The epidemiology of emerging human-infective RNA viruses: discovery, geographical extent, and disappearance

Feifei Zhang

Thesis submitted for the degree of Doctor of Philosophy

University of Edinburgh

2021

Declaration

I declare that the thesis has been composed entirely by myself and that the work has not been submitted for any other degree or professional qualification. I confirm that the work contained herein is my own except where states otherwise by reference or acknowledgment.

Part of this work has been published in the PLOS Pathogens and Nature Medicine.

Feifei Zhang, 2021

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisors, Professor Mark Woolhouse, Dr Margo Chase-Topping, and Dr Amy Pedersen for your advice, support, and guidance throughout my PhD. Mark has taught me how to think insightfully, how to write professionally, and let me understand the importance of validating results repeatedly to get close to the facts. Your invaluable advices and feedback on this thesis have helped me gradually grow up as an independent researcher. I am deeply grateful to Margo for your constant encouragement, statistical support, and for believing in me all throughout. Many thanks to Amy for your expertise that has helped me understand the viral ecology enormously. I would also like to thank the Darwin Trust of Edinburgh for funding my PhD.

Many thanks to the present and past members of Epigroup for their constructive advice and practical help. I would particularly like to thank Liam Brierley—though I have never seen you in person—for sharing your knowledge in virus ecology and providing me many other valuable advices. Thanks to Bram van Bunnik, Gail Robertson, Thibaud Porphyre, and Helen Brown for your statistical assistance and help develop the methods used in the thesis. Thanks to Donald Smith for sharing your knowledge in virology and for validating my data. Thanks to Shengyuan Zhao and Lu Lu who have made my life here in Edinburgh much easier and enjoyable. Thanks to my group members Dishon Muloi, Carlijn Bogaardt, Hannah Lepper, Bryan Wee, Alex

Morgan, Jordan Ashworth, Stefan Rooke, and Thomas Dalhuisen for making a lovely working life in the office—without you this thesis would have never been possible. Many thanks to Kath Tracey who was always willing to provide help.

I am grateful to summer students Melina Beykou, Melissa Taylor, Ruby Tabor, Giles Calder-Gerver, Miles McGibbon, Alexandra Huber, Tara Wagner-Gamble, Samuel Haynes, Alistair Morrison, and Miranda Ferguson for data checking and data collection. Many thanks to colleagues from the WHO Regional Office for Africa, as well as Dr Miriam Njoki, Dr Seth Amanfo, Professor Gordon Awandare, and Professor Francisca Mutapi for helpful discussions on the COVID-19 research in the WHO African Region.

Finally, I would like to thank my parents and sister who always give me support and strength. Thanks to my lovely nephew and niece for cheering me up and helping me get through the difficult times. Special thanks to my partner, Chuan-guo, for always being there for me, particularly when I am in trouble. Thank you so much for being both my life and academic partner!

Abstract

Previous investigations into human infectious diseases have revealed RNA viruses as major etiological agents. Given the recent rate of newly detected human-infective RNA viruses such as severe acute respiratory syndrome coronavirus (SARS-CoV), SARS-CoV-2, Middle East respiratory syndrome coronavirus, and Bundibugyo ebolavirus, targeting virus discovery in high-risk regions, characterizing viruses with the greatest likelihood of spreading and establishing sustained infection in humans would benefit better preparedness for future outbreaks. There is a lack of evidence on determinants of spatio-temporal patterns in the discovery of human-infective RNA viruses, though previous studies have attempted to identify hotspots of emerging infectious diseases caused by various pathogens. There are also no quantitative studies exploring predictors of geographical extent and the disappearance for all currently known human-infective RNA viruses.

This thesis aimed to address the following gaps.

1. Identifying predictors discriminating between areas with and without discovery of human-infective RNA viruses and predicting discovery hotspots, at both global and regional scales. Predictors identified include socio-economic, climatic, land use, and biodiversity variables.

2. Prediction of the geographical extent and the disappearance of human-infective RNA viruses, using features such as taxonomy, virus structure, transmission mode, host range, origin, and clinical presentation.

3. Taking SARS-CoV-2 as an example, investigating how predictors related to demographics, socioeconomics, travel, healthcare, co-morbidities, readiness, geography, COVID-19 testing, and interventions have affected the epidemic of the disease it caused—coronavirus disease 2019 (COVID-19)—between countries in the WHO African Region.

In order to address the gaps outlined above, I firstly geocoded the first reports of 223 human-infective RNA viruses at the global scale. Using a Poisson boosted regression tree (BRT) model, I identified GDP growth, GDP, and urbanization as top predictors of virus discovery, and predicted discovery hotspots including both historical hotspots—eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America, and new hotspots—East and Southeast Asia, India, and Central America. Stratified analyses suggested discovery of vector-borne viruses and strictly zoonotic viruses was more correlated with climatic variables and biodiversity, whereas the discovery of non-vector-borne viruses and human-transmissible viruses was more strongly correlated with GDP and urbanization. Next, I focused on comparisons of the specific predictors of RNA virus discovery in three different regions with different GDP—United States, China, and Africa. A similar methodology as the global analysis was used on each region separately, the results showed that predictors such as GDP and land use continued to be top predictors in three regions, but climate and biodiversity variables were consistently less important predictors than at a global scale.

