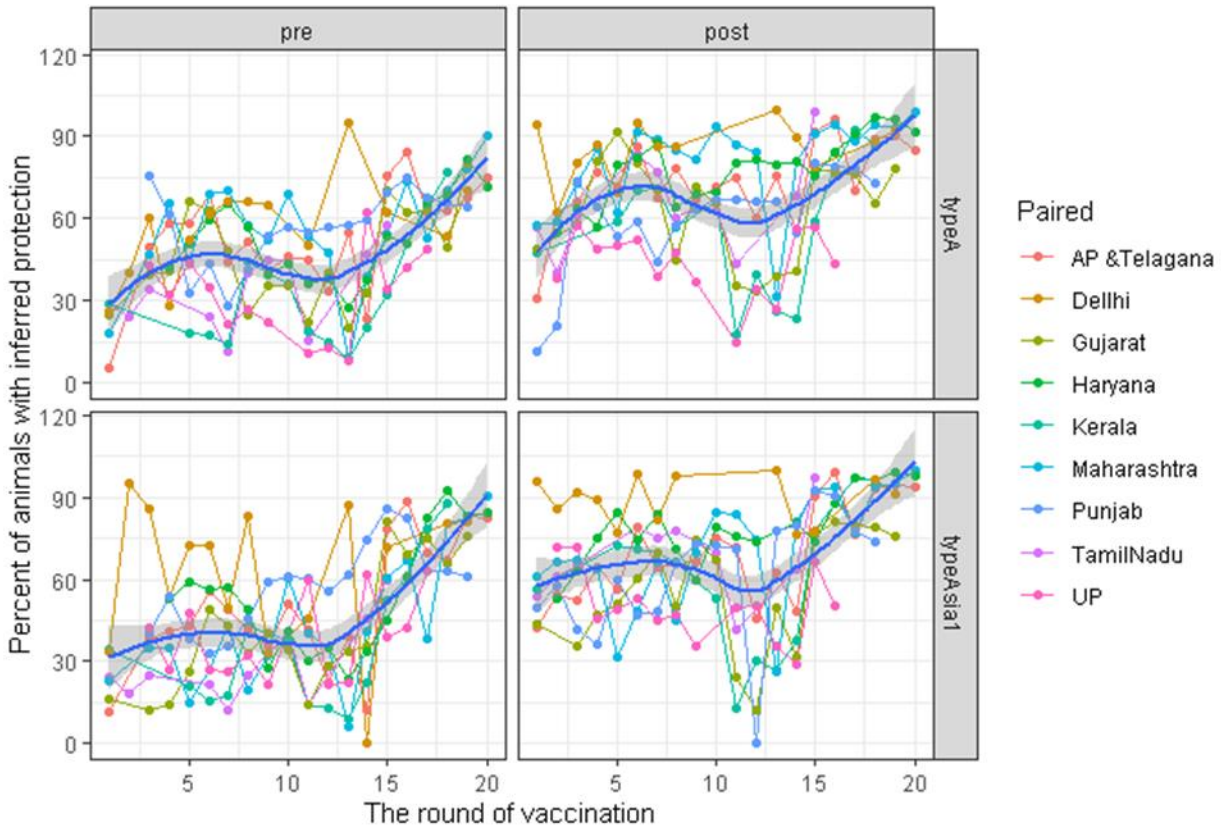


Figure S2.4: Pre- and post-vaccination percent of animals with inferred protection for serotype Asia1 and Serotype A.

Table S2.2: Percent contribution of the model's components to explaining variance in the best fitting model.

Variable	Percentage explained
Unstructured spatial effect	61.9
Structured spatial	14.6
Year random effect	0.5
Year random walk	0.4
Year space interaction	22.6

Table S2.3: Results of the univariate analysis from the spatial only model (coefficients and credible intervals are exponentiated).

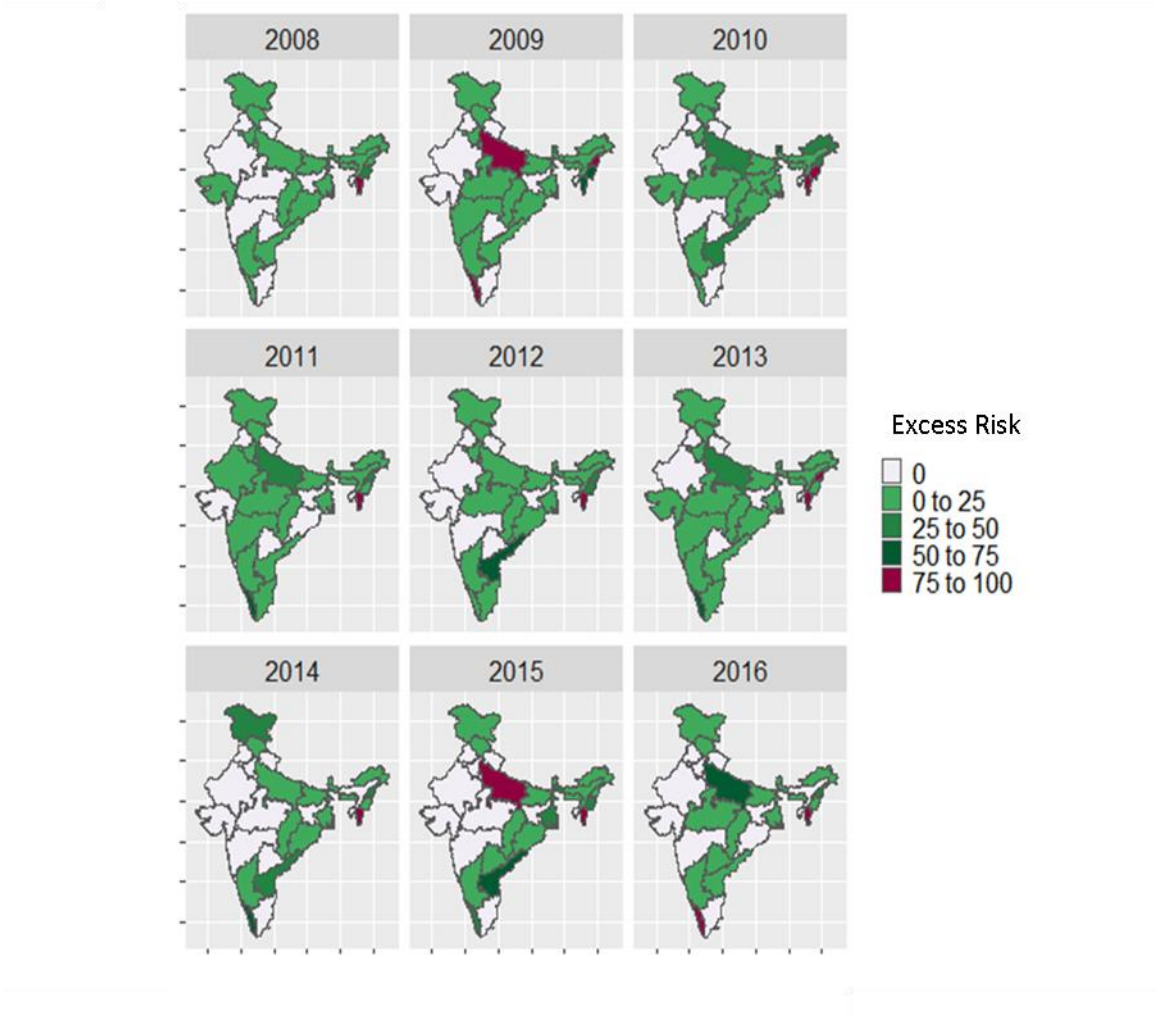
Model with a single covariate	DIC	Credible Interval
Goat density (reference: low)	230.47	1.29 (0.31,5.64)
Forest area density (reference: low)	230.54	3.81 (2.07,7.06)
Pig density (reference: low)	229.81	2.23 (1.06,4.72)
Bordered by a country in pool 1	229.33	6.11 (1.72,21.93)
Bordered by a country in pool 2	229.51	3.15 (0.81,12.54)
Bordered by a country in pool 3	230.61	0.47 (0.07,3.13)
No international border	228.35	0.52 (0.05,0.74)
Road density (reference: low)	230.68	0.60 (0.18,2.04)
Percentage of veterinarians	231.12	1.01 (0.97,1.06)
Water area density (reference: low)	230.06	0.70 (0.19,2.49)

Table S2.4: Results of the univariate analysis from the space time model (coefficients and credible intervals are exponentiated).

Model with a single covariate	DIC	Credible Interval
Annual rainfall average	1223.19	0.91(0.72,1.16)
Annual windspeed average	1227.97	1.13(0.92,1.37)
Annual windspeed variance	1222.66	1.05(0.91,1.21)
Annual temperature average	1224.12	1.09(0.89,1.32)
Annual temperature variance	1223.18	0.96(0.84,1.11)

Table S2.5: Fixed effects identified from the best-fit multivariable spatial-only model (coefficients and credible intervals are exponentiated).

Fixed effect	Coefficient 95% Credible Interval
Intercept	0.93 (0.032, 27.66)
Boarded by a country in pool 1	1.91 (0.34, 10.48)
No international border	0.36 (0.08, 1.55)
Waterbody density (reference: low)	0.92 (0.28, 2.94)
Road density (reference: low)	0.67 (0.21, 2.13)
Pig density 2(reference: low)	1.26 (0.35, 4.58)
Forest area density (reference: low)	5.24 (1.42, 19.64)
Percentage of veterinarians	0.99 (0.95, 1.03)



Supplementary Figure S2.5: Excess risk plot for years 2008-2016 for different states in India.

Chapter 3: Using phylogeography as a proxy for population connectivity for spatial modeling of reported foot-and-mouth disease outbreaks in Vietnam

3.1 Introduction

Space-time risk models are often applied to outbreak data in order to understand patterns and drivers of pathogen spread. In Bayesian space-time models, risk of disease is considered as a spatial process whereby spatial correlations in case counts are captured by accounting for the contiguity or adjacency of spatial units (i.e., states, provinces, etc.)(Berliner, 1996). The resulting spatial structure is a component of a Bayesian hierarchical model that considers latent variables, which include linear predictors and random effects that help account for unexplained variability (Lawson, 2018). Such space-time regressions have been widely used in disease mapping studies to identify high risk areas for both animal and human diseases (Machado et al., 2019,Blangiardo and Cameletti, 2015,Coly et al., 2021).

From our previous studies utilizing a space-time regression for foot and mouth disease (FMD) in India, it was identified that the structured spatial effect only accounted for 14% of variability in reported outbreaks in across states, while the unstructured (random) effect accounted for 61% of variability. One reason for the relatively small contribution of spatial structure to the relative risk of outbreaks may be that spatial adjacency is an imperfect proxy for population connectivity, especially for highly mobile populations of livestock that may frequently be transported long distances. Unfortunately, livestock movement data is often unavailable in many regions where FMDV is endemic, and lack of mobility data is likely a common challenge across many host-pathogen systems. However, patterns of viral dissemination (inferred from phylogeographic models) could serve as a proxy for the underlying connectivity of the host population.

FMDV is a highly contagious disease that is endemic in Southeast Asia affecting pigs, cattle, and buffaloes. The disease is caused by an *Apthovirus* in the *Picornaviridae* family, and serotypes O, A, and Asia 1 have been identified in the region (Le et al., 2016). Clinical signs of the disease such as blisters and pyrexia cause production losses to the farmers. FMD is challenging to control in SEA in part due to diverse animal husbandry practices in the growing economies of the region. In addition, the occurrence of FMD outbreaks in SEA is spatially and temporally variable, and some countries are free from FMD (Paton et al., 2018). There are no studies to identify what exact drivers (ex: human, animal, product) of viral movement, but the virus continues to spread across SEA. Lack of measures to track animal movement between and within countries and undocumented livestock markets are major obstacles to disease preparedness and hinder investigations of patterns of disease occurrence of contagious livestock diseases in the region, including in Vietnam (Rweyemamu et al., 2008,Paton et al., 2018).

Matching movement data with phylogenetic data has been identified as a feasible approach to identify FMD disease transmission among countries and regions as substantiated with the spread of O-PanAsia-2 and O-Ind 2001e to SEA (Paton et al., 2018). In Bankawoska *et al.* (2017), outbreak strains from SEA were phylogenetically matched to lineages arising from India, suggesting that these lineages were introduced from South Asia. Molecular epidemiological studies reveal a pattern of FMDV spreading across SEA and travelling through Vietnam. Serotype O sequences identified in Vietnam had high similarity with the Mya-98 lineage that had spread elsewhere in SEA, and Serotype A sequences from Vietnam belonged to group SEA-97 common to SEA (Le et al., 2016,Brito et al., 2017,de Carvalho Ferreira et al., 2017). The O/Cathay strain first identified in Hong Kong in the 1970s was reintroduced to Taiwan in 1997, and in 2002 to Malaysia, the Philippines, Taiwan, Thailand, and Vietnam. Later in 2002, Vietnam was identified as the recipient of viruses from Malaysia, Thailand, and Hongkong for O/Cathay (Di Nardo et al., 2014). O/ME-SA/Ind-2001e was first identified in SEA during 2013-2017. One of two lineages that evolved from Ind-2001 has been identified in Vietnam, demonstrating the movement of the virus from India to Vietnam (Le et al., 2016,Bachanek-Bankowska et al., 2018,Vu et al., 2017). Similar patterns have identified for O/ME-SA/Pan Asia (2010-2014). Taken together, these studies highlight the transboundary nature of FMDV circulation in SEA.

Considering patterns of O/ME-SA/Pan Asia spread within Vietnam, transmission was identified to be frequent from the South-Central Coast and Northeast to other parts of the country, presumedly due to cattle movement that occurs from central areas of the country to the north and south (Brito et al., 2017). These identified viral movements could help to identify possible dissemination pathways of FMDV within Vietnam and reveal underlying patterns of host movement. Such information could be used to account for host population connectivity.

For space-time risk models of case counts, we hypothesize that historical patterns of viral movement are a better criterion to define population connectedness between spatial units compared to spatial adjacency. Using Vietnam as an example, we construct discrete trait phylogeographic models for FMDV serotype O in Vietnam and SEA. We used the inferred transition rates between each province as the connectivity matrix in a Bayesian space-time regressions and compared this model's ability to explain spatiotemporal variability in relative risk with conventional approaches based on spatial adjacency. We used these models to identify risk factors for FMD such as livestock density, international borders, FMD virus movement between countries, and delineate high-risk areas for FMD in the country. The outcome risk models can be generalized to other infectious diseases that spread in a similar manner that is not present in the country.

3.2 Materials and Methods

3.2.1 Population description

Vietnam is a hub for animal movements in the SEA region, and for that reason, arguably, controlling FMD in Vietnam may be considered prerequisite for controlling FMD in the region. The country is divided into eight major agriculture zones, referred to as Northwest, Northeast, Red River Delta, North Central Coast, South Central Coast, Central Highlands, Southeast, and Mekong River Delta. In addition, there are 63 provinces in Vietnam. Vietnam has had an FMD control program in place since 2006, and at present, Vietnam is in stage 3 of the OIE/FAO progressive control pathway (PCP). Biannual vaccination is conducted free of charge in border provinces and for a fee the other areas of the country (Lee et al., 2020). 85% of livestock farms in Vietnam are small-scale farms (Pham-Thanh et al., 2020). Pig production supplies 80% of the local meat consumption and 74-80% of the total meat production in Vietnam. Although beef production stands at 8% of total livestock production, it is more evenly distributed across the country (The World Bank's Agriculture and Environment and Natural Resources Global Practices).

3.2.2 Data

In this study, we utilize two types of data available from Vietnam: a) data on the reported number of FMD cases per province per year, and b) FMDV VP1 sequences generated from various research and surveillance projects conducted in Vietnam (de Carvalho Ferreira et al., 2017,Di Nardo et al., 2014,Le et al., 2016,Van Diep et al., 2020,Brito et al., 2017,Knowles et al., 2005, Arzt et al., 2017). Because the majority of reported outbreaks and sequences collected were from bovine species (71% of sequences, and 72% of reported clinical cases), we focused our space-time regression analysis on outbreaks reported in cattle and buffalo.

Reported numbers of clinically-infected cattle, buffalo, pigs, and goats, along with estimated outbreak dates, were available from 2007-2017 from the Ministry of Agriculture and Rural Development MARD, Vietnam. Data are collected via passive surveillance at the commune level on a daily/weekly basis at the local sub-Department of Animal Health level. Most cases were clinically diagnosed while some cases were laboratory-confirmed. For cattle, 49306 cases were reported from 1677 outbreaks (2007 -2017) with a mean of 29 (SD 59) infected animals per outbreak. For buffaloes, 70118 cases were reported from 1841 outbreaks (2007-2017) with a mean of 38 (SD 69) infected animals per outbreak. Province-level livestock population data (cattle, buffalo, pigs) for the year 2018 were also available from the General Statistics Office of Vietnam. These data were used to calculate standardized incidence ratios, and to construct space-time regressions (see details below). A polygon shapefile of Vietnam provinces was used

to generate a 0/1 matrix that summarizes which province are adjacent to which (hereafter, referred to as the *spatial adjacency matrix*).

Serotype O sequence data for Vietnam was available from active surveillance of clinical and sub-clinical bovines conducted by our team and from GenBank. In total, 267 FMDV VP1 sequences were available from 53 states representing all eight agriculture zones. 192 sequences collected from farms and slaughterhouses by our collaborative team (see below for details). 113 sequences were downloaded from GenBank. All Vietnam sequences used in the analysis included location (to the state-level) and date meta-data. After removing identical sequences (from the same animal, and same location), a total of 306 sequences were used from Vietnam. All available sequences with date information from adjacent countries during the period of 2000-2017 were also obtained from the GenBank: Thailand (41), Malaysia (37), Laos (32), Cambodia(4), and China(19). For Vietnam, 80 sequences were from pigs, 132 sequences from cattle and 28 sequences from buffaloes. These data were used in Bayesian phylogeographic models (see details below) to infer rates of viral movement between different agricultural zones, which were then used to generate a 0/1 matrix that summarizes which provinces show evidence of population connectivity based on patterns of historical viral dispersal (hereafter, referred to as the *phylogeographic matrix***).**

3.2.3 Sequence acquisition

Oropharyngeal fluid (collected by our team from subclinical animals at farms and slaughterhouses) and epithelium (outbreak) samples were screened for the presence of FMDV using virus isolation (VI), followed by detection of viral RNA in VI supernatant using qRT-PCR as previously described (Stenfeldt et al., 2016,Pacheco et al., 2010). Samples that were positive for viral RNA were subjected to sequencing using one of several methods. Samples from 2013-2015 were sequenced using the Sanger method as previously described (de Carvalho Ferreira et al., 2017) to obtain VP1 sequences, or by next generation sequencing (NGS) to obtain full open reading frame (ORF) sequences. For NGS sequences, overlapping RT-PCR amplicons covering the full ORF were produced using three sets of primers (Brito et al., 2017), and amplicons were sequenced as previously described (Bertram et al., 2019). Samples from 2016-2017 were sequenced by NGS of RT-PCR amplicons covering the P1 region (Xu et al., 2013) as previously described (Bertram et al., 2019). Finally, sequences from 2018-2019 were sequenced by NGS using random and FMDV-specific primers to obtain the complete genome as previously described (Palinski et al., 2019,Bertram et al., 2019). All NGS sequencing was performed using the Illumina NextSeq platform. Read quality filtering, de novo assembly, and assembly to previously published references of regionally endemic lineages were implemented in CLC Genomics Workbench v12

(Qiagen). Sequences of the VP1 region were utilized in this study, as this region was the segment available consistently across years and in GenBank.

3.2.4 Phylodynamic analyses

A total of 397 VP1 sequences from Vietnam and neighboring countries were aligned against the reference strain O/LAO/2/2006 (representing O/PanAsia) using Muscle in MEGA-X. Sequences were checked for recombination using RDP4, and no recombinants were detected. Initial maximum likelihood trees revealed three distinct clades corresponding to lineages O/SEA/Mya-98, O/ME-SA/PanAsia, and Cathay, with the majority of sequences belonging to O/SEA/Mya-98 and O/ME-SA/PanAsia. From the total sequences, 146 sequences were classified as Mya-98, 229 as PanAsia, and 22 sequences as Cathay.

Discrete-space phylogeographic analyses were performed for each lineage separately (O/ME-SA/PanAsia and O/SEA/Mya-98) and for all sequences combined using Bayesian Evolutionary Analysis by Sampling Trees (BEAST v1.10.1.2). A lineage-specific model was not constructed for Cathay due to insufficient data. For these analyses, each agriculture zone within Vietnam (eight zones) and each neighboring country (four countries) were used as discrete traits. Agricultural zones rather than provinces were used within Vietnam in order to achieve adequate numbers of sequences per location for the analysis. Since our intent for this analysis was to identify patterns of FMDV movement within and between countries, all available sequences regardless of host species were included in this analysis.

Maximum likelihood trees of each lineage and of the combined data were analyzed with Tempest (v1.5.3) to establish whether the phylogenies had temporal signals for subsequent analyses, and to root-to-tip regression of genetic distance with the sampling time was used to determine an approximate age of the phylogenetic trees. The HKY nucleotide substitution model was selected as the best nucleotide substitution model using jmodel test (Darriba et al., 2012). The models were run with a lognormal uncorrelated relaxed molecular clock and the Bayesian Skygrid population model. To infer transition rates between discrete locations, we used an asymmetric model with Bayesian Stochastic Variable Selection (BSSVS) to identify non-zero rates of transition in the phylogeographic matrix (Britto et al.,2018). To reconstruct the evolutionary history, two replicate MCMC simulations, with a length of 300 million iterations and sampled every 3000 steps, were performed with computational resources in CIPRES (https://www.phylo.org). Convergence of parameter posteriors were assessed using Tracer 1.6, ensuring effective sample size (ESS) greater than 200 for all parameters. Tree annotator v10.2.2 was used to create maximum clade credibility (MCC) trees, after discarding 10% of iterations as burn-in and using LogCombiner to re-sample to a lower frequency (every 6000 steps) to reduce computational

complexity. Fig Tree (v1.4.3) and ggtree (Yu et al., 2017) were used to visualize the trees, with branches colored according to the country and regions of Vietnam. For discrete traits, Spread3 (v0.9.7) was used to calculate the Bayes factor after disregarding 10% burn-in.

Calculation of adjusted transition rates between agricultural zones

When BSSVS is used to infer transition rates between locations, BEAST MCMC chain output contains two pieces of information regarding transitions between locations: a) the rate coefficient, which is the estimated frequency of transmission between two locations, and b) a 0/1 indicator of whether that transition is included in the model as a particular MCMC iteration. Importantly, when the indicator equals 0, then the rate coefficient is of little meaning, given that it was not included in the model. We processed the raw BEAST output in R such that summaries of posterior of rate coefficients (mean and 95%HPD intervals, summarized by Tracer) were based only on MCMC iterations where the indicator equaled 1, which we refer to as the *adjusted rate*.

Bayes factors were also calculated for all transition rates, with BF>3 interpreted as rates that were non-zero transmissions. However, BF should be interpreted with caution as a measure of population connectivity, as high BF tended to be achieved when there was a clear ancestral location for a discrete transition rate, as might be the case for a rare viral dispersal event between different countries. However, the ancestral location may be difficult to determine in situations where there is frequent, bi-directional transmission between locations (yielding high adjusted transition rates, but low BF support on the ancestral location). For this reason, we utilize adjusted transition rates rather than BFs as a proxy for population connectivity in subsequent analyses. Heatmaps and networks were used to visualize viral movement between discrete locations (based on the adjusted rate), created separately for each lineage and the combined analysis. The adjusted rates were used to create the 0/1 phylogeographic matrices (one each for O/ME-SA/PanAsia,O/SEA/ Mya-98, and combined), thresholded at the median. As the optimal threshold value is unknown, an alternative cut-off value (set at 80% of the median) was additionally investigated. These matrices were used to describe the population connectivity between zones within Vietnam in the space-time regression of reported case counts.

Analysis of interspecific transmission

While our model focused on outbreaks in bovines, we conducted a phylodynamic analysis (as described above) using host-species as a discrete trait to better understand the role of different host species. cattle were identified as transmitting FMDV to pigs and buffaloes. The analysis used 240 sequences from Vietnam (160 from bovines, 80 from pig), and consisted of 7 Cathay , 63 Mya 98 and 169 Pan Asia lineage sequences.

3.2.5 Bayesian space-time risk models

Given that bovine numbers accounted for a majority of reported FMD cases and sequences available, we focus on FMD in bovines for the space-time regression. We first calculated standardized incidence ratios (SIRs) per province per year from 2007 to 2017. Bovine population size data was available for the year 2018 only, and we assumed that these numbers were reasonable stable across the time period assessed here. Population size and nationwide FMD reported case counts were used to calculate the expected number of FMD cases per province per year (e_{it}) if the distribution of FMD cases across space and time was proportional to population size, such that

$$e_{it} = P_{it} \frac{\sum_{it} y_{it}}{\sum_{it} P_{it}}$$

Where P_{it} is the population of the state *i* in year *t*, and y_{it} is the number of FMD cases in province *i* in year *t.* SIR was defined as the observed to the expected ratio (Y_{it}/e_{it}). SIRs were plotted as choropleth maps for all years.

The observed number of cases per province per year was assumed to follow a Poisson distribution $y_{it} \sim Poisson\ (e_{it}, \theta_{it})$, with e_{it} representing the expected number of FMD cases defined as above, and θ_{it} representing the yearly relative risk for each state. This relative risk incorporates both spatially structured (spatial correlation amongst connected provinces) and unstructured (i.e., random variation) effects, such that:

$$\log(\theta_i) = \alpha + \upsilon_s + \nu_s$$

Where α is the intercept representing the overall level of risk in the country, υ_s is the structured spatial effect, and ν_s is the unstructured spatial effect that functions as a random effect for each province. This model is known as the BYM2 model (Riebler et al., 2016). The structured spatial effect incorporates correlations amongst neighboring provinces, conventionally represented by 0/1 spatial adjacency matrix. The model that utilized inferred viral movement (instead of spatial adjacency) as a proxy for population connectivity mirrored the above space-time model, except that the structured effect was designed to account for correlations amongst provinces connected phylogeographically (υ_p) rather than spatially (υ_s). The matrix of inferred movement of the virus among the eight agricultural zones inside Vietnam was projected to create a 0/1 matrix for 63 provinces inside Vietnam (i.e., all provinces within two zones received the same phylogeographic transition rate). Six phylogeographic matrices were considered (PanAsia, Mya-98, combined, each with two cut-off values for dichotomization of adjusted rates). We also explored whether incorporating both spatial and phylogeographic adjacency matrices simultaneous in the same model would improve model fit.

Several model structures exist to incorporate temporal effect: time (year) can be considered as a random effect (ω_t), or a structured effect (γ_t) in which a random walk is used to account for between-year dependencies ($\omega_t + \gamma_t$).