To identify predictors of the geographical extent and the disappearance (no record of infection in the literature for the past ten years or more), I collated information for 223 human-infective RNA viruses on their geographical extents and persistence in causing human infections from peer-reviewed literature. By fitting Bernoulli BRT models, I observed that viral features that predicted wide geographic extent included transmissibility between humans, a +ssRNA genome, narrow host range [i.e. infecting humans only or humans and other non-human primates (NHP) only], and having a reservoir host in a NHP. Viruses were more likely to disappear if they were incapable of transmission between humans, have had a localised geographic extent, a dsRNA genome, were non-pathogenic and non-fatal, were firstly discovered through active discovery programmes rather than passive investigation of the aetiology, and were transmitted by vectors and direct contact. Results for both geographical extent and virus disappearance did not change after factoring out reporting effort. I concluded that multiple characteristics determined the geographical extent and disappearance of human-infective RNA viruses; however, transmission mode and structure were consistently the most important predictors of the geographical extent and disappearance of human-infective RNA viruses. Host range was an important predictor of geographical extent, though less important for disappearance. Geographical extent, clinical presentation and discovery process all contributed to the probability of a virus disappearing.

To understand the differences between epidemics of COVID-19 between countries of the WHO African Region, I selected the timing of the first case and

the mortality rate in the first and second waves as the three outcomes. By applying a series of statistical models including Cox proportional hazards regression models, generalized linear mixed models and multinomial logistic regression models, I found that COVID-19 in Africa arrived earlier and caused greater mortality in countries with more pre-pandemic international connectivity and a more urban population. Mortality was exacerbated by high HIV prevalence. The stringency and timing of government restrictions on behaviour were not associated with a lower per capita mortality rate. A more urban population and a higher infectious disease resilience score were associated with more stringent restrictions and/or a higher per capita mortality rate. The predictor set for the first and second waves were similar, and first wave per capita mortality was a significant predictor of second wave per capita mortality.

In summary, studies in this thesis showed that there were variations in predictors of discovery both between virus types and geographical regions, and identified high-risk regions for virus discovery beyond their historical extent. The studies also provided proof-of-principle for the prediction of attributes such as mortality, geographical extent, and disappearance for new human-infective RNA viruses. These results help identify priority regions for investment in surveillance systems for new human-infective viruses, and to make risk assessments once they have emerged.

Lay summary

The ongoing pandemic of coronavirus disease 2019 (COVID-19), together with previous outbreaks of severe acute respiratory syndrome (SARS), Middle East respiratory syndrome, and novel Influenza A, remind us we are vulnerable to newly emerging RNA viruses. Therefore, understanding where to look for novel human-infective RNA viruses, how wide the viruses can spread, and whether the viruses are able to persist in human populations are important for informing public health preventive decisions in the early stage after the emergence. Despite the increased information gained from investigation of previous emerging RNA virus-related outbreaks, there is little knowledge on where the next novel human-infective RNA viruses are likely to be discovered. Previous models identified what viral characteristics determined transmissibility and virulence in humans, little is known on how viral characteristics will affect the geographical extent and viral disappearance of human-infective RNA viruses. This thesis aimed to identify predictors of discovery of human-infective RNA viruses and predict the discovery hotspots, and identify viral characteristics that affect the geographical extent and the disappearance of human-infective RNA viruses. Predictors of discovery of one single virus—SARS coronavirus 2 (SARS-CoV-2) and the associated mortality in countries of the WHO African Region were also investigated.

Of the 33 predictors related to climate, socioeconomics, land use, and biodiversity, GDP growth, GDP, and urbanization contributed most to virus discovery in the globe. In addition to the historical high-risk areas including

eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America, more viruses were likely to be discovered in three new hotspots—East and Southeast Asia, India, and Central America in 2010–2019. Climatic variables and biodiversity were more important predictors of discovery of vector-borne virus and non-human-transmissible virus, while GDP and urbanization were more important predictors of discovery of non-vector-borne virus and human-transmissible virus. In three different regions including the United States, China and Africa, virus discoveries continued to be strongly associated with GDP and land use, but were less likely to be associated with climate and biodiversity variables than at a global scale.

Geographical extent and disappearance of human-infective RNA viruses were correlated with multiple viral characteristics. Viruses were more likely to spread to a wide geographical extent if they were transmissible between humans, had a +(ss)RNA as genome, and had non-human primates as the only host other than humans. Viruses were more likely to disappear from humans if they were unable to transmit between humans, have been restricted in a localised area, had dsRNA as the genome, were non-fatal and non-pathogenic, have been discovered through surveillance, were vector-borne, and were transmitted by direct contact.

The timing of reporting the first case and the mortality rate of COVID-19 varied greatly across different countries in the WHO African Region. High proportion of urban population, large number of international airports, high volume of international air travel, COVID-19 test capacity and more country borders

contributed to an early detection of the first case. A higher mortality rate was associated with a higher prevalence of HIV. Stringency and timing of government restrictions were not associated with the mortality rate, but countries with a more urban population and high infectious disease resilience score were more likely to have an adverse outcome—more stringent restrictions and/or high per capita mortality. The predictor set for mortality rate in the second wave was similar to that for the first wave, and first wave per capita mortality was positively associated with the second wave per capita mortality.

The hotspots I identified in the thesis direct where the next human-infective RNA virus is likely to be discovered and inform intense surveillance in these regions. My models in the thesis can also help make risk assessment for novel RNA viruses at the early stage of emergence by evaluating their geographical extent and disappearance. The thesis identified risk factors associated with poor outcomes of COVID-19 pandemic in the two waves in the WHO African Region, and help guide future pandemic preparedness planning in this region.

List of abbreviations

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
COVID-19	Coronavirus disease 2019
SARS	Severe acute respiratory syndrome
SARS-CoV	Severe acute respiratory syndrome coronavirus 1
AIDS	Acquired immunodeficiency syndrome
MERS-CoV	Middle East respiratory syndrome coronavirus
WHO	World Health Organization
ICTV	International Committee on Taxonomy of Viruses
BRT	Boosted regression tree
ICC	Infraclass correlation coefficient
IQR	Interquartile range
NHPs	Non-human primates
AUC	Area under the receiver operating characteristic curve
ROCs	Receiver operator curves
HRs	Hazard ratios
CIs	Confidence intervals
AIC	Akaike information criterion
GLMM	Generalized linear mixed model
RRs	Risk ratios
ORs	Odds ratios
DALY	Disability-adjusted life year
CFR	Case fatality rate
HDI	Human development index

List of figures

List of tables

Contents

Chapter 1 General introduction

1.1 Emerging human-infective RNA viruses

On March 11, 2020, the World Health Organization (WHO) declared coronavirus disease 2019 (COVID-19) a world pandemic (Cucinotta & Vanelli, 2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)—a RNA virus identified as the culprit of this pandemic—caused infections in human populations across all six continents in the first two months of 2020 (Gorbalenya et al., 2020; World Health Organisation, 2021c). To tackle the pandemic, an unprecedented global lockdown was implemented (Kucharski et al., 2020). As of early April 2021, SARS-CoV-2 had caused 130.5 million cases and nearly 2.9 million deaths (World Health Organisation, 2021c). In many countries the ongoing waves have ended, but some countries such as Brazil, India, Argentina, Colombia, Iran, and Japan are experiencing their second or even the third and fourth waves (World Health Organisation, 2021c) and the burden of the disease continues to rise.

Historically, pandemics like COVID-19 are not new. Pathogenic infectious diseases have always been substantial global threats to human health (Figure 1.1) (Huremović, 2019; Morens & Fauci, 2020; Piret & Boivin, 2020). Outbreaks including Justinian plague in 541 AD, Black Death in 1348 AD, 1918 influenza pandemic, and acquired immunodeficiency syndrome (AIDS) pandemic starting from 1980s have taken hundreds of millions lives (Morens & Fauci, 2020). Some infectious diseases, such as smallpox (Geddes, 2006),

poliomyelitis (Garon et al., 2015), and yaws (Stamm, 2015) have been eradicated (or nearly eradicated) thanks to good sanitation, vaccines, and antimicrobial agents. However, diseases such as AIDS (Hemelaar et al., 2018) and multidrug-resistant tuberculosis (Sharma & Mohan, 2006) continue to endanger human lives.

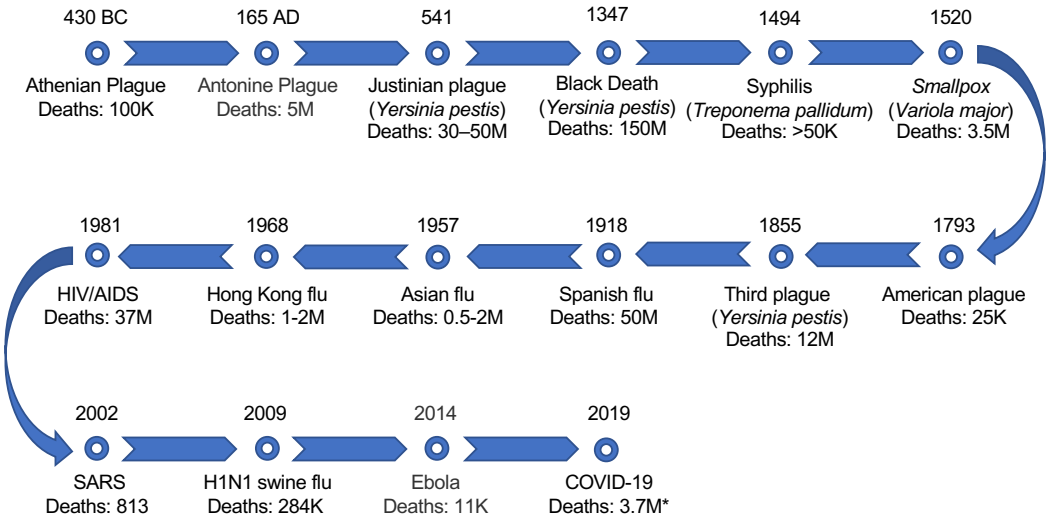


Figure 1.1 Timeline of the pandemics throughout the history

* as of 11 June 2021

Most human pathogens are believed to have emerged from other animals (e.g. SARS-CoV and Ebola virus from bats) (Leroy et al., 2005; Wang et al., 2018) or evolved from closely related viruses in animals [e.g. HIV-1 from simian immunodeficiency virus infecting wild chimpanzees (SIVcpz), HIV-2 from SIV infecting sooty mangabey (SIVsmm)] (Gao et al., 1999), and only a small proportion emerged into humans from completely unknown sources (e.g. Rubella virus and Cosavirus are confined to humans) (Morse et al., 2012).