The best-fit model structure was selected from amongst these models using DIC. This model was then used to evaluate the contribution of hypothesized risk factors in shaping relative risk. Penalized priors were used for all models in the Bayesian analysis, following previous studies (Fuglstad et al., 2019).

Incorporation of risk factors included as fixed effects

To further explain spatiotemporal variation in relative risk, we incorporated factors potentially associated with viral movement as fixed effects into the best-fit spatial and phylogeographic models from above. These included two approaches for accounting for transboundary introductions. The first categorized provinces as to whether they had an international border (0/1). Second, we used the inferred adjusted transition rates from the phylogeographic models to categorize provinces according to whether there was evidence of FMDV introductions from neighboring countries. As before, adjusted transition rates were dichotomized at the median to create one dummy variable each for phylogeographic-inferred transition rates from Cambodia, Malaysia, Laos, China, and Thailand into Vietnam provinces. Given that slaughterhouses are terminal points of supply chains and thus may influence animal movements, the presence of a slaughterhouse in a province was also included as a potential 0/1 risk factor in the model. Slaughterhouses considered here are slaughterhouses that operate under veterinary control and commercial fresh meat establishments registered for export by the national veterinary services. Finally, to account for presence of other hosts that can transmit FMDV, such as goats and pig (Arzt et al., 2011), goat and pig densities were included as potential risk factors.

Prior to model selection, correlations between variables were checked using Pearson's correlation coefficient for continuous variables and chi-square tests for categorical variables. In the latter case, a odds ratio of >8 was considered evidence for collinearity (Dohoo et al., 2009). In the case of collinearity, only the variable with the lowest DIC in its respective univariable model was retained for the multivariable analysis. Backward selection was then performed from a full multivariable model by removing the variables with the widest confidence interval that overlapped zero. From among those different models, the simplest model that was <2 ΔDIC from the model with the lowest DIC value was considered the best-fit model.

Prior sensitivity analysis and evaluation of model fit

We used non-informative penalized complexity priors, which are applicable for a large class of hierarchical models. The penalized priors aim to incorporate minimal information into the inference procedure, and can account overdispersion in the base model (Simpson et al., 2014). A sensitivity analysis was conducted using the different priors**.** We also calculated the correlation between the predicted and observed values (Spearman's correlation) as a measure of model fit.

The final best-fit model was also evaluated based on posterior predictive p-values. Posterior predictive p -values are defined as $p(yi^* \leq y_i|y)$, where y_i^* is the posterior of the predicted distribution from the model. This is interpreted as an approximation of the proportion of the predicted distribution for y_i that is more extreme than the observed value, and values of $p (yi^* \leq yi|y)$ near 0 and 1 indicate poor model fit (Blangiardo and Cameletti, 2015). To better interpret correlations in case counts across different areas, we tabulated bovine case counts per agricultural zone from 2007 to 2017 and created a heatmap showing the Pearson's correlation in case counts between the eight different agriculture zones across time.

3.2.6 Software for space-time regression

All analyses were performed in the R statistical software, using packages tidyverse 1.2.123, spdep 0.7–425 (Bivand et al.,2015), dplyr, stringr, library(beastio) library(boa) library(viridis) library(ggpubr) library(dplyr) library(readr) library(igraph) and ggplot2. For the Bayesian risk models, INLA 19.09.03 (Rue et al., 2017) was used, and model results were processed with INLAOutputs 19.09.03 (Baquero et al., 2018).

3.3 Results

3.3.1 Phylogeographic analyses

We conducted a discrete-space phylogeographic analysis of FMDV serotype O in Vietnam, with separate analyses performed for the 229 O/ME-SA/Pan Asia sequences and 146 O/SEA/Mya-98 sequences, and for the 397 combined serotype O sequences. Across all three phylogeographic models, the mean substitution rate was 0.0069 substitutions/site/year (95% HPD interval: 0.0056-0.0083) for O/ME-SA/Pan Asia, 0.0051 (0.0038-0.0063) for O/SEA/Mya-98, and 0.0062 (0.0053-0.0071) for the combined serotype O analysis. We focus on presenting results from the Pan Asia phylogeographic model, given that inferred viral movements from this model provided the best-fit to the outbreak risk model (see below), but corresponding results for the Mya-98 and combined phylogeographic models are shown in Supplement S3.1 and S3.2. The MCC tree created based on Pan Asia sequences is displayed in Figure 3.1. The heatmap depicting adjusted transition rates between countries and agricultural zones is shown in Figure 3.2A. Of note, there was substantial movement inferred between adjacent zones, as well as more distant zones, such as between the South Central Coast and Southeast to the Northeast. Values were dichotomized at the median adjusted rate (0.575 for Pan Asia, 0.666 for Mya98, and 0.409 for combined sequences), and used to create the 0/1 matrix utilized in the outbreak risk model. A network representation of this matrix is shown in Figure 3.2B, revealing a combination of local and long-distance connections between zones.

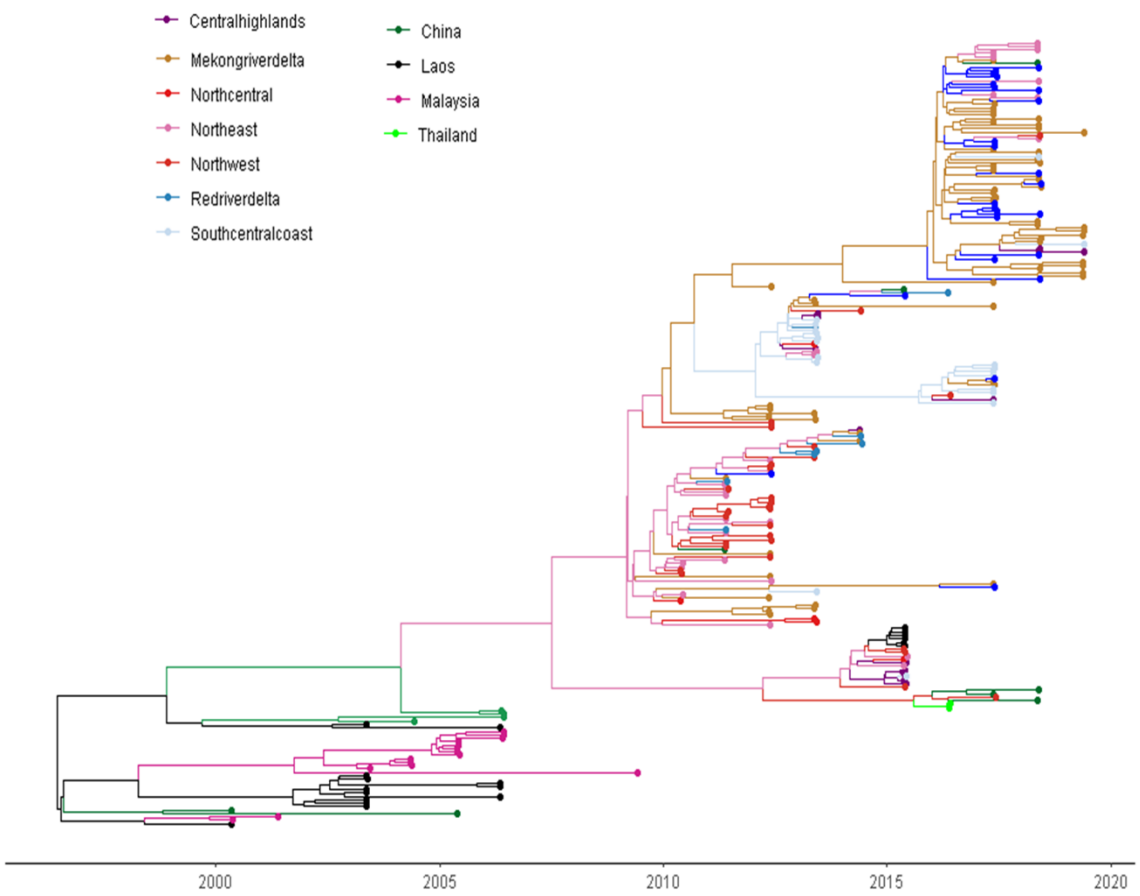
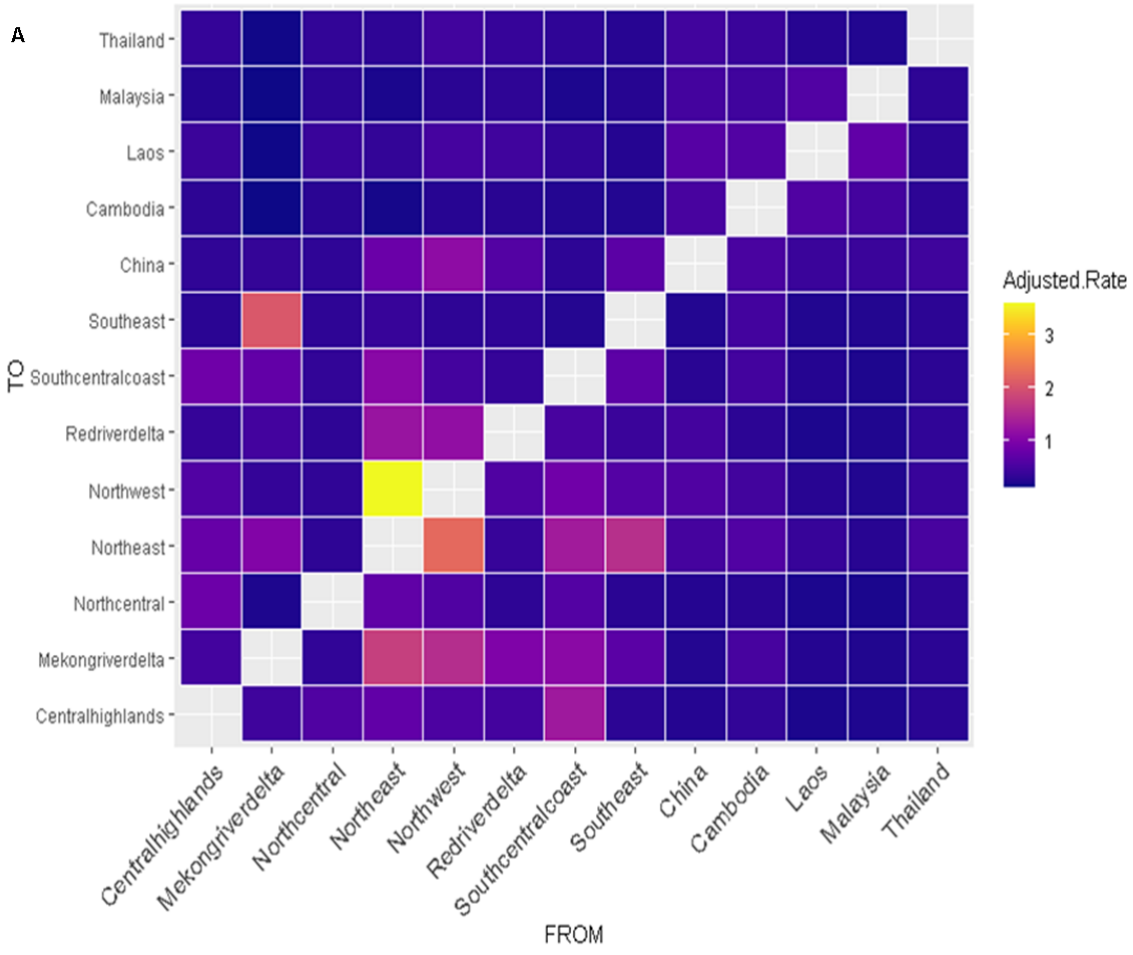
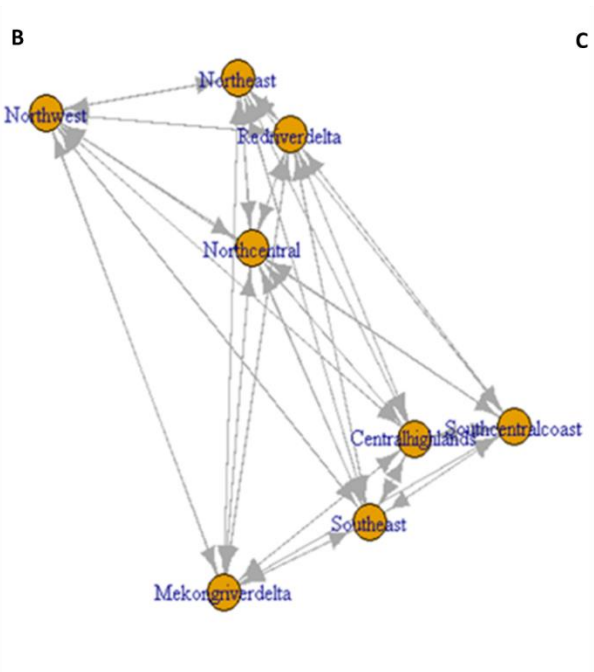


Figure 3.1:Maximum clade credibility FMDV O/ME-SA/PanAsia lineage from agricultural zones of Vietnam and surrounding countries China, Laos, Malaysia, Thailand. Nodes and branches of the tree are colored by location.


A
Thailand
Malaysia
Laos
Cambodia
China
Southeast
Southcentralcoast
Redriverdelta
Northwest
Northeast
Northcentral
Mekongriverdelta
Centralhighlands
TO
Centralhighlands
Mekongriverdelta
Northcentral
Northeast
Northwest
Redriverdelta
Southcentralcoast
Southeast
China
Cambodia
Laos
Malaysia
Thailand
FROM
Adjusted.Rate
3
2
1


B
Northeast
Northwest
Redriverdelta
Northcentral
Centralhighlands
Southcentralcoast
Southeast
Mekongriverdelta


C
N
Northeast
Northwest
Red River Delta
North Central Coast
South
Central
Coast
Central Highlands
Southeast
Mekong
River Delta
0
100
200
400 Kilometers

Figure 3.2: A) The adjusted rate matrix for Pan Asia sequences lineage from agricultural zones of Vietnam and surrounding countries China, Laos, Malaysia, Thailand. B) Network of phylogeographically connected zones inside Vietnam. C) Map of agricultural zones in Vietnam.

While the majority of sequence and reported outbreak data came from bovines, we performed an additional discrete-trait analysis for host species to further explore the potential role of interspecific transmission from pigs in Vietnam. This model showed little evidence that viruses circulating in pigs were transmitted into bovine populations in Vietnam (Bayes factor =0.511), though there was evidence of viral transmission from bovines to pigs (Bayes factor = 77512.17). This result provided additional rationale supporting our decision to focus subsequent space-time risk analyses on reported case counts in bovines.

3.3.2 Space time risk model

The SIR values calculated from 2007 to 2017 are shown in Figure 3.3. Provinces with SIR greater than one can be interpreted as areas with more reported FMD cases than expected given the size of their bovine population.

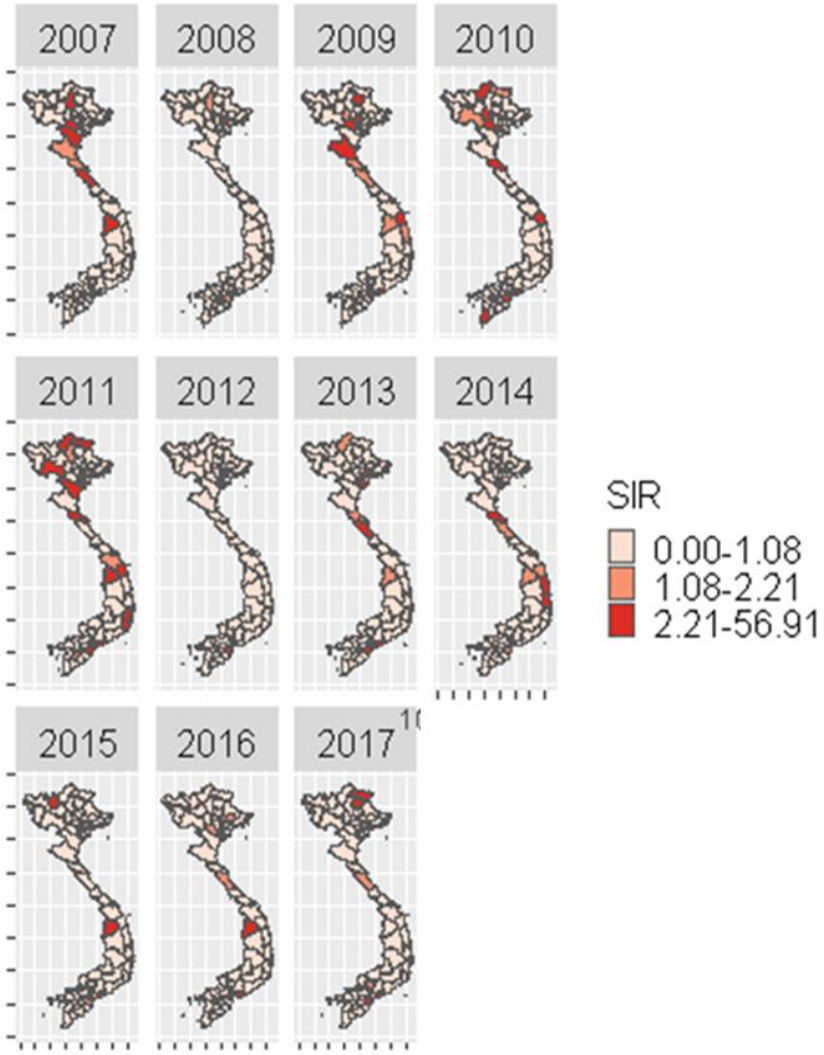


Figure 3.3: Map of annual standardized incidence ratios (SIR) for years 2007-2017 for the different provinces in Vietnam.

Several different models were tested to select the best model structure for the space-time regression (Table 3.1), including different combinations of adjacency matrices (spatial, phylogeographic, or both) and temporal effects. The model with the lowest DIC incorporated time as a random walk and random effect and utilized the Pan Asia phylogeographic matrix (with the threshold for the phylogeographic matrix set at the median adjusted rate, posterior p-values = (79.9,19.2), Table 3.1). This model performed better than the risk models based on Mya-98 or the combined serotype O sequences. Models that utilized the spatial adjacency matrix alone or phylogeographic and spatial matrices in combination produced the higher DIC values.

The best-fit model structures for the conventional space-time and phylogeogrpahic risk models were used as the base of multi-variable models that included additional fixed effects associated

with animal movement (i.e., presence of slaughterhouses, presence of an international border, evidence of phylogeographic links with neighboring countries, and host densities). Fixed variables were screened for correlations using Pearson's correlation coefficients and Chi-Square tests (supplementary table S3.1 and S3.2). If two variables were found to be correlated, we excluded the variable that resulted in higher DICs when assessed in univariable models. For the space-time risk model, viral movement from Laos and China were found to be correlated, and thus only viral movement from Malaysia, Cambodia, Thailand, having an international border, having a slaughterhouse, goat and pig density were retained in the multivariable model. For the phylogeographic risk model, viral movement from Cambodia and Laos were found to be correlated, and thus only viral movement from Malaysia, China, Thailand, having an international border, having a slaughterhouse, goat and pig density were retained in the final model. The best-fit multivariable models are shown in Table 3.2 and 3.3. Both the spatial and phylogeographic risk models identified having a slaughterhouse and having an international border as significant (95% credible intervals do not overlap 0). In the spatial risk model, the inclusion of additional covariates did result in lower DIC, suggesting that these factors explain some variability in case counts. However, the credible intervals for the additional included factors overlapped 0, though the posterior of inferred viral movements from Cambodia only slightly overlapped 0. In the phylogeographic risk model, additional significant factors included pig density and inferred viral movement from China and Malaysia. Overall, the phygeographic multivariable risk model had a lower DIC than the space-time model. High-risk areas identified from both best-fit models are shown in the supplementary figure S3.3 and S3.4. Several areas in north, south, and central regions of the country were identified as high-risk areas through time. A heatmap was used to visualize correlations in case counts between different zones across time (Figure 3.4). Of note, case counts in geographically distant zones, such as the Northeast and the South-Central Coast, are sometimes more correlated with each other than closer regions (Figure 3.4).

Table 3.1: Comparison of different model structures utilized for Bayesian space-time regressions of reported case counts. Phylogenetic matrices were tested with two different cutoffs to dichotomize the phylogeographic adjusted rate matrices. The best-fit structure is marked in bold.

Model group	Temporal effect	Adjacency matrix	Lineage	Cutoff	DIC
Spatial	iid only	spatial	NA	NA	113457.7
	iid and random walk	spatial	NA	NA	112329.1
Phylogeographic	iid and random walk	Phylo	Mya 98	0.67	110658.8

	iid and random walk	Phylo	Mya 98	0.55	110658.8
	iid and random walk	**Phylo**	**PanAsia**	**0.60**	**110607.7**
	iid and random walk	Phylo	PanAsia	0.47	110623.9
	iid and random walk	Phylo	Total sequences	0.41	110622.9
	iid and random walk	Phylo	Total sequences	0.37	110622.9
Phylo and spatial	iid and random walk	phylo and spatial	Mya 98	0.67	112450.0
	iid and random walk	phylo and spatial	Mya 98	0.55	112450.0
	iid and random walk	phylo and spatial	PanAsia	0.60	112710.0
	iid and random walk	phylo and spatial	PanAsia	0.47	112464.6
	iid and random walk	phylo and spatial	Total sequences	0.41	112631.6
	iid and random walk	phylo and spatial	Total sequences	0.37	112631.6
Phylo and spatial (joint)	iid and random walk	phylo and spatial joint	Pan Asia	NA	110611.3

Table 3.2: Results from the final Bayesian space-time model. (DIC 112300).

Fixed effect	**Coefficient (95% Credible Interval)**
Intercept	-2.56 (-3.49, -1.65)

Slaughterhouse	-4.73 (-7.09 , -2.55)
International border	0.82 (0.18, 1.48)
Cambodia	1.78 (-0.62, 4.22)

Table 3.3: Results from the final Bayesian phylogeographic risk model. (DIC 109366.1).

Fixed effect	**Coefficient 95% Credible Interval**
Intercept	-1.35 (-1.53, -1.35)
Pig density	-0.58 (-0.63, -0.53)
Goat density	0.02 (-0.03, 0.06)
Slaughterhouse	-2.70 (-2.88, -2.51)
International border	0.38 (0.34, 0.42)
China	-0.57 (-0.68, -0.49)
Malaysia	0.79 (0.64, 0.94)

3.3.3 Model validation

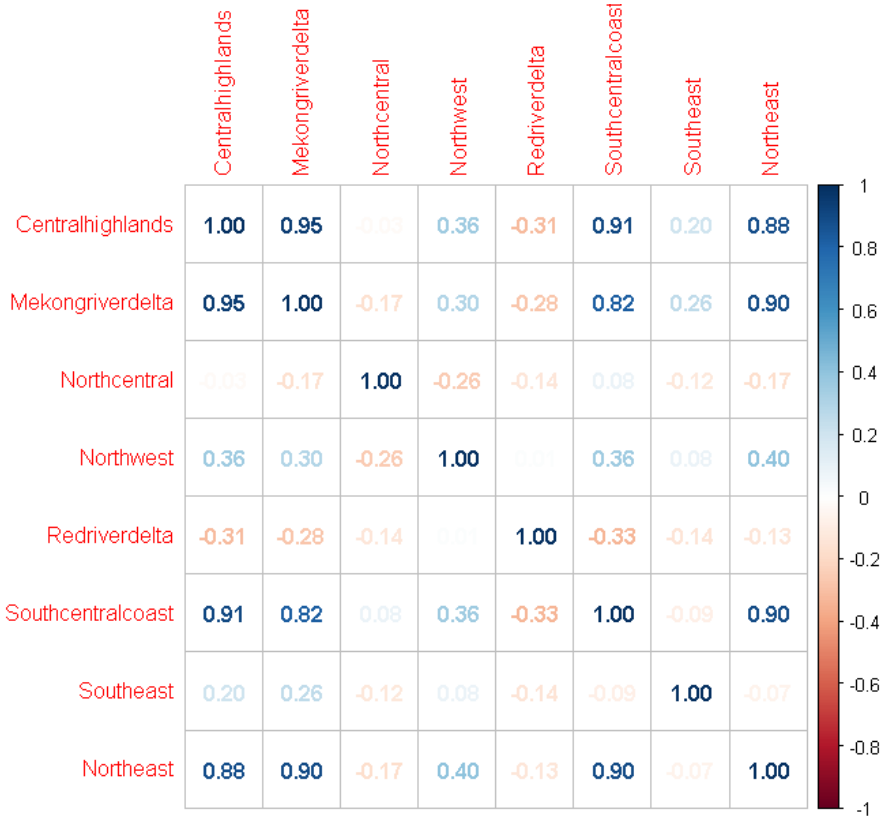


Figure 3.4: Spearman's correlation matrix of reported case counts across agricultural zones from 2007-2017.

Results from the prior sensitivity analysis are included in the supplementary show that results were consistent across different penalized complexity priors. Observed SIR to fitted relative risk values were compared, and Spearman's correlation coefficients were highly significant (p =0.9). For the selected models, the posterior p-values indicated good fits.

3.4 Discussion

Here, we demonstrated that accounting for population connectivity through phylogeographic rate matrices explains more variability in regressions of the risk of reported FMD cases than spatial adjacency amongst provinces. While population connectivity is fundamental to spread of infectious diseases, this finding suggests that it is important for space-time risk models to account for potential long-distance connections amongst different host meta-populations. In the absence of data on host movement, such long-distance connectivity can be inferred from viral movements that are potentially associated with patterns of host movement in the country. Such an approach could be adapted for risk modeling in other host-pathogen systems where pathogen sequence data is more readily available than host movement data. In addition, provinces with high relative

risk of bovine outbreaks were characterized as those with international borders, no bovine slaughterhouses, higher rates of transboundary viral introductions from neighboring countries (particularly Malaysia and Cambodia to a lesser extent), and lower pig densities. We also identified high-risk areas in the northern and southern part of Vietnam where it may be important to implement enhanced surveillance and control measures.

In our analysis, the phylogeographic connectivity network inferred from the sequence data was best able to explain spatial variation in reported clinical case counts, and hence may better capture population connectivity amongst regions than spatial adjacency alone. Similarly, pathogen sequence data from bovine tuberculosis was utilized to infer patterns of host movement in Cameroon. In this case, they had an observed cattle movement network to which the molecular-based network could be compared. They found that the molecular network was a much better approximation of the observed host movements than other methods commonly used in the absence of movement data, such as gravity models (Muwonge et al., 2021). In the United States, patterns of spatial diffusion of porcine reproductive and respiratory syndrome virus were inferred via discrete-space phylogeographic models, and variability in the rate of sector-to-sector diffusion was shown to be associated with the movement of feeder pigs (Makau et al., 2021). Taken together, these studies help validate the use of pathogen molecular data as a proxy for host movement at relatively fine scales.

Because there was uncertainty about how to quantify the phylogeographic data, we tested multiple formulations of the phylogeographic matrix based on different lineages and variable rate thresholds. We tested the model for the two most prevalent lineages separately and combined serotype O sequences, as well as two cut-off values to infer connectivity among regions. We found that the PanAsia phylogeographic matrix provided the best fit in the space-time regression of reported clinical cases. This may be unsurprising given that most detected sequences circulating in Vietnam belonged to this lineage during this period. Furthermore, previous analysis of the phylogeography of PanAsia lineage in Vietnam from 2010-2014 showed that the FMDV circulated throughout the country, with a special emphasize on north and south regions (Brito et al., 2017). In this study, we identified high-risk areas (Supplementary figure S3.3 and S3.4) in northern and southern parts of the country. These align well with spatial hotspots of FMD outbreaks have been previously identified in the Northwest, Northeast, and Red River Delta areas by a Satscan analysis (Lee et al., 2020).

Brito et.al (2017) also demonstrated FMDV movement from the Northeast region of Vietnam to China and Kazakhstan during the period of 2010-2014. We have included older FMDV sequences available from adjacent countries before 2010 (Figure 3.1), showing the introduction

of O/ME-SA/PanAsia into Vietnam from other SEA countries. Unfortunately, more recent FMDV sequences from adjacent countries (after 2010) were not available from GenBank, which prevented us from further examining bidirectional transmission patterns in more recent years. In our outbreak risk model, we found a negative effect of inferred movement from China when incorporated as a fixed effect. Thus, we should be cautious interpreting results from phylogeographic trait analyses. We may hypothesize that there is a positive association between inferred movements and risk, but the directionality of the inferred movement is likely important. Similar to Brito et al. (2017), we found that Vietnam was more likely to be the origin of viral movements into China, rather than the recipient of viral introductions from China, though these results could be affected by minimal data availability from China. This may contribute to the negative effect of viral movements with China in our risk model. Both bovines and pigs are moved north to south in Vietnam and from the north into China (Polly et al., 2015). Animal movements into China are driven by its growing economy increased demand for meat-based diet. From a study done in March to August in 2014 in Vietnam using network analysis , it was identified that animals from the central areas of the country (North Central coast, Central Highlands, and South-Central coast) would move in both directions between the north and south. Initially, animals are transported from surrounding Laos, Thailand and Cambodia to these areas). Cao Bang Province in northern Vietnam was an area where animals are held before crossing to China (Polly et al., 2015), and was also identified as a high-risk area in our study. Nghe An Province, where animal pass to both China and Hanoi from Laos, was not identified as a risk area in our study. Provinces with international borders showed higher relative risk of reported cases, a pattern which held in both the space-time and phylo-time regressions. In addition, our previous study also demonstrated that closely related FMDV sequences can be identified from different parts of the country (Chapter 3), further suggesting epidemiological linkages between distant provinces. For example, one cluster of closely related sequences isolated from clinical FMD outbreaks in northern Vietnam was first detected in slaughterhouses in the south (Gunasekara et al., 2021). This is also demonstrated by correlations in case counts between in geographically distant areas, such as the Northeast and the South-Central Coast, suggesting that distant regions may experience synchronized outbreaks (Figure 3.4). Such correlations amongst distant regions would not be captured by spatial adjacency and are better captured by phylogeographic connectivity, as we observe by the high adjusted rates for Northeast and the South Central Coast in the phylogeographic analysis (Figure 3.2a). On par with the outbreak matrix, most of the virus movement occurred between North West and South-Central Coast. This further indicates that some provinces that are not geographically connected may still be connected due to animal movement-related activities.

Adding additional fixed effects reduced the DIC of both models, indicating that fixed effects were able to explain some observed variability in both models. More specifically, having a connection with neighboring countries Malaysia would have a positive impact on FMD outbreak. Even pig density and bovine density are not highly correlated (ρ 0.45), pig population in Vietnam is higher. In our analysis pig density has a negative effect for occurrence of FMD outbreaks in bovines for this reason. Since animals are moved to slaughterhouses from other areas, we would have expected to have a positive impact from slaughterhouses to the occurrence of FMD outbreaks. Yet, presence of a slaughterhouse in a province had a negative impact on relative risk of clinical cases. This could be a result of the definition of slaughterhouses utilized by OIE PVS analysis, where we retrieved data for this variable. Having only a few designated slaughterhouses may be the reason for the negative effect, particularly if designated slaughterhouses are located in peri-urban areas where livestock densities are low. According to Lemke et al 2008, such slaughterhouses are mainly used for cattle, while pigs are slaughtered in non-designated slaughterhouses without veterinary inspection available in all the provinces in the Vietnam.

From discrete trait analysis, it was identified that the virus moves from cattle populations into buffaloes and pigs, but there was little evidence for transmission in the reverse direction. For pigs, it appears from the limited sequence data available that a single transmission event tends to occur that jumps the species barrier (from cattle to pig), followed by circulation and evolution within the pig population (Supplementary figure S3.5) with relatively little evidence that the virus circulating in pigs spills back to cattle. Bovines live longer compared to pigs and can be carriers, which may contribute to bovines playing a more important role in FMDV persistence at the population level. Using a slightly reduced dataset focusing on a shorter time range, Brito et al. (2017) showed that O/ME-SA/PanAsia could move from pigs to cattle in Vietnam, but this finding was not supported in our study that included almost 100 additional sequences. That being said, analysis of different serotypes and lineages could yield different results, and more research with representative sampling is needed to better quantify cross-species transmission rates. Nonetheless, any role of pig outbreaks in perpetuating spread would not have been captured in risk model focused on bovine case counts, though we did find that pig densities were negatively correlated with the relative risk of bovine case counts.

Limitations

Bayesian space-time risk models provide some shrinkage and spatial smoothing of relative risk compared to raw SIR values (Richardson et al., 2004). One limitation of our approach is that we used the number of animals (reported case counts) as our outcome, which may have more over-dispersion than if we utilized the number of reported outbreaks. That being said, the choice of using case counts or outbreak counts both have limitations; case counts may allow a single large

outbreak to have too much influence on the results, whereas outbreak counts would treat large and small outbreaks equally. However, we chose to use reported case counts, as it is unclear what criteria is used to classify a large group of cases into one large outbreak as opposed to several smaller outbreaks. Thus, using the case counts rather than outbreak counts allowed us to avoid this issue. In either case, we used reported numbers of cases and outbreaks, and it is likely that reported numbers are an underestimate of the true disease incidence. For example, there was a high reported number of cases from the years 2011 and 2012 in Cao Bang province. This could be due to a fast-spreading FMD outbreak or because of increased surveillance in the area during the period. This could have impacted our model, for example, by flattening the impact of the temporal effects in other years with the result of identifying similar predicted risk through time.

When creating the phylogenetic adjacency matrix, we used all the available sequences from all host species up to 2017, and patterns observed may change with the inclusion of additional sequence data. Since sequence data were not available from all provinces, agriculture zones were used to infer patterns of viral movement, which was later transferred to the province level. Despite these limitations, this study provides compatible results with other studies regarding animal and FMDV movement in Vietnam. Although cattle and buffalo production accounts for 8% of total livestock production and pig density is far higher than cattle density in Vietnam, most of the sequence and most of the reported outbreaks were from the cattle. This may be because FMD is severely under-reported and subject to limited surveillance activities in small holder pig production. This study may not be representative of what is happening in pig populations. Results from the discrete trait analysis where it is identified that the cattle are the transmitter of the virus to pigs and buffaloes can change with availably of more sequence data from the considered species as our analysis mainly consist of sequences obtained from cattle.

Conclusions

We have used both spatial connectivity and phylogenetic connectivity matrices to determine FMD risk in Vietnam. Due to its ability to account for FMDV movement, we consider using phylogenetic data trait analysis created matrix would better account for unexplained variability for contagious animal diseases as much as spatial adjacency matrix in Bayesian Poisson regression models considering highly mobile host densities. Vietnam is in stage 3 of the PCP pathway where it is necessary to perform zoning based on FMD risk areas in the country with vaccination. High-risk areas that exceed the constant relative risk areas were identified in the north and south part of the country through consecutive years in Kon Tum, Lang Son, Cao Bang, and Bac Lieu provinces in the south. We also identified having international border and transboundary animal movement from some adjacent countries would increase the risk of FMD. To control FMD

outbreaks, incorporating genomic surveillance plans when creating a strategic FMD control plan would support Vietnam to progress in the PCP pathway.

Table S3.1: Variable selection from the univariate analysis.

Variable 1	Variable 2	p value	Chi Square	OR
Malaysia	Cambodia	1.45E-07	27.65	inf
Malaysia	Laos	2.20E-16	366.34	inf
Malaysia	China	1	0	1
Malaysia	Thailand	2.20E-16	283.3	inf
Malaysia	International border	2.20E-16	138.71	NA
Malaysia	Slaughterhouse	2.50E-04	13.36	0.43
Cambodia	Laos	3.71E-13	52.79	inf
Cambodia	China	5.96E-14	56.38	inf
Cambodia	Thailand	2.20E-16	68.5	inf
Cambodia	International border	2.20E-16	82.3	NA
Cambodia	Slaughterhouse	9.60E-04	10.9	0
Laos	China	2.20E-16	70.88	3.85
Laos	Thailand	3.13E-15	62.18	3.53
Laos	International border	2.20E-16	177.12	NA
Laos	Slaughterhouse	0.8142	0.05	0.92
China	Thailand	2.20E-16	92.96	0.2
China	International border	2.20E-16	76.36	NA
China	Slaughterhouse	1.25E-12	50.41	7.97
Thailand	International border	2.20E-16	186.9	NA
Thailand	Slaughterhouse	1.61E-09	3.64E+01	0.22
International border	Slaughterhouse	6.56E-08	36.27	NA

Table S3.2: Results from the univariate analysis.

Variable	Space-time DIC	Phylo-time DIC
Malaysia	112610.45	110374.64
Cambodia	112429.72	110446.79
Laos	112829.54	110547.01
China	112640.54	110164.08
Thailand	112580.43	110257.4
International border	112846.95	110142.43
Slaughterhouse	112382.32	110074.45
Goat density	112758.47	110572.4
Pig density	112537.39	110521.59

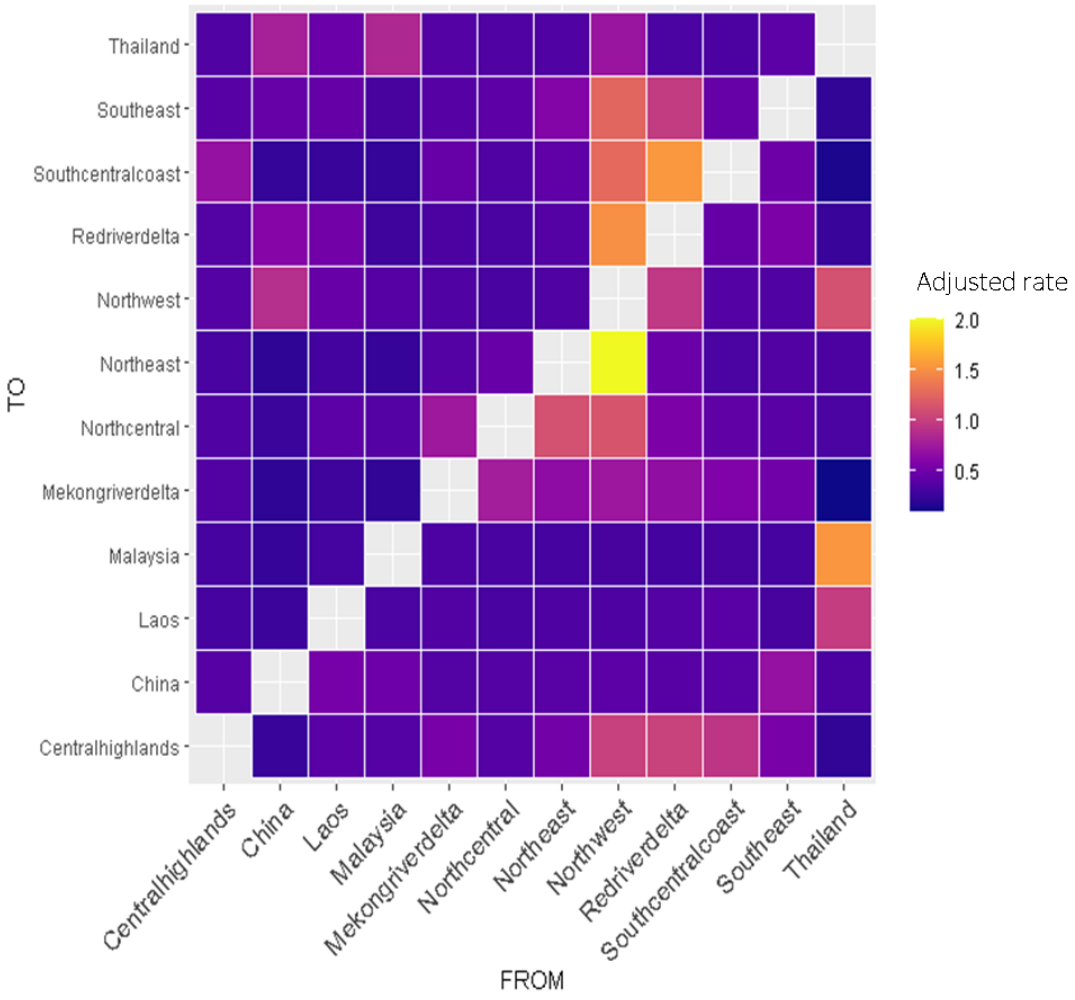


Figure S3.1: The adjusted rate matrix for Mya 98 sequences from agricultural zones of Vietnam and surrounding countries China, Laos, Malaysia, Thailand.

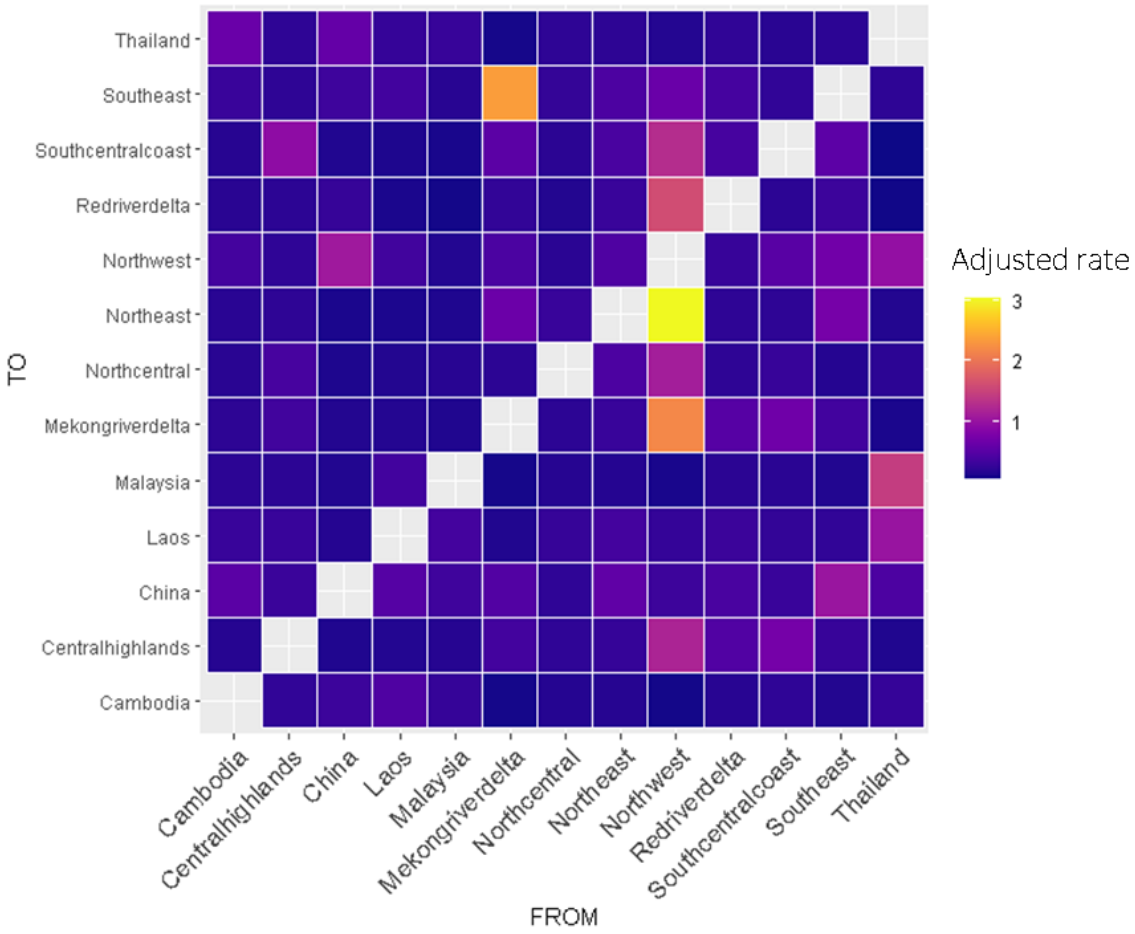


Figure S3.2: The adjusted rate matrix for total sequences from agricultural zones of Vietnam and surrounding countries China, Laos, Malaysia, Thailand.

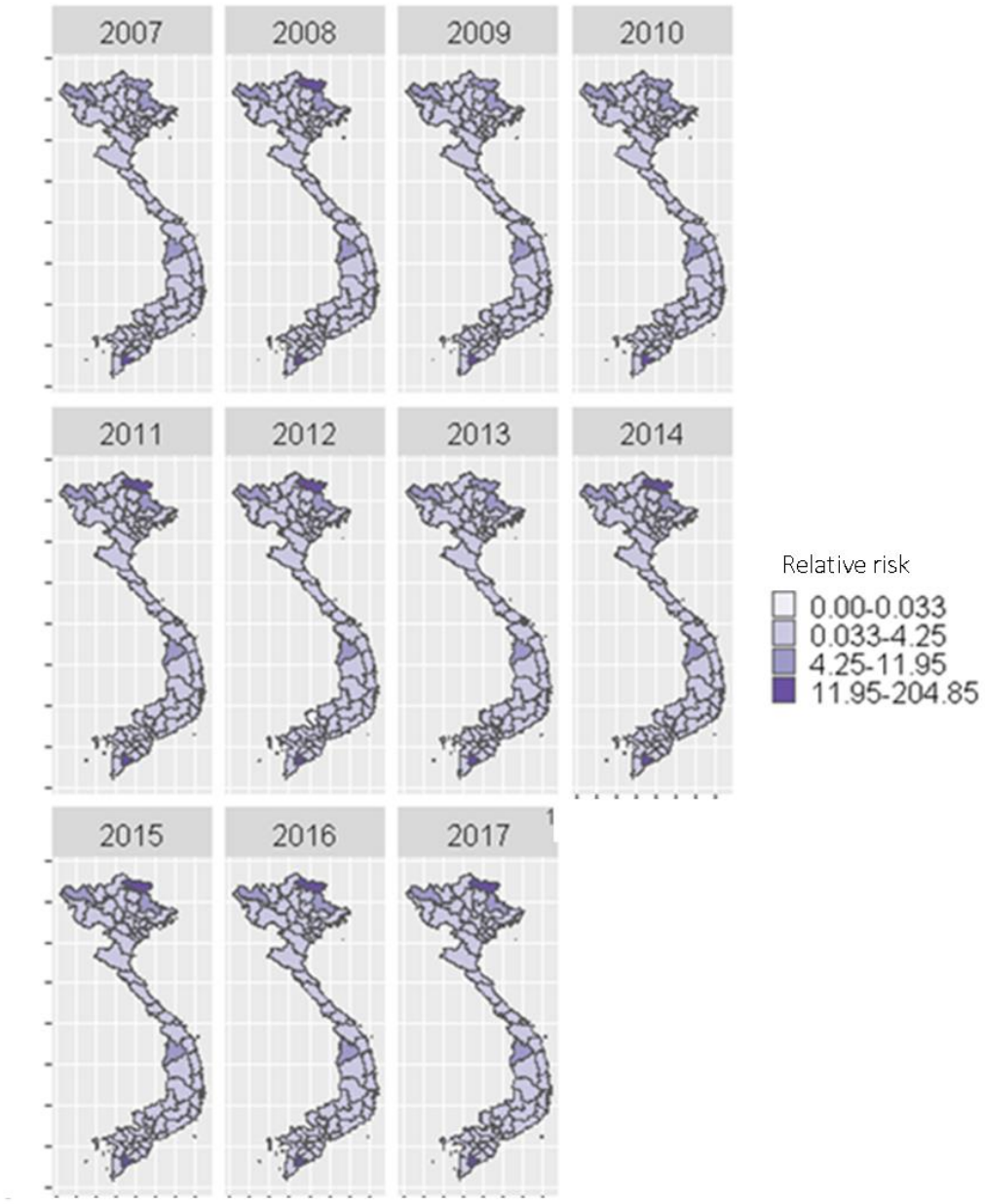


Figure S3.3 : Fitted relative risk of outbreaks for each state from the best-fit phylo time multivariable risk model.

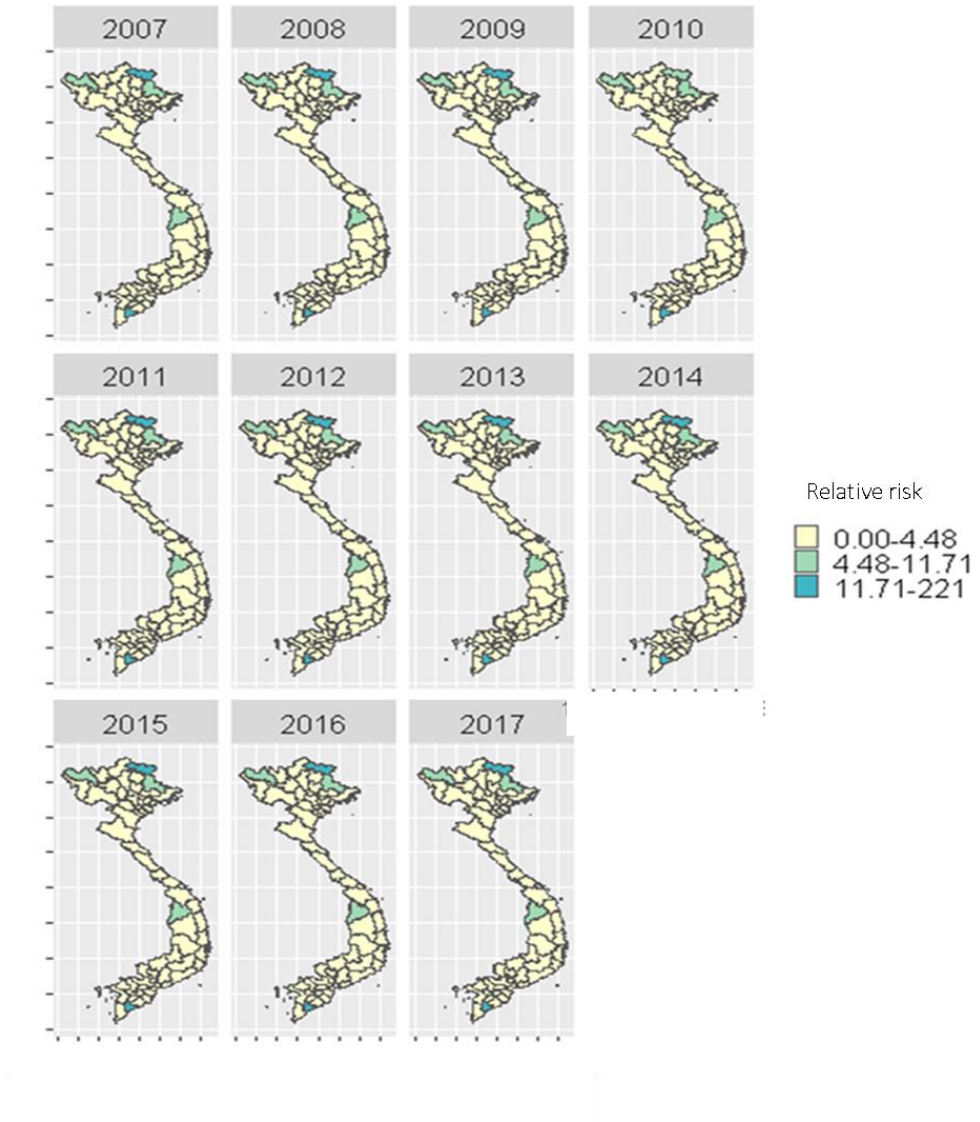


Figure S3.4: Fitted relative risk of outbreaks for each state from the best-fit space time multivariable risk model.

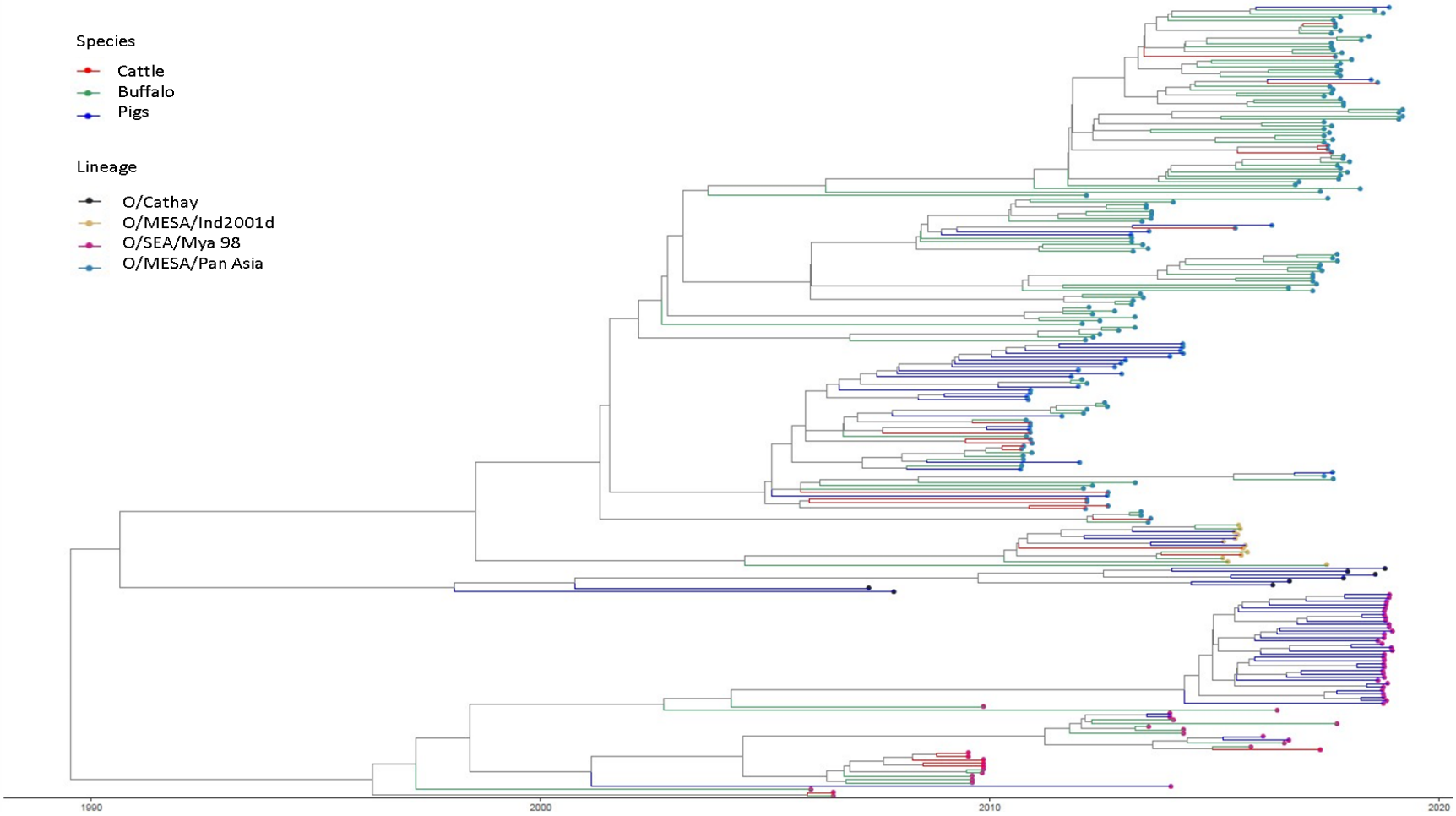


Figure S3.5: Species discrete trait analysis Maximum clade credibility FMDV. Branches related to cattle, buffalo and pigs are shown in green, red and purple colors.