Emerging pathogens are categorised as either newly emerged pathogens [e.g. Middle East respiratory syndrome coronavirus (MERS-CoV), SARS-CoV, Lujo virus] or previously recognized pathogens (namely re-emerging pathogens) with increased incidence or extended geographic range (e.g. Ebola virus, Chikungunya virus, Zika virus) (Hassell et al., 2017). The consequence of emerging pathogens ranges from one-off infections (i.e. humans as the dead-end host) or small epidemics/outbreaks, to global pandemics like COVID-19 (Morens & Fauci, 2020). In this thesis, I concentrated on the viral pathogens, specifically on RNA viruses, that have emerged to humans at least once under natural circumstance, therefore excluding emerging pathogens associated with bioterrorism, and those created by humans in laboratories that are released accidentally (Morens & Fauci, 2020).

An RNA virus is a virus that has RNA (ribonucleic acid) as the genetic material as opposed to DNA (deoxyribonucleic acid). The RNA may be either single-stranded [including +(ss)RNA, -(ss)RNA, and ssRNA-RT], or double-stranded (dsRNA) (Aris-Brosou et al., 2019; Woolhouse & Brierley, 2018). RNA viruses have a high rate of adaptive mutation, evolution, and re-assortment which makes them well adapted to new hosts (Dolan et al., 2018; Woolhouse et al., 2016). Previous investigations into human infectious diseases have revealed RNA viruses as major etiological agents (Taylor et al., 2001). For example, recently identified infections, such as SARS, COVID-19, MERS, and novel influenza subtypes (Kile et al., 2017; Ksiazek et al., 2003; Mackay & Arden, 2015; World Health Organisation, 2020b, 2021c), have all had an impact on population health, causing millions of deaths worldwide.

A data set containing a subset of all human-infective RNA viruses ever discovered was first published in 2001 (Taylor et al., 2001). In 2018, the dataset was updated and published (Woolhouse & Brierley, 2018), along with epidemiological characteristics of each virus including the date the virus was first reported in humans, transmissibility between humans, transmission routes, and host range. The most recent data set was published in 2020 and included 223 human-infective RNA viruses (Zhang et al., 2020). All virus species are recognised by the International Committee on Taxonomy of Viruses (ICTV) and are classified into 57 genera and 23 families (Zhang et al., 2020).

1.2 Emergence comes before discovery

The ongoing pandemic of COVID-19 emphasises the need for understanding how novel pathogens start to infect humans, how to prioritize surveillance and to predict the future emergence risk. While capturing the emergence of a virus would be more relevant to public health, it is often not possible because most viruses have existed in nature for many years before they emerged in humans and were discovered (Woolhouse & Gaunt, 2007). As such, we do not know exactly when and where emergence took place, only that it must have preceded discovery (Zhang et al., 2020). Virus discovery is a complicated process involving many factors (Morse et al., 2012; Parrish et al., 2008). Some factors affect discovery directly via relying on the advanced diagnostic technology and sufficient testing resources, whereas others affect discovery indirectly via the intermediate step of emergence (Zhang et al., 2020). Presently, we lack the ability to distinguish all the drivers for emergence and

discovery, though we can distinguish some by correcting for the discovery effort and improving the study design (e.g. longitudinal studies).

Factors that drive the emergence of pathogens, particular zoonotic pathogens, have been described comprehensively in the literature (Morse et al., 2012; Parrish et al., 2008). In general, evidence has come from three forms of analyses: analysis of single emergence event (Parrish et al., 2008), quantifying the spillover (or host switching/cross-host transmission) risk using traits of both hosts and viruses (C. K. Johnson et al., 2015; Olival et al., 2017; Pulliam & Dushoff, 2009), and record of first emergence event in humans globally over time (Allen et al., 2017; Jones et al., 2008). There are strengths and shortcomings associated with each form of evidence. Analysis of a single emergence event (e.g. SARS-CoV, HIV-1; analysis form 1) helps to understand causal relations between emergence and drivers (Parrish et al., 2008), but these causal relations often cannot be generalised to other emergence events. Quantifying the spillover (analysis form 2) and recording the first emergence (analysis form 3) cannot infer the causal relations, but can reveal the hotspots of emerging pathogens and/or predict which mammal species likely harbor the next viral pathogen with pandemic potential. Of note, although all these studies called them the emerging pathogens, by definition they are actually novel discovered pathogenic/non-pathogenic microbes.