Chapter 4: Slaughterhouses as sentinels for genomic surveillance of subclinical foot-and-mouth disease virus infections in Vietnam

4.1 Introduction

Foot-and-mouth disease (FMD) is a contagious disease affecting cloven-hoofed mammals that causes recurrent outbreaks, subclinical infection, and substantial economic losses in infected regions (Stenfeldt and Arzt, 2020). Foot-and-mouth disease virus (FMDV) is endemic in many developing countries of Asia and Africa, where limited veterinary resources create a need for cost-effective surveillance measures. Surveillance for transboundary animal diseases, such as FMD, typically relies on passive surveillance through outbreak reporting, which sometimes leads to delayed control measures and greater disease spread. Early detection of outbreaks is important to enforce preventive measures and mitigate the impact of the disease, particularly for rapidly evolving RNA viruses, such as FMDV, that have a broad genetic and antigenic diversity. Sampling of animals across the host population to ascertain the prevalence of infection (with or without evidence of clinical signs) is referred to as active sampling, and can be performed in farms, animal markets, or slaughterhouses to provide a more timely indicator of infection prevalence in a population, particularly if coupled with sequencing to detect emerging variants (Cameron et al., 2020,Armson et al., 2020, Thumbi et al., 2019).

Farm-based active surveillance through randomized sampling would be considered the benchmark of understanding the prevalence and distribution of diseases. Various studies have reported farm-based genomic surveillance of subclinical FMDV strains in endemic regions (de Carvalho Ferreira et al., 2017,Bertram et al., 2020,Omondi et al., 2020,Farooq et al., 2018). However, routine farm-based surveillance is often impractical based upon logistical complexity and expense, particularly in rural settings with sub-optimal infrastructure. Slaughterhouses are concentration points where animals from many farms aggregate, and can potentially serve as a convenient, quasi-representative sample of animals from the surrounding host population (Innocent et al., 2017,Willeberg et al., 2018,Arguello et al., 2013). This strategy is employed in veterinary public health to detect diseases or zoonoses of public health concern, such as *Fasciola* or bovine tuberculosis (Innocent et al., 2017,Kao et al., 2018,Willeberg et al., 2018). Slaughterhouse data alone, and in combination with other variables, have also been utilized for determining the risk factors associated with preserving the quality of meat, and evaluating antibiotic usage in farm animals (Peterson et al., 2017,Savin et al., 2020). In most countries, only visual inspections of carcasses are performed in slaughterhouses, though depending on the pathogen, effective disease surveillance can be achieved at slaughterhouses by combining laboratory testing with visual inspection (Fertner et al., 2017,Willeberg et al., 2018). For example, routine slaughterhouse surveillance and laboratory testing to detect emerging diseases is

conducted in the European Union (EFSA and ECDC)(Bonardi, 2017) and the USA (USDA and APHI)(Peterson et al., 2017).

Slaughterhouse-based surveillance is typically passive in nature and is employed for diseases with poor antemortem diagnostic options, and slow-spreading pathogens and parasites that do not require a rapid response; hence it is rarely used for rapidly spreading diseases such as FMD. However, there is substantial and often sub-clinical spread of FMD in endemic countries (Stenfeldt and Arzt, 2020) that is not captured by passive surveillance of reported outbreaks. Active surveillance at slaughterhouses, defined here as the laboratory testing of randomly or purposively selected samples at the slaughterhouse, may provide a cost-effective approach to identifying undetected viral circulation and identifying prevalent or emerging strains. The utility of a slaughterhouse-based genomic surveillance system has not been evaluated for FMDV but could be valuable to improve genomic surveillance in endemic regions for early detection and selection of appropriate vaccines.

Most countries in Southeast Asia (SEA) are FMDV endemic. In Vietnam, serotypes O and A currently circulate in the country (de Carvalho Ferreira et al., 2017). Serotype O causes 80% of outbreaks, with four distinct lineages present: ME-SA (Mya-98), SEA (PanAsia), O-Ind2001, and Cathay. The PanAsia lineage is currently dominant, having been introduced in 2006 (Le et al., 2016). O/Ind 2001d was introduced into the Southern part of the country in 2015 and is currently circulating along with the PanAsia lineage (Vu et al., 2017). In addition, the Mya-98 lineage was first identified in Vietnam in 1998 and continues to cause sporadic outbreaks (Van Diep et al., 2020). Serotype A FMDVs identified in the country belong to the SEA/97, genotype IX and are closely related to strains from Laos and Thailand (de Carvalho Ferreira et al., 2017,Vp et al., 2010). From these observations, it is apparent that FMDV dynamics within Vietnam are characterized by the periodic introduction or emergence of new variants of both serotypes, some of which may become widespread within the country. To develop appropriate control measures or inform vaccine selection, it is important to identify emerging lineages as early as possible. Active surveillance rather than passive outbreak surveillance could provide this opportunity.

The objective of this study was to evaluate sampling of clinically normal ruminant livestock at slaughterhouses as a strategy for genomic surveillance of FMDV under endemic conditions. Specifically, we investigated the extent to which viruses recovered from slaughterhouses reflect the diversity found in the source population (inferred by farm sampling), and whether they can serve as sentinels for the early detection of outbreak strains identified through passive surveillance.

4.2 Methods

4.2.1 Study populations and sampling design

Farm-based sampling

Cattle and buffalo farms from eight provinces in northern (Lang Son, Phu Tho, Bak Kan, Ha Tinh) and southern (Ninh Thuan, Dong Thap, Dak Lak, Binh Phuoc) Vietnam were selected for this study based on the recent outbreak history and their identification as FMD hotspots (Lee et al., 2019,Bertram et al., 2018). Selected provinces bordering China (Bak Kan, Lang Son), Laos (Ha Tinh) and Cambodia (Dak Lak, Binh Phouc, Dong Thap) were selected to capture the potential introduction of FMDVs through transboundary movement. A serial cross-sectional study was carried out across these provinces. Briefly, in each province, 70 to 450 farms (average herd size =3 animals) were serially sampled from 2015 to 2019. Sera and oropharyngeal fluid (OPF) were collected from 30 to 250 animals per province per time point (Table 4.1). Animals that were seropositive for FMDV non-structural proteins (NSP) on the first round of sampling were re-sampled in consecutive rounds. The number of animals tested from each farm was variable across time, as was the time point in which farms were first initiated into the study.

Table 4.1: Descriptive characterization of longitudinal farm sample screening for FMDV NSP-serology, detection of FMDV RNA in oropharyngeal fluid (OPF), and sequence isolation.

	Province	Sampling Dates	No. of farms	NSP Serology (positive/total); Percent positive	RNA Detection in OPF Samples (positive/total); Percent positive	No. VP1 Seq. Obtained
Southern Provinces	Ninh Thuan	Oct 2016 Jun – Sep 2017 Jun – Sep 2018 Jan – Feb 2019	69	(1010/1290); 78.3%	(72/1003); 7.2%	23
	Dong Thap	Aug 2015 Oct 2016 Jun, Sep – Nov 2017 Jun – Aug 2018 Jan – Feb 2019	135	(888/1965); 45.2%	(197/882); 22.3%	98

Northern Provinces	Dak Lak	Aug 2015 Aug 2017 Jun – Oct 2018 Jan – Feb 2019	212	(1233/2173); 56.7%	(97/1230); 7.8%	72
	Binh Phuoc	Sep 2015	160	(84/514); 16.3%	(2/80); 2.5%	0
	Lang Son	2015 2016 Jun – Sep 2017 May – Aug 2018	227	(208/1387); 15%	(3/223); 1.3%	1
	Phu Tho	2015 2016 Aug – Nov 2017 Jun – Sep 2018 Jan – Feb 2019	442	(269/1256); 21.4%	(2/274); 0.8%	0
	Bak Kan	Oct 2016 Aug – Nov 2017 Jun – Sep 2018 Jan – Feb 2019	303	(1264/2790); 45.3%	(73/1241); 5.8%	22
	Ha Tinh	Aug 2015	274	(86/500); 17.2%	(0/112); 0%	0

Slaughterhouse-based sampling

Two cattle and buffalo slaughterhouses in Long An and Tay Ninh provinces in southern Vietnam were selected as pilot locations for genomic surveillance (Table 4.2). These slaughterhouses were selected partly because of their proximity to Cambodia, in order to investigate transboundary movements of FMDVs between these countries and due to animal movement from northern to southern Vietnam (Brito et al., 2017). Animals collected from several farmers in

surrounding communes were typically brought to the slaughterhouses by middlemen. Serial cross-sectional sampling was carried out from 2017 to 2019 every 15 days. Approximately 30 animals were sampled (serum and OPF) from each slaughterhouse in each round of sampling.

Table 4.2: Descriptive characterization of slaughterhouse sample screening from two slaughterhouses in southern Vietnam.

Province	Sampling Dates	NSP Serology (positive/total); Percent positive	RNA Detection in OPF Samples (positive/total); Percent positive	No.VP1 Sequences Obtained
Long An	Oct 2017 – May 2018 Jan – Feb 2019	(179/480); 37.3%	(51/480); 10.6%	51
Tay Ninh	Oct 2017 – Jun 2018 Jan – Feb 2019	(277/480); 57.7%	(71/480); 14.8%	71

Outbreak virus sequences

Outbreak sequences from across the country were also included in this study to quantify the genetic diversity of FMDV captured by passive surveillance activities. Sampling of outbreaks typically occurs after an outbreak (i.e., clinical signs in one or more animals) is reported and followed up. Sampling is usually conducted by the Ministry of Agriculture and Rural Development (MARD), Vietnam, sometimes in collaboration with the United States Department of Agriculture (USDA). Not all outbreaks are reported, and not all reported outbreaks are sampled. Outbreak sequences (VP1 region) from cattle, buffalo and pigs were generally obtained through sampling epithelium and oropharyngeal fluid. 80 and 26 serotype O and A outbreak sequences, respectively, were available from 2009 to 2019 from MARD, USDA, and GenBank, which were assumed to represent outbreak samples collected as part of passive surveillance.

4.2.2 Laboratory analysis

Serum samples were screened for the presence of antibodies against FMDV non-structural proteins (NSP) using a 3ABC ELISA (Priocheck®, Prionics, Netherland) following manufacturers' instructions as previously described (Fereirra et al., 2017). OPF and epithelium (outbreak) samples were screened for the presence of FMDV using virus isolation (VI), followed by detection of viral RNA in VI supernatant using qRT-PCR as previously described (Stenfeldt et al., 2016, Pacheco et al., 2010). Samples that were positive for viral RNA were subjected to sequencing

using one of several methods. Samples from 2013-2015 were sequenced using the Sanger method as previously described (de Carvalho Ferreira et al., 2017) to obtain VP1 sequences, or by next generation sequencing (NGS) to obtain full open reading frame (ORF) sequences. For NGS sequences, overlapping RT-PCR amplicons covering the full ORF were produced using three sets of primers (Brito et al., 2017), and amplicons were sequenced as previously described (Bertram et al., 2019). Samples from 2016-2017 were sequenced by NGS of RT-PCR amplicons covering the P1 region (Xu et al.,2013) as previously described (Bertram et al., 2019). Finally, sequences from 2018-2019 were sequenced by NGS using random and FMDV-specific primers to obtain the complete genome as previously described (Palinski et al., 2019,Bertram et al., 2019). All NGS sequencing was performed using the Illumina NextSeq platform. Read quality filtering, *de novo* assembly, and assembly to previously published references of regionally endemic lineages were implemented in CLC Genomics Workbench v12 (Qiagen). Sequences of the VP1 region were utilized in this study.

4.2.3 Analysis of diagnostic data

The proportions of NSP-positive and rRT-PCR positive animals were calculated for each province for each year for farm-based sampling and for each round of slaughterhouse sampling. To determine whether slaughterhouses are a good indicator of infection prevalence in the surrounding population, we compared apparent seroprevalence and percent positive on rRT-PCR (OPF sampling) at slaughterhouses and from farms in neighboring provinces during the same time period.

4.2.4 Phylogenetic analysis

Identification of circulating clusters

In order to document the effectiveness of slaughterhouse surveillance as a vehicle for genomic surveillance, we first classified sequences into genetic clusters of closely related viruses. Delineation of different clusters enabled tabulation of when and where distinct FMDV variants were detected.

Using the sequence data for the VP1 region of FMDV, we used a discriminant analysis of principle components (DAPC) to find the optimal clustering of sequences that minimized within-cluster genetic variation and maximized between-cluster distance, following (Jombart et al., 2010). Resulting clusters correspond to clades on a phylogenetic tree. Nine principal components were able to explain 90% of the variability in the genetic data and were used for the discriminatory clustering analysis for both Serotype O and A. The Bayesian information criterion (BIC) was used to determine the parsimonious number of clusters. This analysis was performed with the R package *adegenet* (Jombart, 2008).

Sequences from each cluster were blasted against NCBI and WRLFMD prototype lineages to identify the lineage to which each cluster belonged. The clusters were also compared with the currently used vaccine strains in a maximum-likelihood phylogenetic tree. For large clusters identified by DAPC (>10 sequences), the locations and time of appearance of sequences in different parts of Vietnam were mapped using ESRI ArcGIS.

Time-scaled phylogenies

In order to identify the emergence of different viral clusters through time and document the timeliness of slaughterhouse surveillance in detecting new clusters, a time-scaled phylogenetic analysis was performed using the Bayesian Evolutionary Analysis Sampling Tree (BEAST v1.10.4) software for both serotype A (132 sequences) and O (193 sequences). For serotype O, a total of 72 sequences from farm-based sampling, 41 sequences from slaughterhouses, and 80 sequences from outbreaks were included in the analysis. For serotype A, 30 sequences from farm-based sampling, 16 sequences from slaughterhouses, and 86 sequences from outbreaks were included. Because farm sampling was longitudinal, in some cases, the same animal was consecutively sampled at different rounds, resulting in nearly identical sequences from the same animal. In such instances, only the first sequence per animal was included. All available outbreak and slaughterhouse sequences were used. Sequences were screened for recombination prior to further analysis using RDP4 software (Martin et al., 2015) and aligned using MUSCLE algorithm (Edgar, 2004). The best-fit nucleotide substitution model was the HKY model, which was identified through JMODEL test (Darriba et al., 2012).

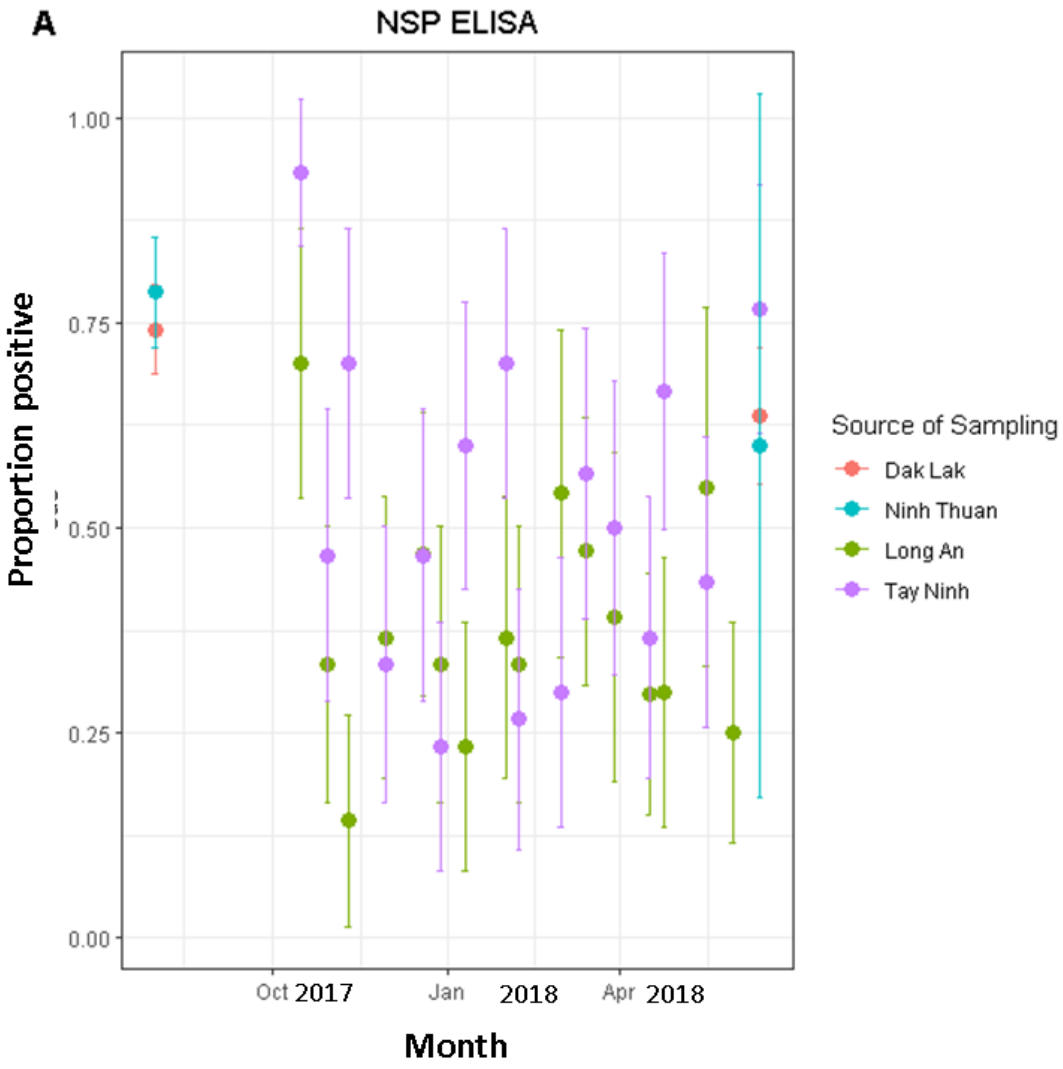
A relaxed uncorrelated log-normal molecular clock was tested with four different population models (constant, expansion, exponential, and Bayesian Skygrid), with the marginal likelihood of each candidate model compared using path-sampling and stepping-stone estimators (Baele et al., 2016). Each model was run for 200 million iterations on CIPRES (Miller et al., 2010). Tracer 1.7.1 was used to assess the conversion of the chains visually and for effective sample sizes of >200 (Rambut et al., 2007). A relaxed clock coalescent Skygrid model was selected for both serotype O and A. A maximum clade credibility (MCC) was constructed from 10,000 posterior samples of trees (discarding 10% burn-in), and annotated using *ggtree* (Rambut et al., 2006, Yu et al., 2017). Time to most common recent ancestor (tMRCA) of each cluster and 95% highest posterior densities (95%HPD) were obtained from the MCC tree.

4.3 Results

4.3.1 Descriptive data (sample screening)

A total of 11,875 serum samples and 5,045 OPF samples from farms were tested via NSP-ELISA and rRT-PCR, respectively, and 216 VP1 sequences were obtained (Table 4.1). Overall, 42.4% (95%CI: 32.2-52.1%) of serum samples were NSP-positive, and 8.8% (95%CI: 3.4-15.1%) of OPF were rRT-PCR-positive; 1200 serum samples and 1200 OPF samples were collected from slaughterhouses, and 95 sequences were obtained (Table 4.2). Across 16 rounds of sampling, 37.3% (95%CI:32.9-41.7%) of serum samples were NSP-positive and 10.6% (95%CI: 4.1-16%) were rRT-PCR-positive in the Long An slaughterhouse, whereas 51.8% (95%CI: 47.3-56.4%) of serum samples were NSP-positive and 16.7% (95%CI: 9.6-24%) were rRT-PCR-positive in the Tay Ninh slaughterhouse. Detailed summaries of diagnostic results by year and province are reported in Supplementary Tables S4.1-4.7.

The proportion NSP sero-positive in both slaughterhouses had substantial variability across samplings, and confidence intervals were quite wide due to relatively low sample size per time point (Figure 4.1a). Thus, it was difficult to pinpoint differences between the two slaughterhouses or discern temporal trends. Farm sampling data were available from two provinces (DakLak and Ninh Thuan) located in the same region as the slaughterhouses, sampled at approximately similar time points. In these provinces, on-farm prevalence was similar to what was determined in the slaughterhouses, but confidence intervals were wide (Figure 4.1 b). Amongst NSP-positive animals at slaughterhouses (Long An: n = 167; TayNinh: n = 231), 30.5% (95% CI: 20-38%) and 30.7% (95% CI:22-40%) were rRT-PCR positive, respectively (Figure 4.1 b).


A
NSP ELISA
Proportion positive
1.00
0.75
0.50
0.25
0.00
Oct 2017
Jan 2018
Apr 2018
Month
Source of Sampling
Dak Lak
Ninh Thuan
Long An
Tay Ninh

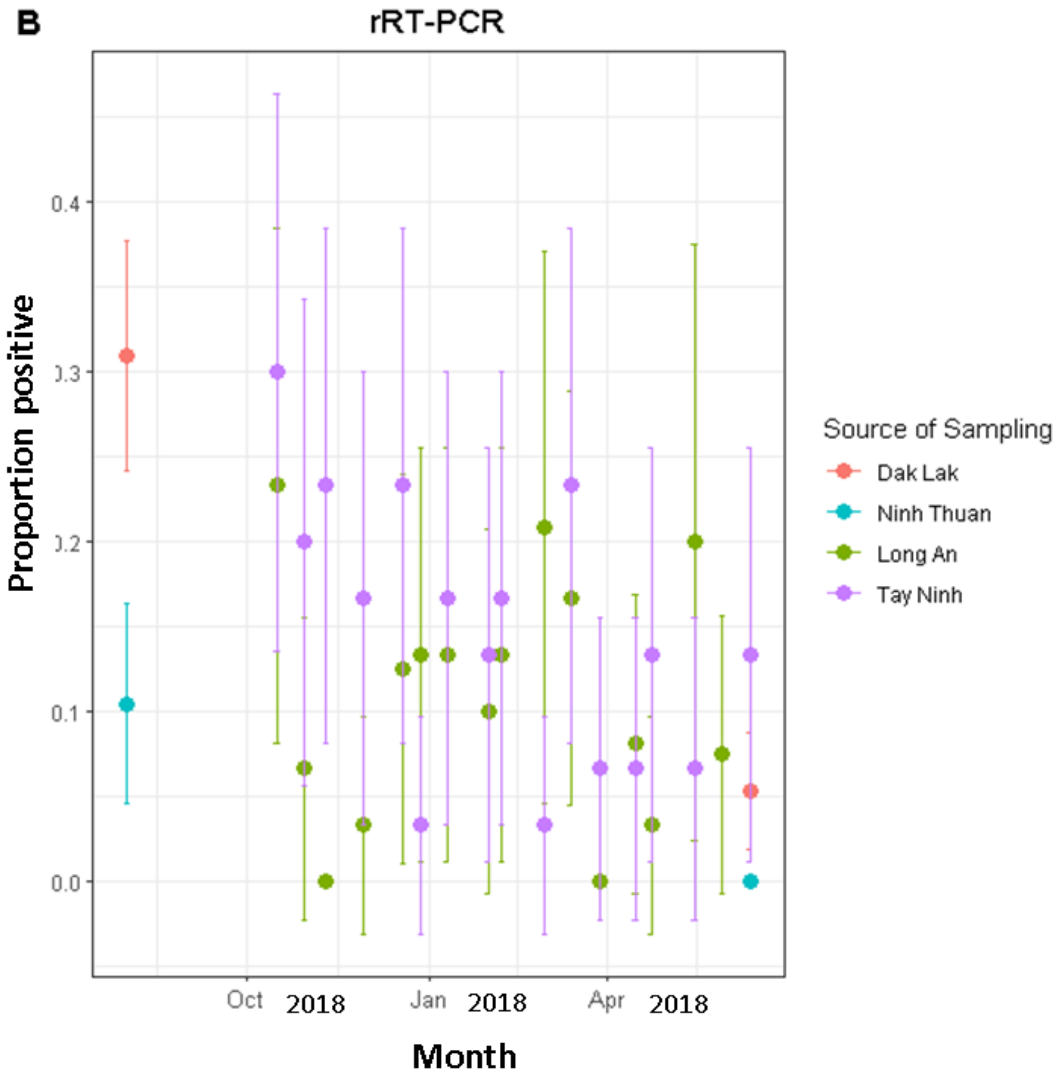
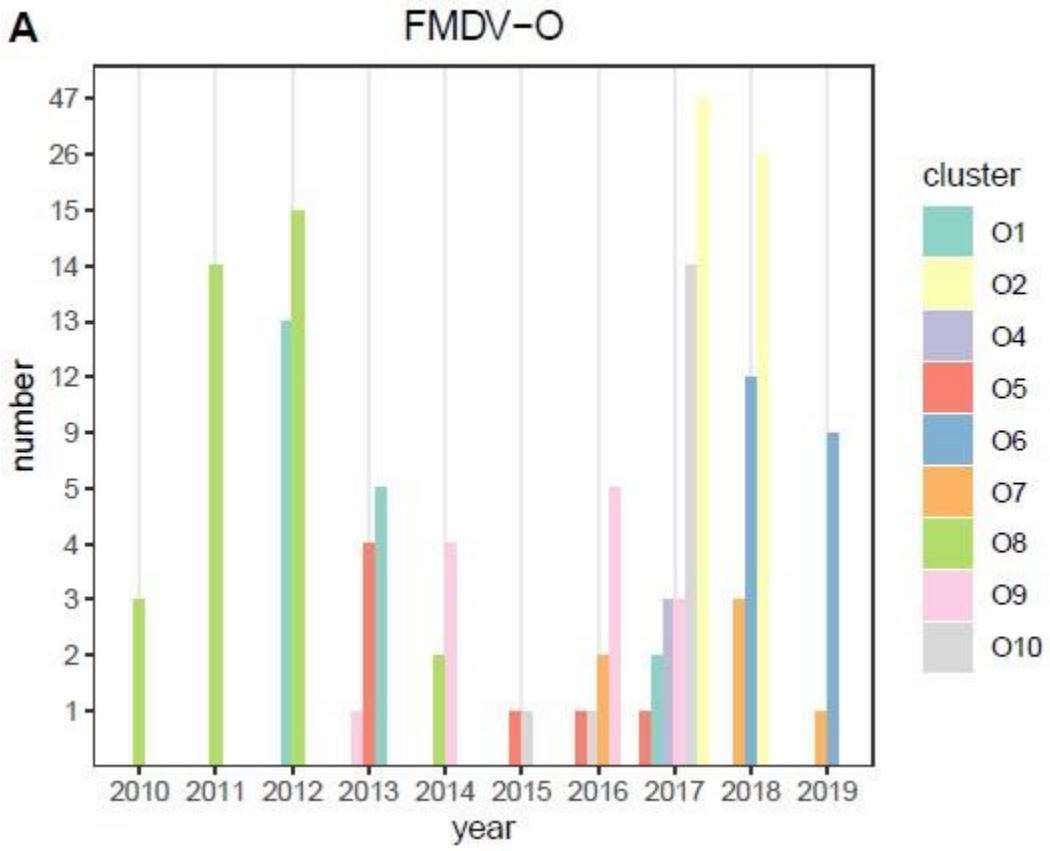


Figure 4.1: A) Proportion of animals NSP-positive in farms and slaughterhouses from August 2017 to June 2018. B) rRT-PCR detection rate of FMDV RNA in oropharyngeal fluid from farms and slaughterhouses from August 2017 to June 2018. Error bars represent 95% confidence intervals. Slaughterhouses were in Long An and Tay Ninh. Farms were in Ninh Thuan and Dak Lak.

4.3.2 Cluster analysis

For both serotypes, the first nine principal components accounted for 90% of the variability in the genetic data. Through application of DAPC using these nine components, nine clusters were identified based on genetic diversity within serotype O and eight clusters were identified within serotype A. For serotype O, seven clusters belonged to the MESA-PanAsia lineage and two of the clusters belonged to Mya-98 and Cathay lineages. For Serotype A, all clusters belonged to the SEA/97 lineage (Figure 4.4 and Supplementary Figure S4.2). Six of nine and four of eight

serotype O and A clusters, respectively, had >10 sequences, each with an average within-cluster genetic distance of 1.0 – 6.6% in the VP1 region. Supplementary tables 8 and 9 show details of clusters with more than ten sequences, including the lineage to which they belong, place of isolation across years, species, and within- and between-group genetic distances for both serotype O and A. An examination of the number of sequences isolated per cluster through time reveals a pattern whereby a particular cluster emerges (or is first detected), peaks, and subsequently declines in frequency through time (Figure 4.2 A and B).



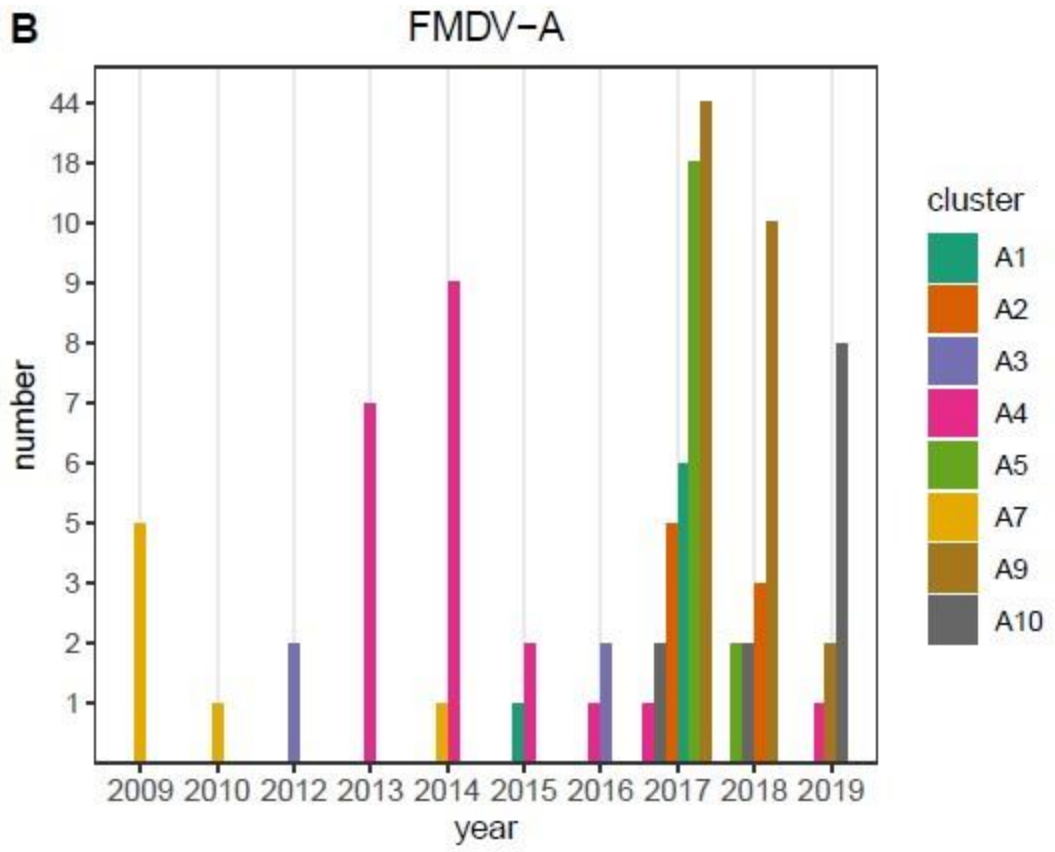


Figure 4.2: Number of sequences isolated per genetic cluster per year for A) serotype O and B) serotype A from the year 2010 to 2019. Serotype O cluster 6 and 9 belonged to the Mya-98 lineage, cluster 7 to the Cathay lineage, and all the other clusters belonged to the Pan Asia lineages . All serotype A clusters belonged to the SEA/97 lineage.

Some (17/56, 30.1%) sequences in serotype A-cluster 9 were identified as recombinant sequences within a different study analyzing full-length sequences (Bertram et al., 2021). Although the VP1 portion of these viruses is not recombinant and belongs to A/Sea-97, other parts of the genome belong to O/ME-SA/Pan-Asia. Due to the phylogenetic clustering of these 18 sequences with other sequences for which full-length genomes were not available, it is likely that all sequences within this cluster were the same A-O recombinant.

4.3.3 Phylogenetic data analysis

To evaluate the utility and timeliness of slaughterhouse surveillance, we focused only on the large clusters (>10 sequences per cluster) that were identified in the southern part of the country during the time period in which active sampling was conducted at slaughterhouses in this region (2017-2019). Four and three clusters met these criteria for serotypes O and A, respectively. Of these seven serotype O and A clusters circulating in southern Vietnam at this time, six were detected at

slaughterhouses, which suggests that slaughterhouse sampling is effective for revealing the diversity of circulating FMDVs in the host population (Figures 4.3-4.4, Supplementary Tables S4.8-4.9). The one cluster which was not detected at slaughterhouses was one that only contained outbreak sequences from pigs (Serotype O-Mya-98 Cluster 6), which were not sampled within as part of farm-based or slaughterhouse surveillance efforts.

For one of the six clusters detected at slaughterhouses (Serotype O cluster 2), detection through active slaughterhouse surveillance preceded passive outbreak surveillance by 4-6 months (Figure 4.3). Specifically, the serotype O cluster 2 sequences associated with outbreaks in the northern Vietnam in 2018 was detected in slaughterhouses in southern Vietnam in 2017 (Figure 4.5). For three clusters in serotype O (clusters 8, 9, 10) and one cluster in serotype A (cluster 4), clusters were detected in outbreak samples before appearing in active farm and slaughterhouse samples. However, the outbreak samples occurred during time periods during which no active surveillance was conducted for four of these clusters.

Table 4.3: Summary of clusters with >10 sequences for both serotype O and A. Sequences were obtained from outbreaks (OB), farms (FA), and slaughterhouses (SH). Regions of the country are divided as northern, central and southern Vietnam. †Clusters that were circulating in southern Vietnam during period of slaughterhouse sampling.

Serotype/ cluster ID	**source**	**Number of sequences per source**	**Total number of sequences**	**Region of first detection**	**Earliest date detected**
O-1	OB	1	20	South (FA)	2012-04-13
	FA	19			
O-2†	OB	9	73	South (SH)	2017-01-10
	FA	22			
	SH	42			
O-6†	OB	21	21	South (OB)	2018-02-07
O-8	OB	25	34	North (FA)	2010-12-22
	FA	9			
O-9†	OB	10	13	Central (OB)	2013-10-07
	FA	2			
	SH	1			
O-10†	OB	2	16	South (OB)	2015-09-10
	FA	3			

	SH	9			
A-4†	OB	19	21	Central (OB)	2013-10-09
	FA	1			
	SH	1			
A-5†	FA	6	20	Central (FA)	2017-01-08
	OB	5			
	SH	9			
A-9	FA	5	56	Central (FA)	2017-01-08
	OB	50			
A-10†	FA	6	12	South (FA)	2018-10-03
	SH	6			

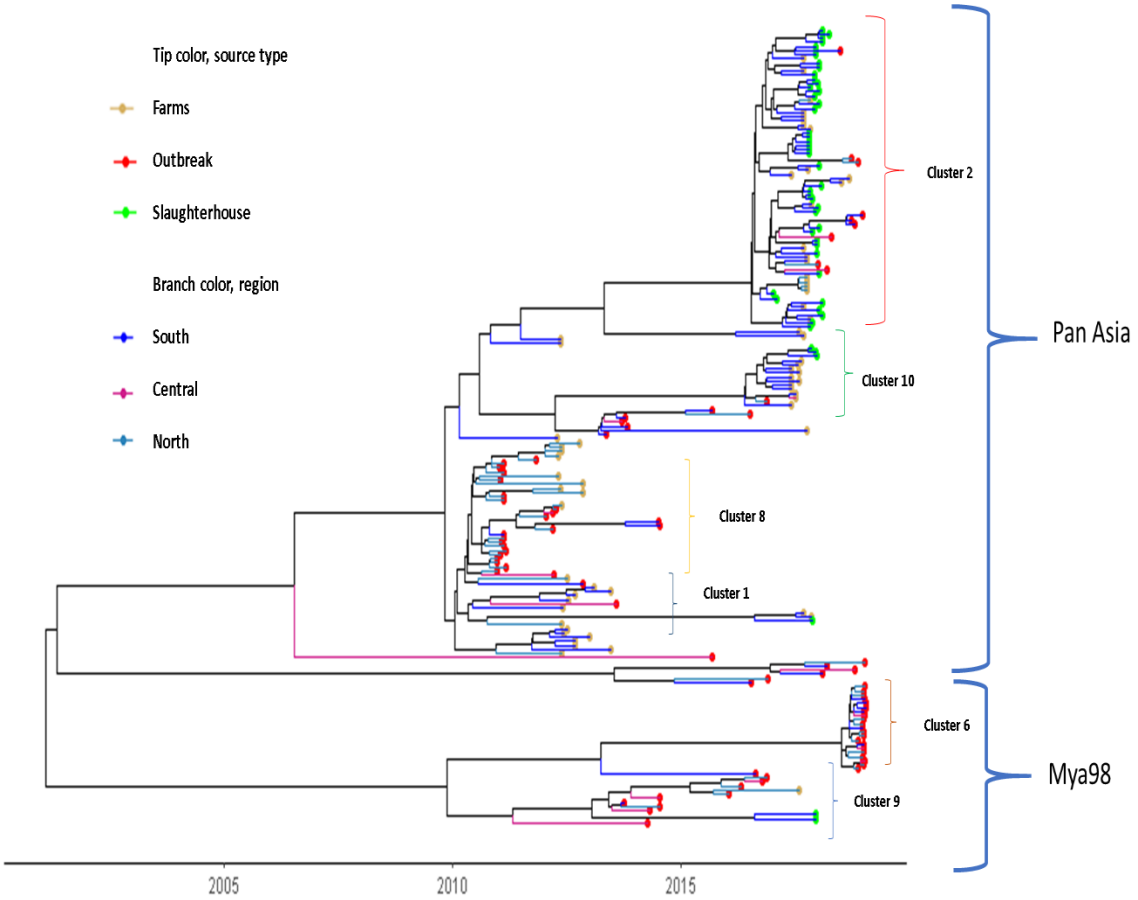


Figure 4.3: Time-scaled phylogeny for serotype O sequences isolated in Vietnam. Tip color indicates the source type of data (slaughterhouse, farm and outbreak). Different branch colors shows the region of the country where sequences were isolated. Small brackets identify the clusters and the large brackets identify the lineages each cluster belongs to.

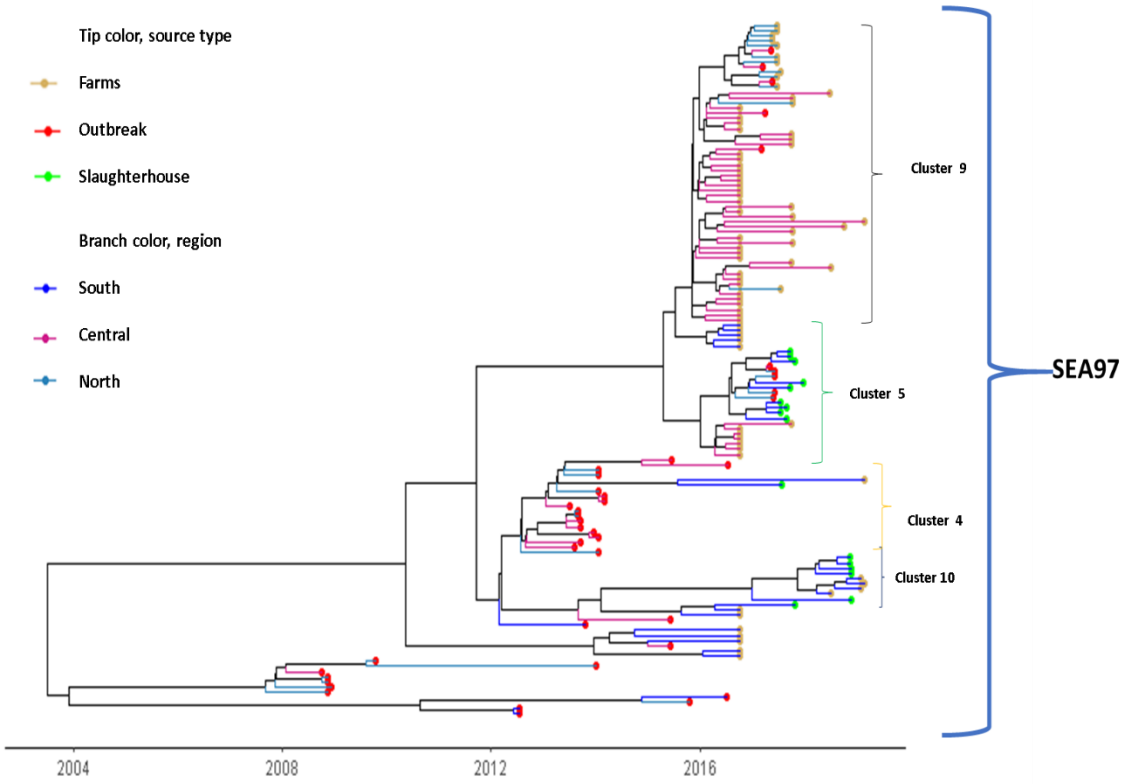


Figure 4.4: Time-scaled phylogeny for serotype A sequences isolated in Vietnam. All isolates belonged to the SEA-97 lineage. Tip color indicates the source type of data (slaughterhouse, farm and outbreak). Different branch colors show the region of the country where sequences were isolated. Small brackets identify the clusters and the large brackets identify the lineage each cluster belongs to.

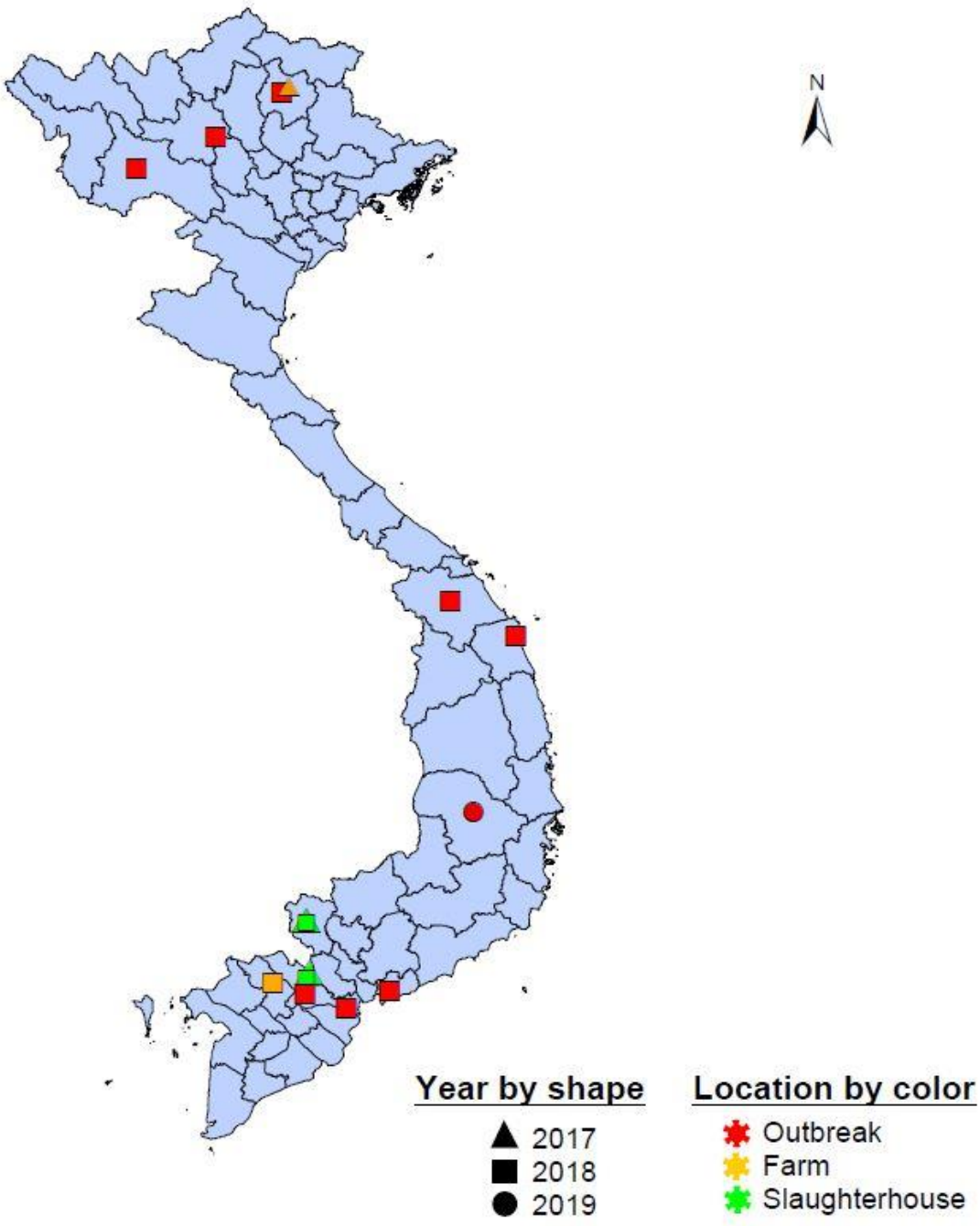


Figure 4.5: Spatial distribution of sequences in serotype O cluster 2. Outbreak samples are shown in red, slaughterhouse samples in green, and farms samples in orange. Shape indicates year of sampling.

4.5 Discussion

This study provides evidence that, in endemic settings, active surveillance of asymptomatic animals at slaughterhouses can be an effective means of genomic surveillance for FMDV. We identified six distinct serotype O and four serotype A genetic clusters through sequencing FMDVs recovered from serial cross-sectional sampling at selected slaughterhouses in southern Vietnam, active surveillance at farms in the same region, and passive surveillance based on outbreak reporting throughout the country. The data herein indicate that most (6 out of 7) large clusters circulating in southern Vietnam between 2017-19 were detected at least once at slaughterhouses. In addition, our results suggest that slaughterhouse-based surveillance can provide more timely information on circulating or emerging FMDV variants as compared to passive detection through outbreaks. Specifically, some variants were detected at slaughterhouses four to six months prior to their association with reported outbreaks. These results demonstrate the potential utility of systematic genomic surveillance across a network of slaughterhouses in endemic countries.

While slaughterhouse surveillance was able to capture the underlying diversity documented in farms of the same region, proportion positivity for FMDV RNA detection (rRT-PCR) and sero-reactivity (NSP-ELISA) were highly variable through time which precluded making any conclusions about the representativeness of slaughterhouse samples for estimating prevalence. This was further complicated by the difference in the time schedule of sampling at slaughterhouses and farms, and insufficient sample sizes per time point. Both sampling efforts were not truly random. Additionally, because these slaughterhouses were in border provinces, some animals may have arrived through transboundary animal movements, which may not be representative of seroprevalence in farms in the region. Despite these challenges, this study demonstrates that sentinel surveillance at slaughterhouses is convenient and inexpensive, and while slaughterhouse-based sampling may not provide precise estimates of prevalence, routine genomic surveillance at slaughterhouses may be effective for early detection of novel FMDV variants.

Within the scope of this study, circulating viruses in Vietnam were associated with the serotype A SEA/97 lineage and the serotype O Cathay, Pan Asia and Mya-98 lineages, with Pan Asia being the most common. This finding is consistent with other recent molecular epidemiology studies in Vietnam (de Carvalho Ferreira et al., 2017, Bertram et al., 2019, Brito et al., 2017). Analysis of 325 viral sequences collected from slaughterhouses, farms, and outbreaks revealed nine genetic clusters within these lineages. These genetic clusters do not correspond to the spatial clustering of outbreaks reported in different parts of Vietnam (Lee et al., 2020). For example, the 73 sequences belonging to serotype O-cluster 2 were found throughout the country (Figure 4.5).

Viruses isolated from slaughterhouses clustered together with viruses recovered from farms during the same period, indicating that slaughterhouses are representative of FMDV circulation at the farm level. Indeed, six out of seven clusters identified in southern Vietnam from 2017-2019 were detected at least once at these two slaughterhouses. The one cluster not detected in slaughterhouses was comprised exclusively of outbreak samples from pigs, demonstrating a limitation of the active surveillance schemes in this study (sampling asymptomatic/carrier ruminants at slaughterhouses misses lineages with tropism for pigs). Nonetheless, the diversity of FMDVs detected at slaughterhouses was largely representative of the diversity identified in the general population, as quantified from farm-based sampling and passive surveillance.

Sequences identified from Vietnam were closely related to viruses isolated from adjacent countries, indicating a role of transboundary animal movement for FMDV spread and highlighting the importance of regional approach to control FMD in Vietnam (Brito et al., 2017). In order to identify and control incursions of novel FMDV variants promptly, it is important to incorporate genomic surveillance as a part of routine surveillance at key locations. Our results demonstrate how monitoring slaughterhouses in southern Vietnam, bordering Cambodia, was able to provide early detection of novel variants that could potentially have been introduced from neighboring countries. Rather than implementing slaughterhouse surveillance across the entire country, it could be cost-effective to employ a “risk-based” approach whereby a network of sentinel slaughterhouses could be strategically established with consideration to transboundary animal movement and outbreak hotspots (Lee et al.,2015). Our results suggest that such a network could identify new FMDV variants in a similar timeframe and in some cases earlier compared to the current status quo of passive surveillance. Such early warning could provide more time for authorities to decide on appropriate control measures and vaccine selection.

Slaughterhouse sampling did not result in earlier detection of genetic clusters in all cases. For clusters that were detected through outbreak sampling (passive surveillance) prior to subclinical detection (active surveillance at slaughterhouses), the outbreak data was not aligned spatially or temporally with the period in which slaughterhouse sampling was conducted. Thus, the apparent delay in detection at slaughterhouses relative to outbreak reporting may reflect that the cluster was not circulating in populations near the slaughterhouses during the period of sampling. However, a larger network of slaughterhouse-based surveillance throughout the country may have detected such clusters earlier.

Time-scaled phylogenies illustrated that closely related viruses were identified in farms both before and after they are detected as associated with an outbreak. Infectious FMDV was isolated from OPF samples in animals sampled 6-12 months after the outbreak-associated sequence.

These animals sampled in slaughterhouses and farms did not have clinical signs of FMD at the time of sampling, and thus detection of virus in such animals represented either persistent infections in carrier animals or early (acute) sub-clinical (neoteric) infections (Stenfeldt and Arzt, 2020). Related to this, the recovery of viruses in OPF samples collected from persistently infected carriers introduces some uncertainty in the dating of the incidence of infection, as the sample collection date is surely later than the infection date (Bronsvoort et al., 2016,Hayer et al., 2018,Bertram et al., 2020). This could potentially have impacted the date estimates in the time-scaled phylogenies, though we do not think that it changes our general conclusions in relation to the representativeness and timeliness of slaughterhouse-based surveillance.

It is apparent from our data that genetic clusters emerged and disappeared across time. Unfortunately, the nature of this study did not allow for examination of drivers of cluster emergence. Because cross-protection amongst related strains may only be partial, immune-driven interactions among co-circulating viruses at the population level could lead to the replacement of existing clusters with new clusters. Cross-protection may result in clinical protection from a different strain of the same serotype, but still may allow for viral replication, transmission, and immune-mediated selection, thus creating ecological or evolutionary selection pressures for viral evolution and cluster turnover. A similar phenomenon of serial subclinical infections with distinct heterologous and homologous strains of FMDV was demonstrated in Asian buffalo in Pakistan (Farooq et al, 2018). Alternatively, FMDV evolution and circulation of specific genetic clusters in endemic settings may be a product of stochastic spatiotemporal processes (e.g., founder effects) within heterogeneously structured host populations (Fournié et al., 2018), which combine to generate a pattern of introduction, spread, and fade out of clusters over time.