Some previous studies supported dividing the process of emergence into multiple stages (Morse et al., 2012; Parrish et al., 2008; Wolfe et al., 2007), and some studies supported assigning pathogenic/non-pathogenic microbes

into pyramid levels according to which stage they can reach (Figure 1.2) (Woolhouse et al., 2016). I herein divided the process of emergence into three stages as defined by Morse SS et al (Morse et al., 2012): Stage 1. pre-emergence when there is no human infection; Stage 2. localised emergence including both single infections without interhuman transmission and small-scale outbreaks with interhuman transmission; Stage 3. pandemic emergence with widespread transmission and dissemination. Most known pathogenic/non-pathogenic microbes stay in stage 1, and only a few can enter stage 2 or stage 3 (Parrish et al., 2008; Woolhouse et al., 2016). These microbes can be discovered at any stage, whereas first discovery of human infection mostly occurs in stage 2 because local surveillance systems are usually able to capture them before stage 3 (Morse et al., 2012). Therefore, discovery of novel microbes is mainly driven by factors before stage 2.

Well-studied examples of single emergence events (analysis form 1) have linked the emergence to ecological and environmental changes, many of which change the contact rates between animal hosts and humans and provide new opportunities for pathogens to emerge to human populations. At stage 1, a variety of pathogens exist in animal reservoirs and vectors naturally. Mammal species that exist in human environments, such as rodents and bats, harbour a large proportion of zoonotic pathogens and constitute a large zoonotic pool (Hassell et al., 2017; Olival et al., 2017). For example, lethal Ebola virus, SARS-CoV, and MERS-CoV naturally existed in bats before emerging in humans (Olival et al., 2017; Parrish et al., 2008). Anthropogenic activities, such as encroachment on natural environments and urbanization, exploitation of

previously unsettled geographic areas have led to alterations of the ecological system and species assemblage, as well as the contact rates between humans and animal hosts (Hassell et al., 2017; Mackey et al., 2014). These changes would alter the natural locations of pathogens, and some pathogens successfully cross the species barrier and establish infections in humans. Three good examples of this include SARS-CoV, HIV, and Nipah virus. SARS-CoV and HIV spilled over to humans when preparing/handling/eating animals (bats and civets for SARS-CoV, and chimpanzees for HIV). Changes in agricultural practices—the movement of infected pigs into the new outbreak areas facilitated the emergence of Nipah virus in Malaysia (Patz et al., 2004). On 4 September 2021, the fifth outbreak of the Nipah virus disease occurred in India, and the patient may have contracted the virus after contacting with the natural host—Pteropus fruit bats. Measures to prevent the next outbreak in India include avoiding exposure to infected bats and fruits contaminated by the bats (World Health Organisation, 2021b).

Socio-economic factors including population growth and globalization of growing travel and trade promote human exposure to novel pathogens and affect the emergence of pathogens in new areas. It is estimated that 90% of population growth is predicted to occur in cities in developing regions, at an unprecedented rate (Hassell et al., 2017). By-product impacts such as human migration, sanitary conditions, and informal settlements can further facilitate the spread of pathogens, especially for cities in developing countries with fast population growth and high density (Hassell et al., 2017). In a highly globalised world, pathogens carried by infected travellers, animals, vectors, and other

commercial goods have more opportunities to overcome their original geographical boundaries, and to connect with new and growing host populations (P. T. Johnson et al., 2015; Li et al., 2014; Lipkin, 2013). For example, four pathogens have been imported to China in 2000–2016, including HIV-2 from Cote d'Ivoire, yellow fever virus and Rift Valley fever phlebovirus from Angola, and Zika virus from Venezuela (Chen & Lu, 2016; Yan et al., 2000) (Liu et al., 2016; Sun et al., 2016).

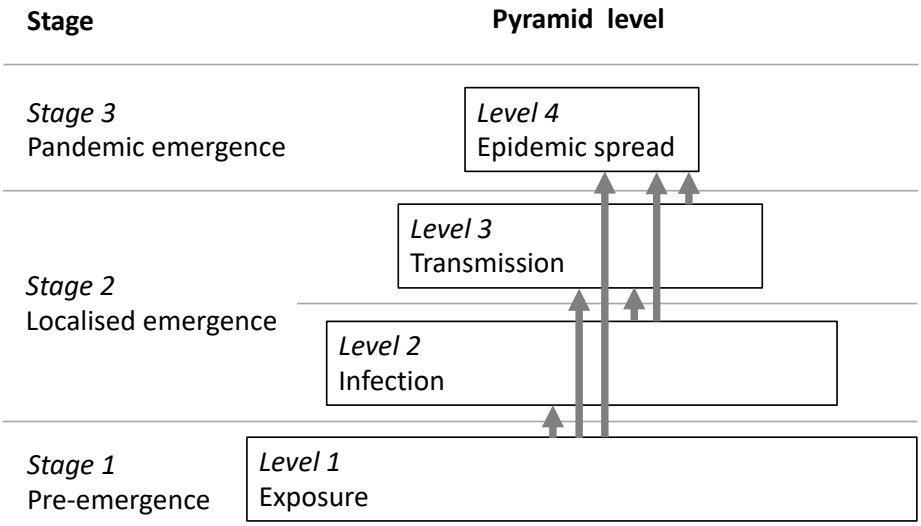


Figure 1.2 Emerging process of pathogens to humans

Adapted from Morse SS et al (Morse et al., 2012) and Woolhouse et al (Woolhouse et al., 2016). The process of emergence is divided into three stages (left, see text for details). Pathogens are assigned into pyramid levels according to which stage they can reach (right). Level 1 denotes viruses do not infect humans after exposure. Level 2 denotes viruses that are capable of infecting humans but cannot be transmitted between humans. Level 3 denotes viruses that can be transmitted between humans but are restricted to self-limiting outbreaks. Level 4 denotes viruses that can cause epidemics in humans, with sustained human-to-human transmission. Arrows indicate transitions between different levels.

It is hard to ascertain what the role of climate change in the emergence of viral pathogens, though the climatic and weather conditions have influenced

arboviruses dynamics (Mackey et al., 2014). Climate is seemingly a less frequent driver for emerging infectious diseases; one study showed that the climatic condition ranked the fifth as the driver for the emerging vector-borne zoonotic diseases, after land use, war and famine, international travel and commerce, and human susceptibility (Swei et al., 2020). In response to climate change, infected vectors could bring the viral pathogens to humans in new areas where the climatic conditions are suitable for vectors' survival (Baylis, 2017).

A second form of analysis quantifying the spillover risk using traits of both hosts and viruses has revealed some evolutionary mechanisms. From the host aspect, zoonotic virus richness in wildlife (proportion of zoonotic viruses per wild mammal species) was associated with phylogenetic distance from humans, host taxonomy (e.g. bats harbour a higher proportion of zoonotic viruses than all other mammals), and human population within a species range (likely reflect human-wildlife contact) (Olival et al., 2017). This theory was revised, however, after extending the analysis to include all avian and mammalian reservoir hosts and to consider the animal hosts at the order level [(Olival et al., 2017) study considered animal hosts at the species level]. New results suggested that all host orders were created equal with respect to the number of human-infective viruses they harbour, i.e. no certain animal species disproportionately maintained more zoonoses (Mollentze & Streicker, 2020). The number of zoonoses per animal species maintained was positively associated with their group size, for example, bats and rodents maintained a

larger number of zoonotic viruses because of their larger group size relative to other mammalian hosts (Mollentze & Streicker, 2020).

From the viral perspective, the probability of a virus being zoonotic was predicted by phylogenetic host breadth (a viral feature that measures the phylogenetic range from all known mammal hosts, excluding humans) and other viral traits (whether or not a virus replicates in the cytoplasm without nuclear entry which likely affect a barrier to viral replication, vector-borne or non-vector-borne) (Olival et al., 2017; Pulliam & Dushoff, 2009). Similarly, another study suggested viruses capable of infecting a wide range host (i.e. high host plasticity) were more likely to spread to humans (C. K. Johnson et al., 2015). Zoonotic potential for a given animal virus can also be assessed using viral genome sequences that are independent of virus taxonomy (Mollentze et al., 2021); the genomic models suggested there are generalizable features from viral genomes that can inform viruses' ability to infect humans and have been used to identify 50% potential zoonoses that were in high or very high-risk categories out of 645 animal viruses.

Although an emerging virus can occur anywhere there do seem to be regions where emerging zoonoses are most likely to occur (so-called hotspots). Spatial models of the first emergence event in humans (analysis form 3), have connected the distribution of emergence of infectious diseases across the globe to ecological, environmental, and socio-economic factors, and predicted the high-risk areas of emerging zoonoses caused by a mixture of pathogens (Allen et al., 2017; Jones et al., 2008). However, no previous study has

investigated the predictors and hotspots for novel human-infective RNA viruses. In chapter 2 and chapter 3, I developed the third form of analysis (i.e. spatial models) to describe the spatiotemporal pattern of the first discovery of human-infective RNA viruses, identified predictors, and predicted the future hotspots for their discoveries. The predictors I included were based on the first two forms of studies, i.e. previous single emergence event studies as well as evolutional studies. I included predictors related to land use, climate, and biodiversity to reflect their role in promoting the emergence of viruses and then in virus discovery.

There are direct factors associated with virus discovery that are independent of emergence. After emergence, some viruses were discovered via investigation of the aetiology for epidemics and outbreaks (e.g. SARS-CoV and MERS-CoV) (Peiris et al., 2003; Zaki et al., 2012); some viruses were discovered via active discovery programmes (Cali mammarenavirus, Alphacoronavirus 1) (Buchmeier et al., 1974; Terao et al., 2007); some viruses were discovered indirectly when investigating other viruses, e.g. Zika virus when exploring aetiology of yellow fever (Gubler et al., 2017) and Highlands J virus when investigating the epidemic of Saint Louis encephalitis virus in Florida (Meehan et al., 2000). Regardless of the discovery method typically, richer, more developed areas with more research funding, better access to technologies for virus detection and more effective surveillance system play a major role in discovering new microbes (Jones et al., 2008; Morse et al., 2012). In chapter 2 and chapter 3, I included socio-economic factors including GDP and university count as a proxy for discovery effort of novel viruses.