Active surveillance plays a key role in controlling contagious diseases such as FMD (Henritzi et al., 2020,Shapshak et al., 2015). The effectiveness of such surveillance is dependent upon early detection of viral variants using appropriate molecular tools combined with sensibly executed surveillance systems. In this study, we demonstrate that active surveillance in sentinel slaughterhouses can capture much of the genetic diversity of circulating endemic FMDVs. Our results suggest that routine genomic surveillance in slaughterhouses would provide representative and timely data on both established and emerging genetic variants, in some cases detecting novel variants four to six months prior to their detection via passive surveillance. These results underscore the potential utility of systematic genomic surveillance for FMDV and other pathogens in slaughterhouses in endemic countries.

Table S4.1: Serology and OPF screening from farms in each province in 2015.

Province	Serum Tested	Serum positive	Seropositive percentage of samples	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
Hà Tĩnh	500	86	17.2 (13,21)	112	0	0
Lạng Sơn	484	20	4 (2,5)	33	0	0
Phú Thọ	500	50	10 (7,12)	62	0	0
Bắc Kạn	*na*	*na*	*na*	*na*	*na*	*na*
Ninh Thuận	*na*	*na*	*na*	*na*	*na*	*na*
Đồng Tháp	485	149	31 (26,34)	139	36	25.8
Bình Phước	514	84	16 (13,19)	80	2	2.5
Đăk Lăk	504	142	28 (24,32)	142	3	2.1
TOTAL	2,987	531		568	41	7.2

Table S4.2: Serology and OPF screening from farms in each province in 2016.

Province	Serum Tested	Serum positive	Seropositive percentage of samples	OPF Screened (rRT-PCR)	OPF positive	Percentage OPF positive
Hà Tĩnh	*na*	*na*	*na*	*na*	*na*	*na*
Lạng Sơn	236	31	13 (8,17)	33	2	6
Phú Thọ	253	34	13 (9,17)	32	2	6.2
Bắc Kạn	496	167	34 (29,37)	165	32	19.4
Ninh Thuận	250	177	71 (65,76)	177	20	11.2
Đồng Tháp	254	100	39 (33,45)	102	33	32.3
Bình Phước	*na*	*na*	*na*	*na*	*na*	*na*
Đăk Lăk	*na*	*na*	*na*	*na*	*na*	*na*
TOTAL	1,489	509		509	89	17.4

Table S4.3: Serology and OPF screening from farms in each province in 2017.

Province	Serum Tested	Serum positive	Seropositive percentage of samples	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
Hà Tĩnh	*na*	*na*	*na*	*na*	*na*	*na*
Lạng Sơn	316	92	29 (24,34)	92	0	0
Phú Thọ	177	82	46 (39,53)	77	0	0
Bắc Kạn	982	447	46 (42,48)	443	33	7.4
Ninh Thuận	585	467	80 (76,83)	461	49	10.6
Đồng Tháp	369	210	57 (51,61)	212	85	40
Bình Phước	*na*	*na*	*na*	*na*	*na*	*na*
Đăk Lăk	252	182	72 (66,77)	181	56	31
TOTAL	2,681	1,480		1,466	223	

Table S4.4: Serology and OPF screening from farms in each province in 2018.

Province	Serum Tested	Serum positive	Seropositive percentage of samples	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
Hà Tĩnh	*na*	*na*	*na*	*na*	*na*	*na*
Lạng Sơn	351	65	19 (14,22)	65	1	1.53
Phú Thọ	149	48	32 (24,39)	48	0	0
Bắc Kạn	752	363	48 (44,51)	347	3	0.8
Ninh Thuận	319	257	81 (76,84)	256	3	1.1
Đồng Tháp	540	284	53 (48,56)	284	31	11
Bình Phước	*na*	*na*	*na*	*na*	*na*	*na*
Đăk Lăk	942	605	64 (61,67)	605	27	4.4
TOTAL	3,053	1,622		1,605	65	

Table S4.5: Serology and OPF screening from farms in each province in 2019.

Province	Serum Tested	Serum positive	Seropositive percentage of samples	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
Hà Tĩnh	*na*	*na*	*na*	*na*	*na*	*na*
Lạng Sơn	*na*	*na*	*na*	*na*	*na*	*na*
Phú Thọ	117	55	47 (38,56)	55	0	0
Bắc Kạn	560	287	51 (47,55)	286	5	1.74
Ninh Thuận	136	109	80 (73,86)	109	0	0
Đồng Tháp	317	145	46 (40,51)	145	12	8.27
Bình Phước	*na*	*na*	*na*	*na*	*na*	*na*
Đăk Lăk	475	304	64 (59,68)	302	11	3.64
TOTAL	1,605	900		897	28	

Table S4.6: Slaughterhouse Serology and OPF sample collection summary Long An.

Sampling Round	Serum Tested	Number of positive 3ABC	Percentage 3ABC ELISA positive	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
1	30	21	70.0 (56,86)	30	7	23.3
2	30	10	33.3 (16.5, 50)	30	2	6.6
3	28	4	14.3 (1.3,27)	28	0	0
4	30	11	36.7 (19,53)	30	1	3.3
5	32	15	46.9 (29,64)	32	4	12.5
6	30	10	33.3 (16.5,50)	30	4	13.3
7	30	7	23.3 (8.2, 38)	30	4	13.3
8	30	11	36.7 (19.5,53)	30	3	10
9	30	10	33.3 (16.5, 50)	30	4	13.3
10	24	13	54.2 (34.3, 74)	24	5	20.8
11	36	17	47.2 (31,63)	36	6	16.6

12	23	9	39.1 (19.2,56)	23	0	0
13	37	11	29.7 (15,44)	37	3	8.1
14	30	9	30.0 (13.6,46)	30	1	3.3
15	20	11	55.0 (33,76)	20	4	20
16	40	10	25.0 (11,38)	40	3	7.5
Grand Total	480	179	37.3	480	51	

Table S4.7: Slaughterhouse Serology and OPF sample collection summary Tay Ninh.

Sampling Round	Serum Tested	Number of positive 3ABC	Percentage 3ABC ELISA positive	OPF Screened (rRT-PCR)	OPF positive for FMDV RNA	Percentage OPF positive
1	30	28	93.3 (84, 100)	30	9	30
2	30	14	46.7 (29,64)	30	6	20
3	30	21	70.0 (53,86)	30	7	23.3
4	30	10	33.3 (16.5,50)	30	5	16.6
5	30	14	46.7 (29,64)	30	7	23.3
6	30	7	23.3 (8,38)	30	1	3.3
7	30	18	60.0 (42,77)	30	5	16.6
8	30	21	70.0 (53,86)	30	4	13.3
9	30	8	26.7 (11,42)	30	5	16.6
10	30	9	30.0 (13,46)	30	1	3.3
11	30	17	56.7 (39,74)	30	7	23.3
12	30	15	50.0 (32,74)	30	2	6.6
13	30	11	36.7 (19.5, 53.8)	30	2	6.6
14	30	20	66.7 (49.8,83)	30	4	13.3
15	30	13	43.3 (25.6, 70)	30	2	6.6
16	30	23	76.7 (61,91)	30	4	13.3
Grand Total	480	277	57.7	480	71	

Table S4.8: Large cluster information of serotype O.

	Type	Area	Dates Detected	Species	MRCA	Number of sequences	Closest lineage	Within Cluster GD
cluster 1	Farm	North	2012-05-09 to 2012-10-11	Cattle, Buffalo	2010.1 (2001.77, 2028.89)	20	O/ME-SA/PANASIA	0.034
	Farm	Central	2012-04-23					
	Farm and Outbreak	South	2012-04-13 to 2017-01-09*					
cluster 2†	Farm and Outbreak	North	2017-10-01 to 2018-11-11	Pig, Cattle, Buffalo, Goat	2016.5 (2014.4,2018.8)	73	O/ME-SA/PANASIA	0.011
	Outbreak	Central	2018-04-13	Cattle				
	Farm and Slaughterhouse	South	2017-01-10 to 2018-07-01	Cattle, Buffalo				
cluster 6†	Outbreak	North	2018-11-13 to 2019-01-15	Pig	2017.9 (2017.5, 2019.1)	21	Mya-98	0.021
	Outbreak	Central	2018-12-27 to 2019-01-05					
	Outbreak	South	2018-11-13 to 2019-01-11					
cluster 8	Farm and Outbreak	North	2010-12-22 to 2014-07-14	Pig, Cattle, Buffalo	2009.5 (2001.1,2018.9)	34	O/ME-SA/PANASIA	0.021
	Farm and Outbreak	Central	2010-12-26 to 2012-05-23	Pig, Buffalo				
	Outbreak	South	2010-12-21 to 2014-07-10	Pig, Cattle, Buffalo				
cluster 9†	Outbreak	North	2016-01-16 to		2013.1	11	Mya-98 B	0.066

			2016-11-18	Cattle	(2007.2,2019.2)			
	Outbreak	Central	2013-10-07 to 2016-10-14	Pig (2016-11-18)				
	Farm and Slaughterhouse	South	2016-08-18 to 2017-12-15					
cluster 10†	Outbreak	North	2016-11-16	Cattle Pig (2015-09-10)	2016.2	16	O/ME-SA/PANASIA	0.027
	Farm, Outbreak and Slaughterhouse	South	2015-09-10 to 2017-08-01		(2014.1,2019.6)			

† Clusters that were circulating in southern Vietnam during period of slaughterhouse sampling

* 2017 sequence does not cluster with other sequences in cluster 1, thus this cluster was not considered to be circulating after 2012.

Table S4.9: Large cluster information for serotype A.

Name	Type	Area	Dates detected	Species	MRCA	Number of sequences	Closest lineage	Within cluster GD
cluster9	Farm	North	2017-08-25 to 2018-01-08	Cattle	2015.3 (2012.2, 2019.5)	56	SEA/97	0.01
	Farm and outbreak	Central	2017-01-08 to 2019-01-05					
cluster5†	Outbreak	North	2017-08-29 to 2017-09-11	Cattle	2015.8 (2013.1,2019.5)	20	SEA/97	0.01
	Farm and outbreak	Central	2017-01-08 to 2017-08-05					
	Slaughterhouse	South	2017-10-17 to 2018-03-29					
cluster4†	Outbreak	North	2013-12-01 to 2014-04-24	Cattle Pig (2014-04-24)	2012.4 (2006.4,2019.4)	21	SEA/97	0.02
	Outbreak	Central	2013-10-09 to 2016-10-13	Cattle Pig (2015-09-10)				
	Outbreak and Slaughterhouse	South	2017-10-31 to 2019-06-01	Buffalo				
cluster10†	Farm	South	2017-01-06 to 2019-06-01	Cattle	2016.8 (2015.5,2020.1)	12	SEA/97	0.04
	Slaughterhouse		2018-01-24 to 2019-02-27					

† Clusters that were circulating in southern Vietnam during period of slaughterhouse sampling

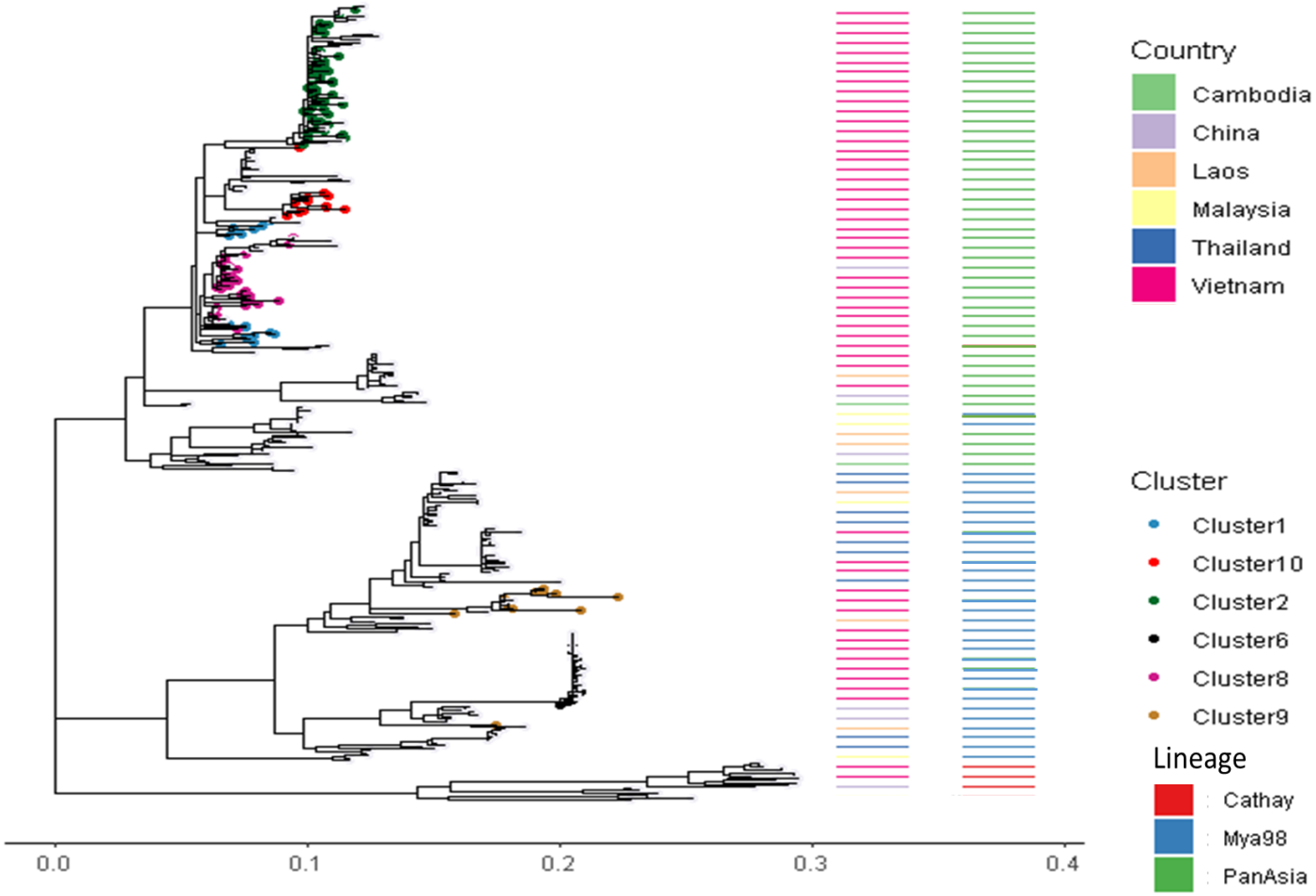


Figure S4.1: Serotype O clusters with other sequences from South East Asian countries Cambodia, Laos, Malaysia, Thailand.

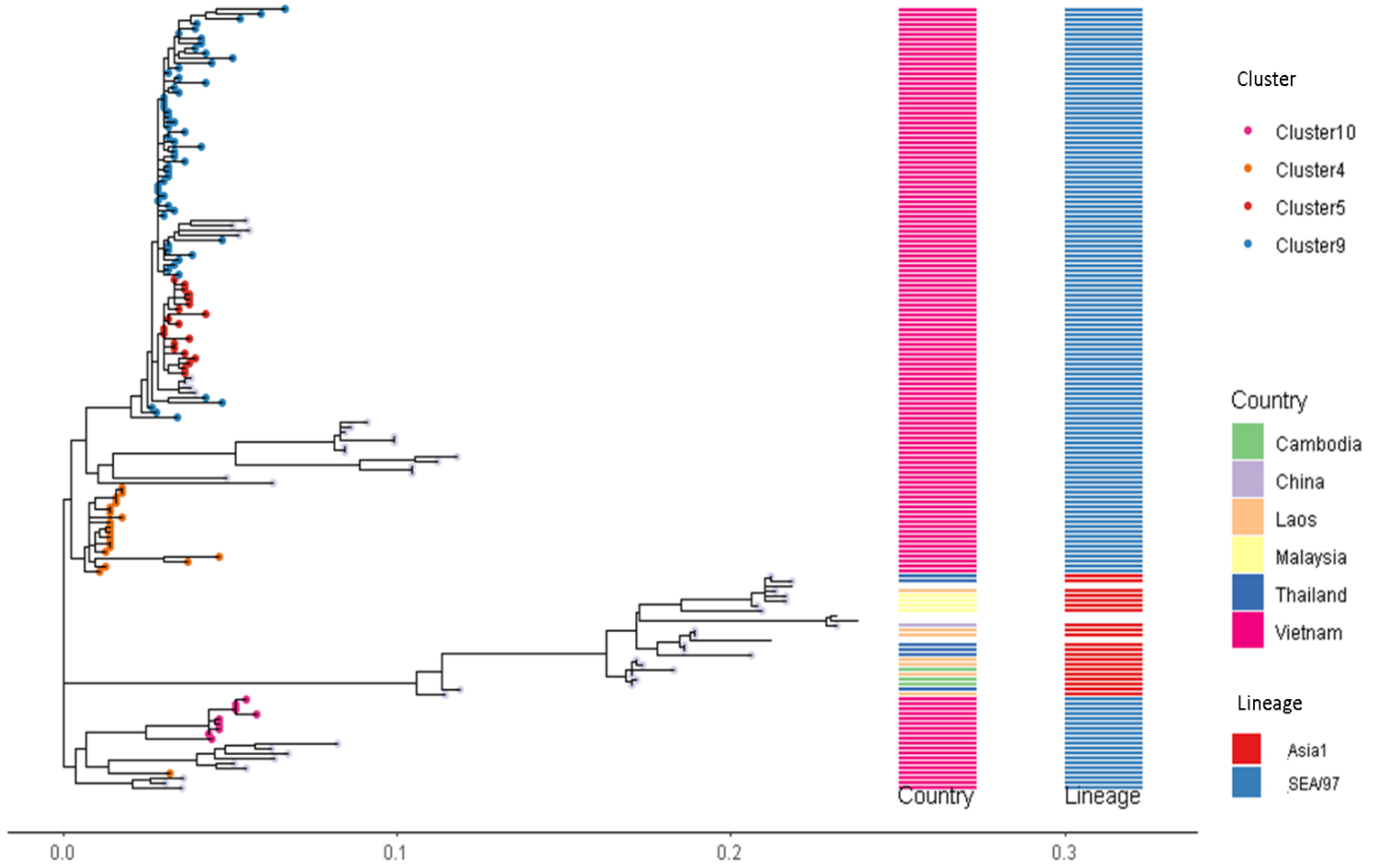


Figure S4.2: Serotype A clusters with other sequences from South East Asian countries Cambodia, Laos, Malaysia, Thailand.

Chapter 5: Conclusion

5.1 Overview

Due to the central locations of Vietnam and India in their respective FMDV pools, control of FMDV in these countries is important for broader control of the disease in South and Southeast Asia. This situation is exemplified by the emergence of several new transboundary lineages of FMDV from India (Bachanek-Bankowska et al., 2018,Di Nardo et al., 2021), as well as the detection of many of these lineages in Vietnam, which is a hub for animal movements in Southeast Asia (de Carvalho Ferreira et al., 2017,B. Brito et al., 2017,Di Nardo et al., 2011). The Progressive Control Pathway (PCP) provides a risk-based approach for FMD endemic countries to control FMD. For a given country to move from one stage to the other, as an assessment of each step, an epidemiological evaluation is carried out. However, the PCP provides limited guidelines on the types of epidemiological analysis a country can implement. Vietnam and India are in stage 3, and I have used data from these two countries to demonstrate the application of how these cutting-edge epidemiological approaches can be employed to support PCP. Table 5.1 shows an overview of what steps of PCP stage 1-3 can be supported by the epidemiological approaches utilized in this dissertation. In this dissertation, I have used Bayesian spatio-temporal Poisson regression models to identify high risk areas for FMD using available outbreak information and elucidate dynamics of virus circulation (Chapter 2 and 3). I also explored the possibility of using phylogenetic data to improve Bayesian Poisson regression models (Chapter 3). In Chapter 4, I have explored the potential use of sentinel genomic surveillance measures in Vietnam using molecular data collecting at slaughterhouses. Findings from these studies are important for both countries to progress in the PCP pathway to reach towards freedom from FMD. Having said that, not every country wants to reach freedom from FMD given the balance between cost of control measures versus impact of the disease, and depending on this economic balance, ongoing disease management and mitigation may be a desirable position for some endemic countries.

5.2 Implications for Policy and Practice

Since both Vietnam and India are endemic to FMD, it is difficult and potentially not cost-effective to implement control measures countrywide (Adamchick et al., 2021). A more feasible approach would be a risk-based control plan to gradually and systematically reduce risk and burden of disease. As outlined by the PCP, countries in stage 3 are recommended to focus on zonal freedom. However, there are a number of stage 1 and 2 activities that continue to need attention in stage 3 as well. Stage 1 (“pcp-26012011.pdf,”) includes a value chain analysis, estimates of the socio-economic impact of the disease, and enabling an environment for regional corporation.

India has achieved most of the steps in the stage1 and stage 2. This include the value chain analysis (Singh et al., 2013), stakeholder identification, identification of a hypothesis of virus circulation (http://www.pdfmd.ernet.in/index_files/Annual_Reports.html), identification of serotypes and topotypes of the virus (Subramaniam et al., 2015,Biswal et al., 2015,Mohanty et al., 2015, Mahapatra et al., 2015a), enabling the control environment for control by strengthening the veterinary services ("25022019_India_PVS_Evaluation_report_final.pdf,") and participation in regional corporation programs (OIE/FAO Global network of FMD Reference Laboratory). However, prior to this dissertation, the spatio-temporal dynamics of outbreaks had not been examined, thus Chapter 2 helped to identify the risk hot spot where the control measures should be mainly focused as well as state-level risk factors for FMD outbreaks in India, which will contribute to a better understanding of virus circulation. In PCP stage 2, ongoing monitoring of the FMD control programs should be conducted. In Chapter 2, we have shown how data from ongoing serological and outbreak monitoring can be analyzed in a spatio-temporal framework to quantify the extent to which standardized incidence of reported FMD outbreaks have reduced over time with the implementation of the control programs. We also show that relative risk of (reported) outbreaks is higher in states with low LPB-ELISA sero-positivity, and in the provinces with an international border. We also document substantial spatiotemporal heterogeneity in the percent of animals with inferred protection pre- and post-vaccination, which could lead to gaps in population immunity in between biannual vaccination. Which could be related to lack of a prime-boost vaccine regime, presence of stray cattle, or potential issues with the potency or delivery of the vaccine itself. During the time period studied, the vaccination program in India did not follow the exact guidelines of OIE whereby animals should be tested in an age stratified manner to detect the protective antibody titers, and this made it challenging to interpret the variability in sero-prevalence. Age-stratified sampling has since implemented in 2020.

Now in the stage 3, India will continue to develop risk-based strategic plans with the focus on high-risk areas. In stage 3, application of risk-based control measures in the targeted zones with the target of obtaining freedom from FMD in at least one zone should be conducted. There should also be evidence that the control measures have reduced the number of outbreaks. Chapter 2 provides some evidence that control efforts have reduced outbreaks.

Vietnam is also in the stage 3 of the PCP pathway proposed by OIE. As a part of stage 1 of the PCP, the distribution of the virus in the country is well described and understood (B. Brito et al., 2017,Le et al., 2016,de Carvalho Ferreira et al., 2017). FMD hotspot are identified in northern and southern part of the country (Lee et al., 2020), and it is important to continually monitor the distribution of the virus as a part of routine surveillance in the official FMD control program. In Vietnam, we were unable to obtain any vaccine coverage, vaccination, or serological monitoring

data. While Vietnam does have a vaccination program in place, monitoring of protective antibody titers of animals (a requirement of PCP stage 2) was not available for our study. Data from such a program is key to measuring the progress of the vaccination program. If not for the whole country, this monitoring could be implemented in areas where the country plan to achieve zonal freedom. One of the major lessons learned from spatio-temporal modeling of FMD outbreaks in Chapter 2 was that the spatial adjacency between states did not account for a substantial amount of variability in the relative risk of outbreaks, suggesting that spatial adjacency (at least at the state-level scale) was not a strong driver of virus circulation. An alternative hypothesis of disease circulation is that it is driven by population connectivity via animal movements, which was not accounted for in the India work. Unfortunately, many endemic countries do not have digital databases tracking animal movements, which limits our ability to incorporate animal movement in hypotheses of virus circulation. However, molecular data combined with phylogeographic analysis can help reconstruct past viral movements, which are likely to be correlated with host movement. In Chapter 3, we explored incorporating phylogeographic inference with outbreak space-time regressions to produce information about high-risk areas of reported FMD outbreaks in Vietnam. From these models we identified the high-risk areas for reported FMD outbreaks in the northern and southern part of the country, including areas where transboundary animal movement is common particularly from Cambodia and Malaysia. We also identified that accounting for virus movement through phylogeographic analysis served as a useful proxy for population connectivity in spatial-temporal risk models.

According to the PCP pathway, a country aiming towards FMD freedom should have a strategic control plan in place by the stage 2. The identified high-risk areas, risk factors and circulation of the virus can provide support for establishing and improving strategic control plans for Vietnam and India. Since India and Vietnam have porous borders with many other countries, conducting country-specific control programs could be more effective for transboundary diseases such as FMD if implemented alongside regional control plans proposed by FAO/OIE, sharing resources and strengthening collaborations with neighboring countries. In both countries, surveillance in border states/provinces would be beneficial, with enforced regulations and record-keeping on between-country animal movements. In practical situation, this is difficult due to noncompliance. Participatory studies should be conducted to determine the value chains at the border provinces to determine best regulations that people are likely to comply with.

Across all stages 1-3 of the PCP, circulating FMD strains in the country should be identified, with sampling representative of different production systems and geographic regions and laboratory testing conducted locally and by sending samples to the OIE reference laboratories. Chapter 4 of this thesis provides a proof-of-concept of genomic surveillance at slaughterhouses to detect

circulating strains of the FMDV that might lead to future outbreaks in Vietnam. Disease surveillance is defined as the systematic, active, ongoing observations of occurrence and the distribution of the disease (Murray and Cohen, 2017). Surveillance provides us the magnitude of the problem and a glimpse of population level virus circulation. This is particularly important for viral diseases, as early detection of emerging virulent viral strains is critical for control. In addition, as evidenced by the SARS-COV2 pandemic, genomic surveillance has become a much more prominent component of epidemic response, with important applications related to the immunology/vaccinology of the disease, pathology and clinical practice, and epidemiological understanding.

For animal diseases, surveillance can be carried out in farms, slaughterhouses, or other points of aggregation. Farm surveillance is not cost-effective due to increased biosecurity hazards, ethical aspects, and increased manpower costs. For FMD, clinical signs are not pathognomic and farmers may be unaware during the early period of spread in their farms, thus early detection of diseases is challenging in farm settings (Schirdewahn et al., 2021). Routine collection of samples provides better evidence, though costly. Active molecular surveillance for FMDV is widely used in both India and Vietnam (Mahapatra et al., 2015,Gibbens et al., 2001,B. P. Brito et al., 2017,Vu et al., 2017,B. Brito et al., 2017), but typically relies on farm-based active surveillance or sampling of reported outbreaks on farms. These methods can be further optimized through cost-efficient active sampling at points of aggregation, such as slaughterhouses, to determine widely circulating or new introduced strains that may cause outbreaks.