1.3 Post-discovery manifestation for emerging pathogens in humans

Understanding how pathogenic/non-pathogenic microbes manifest in humans after discovery is important for allocating global public health resources, given emerging microbes differ greatly in terms of their impact on global health and economy. With respect to influenza viruses, only Influenza A viruses have caused pandemics in humans (e.g. 1918–1919 Spanish flu, 1957–58 Asian flu, 1968–1969 Hong Kong flu, and pandemic H1N1 2009), whereas Influenza B viruses only cause seasonal epidemics and Influenza C viruses cause no epidemics (Gatherer, 2009; WHO Regional Office for South-East Asia, 2009). Therefore, understanding predictors of post-discovery manifestations enable pre-emptive responses and reduce the adverse effect that new microbes would bring.

In the last ten years, more efforts have been directed towards quantifying the transmissibility (Geoghegan et al., 2016; C. K. Johnson et al., 2015; Walker et al., 2018) and virulence (Brierley et al., 2019; Guth et al., 2019) of new viral pathogens using both viral ecological traits and predictors from their animal hosts. These studies have focused almost exclusively on zoonotic viruses, emphasising their importance among emerging pathogens. After spill-over to humans, viruses capable of transmission among humans were well predicted by viral traits including viral structure (segmented or non-segmented; enveloped or non-enveloped), transmission mode (vector-borne or non-vector-borne), host range (narrow or wide), clinical presentations (host mortality or

virulence; duration of infection), and viral tissue presence (nervous system, respiratory tract, and gastrointestinal system etc.) (Geoghegan et al., 2016; C. K. Johnson et al., 2015; Walker et al., 2018; Woolhouse et al., 2016). Apart from transmission mode and host range, virulence in humans was additionally predicted by other viral traits including viral taxonomy, tissue tropism (systemic, neural tropism, and renal tropism etc.), and transmissibility between humans (Brierley et al., 2019; Guth et al., 2019).

The evolutionary trade-off between virulence and transmissibility (i.e. virulence constrains transmission between human populations) has been a matter of controversy for many years (Brierley et al., 2019; Guth et al., 2019). Comparison of virulence of RNA viruses with different transmission level in humans (Figure 1.2) suggested level 4 viruses capable of epidemic spread had a lower virulence than level 3 viruses that are transmissible but are restricted to self-limiting outbreaks, though level 2 viruses incapable of transmission did not show higher virulence (Brierley et al., 2019). A common view of pathogen evolution, which supports the trade-off theory, claims that a cost of efficiency on inter-host transmission would be incurred if natural selection favours pathogens of higher virulence within host, given the higher virulence would greatly shorten the infectious period (Brierley et al., 2019; de Roode et al., 2008; Geoghegan et al., 2016). Therefore, pathogens need to balance between transmissibility and virulence, e.g. both at optimal level, and achieve the maximum transmission rate in hosts (Moya et al., 2004). A good example is the two SARS coronaviruses: SARS-CoV and SARS-CoV-2. In comparison to SARS-CoV, SARS-CoV-2 was found to be more transmissible,

but less pathogenic (Chen, 2020). The consequence is SARS-CoV disappeared in 2004 while SARS-CoV-2 causes the ongoing pandemic with sustained human-to-human infection (Petersen et al., 2020). However, the adaption of virulence involves complex interactions between intra-host replication and inter-transmission (Lipsitch & Moxon, 1997), e.g. diseases virulence is likely to differ between vertically and horizontally transmitted pathogens (Bergstrom et al., 1999).

The opposite correlations between virulence or transmissibility and a third viral feature can partly provide support for the trade-off theory (Guth et al., 2019). For example, vector-borne viruses were found to have a lower transmissibility among humans than its non-vector-borne counterparts, mostly because of the requirement to replicate in both invertebrate vectors and vertebrate hosts (Geoghegan et al., 2016; Walker et al., 2018). For virulence, previous studies suggested vector-borne pathogens had a higher virulence, particularly compared to those that were directly transmitted (Day, 2002; Ewald, 1983), though one study found vector-borne viruses to have a lower probability of severe virulence than non-vector-borne viruses (Brierley et al., 2019).

The link between virulence or transmissibility and host range was also explored. Zoonotic viruses carried by distantly related hosts (wide host range) were found to be more virulent but less likely to be capable of establishing transmission among humans, whereas hosts most closely related to humans such as non-human primates harboured zoonotic viruses with higher transmissibility (Guth et al., 2019; Walker et al., 2018). Theoretically, non-

human primates have a shorter phylogenetic distance with humans, and the pathogens they carry may be pre-adapted to humans and then transmit efficiently (Walker et al., 2018). Meanwhile, new pathogens originating from hosts with shorter phylogenetic distance with humans usually show similar levels of virulence in humans (Longdon et al., 2015). However, one study suggested viruses detected in a large number of animal host orders (wide host range) were more likely to transmit between humans than viruses found in a single animal host order (C. K. Johnson et al., 2015). The possible explanation is that some factors, such as more frequent contact rates between humans and distantly related host species, has promoted the evolutionary co-selection for viruses with both greater abilities to adapt to new hosts and to transmit in the new hosts (C. K. Johnson et al., 2015; Walker et al., 2018).