With the high rate of emergence of novel FMDVs as well as frequent transboundary introductions, early detection of new strains with the help of molecular tools provide better opportunities to control the disease. Currently circulating sequences in Vietnam are related to O/ME-SA/Ind-2001, O/ME-SA/Pan Asia and O/SEA/Mya-98; O/ME-SA/Pan Asia being most common. Within these lineages, our results show that several different genetic clusters emerge, spread, and disappear with time. In recent years, the ability to generate and analyze genetic sequence data is accelerating, and Chapter 4 explore ways to efficiently collect field samples for the purposes of genomic surveillance. We have shown that, in some cases, it’s possible to identify clusters associated with outbreak sequences earlier by conducting genomic surveillance at the slaughterhouses and that the genetic diversity of FMDVs isolated from apparently healthy animals at slaughterhouse are representative of the diversity of FMDVs isolated from the source population (i.e., farms) from the same region. This step would be important for rapid outbreak detection in FMD free zones when a country in stage 3 transition to stage4 of PCP.

Table 5.1: General steps in PCP stage 1-3 that can be supported to be evaluated by epidemiological approaches demonstrated in this dissertation. Shaded boxes indicates where these proposed interventions would be contributing.

Progressive Control Pathway		**Chapter 2 India**	**Chapter 3 Vietnam**	**Chapter 4 Vietnam**
Stage 1 **Activities to understand FMD risk**	Implement risk-based approach to reduce FMD			
	Hypothesis of how the FMDV circulate in the country including currently circulating strains			
	Value chain analysis			
	Socio economic impact of FMD			
	Evaluation based on OIE-PVS pathway			
	Regional collaboration			
	FMD risk hot spots are identified			
	Identify strategies to control FMD	Identify factors underpinning high-risk areas		
Stage 2 **Risk based strategic plan**				
	Ongoing monitoring of the circulating strains			
	Implement risk-based zone targeted control measures	High risk area identification		
	Show impact of FMD reduced with control measures	Evaluation of the vaccine program		
	Allocation of sufficient resources			
Stage 3 **Reduction of outbreak incidence and virus circulation in at least on zone of the country**				
	Evaluate the incursion of new serotypes			Enhanced surveillance
	Sustainable veterinary services			
	Legal framework for animal identification			
	Analysis of virological and outbreak data and analysis of serological survey			
	FMD contingency plan			
	Strengthening the veterinary service of the country for own epidemiological investigations			
	Rapid detection of outbreaks at least in a one zone of the country			Enhanced surveillance
	Endorsement of the official control program by OIE			

5.3 Limitations

General limitations of chapter 2 and 3: As true of many observational epidemiological studies, our study has some limitations with data availability and study design. The Bayesian

spatiotemporal models that have been used in chapter 2 and 3 used reported outbreak data, which may not accurately represent all outbreaks within a country. Apart from the research studies where active sampling is performed, both countries rely on passive surveillance to report outbreaks. According to the PVS analysis ("25022019_India_PVS_Evaluation_report_final.pdf," n.d.), India has a paper-based system in place. In Vietnam, there is a need of establishing a comprehensive national database for passive surveillance (Lee et al., 2020) and there is no compliance of notification of disease outbreaks. The veterinarians conduct a passive surveillance programs, and samples are collected and sent to the laboratories. Thus, reported outbreak data are coming from passive surveillance and may depend on many factors such as the quality of the veterinary service and the established disease reporting and surveillance systems in the country. This could be a particular issue if the reliability or completeness of outbreak reporting varies across regions or years. In addition, for India, we did not have the number of infected animals per outbreak.

We were also not able to account for structured spatial effects in border state/provinces since we do not have outbreak data from the provinces in the surrounding countries. This could influence our results by introducing edge effects. For both the Vietnam and India models, we did introduce a fixed effect representing whether a province/state had an international border and found that relative risk was generally increased by these borders. However, we could not match case counts in such states/provinces with outbreaks occurring in the border regions of the neighboring country. For the Vietnam model, this may be somewhat mitigated by the use of the phylogeographic adjacency matrix in place of spatial adjacency matrix, as this more explicitly accounts for evidence of viral movement from other countries into specific states/provinces, but our ability to uncover such linkages is impacted by sequence data availability in both countries.

For both studies, there were limitations related to the spatial scale of available data. Based on the available data, we had to focus on macro-scale risk factors such as environmental factors, host densities, international borders, etc. However, analyzing these factors at a province/state level likely introduced some degree of ecological fallacy. As an example, host densities were summarized at the province/state, but there is likely substantial heterogeneity in host density at smaller scales; thus, not all hosts equally were equally impacted by a particular risk factor. This may be one reason that many of the assessed risk factors were not significant, particularly for the India model. However, the lack of significance at the course scale of our analysis cannot be taken as a true lack of importance of a particular risk factor given the mismatch in the scale of the analysis to the scale of affect. In addition, the scale of analysis also did not allow us to assess factors that would be important for risk at a farm-level. As with any contagious disease, adhering to the appropriate biosecurity measures at the farm-level would be important for minimizing risk.

We did not consider any biosecurity measures or other farm-level factors due to unavailability of data and the course scale of the analysis. Conducting farmer-based questionnaires and sampling would allow for a more finer scale analysis of risk factors. It would also be beneficial for both countries to report FMD outbreaks at a smaller administrative division level, which could potentially provide a better identification of spatial risk factors and avoid potential ecological fallacies.

Chapter 2 limitations: In India, outbreak reports did not include number of animals affected. In addition, outbreak numbers were reported annually at a state level. Thus, this analysis may have ecological fallacies, as described more generally above. India is a huge country and states consist of many districts. While some areas within a state may have high livestock populations, some areas may be highly urban. Outbreak numbers summarized at the whole state would not account for the fact that some districts may have zero outbreaks. Even if we see an association between outbreaks and a risk factor at the state level, this may not be true for individual districts. However, we believe that this would result in a failure to detect the importance of some factors (Type II error) rather than the incorrect attribution of risk (Type I error). For example, none of the environmental factors were included in the final model, likely because the inclusion of a yearly average summarized across the state masked much of the spatiotemporal variability that may have impacted transmission on a smaller spatial and shorter time scale. We tried to further mitigate this in Chapter 3 by not using environmental risk factors.

Finally, at the time of this study, the most current livestock census in India was conducted in 2012. When calculating the SIR, we had to use the same value from 2012 for all the years from 2008-2016. In 2019, the new census data became available to the public, and livestock numbers across the two censuses were not substantially different. We also did not have animal movement data, but we did include potential proxy parameters such as road density and the presence of grazing areas. In Chapter 3, we attempted to overcome this limitation by including more well-supported proxies for host movement based on phylogeographic models.

Chapter 3 limitations: For the phylogeographic analysis, we inferred FMDV movement based on the phylogeographic trait analysis. However, there was no data available to substantiate that FMDV movement represent the animal movement except using questionnaires, participatory approach with the stakeholders (Polly et al., 2013), though one could argue that viral movement by any means (movement of animals, fomites, or contaminated products) would be relevant to include in space-time models of outbreak data. It may be more appropriate to conceptualize the rates of phylogeographic movement as representing viral population connectivity as opposed to host population connectivity.

For the discrete space phylogeographic analysis, we used all serotype O sequences available from Vietnam and surrounding countries. Some of these sequences are from active surveillance and may be from carrier animals. There is still uncertainty in regards to how subclinical infections relate to clinical outbreaks in our phylogeographic models, since limited evidence exists that that subclinical infections spawn FMD outbreaks (Bertram et al., 2018). That being said, sequences from subclinical infections may not represent true persistent infections (so-called carriers) but rather acute subclinical infections or recently convalescent animals. Furthermore, the presence of such animals in our analysis does not mean that they played any role in onward transmission; rather, carrier animals may simply harbor persistent infections with viruses that were previously circulating and transmitting in an acute form.

In Vietnam, pigs account for the majority of livestock, but limited sequence data and outbreak data were available for pigs compared to cattle and buffaloes. We did not find strong evidence for pig-to-bovine transmission, which we used as rationale to focus solely on bovine outbreaks, but limited data may have affected this result. A model that focuses particularly on pigs would be of interest to Vietnam.

For adjacent countries, GenBank data did not include the exact location where a FMDV sequence was isolated. Therefore, the center point of the country was used as the origin of the sequence. Within Vietnam, sequence data were available for only some of the provinces. Sequence data available in GenBank for both Vietnam and adjacent countries could be the result of purposive sampling conducted as a part of different research studies, making these regions over-represented in our analysis and consequently overestimating the role of these regions in viral population dynamics. That being said, many research studies (including ours) focus on areas that are considered disease hotspots, which means that the greater number of sequences available from such regions might actually reflect greater disease circulation, though we cannot confirm this.

Chapter 4 limitations: Data for chapter 4 are from a longitudinal study sampling farms and a serial cross-sectional study sampling slaughterhouse. Both studies were purposive sampling, meaning that the study areas were selected to be areas with known recent FMDV outbreaks and transboundary animal movement. To ideally represent what is happening in the population, it would have been appropriate to select farms randomly for the longitudinal study or to include slaughterhouses from multiple FMD hotspots across the country. For example, both northern and southern Vietnam have been identified as high risk areas (Lee et al., 2020), however our study is more reflective of southern Vietnam (both slaughterhouses were located in this part of the

country) rather than representing the whole country. That being said, this study demonstrated the potential utility of slaughterhouse-based surveillance, and a broader network of slaughterhouses could be employed as part of a national-level surveillance strategy.

5.4 Moving from Science to Policy and Practice

This study was aimed at demonstrating how the epidemiological data can support FMD endemic countries to progress in PCP at different stages. For the findings of these particular studies, the next step would be to convey our findings to relevant animal health authorities and policymakers. Through discussions with these stakeholders, findings from this study become more valuable if they are adjusted according to the input from the stakeholders in the respective countries. One step we have taken to ensure that these results reach relevant animal health authorities is the inclusion of collaborators and co-authors from each country. However, as true for most epidemiological research, efforts should be made to not just publish in scientific journals, but also to disseminate findings more locally, for example by contributing summary reports to ICAR, India and MARD, Vietnam. In addition, one of the purposes of this dissertation is to outline possible epidemiological approaches that can be taken to support PCP. Possible mechanisms to disseminate these approaches more broadly to stakeholders engaged in PCP planning include online programs conducted by EuFMD that are focused on capacity building for veterinarians participating from FMD endemic countries. With a more generalizable approach, findings from this study can be incorporated as a part of such outreach activities, including as part of their online education platform. There are other programs, at least conducted in Europe (ex: at -risk program) (Kostoulas et al., 2019), to train field veterinarians with user friendly tools that have been developed to incorporate Bayesian analysis methods to determine FMD risk. As this study specifically focus on FMD endemic countries, such methods can be expanded to FMD endemic countries by training veterinarians with a similar approached utilized in this research. Another method to share our findings and epidemiological approaches is by presenting to the FMD research community by participating in the research conferences for FMD, such as the Global FMD Research Alliance (GFRA) and open session of the EuFMD, both of which attract many researchers and professionals from FMD endemic countries meet. Taken together, these steps are necessary to move epidemiological approaches in support of FMD control from the academic research arena to policy and practice.

Bibliography

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data 5, 170191. https://doi.org/10.1038/sdata.2017.191

Adamchick, J., Rich, K.M., Perez, A.M., 2021. Self-reporting of risk pathways and parameter values for foot and mouth disease in slaughter cattle from alternative production systems by Kenyan and Ugandan veterinarians.

Alexandersen, S., Zhang, Z., Donaldson, A.I., Garland, A.J.M., 2003. The pathogenesis and diagnosis of foot-and-mouth disease. J Comp Pathol 129, 1–36. https://doi.org/10.1016/s0021-9975(03)00041-0

Arjkumpa, O., Sansamur, C., Sutthipankul, P., Inchaisri, C., Na Lampang, K., Charoenpanyanet, A., Punyapornwithaya, V., 2020. Spatiotemporal analyses of foot and mouth disease outbreaks in cattle farms in Chiang Mai and Lamphun, Thailand. BMC Vet Res 16, 170. https://doi.org/10.1186/s12917-020-02392-6

Arguello, H., Alvarez-Ordoñez, A., Carvajal, A., Rubio, P., Prieto, M., 2013. Role of slaughtering in Salmonella spreading and control in pork production. J. Food Prot. 76, 899–911. https://doi.org/10.4315/0362-028X.JFP-12-404

Armson, B., Gubbins, S., Mioulet, V., Qasim, I.A., King, D.P., Lyons, N.A., 2020. Foot-and-Mouth Disease Surveillance Using Pooled Milk on a Large-Scale Dairy Farm in an Endemic Setting. Front. Vet. Sci. 7. https://doi.org/10.3389/fvets.2020.00264

Arzt, Jonathan, Baxt, B., Grubman, M.J., Jackson, T., Julleff, N., Rhyan, J., Rieder, E., Waters, R., Rodriguez, L.L., 2011. The Pathogenesis of Foot-and-Mouth Disease II: Viral Pathways in Swine, Small Ruminants, and Wildlife; Myotropism, Chronic Syndromes, and Molecular Virus–Host Interactions. Transboundary and Emerging Diseases.

Arzt, J., Fish, I., Pauszek, S.J., Johnson, S.L., Chain, P.S., Rai, D.K., Rieder, E., Goldberg, T.L., Rodriguez, L.L., Stenfeldt, C., 2019. The evolution of a super-swarm of foot-and-mouth disease virus in cattle [WWW Document]. PLoS ONE. URL https://link.galegroup.com/apps/doc/A583563657/AONE?sid=lms (accessed 9.18.19).

Arzt, J., Brito, B., Pauszek, S.J., Hartwig, E.J., Smoliga, G.R., Vu, L.T., Vu, P.P., Stenfeldt, C., Rodriguez, L.L., Long, N.T., Dung, D.H., 2017. Genome Sequence of Foot-and-Mouth Disease Virus Serotype O Lineage Ind-2001d Collected in Vietnam in 2015. Genome Announc 5, e00223-17. https://doi.org/10.1128/genomeA.00223-17

Arzt, J., Juleff, N., Zhang, Z., Rodriguez, L.L., 2011. The pathogenesis of foot-and-mouth disease I: viral pathways in cattle. Transbound Emerg Dis 58, 291–304. https://doi.org/10.1111/j.1865-1682.2011.01204.x

Bachanek-Bankowska, K., Di Nardo, A., Wadsworth, J., Mioulet, V., Pezzoni, G., Grazioli, S., Brocchi, E., Kafle, S.C., Hettiarachchi, R., Kumarawadu, P.L., Eldaghayes, I.M., Dayhum, A.S., Meenowa, D., Sghaier, S., Madani, H., Abouchoaib, N., Hoang, B.H., Vu, P.P., Dukpa, K., Gurung, R.B., Tenzin, S., Wernery, U., Panthumart, A., Seeyo, K.B., Linchongsubongkoch, W., Relmy, A., Bakkali-Kassimi, L., Scherbakov, A., King, D.P., Knowles, N.J., 2018. Reconstructing the evolutionary history of pandemic foot-and-mouth disease viruses: the impact of recombination within the emerging O/ME-SA/Ind-2001 lineage. Sci Rep 8, 14693. https://doi.org/10.1038/s41598-018-32693-

Bartels, C., Afonso, J., Sieng, S., McLaws, M., n.d. Risk analysis on incursion of exotic FMD viruses into South-East Asia 59

Baele, G., Lemey, P., Suchard, M.A., 2016. Genealogical Working Distributions for Bayesian Model Testing with Phylogenetic Uncertainty. Syst Biol 65, 250–264. https://doi.org/10.1093/sysbio/syv083

Berliner, L.M., 1996. Hierarchical Bayesian Time Series Models, in: Hanson, K.M., Silver, R.N. (Eds.), Maximum Entropy and Bayesian Methods, Fundamental Theories of Physics. Springer Netherlands, Dordrecht, pp. 15–22. https://doi.org/10.1007/978-94-011-5430-7_3

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 43, 1–20. https://doi.org/10.1007/BF00116466

Bertram, M.R., Palinski, R.M., Ranjan, R., Biswal, J.K., Pauszek, S.J., Hartwig, E.J., Smoliga, G.R., Fish, I.H., Vierra, D., Subramaniam, S., Mohapatra, J.K., Das, B., Pattnaik, B., Arzt, J., 2019. Genome Sequences of 18 Foot-and-Mouth Disease Virus Outbreak Strains of Serotype O Sublineage Ind2001d from India, 2013 to 2014. Microbiol Resour Announc 8. https://doi.org/10.1128/MRA.00776-19

Bertram, M.R., Brito, B., Palinski, R.M., Fish, I.H., Pauszek, S.J., Hartwig, E.J., Smoliga, G.R., Vu, L.T., Hoang, B.H., Phuong, N.T., Hung, V.V., Vu, P.P., Dung, N.K., Tien, N.N., Dong, P.V., Dung, D.H., Arzt, J., 2021. Novel Recombinant Foot-and-Mouth Disease Virus Circulating in Vietnam. Microbiol Resour Announc 10. https://doi.org/10.1128/MRA.01263-20

Bertram, M.R., Vu, L.T., Pauszek, S.J., Brito, B.P., Hartwig, E.J., Smoliga, G.R., Hoang, B.H., Phuong, N.T., Stenfeldt, C., Fish, I.H., Hung, V.V., Delgado, A., VanderWaal, K., Rodriguez, L.L., Long, N.T., Dung, D.H., Arzt, J., 2018. Lack of Transmission of Foot-and-Mouth Disease Virus From Persistently Infected Cattle to Naïve Cattle Under Field Conditions in Vietnam. Front. Vet. Sci. 5. https://doi.org/10.3389/fvets.2018.00174

Bertram, M.R., Bravo de Rueda, C., Garabed, R., Dickmu Jumbo, S., Moritz, M., Pauszek, S., Abdoulkadiri, S., Rodriguez, L.L., Arzt, J., 2018a. Molecular Epidemiology of Foot-and-Mouth Disease Virus in the Context of Transboundary Animal Movement in the Far North Region of Cameroon. Front. Vet. Sci. 5. https://doi.org/10.3389/fvets.2018.00320

Bertram, M.R., Yadav, S., Stenfeldt, C., Delgado, A., Arzt, J., 2020. Extinction Dynamics of the Foot-and-Mouth Disease Virus Carrier State Under Natural Conditions. Front Vet Sci 7, 276. https://doi.org/10.3389/fvets.2020.00276

Biswal, J.K., Ranjan, R., Subramaniam, S., Mohapatra, J.K., Patidar, S., Sharma, M.K., Bertram, M.R., Brito, B., Rodriguez, L.L., Pattnaik, B., Arzt, J., 2019. Genetic and antigenic variation of foot-and-mouth disease virus during persistent infection in naturally infected cattle and Asian buffalo in India. PLoS ONE 14, e0214832. https://doi.org/10.1371/journal.pone.0214832

Biswal, J.K., Bisht, P., Subramaniam, S., Ranjan, R., Sharma, G.K., Pattnaik, B., 2015. Engineering foot-and-mouth disease virus serotype O IND R2/1975 for one-step purification by immobilized metal affinity chromatography. Biologicals 43, 390–398. https://doi.org/10.1016/j.biologicals.2015.06.001

Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-temporal Bayesian Models with R - INLA. John Wiley & Sons.

Bertram, M.R., Bravo de Rueda, C., Garabed, R., Dickmu Jumbo, S., Moritz, M., Pauszek, S., Abdoulkadiri, S., Rodriguez, L.L., Arzt, J., 2018. Molecular Epidemiology of Foot-and-Mouth Disease Virus in the Context of Transboundary Animal Movement in the Far North Region of Cameroon. Front Vet Sci 5, 320. https://doi.org/10.3389/fvets.2018.00320

Branscum, A.J., Perez, A.M., Johnson, W.O., Thurmond, M.C., 2008. Bayesian spatiotemporal analysis of foot-and-mouth disease data from the Republic of Turkey. Epidemiology & Infection 136, 833–842. https://doi.org/10.1017/S0950268807009065

Bravo de Rueda, C., de Jong, M.C., Eblé, P.L., Dekker, A., 2015. Quantification of transmission of foot-and-mouth disease virus caused by an environment contaminated with secretions and excretions from infected calves. Vet Res 46, 43. https://doi.org/10.1186/s13567-015-0156-5

Brito, B., Pauszek, S.J., Eschbaumer, M., Stenfeldt, C., de Carvalho Ferreira, H.C., Vu, L.T., Phuong, N.T., Hoang, B.H., Tho, N.D., Dong, P.V., Minh, P.Q., Long, N.T., King, D.P., Knowles, N.J., Dung, D.H., Rodriguez, L.L., Arzt, J., 2017. Phylodynamics of foot-and-mouth disease virus O/PanAsia in Vietnam 2010–2014. Veterinary Research 48, 24. https://doi.org/10.1186/s13567-017-0424-7 1.

Brito, B.P., Mohapatra, J.K., Subramaniam, S., Pattnaik, B., Rodriguez, L.L., Moore, B.R., Perez, A.M., 2018. Dynamics of widespread foot-and-mouth disease virus serotypes A, O and Asia-1 in southern Asia: A Bayesian phylogenetic perspective. Transbound Emerg Dis 65, 696–710. https://doi.org/10.1111/tbed.12791

Brito, B.P., Perez, A.M., Cosentino, B., Rodriguez, L.L., König, G.A., 2011. Factors associated with within-herd transmission of serotype A foot-and-mouth disease virus in cattle, during

the 2001 outbreak in Argentina: a protective effect of vaccination. Transbound Emerg Dis 58, 387–393. https://doi.org/10.1111/j.1865-1682.2011.01217.x
Brito, B.P., Rodriguez, L.L., Hammond, J.M., Pinto, J., Perez, A.M., 2017b. Review of the Global Distribution of Foot-and-Mouth Disease Virus from 2007 to 2014. Transboundary and Emerging Diseases 64, 316–332. https://doi.org/10.1111/tbed.12373
Bonardi, S., 2017. Salmonella in the pork production chain and its impact on human health in the European Union. Epidemiology & Infection 145, 1513–1526. https://doi.org/10.1017/S095026881700036X
Bronsvoort, B.M. deC, Handel, I.G., Nfon, C.K., Sørensen, K.-J., Malirat, V., Bergmann, I., Tanya, V.N., Morgan, K.L., 2016. Redefining the "carrier" state for foot-and-mouth disease from the dynamics of virus persistence in endemically affected cattle populations. Scientific Reports 6, 29059. https://doi.org/10.1038/srep29059
Cameron, A.R., Meyer, A., Faverjon, C., Mackenzie, C., 2020. Quantification of the sensitivity of early detection surveillance. Transbound Emerg Dis. https://doi.org/10.1111/tbed.13598
Chen, J., Wang, J., Wang, M., Liang, R., Lu, Y., Zhang, Q., Chen, Q., Niu, B., 2020. Retrospect and Risk Analysis of Foot-and-Mouth Disease in China Based on Integrated Surveillance and Spatial Analysis Tools. Front Vet Sci 6, 511. https://doi.org/10.3389/fvets.2019.00511
Chhetri, B.K., Perez, A.M., Thurmond, M.C., 2010. Factors associated with spatial clustering of foot-and-mouth disease in Nepal. Trop Anim Health Prod 42, 1441–1449. https://doi.org/10.1007/s11250-010-9573-7
Coly, S., Garrido, M., Abrial, D., Yao, A.-F., 2021. Bayesian hierarchical models for disease mapping applied to contagious pathologies. PLOS ONE 16, e0222898. https://doi.org/10.1371/journal.pone.0222898
Dahiya, S.S., Subramaniam, S., Biswal, J.K., Das, B., Prusty, B.R., Ali, S.Z., Khulape, S.A., Mohapatra, J.K., Singh, R.K., 2020. Genetic characterization of foot-and-mouth disease virus serotype O isolates collected during 2014-2018 revealed dominance of O/ME-SA/Ind2001e and the emergence of a novel lineage in India. Transbound Emerg Dis. https://doi.org/10.1111/tbed.13954
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9, 772–772. https://doi.org/10.1038/nmeth.2109
de Carvalho Ferreira, H.C., Pauszek, S.J., Ludi, A., Huston, C.L., Pacheco, J.M., Le, V.T., Nguyen, P.T., Bui, H.H., Nguyen, T.D., Nguyen, T., Nguyen, T.T., Ngo, L.T., Do, D.H., Rodriguez, L., Arzt, J., 2017. An Integrative Analysis of Foot-and-Mouth Disease Virus Carriers in Vietnam Achieved Through Targeted Surveillance and Molecular Epidemiology. Transbound Emerg Dis 64, 547–563. https://doi.org/10.1111/tbed.12403
Di Nardo, A., Ferretti, L., Wadsworth, J., Mioulet, V., Gelman, B., Karniely, S., Scherbakov, A., Ziay, G., Özyörük, F., Parlak, Ü., Tuncer-Göktuna, P., Hassanzadeh, R., Khalaj, M., Dastoor, S.M., Abdollahi, D., Khan, E.-H., Afzal, M., Hussain, M., Knowles, N.J., King, D.P., 2021. Evolutionary and Ecological Drivers Shape the Emergence and Extinction of Foot-and-Mouth Disease Virus Lineages. Molecular Biology and Evolution. https://doi.org/10.1093/molbev/msab172
Di Nardo, A., Knowles, N.J., 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. Rev Sci Tech 30, 63–85. https://doi.org/10.20506/rst.30.1.2022
Donaldson, A.I., 1972. The influence of relative humidity on the aerosol stability of different strains of foot-and-mouth disease virus suspended in saliva. J. Gen. Virol. 15, 25–33. https://doi.org/10.1099/0022-1317-15-1-25
Dohoo, I., Martin, S., Stryhn, H., 2009. Veterinary Epidemiologic Research.
Dukpa, K., Robertson, I., Edwards, J., Ellis, T., Tshering, P., Rinzin, K., Dahal, N., 2011. Risk factors for foot-and-mouth disease in sedentary livestock herds in selected villages in four regions of Bhutan. New Zealand Veterinary Journal 59, 51–58. https://doi.org/10.1080/00480169.2011.552852
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Elena, S.F., Sanjuán, R., 2005. Adaptive Value of High Mutation Rates of RNA Viruses: Separating Causes from Consequences. Journal of Virology 79, 11555–11558. https://doi.org/10.1128/JVI.79.18.11555-11558.2005

Farooq, U., Ahmed, Z., Naeem, K., Bertram, M., Brito, B., Stenfeldt, C., Pauszek, S.J., LaRocco, M., Rodriguez, L., Arzt, J., 2018. Characterization of naturally occurring, new and persistent subclinical foot-and-mouth disease virus infection in vaccinated Asian buffalo in Islamabad Capital Territory, Pakistan. Transboundary and Emerging Diseases 65, 1836–1850. https://doi.org/10.1111/tbed.12963

Fertner, M., Denwood, M., Birkegård, A.C., Stege, H., Boklund, A., 2017. Associations between Antibacterial Treatment and the Prevalence of Tail-Biting-Related Sequelae in Danish Finishers at Slaughter. Front. Vet. Sci. 4. https://doi.org/10.3389/fvets.2017.00182

Fournié, G., Waret-Szkuta, A., Camacho, A., Yigezu, L.M., Pfeiffer, D.U., Roger, F., 2018. A dynamic model of transmission and elimination of peste des petits ruminants in Ethiopia. Proc Natl Acad Sci U S A 115, 8454–8459. https://doi.org/10.1073/pnas.1711646115

Fuglstad, G.-A., Simpson, D., Lindgren, F., Rue, H., 2019. Constructing Priors that Penalize the Complexity of Gaussian Random Fields. Journal of the American Statistical Association 114, 445–452. https://doi.org/10.1080/01621459.2017.1415907

G, G., B, G.K., A, K., Hegde, R., Kumar, N., Prabhakaran, K., Wadhwan, V.M., Kakker, N., Lokhande, T., Sharma, K., Kanani, A., Limaye, null, K, N., Pn, A., De, A.K., Khan, T.A., Misri, J., Dash, B.B., Pattnaik, B., Habibur, R., 2021. Foot and Mouth Disease (FMD) incidence in cattle and buffaloes and its associated farm-level economic costs in endemic India. Prev Vet Med 190, 105318. https://doi.org/10.1016/j.prevetmed.2021.105318

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

Gupta, S., Ferguson, N., Anderson, R., 1998. Chaos, Persistence, and Evolution of Strain Structure in Antigenically Diverse Infectious Agents. Science 280, 912–915. https://doi.org/10.1126/science.280.5365.912

Gunasekera, U.C., Sivasothy, A., Wedasingha, N., Thayaparan, S., Rotewewa, B., Muralithas, M., Baumann, M.P.O., Punyapornwithaya, V., 2017. Analyzing the Foot and Mouth Disease outbreak as from 2008 to 2014 in cattle and buffaloes in Sri Lanka. Prev. Vet. Med. 148, 78–88. https://doi.org/10.1016/j.prevetmed.2017.10.008

Gunasekara, U., Bertram, M.R., Dung, D.H., Hoang, B.H., Phoung, N.T., Hung, V.V., Long, N.V., Phan, M., Vu, L.T., Dong, P.V., Perez, A., VanderWaal, K., Artz, J., 2021. Use of Slaughterhouses as Sentinel Points for Genomic Surveillance of Subclinical Foot-and-mouth Disease Virus in Vietnam.

Hayer, S.S., VanderWaal, K., Ranjan, R., Biswal, J.K., Subramaniam, S., Mohapatra, J.K., Sharma, G.K., Rout, M., Dash, B.B., Das, B., Prusty, B.R., Sharma, A.K., Stenfeldt, C., Perez, A., Delgado, A.H., Sharma, M.K., Rodriguez, L.L., Pattnaik, B., Arzt, J., 2018. Foot-and-mouth disease virus transmission dynamics and persistence in a herd of vaccinated dairy cattle in India. Transbound Emerg Dis 65, e404–e415. https://doi.org/10.1111/tbed.12774

H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models using inte-grated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society, Series B,71(2):319{392, 2009.

Hegde, R., Gomes, A.R., Giridhar, P., Kowalli, S., Shivashankar, B.P., Sudharshana, K.J., Nagaraj, K., Sesharao, R., Mallinath, K.C., Shankar, B.P., Nagaraj, D., Seema, C.M., Khan, T.A., Nagaraj, G.V., Srikala, K., Dharanesh, N.K., Venkatesha, M.D., Renukaprasad, C., 2014. Epidemiology of foot and mouth disease in Karnataka state, India: a retrospective study. VirusDis. 25, 504–509. https://doi.org/10.1007/s13337-014-0239-3

Hemme, T., Garcia, O., Saha, A., 2003. A Review of Milk Production in India with Particular Emphasis on Small-Scale Producers (PPLPI Working Paper No. 23777). Food and Agriculture Organization of the United Nations, Pro-Poor Livestock Policy Initiative.

Henritzi, D., Petric, P.P., Lewis, N.S., Graaf, A., Pessia, A., Starick, E., Breithaupt, A., Strebelow, G., Luttermann, C., Parker, L.M.K., Schröder, C., Hammerschmidt, B., Herrler, G., Beilage, E.G., Stadlbauer, D., Simon, V., Krammer, F., Wacheck, S., Pesch, S., Schwemmle, M., Beer, M., Harder, T.C., 2020. Surveillance of European Domestic Pig Populations Identifies an Emerging Reservoir of Potentially Zoonotic Swine Influenza A Viruses. Cell Host Microbe 28, 614-627.e6. https://doi.org/10.1016/j.chom.2020.07.006

Innocent, G.T., Gilbert, L., Jones, E.O., McLeod, J.E., Gunn, G., McKendrick, I.J., Albon, S.D., 2017. Combining Slaughterhouse Surveillance Data with Cattle Tracing Scheme and Environmental Data to Quantify Environmental Risk Factors for Liver Fluke in Cattle. Front. Vet. Sci. 4. https://doi.org/10.3389/fvets.2017.00065

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11, 94. https://doi.org/10.1186/1471-2156-11-94

Kao, S.-Y.Z., VanderWaal, K., Enns, E.A., Craft, M.E., Alvarez, J., Picasso, C., Wells, S.J., 2018. Modeling cost-effectiveness of risk-based bovine tuberculosis surveillance in Minnesota. Prev Vet Med 159, 1–11. https://doi.org/10.1016/j.prevetmed.2018.08.011

Knowles, N.J., Samuel, A.R., Davies, P.R., Midgley, R.J., Valarcher, J.-F., 2005. Pandemic Strain of Foot-and-Mouth Disease Virus Serotype O. Emerg Infect Dis 11, 1887–1893. https://doi.org/10.3201/eid1112.050908

Knight-Jones, T.J.D., Gubbins, S., Bulut, A.N., Stärk, K.D.C., Pfeiffer, D.U., Sumption, K.J., Paton, D.J., 2016. Mass vaccination, immunity and coverage: modelling population protection against foot-and-mouth disease in Turkish cattle. Scientific Reports 6, 22121. https://doi.org/10.1038/srep22121

Knight-Jones, T.J.D., Rushton, J., 2013. The economic impacts of foot and mouth disease - what are they, how big are they and where do they occur? Prev Vet Med 112, 161–173. https://doi.org/10.1016/j.prevetmed.2013.07.013

Knowles, N.J., Bachanek-Bankowska, K., Wadsworth, J., Mioulet, V., Valdazo-González, B., Eldaghayes, I.M., Dayhum, A.S., Kammon, A.M., Sharif, M.A., Waight, S., Shamia, A.M., Tenzin, S., Wernery, U., Grazioli, S., Brocchi, E., Subramaniam, S., Pattnaik, B., King, D.P., 2016. Outbreaks of Foot-and-Mouth Disease in Libya and Saudi Arabia During 2013 Due to an Exotic O/ME-SA/Ind-2001 Lineage Virus. Transbound Emerg Dis 63, e431-435. https://doi.org/10.1111/tbed.12299

Kostoulas, P., Giovannini, A., Alba, A., Comin, A., Meletis, E., Iannetti, S., Napp, S., Lindberg, A., Solomakos, N., 2019. Harmonisation Of Transmissible disease Interpretation in the EU (HOTLINE). EFSA Supporting Publications 16, 1678E. https://doi.org/10.2903/sp.efsa.2019.EN-1678

Kumar, A., 2012. Kumar, Anjani, Shiv Jee and Chitra Yadav (2012). Export of Buffalo Meat in India: Performance and Prospects. The Indian Journal of Animal Sciences. 82(12):1578-83. The Indian journal of animal sciences 82, 1578–83.

Lawson, A.B., 2018. Bayesian latent modeling of spatio-temporal variation in small-area health data. WIREs Computational Statistics 10, e1441. https://doi.org/10.1002/wics.1441

Le, V.P., Vu, T.T.H., Duong, H.-Q., Than, V.T., Song, D., 2016. Evolutionary phylodynamics of foot-and-mouth disease virus serotypes O and A circulating in Vietnam. BMC Vet Res 12. https://doi.org/10.1186/s12917-016-0896-0

L, M., es, Melton, A., Edwards, S., n.d. From Where the Buffalo Roam: India's Beef Exports [WWW Document]. URL http://www.ers.usda.gov/publications/pub-details/?pubid=37673 (accessed 12.24.20).

Lee, H.S., Pham, T.L., Wieland, B., 2020. Temporal patterns and space-time cluster analysis of foot-and-mouth disease (FMD) cases from 2007 to 2017 in Vietnam. Transbound Emerg Dis 67, 584–591. https://doi.org/10.1111/tbed.13370

Machado, G., Korennoy, F., Alvarez, J., Picasso-Risso, C., Perez, A., VanderWaal, K., 2019. Mapping changes in the spatiotemporal distribution of lumpy skin disease virus. bioRxiv 531343. https://doi.org/10.1101/531343

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution 1. https://doi.org/10.1093/ve/vev003
Mahapatra, M., Yuvaraj, S., Madhanmohan, M., Subramaniam, S., Pattnaik, B., Paton, D.J., Srinivasan, V.A., Parida, S., 2015. Antigenic and genetic comparison of foot-and-mouth disease virus serotype O Indian vaccine strain, O/IND/R2/75 against currently circulating viruses. Vaccine 33, 693–700. https://doi.org/10.1016/j.vaccine.2014.11.058
Makau, D.N., Alkhamis, M.A., Paploski, I. a. D., Corzo, C.A., Lycett, S., VanderWaal, K., 2021. Integrating animal movements with phylogeography to model the spread of PRRSV in the USA. Virus Evolution. https://doi.org/10.1093/ve/veab060
Miller, M., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE). https://doi.org/10.1109/GCE.2010.5676129
Mikkelsen, T., Alexandersen, S., Astrup, P., Champion, H.J., Donaldson, A.I., Dunkerley, F.N., Gloster, J., Sørensen, J.H., Thykier-Nielsen, S., 2003. Investigation of airborne foot-and-mouth disease virus transmission during low-wind conditions in the early phase of the UK 2001 epidemic. Atmospheric Chemistry and Physics 3, 2101–2110. https://doi.org/10.5194/acp-3-2101-2003
Mohanty, N.N., Subramaniam, S., Rout, M., Sarangi, L.N., Bisht, P., Pandey, L.K., Mohapatra, J.K., Panda, H.K., 2015. Serosurveillance of foot-and-mouth disease in ruminant population of Coastal Odisha, India. Beni-Suef University Journal of Basic and Applied Sciences 4, 279–283. https://doi.org/10.1016/j.bjbas.2015.11.002
Mohapatra, J.K., Pandey, L.K., Sanyal, A., Pattnaik, B., 2011. Recombinant non-structural polyprotein 3AB-based serodiagnostic strategy for FMD surveillance in bovines irrespective of vaccination. J. Virol. Methods 177, 184–192. https://doi.org/10.1016/j.jviromet.2011.08.006
Munsey, A., Mwiine, F., Ochwo, S., Velazquez-Salinas, L., Ahmed, Z., Maree, F., Rodriguez, L., Rieder, E., Perez, A., Dellicour, S., VanderWaal, K., 2021. Phylogeographic analysis of foot-and-mouth disease virus serotype O dispersal and associated drivers in East Africa. Molecular Ecology 30. https://doi.org/10.1111/mec.15991
Muroga, N., Kobayashi, S., Nishida, T., Hayama, Y., Kawano, T., Yamamoto, T., Tsutsui, T., 2013. Risk factors for the transmission of foot-and-mouth disease during the 2010 outbreak in Japan: a case–control study. BMC Veterinary Research 9, 150. https://doi.org/10.1186/1746-6148-9-150
Muwonge, A., Bessell, P.R., Porphyre, T., Motta, P., Rydevik, G., Devailly, G., Egbe, N.F., Kelly, R.F., Handel, I.G., Mazeri, S., Bronsvoort, B.M. deC, 2021. Inferring livestock movement networks from archived data to support infectious disease control in developing countries. https://doi.org/10.1101/2021.03.18.435930
Murray, J., Cohen, A.L., 2017. Infectious Disease Surveillance. International Encyclopedia of Public Health 222–229. https://doi.org/10.1016/B978-0-12-803678-5.00517-8
Omondi, G.P., Gakuya, F., Arzt, J., Sangula, A., Hartwig, E., Pauszek, S., Smoliga, G., Brito, B., Perez, A., Obanda, V., VanderWaal, K., 2020. The role of African buffalo in the epidemiology of foot-and-mouth disease in sympatric cattle and buffalo populations in Kenya. Transboundary and Emerging Diseases 67, 2206–2221. https://doi.org/10.1111/tbed.13573
Pacheco, J.M., Arzt, J., Rodriguez, L.L., 2010. Early events in the pathogenesis of foot-and-mouth disease in cattle after controlled aerosol exposure. The Veterinary Journal 183, 46–53. https://doi.org/10.1016/j.tvjl.2008.08.023
Paton, D.J., Gubbins, S., King, D.P., 2018. Understanding the transmission of foot-and-mouth disease virus at different scales. Current Opinion in Virology, Emerging viruses: intraspecies transmission • Viral Immunology 28, 85–91. https://doi.org/10.1016/j.coviro.2017.11.013
Palinski, R.M., Bertram, M.R., Vu, L.T., Pauszek, S.J., Hartwig, E.J., Smoliga, G.R., Stenfeldt, C., Fish, I.H., Hoang, B.H., Phuong, N.T., Hung, V.V., Vu, P.P., Dung, N.K., Dong, P.V., Tien, N.N., Tho, N.D., Dung, D.H., Arzt, J., 2019. First Genome Sequence of Foot-and-

Mouth Disease Virus Serotype O Sublineage Ind2001e from Southern Vietnam. Microbiol Resour Announc 8. https://doi.org/10.1128/MRA.01424-18
Pattnaik, B., Subramaniam, S., Sanyal, A., Mohapatra, J.K., Dash, B.B., Ranjan, R., Rout, M., 2012. Foot-and-mouth Disease: Global Status and Future Road Map for Control and Prevention in India. Agric Res 1, 132–147. https://doi.org/10.1007/s40003-012-0012-z
Peterson, E., Remmenga, M., Hagerman, A.D., Akkina, J.E., 2017. Use of Temperature, Humidity, and Slaughter Condemnation Data to Predict Increases in Transport Losses in Three Classes of Swine and Resulting Foregone Revenue. Front. Vet. Sci. 4. https://doi.org/10.3389/fvets.2017.00067
Savin, M., Bierbaum, G., Hammerl, J.A., Heinemann, C., Parcina, M., Sib, E., Voigt, A., Kreyenschmidt, J., 2020. Antibiotic-resistant bacteria and antimicrobial residues in wastewater and process water from German pig slaughterhouses and their receiving municipal wastewater treatment plants. Science of The Total Environment 727, 138788. https://doi.org/10.1016/j.scitotenv.2020.138788
Pham-Thanh, L., Magnusson, U., Can-Xuan, M., Nguyen-Viet, H., Lundkvist, Å., Lindahl, J., 2020. Livestock Development in Hanoi City, Vietnam—Challenges and Policies. Front Vet Sci 7, 566. https://doi.org/10.3389/fvets.2020.00566
Polly, P., Bourgeois Lüthi, N., Huachun, L., Naing Oo, K., Phonvisay, A., Premashthira, S., Abila, R., Widders, P., Kukreja, K., Miller, C., 2015. Movement pathways and market chains of large ruminants in the Greater Mekong Sub-region. O.I.E (World Organisation for Animal Health). https://doi.org/10.20506/standz.2782
Polly, P.D., Lawing, A.M., Fabre, A.-C., Goswami, A., 2013. Phylogenetic Principal Components Analysis and Geometric Morphometrics. Hystrix It. J. Mamm. 24, 33–41. https://doi.org/10.4404/hystrix-24.1-6383
Ranabijuli, S., Mohapatra, J.K., Pandey, L.K., Rout, M., Sanyal, A., Dash, B.B., Sarangi, L.N., Panda, H.K., Pattnaik, B., 2010. Serological Evidence of Foot-and-Mouth Disease Virus Infection in Randomly Surveyed Goat Population of Orissa, India. Transboundary and Emerging Diseases 57, 448–454. https://doi.org/10.1111/j.1865-1682.2010.01161.x
Ranjan, R., Biswal, J.K., Subramaniam, S., Singh, K.P., Stenfeldt, C., Rodriguez, L.L., Pattnaik, B., Arzt, J., 2016. Foot-and-Mouth Disease Virus-Associated Abortion and Vertical Transmission following Acute Infection in Cattle under Natural Conditions. PLOS ONE 11, e0167163. https://doi.org/10.1371/journal.pone.0167163
Richardson, S., Thomson, A., Best, N., Elliott, P., 2004. Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. Environ Health Perspect 112, 1016–1025. https://doi.org/10.1289/ehp.6740
Riebler, A., Sørbye, S., Simpson, D., Rue, H., 2016. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Statistical methods in medical research. https://doi.org/10.1177/0962280216660421
Rue, H., Riebler, A., Sørbye, S.H., Illian, J.B., Simpson, D.P., Lindgren, F.K., 2017. Bayesian Computing with INLA: A Review. Annual Review of Statistics and Its Application 4, 395–421. https://doi.org/10.1146/annurev-statistics-060116-054045
Rweyemamu, M., Roeder, P., Mackay, D., Sumption, K., Brownlie, J., Leforban, Y., Valarcher, J.-F., Knowles, N.J., Saraiva, V., 2008. Epidemiological Patterns of Foot-and-Mouth Disease Worldwide. Transboundary and Emerging Diseases 55, 57–72. https://doi.org/10.1111/j.1865-1682.2007.01013.x
Rahman, A.K.M.A., Islam, S.K.S., Sufian, M.A., Talukder, M.H., Ward, M.P., Martínez-López, B., 2020. Foot-and-Mouth Disease Space-Time Clusters and Risk Factors in Cattle and Buffalo in Bangladesh. Pathogens 9, E423. https://doi.org/10.3390/pathogens9060423
Schirdewahn, F., Lentz, H.H.K., Colizza, V., Koher, A., Hövel, P., Vidondo, B., 2021. Early warning of infectious disease outbreaks on cattle-transport networks. PLoS One 16, e0244999. https://doi.org/10.1371/journal.pone.0244999
Shanafelt, D.W., Perrings, C.A., 2017. Foot and mouth disease: the risks of the international trade in live animals. Rev Sci Tech 36, 839–865. https://doi.org/10.20506/rst.36.3.2719
Shapshak P, Sinnott J, Somboonwit C, Kuhn J, Feldblyum T, Segal D. Seasonal and Pandemic Influenza Surveillance and Disease Severity Global Virology I - Identifying and Investigating Viral Diseases. 2015 May 761-789. PMCID: PMC7121762.

Sharma, G.K., Mohapatra, J.K., Mahajan, S., Matura, R., Subramaniam, S., Pattnaik, B., 2014. Comparative evaluation of non-structural protein-antibody detecting ELISAs for foot-and-mouth disease sero-surveillance under intensive vaccination. J. Virol. Methods 207, 22–28. https://doi.org/10.1016/j.jviromet.2014.06.022

Simpson, D.P., Rue, H., Martins, T.G., Riebler, A., Sørbye, S.H., 2014. Penalising model component complexity: A principled, practical approach to constructing priors.

Singh, B., Prasad, S., Sinha, D., Verma, M.R., 2013. Estimation of economic losses due to foot and mouth disease in India [WWW Document]. ResearchGate. URL https://www.researchgate.net/publication/270215319_Estimation_of_economic_losses_due_to_foot_and_mouth_disease_in_India (accessed 9.26.19).

Smith, P., Lüthi, N.B., Hua-chun, L., Oo, K., 2016. Movement pathways and market chains of large ruminants in the Greater Mekong Sub-region [WWW Document]. URL /paper/Movement-pathways-and-market-chains-of-large-in-the-Smith-L%C3%BCthi/503b0ec591f45a3defeccc5bf7fba8449873fc25 (accessed 12.24.20).

Souley Kouato, B., De Clercq, K., Abatih, E., Dal Pozzo, F., King, D.P., Thys, E., Marichatou, H., Saegerman, C., 2018. Review of epidemiological risk models for foot-and-mouth disease: Implications for prevention strategies with a focus on Africa. PLoS One 13, e0208296. https://doi.org/10.1371/journal.pone.0208296

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. J Royal Statistical Soc B 64, 583–639. https://doi.org/10.1111/1467-9868.00353

Stenfeldt, C., Pacheco, J.M., Smoliga, G.R., Bishop, E., Pauszek, S.J., Hartwig, E.J., Rodriguez, L.L., Arzt, J., 2016. Detection of Foot-and-mouth Disease Virus RNA and Capsid Protein in Lymphoid Tissues of Convalescent Pigs Does Not Indicate Existence of a Carrier State. Transboundary and Emerging Diseases 63, 152–164. https://doi.org/10.1111/tbed.12235

Stenfeldt, C., Arzt, J., 2020. The Carrier Conundrum; A Review of Recent Advances and Persistent Gaps Regarding the Carrier State of Foot-and-Mouth Disease Virus. Pathogens 9, 167. https://doi.org/10.3390/pathogens9030167

Subramaniam, S., Mohapatra, J.K., Sharma, G.K., Biswal, J.K., Ranjan, R., Rout, M., Das, B., Dash, B.B., Sanyal, A., Pattnaik, B., 2015. Evolutionary dynamics of foot-and-mouth disease virus O/ME-SA/Ind2001 lineage. Vet Microbiol 178, 181–189. https://doi.org/10.1016/j.vetmic.2015.05.015

Subramaniam, S., Mohapatra, J.K., Das, B., Sanyal, A., Pattnaik, B., 2015. Genetic and antigenic analysis of foot-and-mouth disease virus serotype O responsible for outbreaks in India during 2013. Infect Genet Evol 30, 59–64. https://doi.org/10.1016/j.meegid.2014.12.009

Thumbi, S.M., Njenga, M.K., Otiang, E., Otieno, L., Munyua, P., Eichler, S., Widdowson, M.-A., McElwain, T.F., Palmer, G.H., 2019. Mobile phone-based surveillance for animal disease in rural communities: implications for detection of zoonoses spillover. Philosophical Transactions of the Royal Society B: Biological Sciences 374, 20190020. https://doi.org/10.1098/rstb.2019.0020

Van Diep, N., Ngoc, T.T.B., Hoa, L.Q., Nga, B.T.T., Kang, B., Oh, J., Lan, N.T., Le, V.P., 2020. O/SEA/Mya-98 lineage foot-and-mouth disease virus was responsible for an extensive epidemic that occurred in late 2018 in Vietnam. Arch Virol 165, 2487–2493. https://doi.org/10.1007/s00705-020-04763-8

Vp, L., T, N., Kn, L., Yj, K., Hs, L., Vc, N., Td, M., Th, D., Sm, K., Is, C., Jh, P., 2010. Molecular characterization of serotype A foot-and-mouth disease viruses circulating in Vietnam in 2009. Vet Microbiol 144, 58–66. https://doi.org/10.1016/j.vetmic.2009.12.033

Vu, L.T., Long, N.T., Brito, B., Stenfeldt, C., Phuong, N.T., Hoang, B.H., Pauszek, S.J., Hartwig, E.J., Smoliga, G.R., Vu, P.P., Quang, L.T.V., Hung, V.V., Tho, N.D., Dong, P.V., Minh, P.Q., Bertram, M., Fish, I.H., Rodriguez, L.L., Dung, D.H., Arzt, J., 2017. First detection of foot-and-mouth disease virus O/Ind-2001d in Vietnam. PLoS One 12. https://doi.org/10.1371/journal.pone.0177361

Wajid, A., Chaudhry, M., Rashid, H.B., Gill, S.S., Halim, S.R., 2020. Outbreak investigation of foot and mouth disease in Nangarhar province of war-torn Afghanistan, 2014. Sci Rep 10, 13800. https://doi.org/10.1038/s41598-020-70489-x

Willeberg, P.W., McAloon, C.G., Houtsma, E., Higgins, I., Clegg, T.A., More, S.J., 2018. The Herd-Level Sensitivity of Abattoir Surveillance for Bovine Tuberculosis: Simulating the Effects of Current and Potentially Modified Meat Inspection Procedures in Irish Cattle. Front. Vet. Sci. 5. https://doi.org/10.3389/fvets.2018.00082

Xu, L., Hurtle, W., Rowland, J.M., Casteran, K.A., Bucko, S.M., Grau, F.R., Valdazo-González, B., Knowles, N.J., King, D.P., Beckham, T.R., McIntosh, M.T., 2013. Development of a universal RT-PCR for amplifying and sequencing the leader and capsid-coding region of foot-and-mouth disease virus. J Virol Methods 189, 70–76. https://doi.org/10.1016/j.jviromet.2013.01.009

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28–36. https://doi.org/10.1111/2041-210X.12628

Rambaut A, Drummond A (2007) Tracer version 1.7.1, Computer program and documentation distributed by the author. http://tree.bio.ed.ac.uk/software/tracer/. Accessed 12/21/20Rambault A (2006–2009) Fig Tree Tree Figure Drawing Tool, Version 1.4.4 http://tree.bio.ed.ac.uk/software/figtree/. Accessed 12/21/20.

Endorsed official control programme for FMD, List of Members with endorsed official control programme for FMD, [WWW Document] URL https://www.oie.int/en/animal-health-in-the-world/official-disease-status/fmd/fmd-official-control-programme, last accessed 01/06/2021

Glossary OIE, Aquatic health code, 2010 [WWW Document] URL https://www.oie.int/fileadmin/Home/eng/Health_standards/aahc/2010/en_glossaire.htm., last accessed 01/06/2021.

OIE PVS Gap analysis evaluation mission report India pdf 2018 [WWW Document] URL (https://www.oie.int/solidarity/pvs-gap-analysis/pvs-gap-analysis-reports/) last accessed 01/01/2019).

Department of animal husbandry and dairying, 19th and 18th Livestock census pdf [WWW Document] URL (http://dahd.nic.in/documents/statistics/livestock-census), last accessed 01/06/2021.

Annual reports, ICAR directorate of Foot and Mouth Disease [WWW Document] http://www.pdfmd.ernet.in/index_files/Annual_Reports.html, last accessed 01/06/2021.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open-Source Software, 4(43), 1686, [WWW Document] https://doi.org/10.21105/joss.01686, last accessed 12/26/2020.

pcp-26012011.pdf

25022019_India_PVS_Evaluation_report_final.pdf