In addition to virulence, transmission mode, and host range, transmissibility was also associated with viral structure, duration of infection, and viral tissue presence. For example, non-segmented viruses were more likely to be transmissible than segmented viruses among humans (Geoghegan et al., 2016). Geoghegan et al (Geoghegan et al., 2016) indirectly attributed this to a simpler replication process of +(ss)RNA viruses, given all +(ss)RNA viruses in their dataset had non-segmented genomes. Higher transmissibility of non-enveloped viruses among humans may be because non-enveloped viruses are more environmentally stable than enveloped ones and have a higher probability of transmitting via direct or indirect contact (Geoghegan et al., 2016).

Viruses causing chronic infections have been found to have a higher transmissibility in humans—the possible reason is that longer duration of infection and viral shedding facilitate the probability of transmission to a new host (e.g. measles virus, HIV-1, hepatitis C virus) (Geoghegan et al., 2016). Transmissibility of viral zoonotic viruses in humans was also associated with viral tissue presence, with virus isolated from respiratory tract and central nervous system having higher transmissibility (Walker et al., 2018).

Viral taxonomy and tissue tropism were additionally associated with virulence of novel viral pathogens. Difference in virulence of human-infective RNA viruses across virus families can be partly explained by the variation in viral traits such as host range, transmission mode, and structure (Brierley et al., 2019). For example, most or all RNA viruses in families *Flaviviridae*, *Peribunyaviridae*, *Phenuviridae*, and *Togaviridae* were vector-borne (Woolhouse & Brierley, 2018). Viruses were more likely to have higher virulence if they caused systemic infections and have neural or renal tropism (Brierley et al., 2019). The possible reason is that viruses have higher levels of replication and are well adapted in these targeted tissue types, and cause greater damage to the human hosts (Longdon et al., 2015).

Transmissibility and virulence are two important post-discovery manifestations of virus emergence. However, less is known about the predictors that enable pathogens to move from emergence to epidemic status. Some viruses are able to transmit between humans, but do not cause epidemics (e.g. Brazilian mammarenavirus, Nelson Bay orthoreovirus) or only cause single infections

(e.g. Lujo virus, Bas-Congo virus) (Woolhouse & Brierley, 2018). In chapter 4, I focused on another two post-discovery manifestations: geographical extent and disappearance. In spite of some controversies from previous studies on transmissibility and virulence and the associated factors, I hypothesised that 1) transmissible viruses and viruses with lower virulence are more likely to establish wide geographic extent and less likely to disappear; 2) Non-vector-borne viruses and viruses with narrow host range are more likely to establish wide geographic extent and less likely to disappear in humans, given that they are highly likely to be transmissible and have lower virulence.

1.4 Predictors associated with COVID-19 first case and mortality

The ICTV Coronaviridae Study Group determined that SARS-CoV-2 is a subspecies of the existing species—SARS-CoV (Gorbalenya et al., 2020), but the two viruses behave very differently with respect to importation risk, virulence and epidemic potential (Petersen et al., 2020). Therefore, it is important to look at predictors of the epidemiology of SARS-CoV-2 related disease—COVID-19. In chapter 5, I studied the epidemiology of COVID-19 and investigated how predictors related to demographics, socioeconomics, travel, healthcare, co-morbidities, readiness, geography, COVID-19 testing, and interventions have affected the timing of the first case (discovery) and the mortality rate (post-discovery manifestation) in the WHO African Region.

Few studies have explored risk factors for introduction of the first COVID-19 case. One cross-country study found that a higher risk of importation (based on flight connectivity to China) was associated with early introduction of the first case (Haider et al., 2020).

As of 29 January 2021, 83 cross-sectional ecological studies have explored the mortality/fatality/death-related factors of COVID-19 at national/subnational levels. Of these, 35 were cross-country studies and 48 were within-country studies. Of the 35 cross-country studies, 20 included WHO African Region countries, one of which focused on 53 African countries (Arsalan et al., 2020; Asfahan et al., 2020; Carrillo-Larco & Castillo-Cara, 2020; Chakraborti et al., 2020; Chakraborty & Ghosh, 2020; Chaudhry et al., 2020; Ergönül et al., 2021; Fountoulakis et al., 2020; Hashim et al., 2020; Hradsky & Komarek, 2021; Kranjac & Kranjac, 2020; Lawal, 2021; Leffler et al., 2020; Li et al., 2020; Medeiros de Figueiredo et al., 2020; Okeahalam et al., 2020; Olivieri et al., 2021; Pan et al., 2020; Sannigrahi et al., 2020; Sornette et al., 2020). This African study found that the number of nursing and midwifery personnel was negatively associated with COVID-19 mortality, while universal healthcare index of service coverage and prevalence of insufficient physical activity among adults aged 18 years or older was positively associated with COVID-19 mortality (Okeahalam et al., 2020). However, this study only included health-related factors, and missed other factors related to demographics, socioeconomics, and government responses. Studies that included WHO African Region countries identified risk factors for COVID-19 mortality related to demographics (the proportion of elderly/population mean age/life

expectancy, population density, the number of tourists), socioeconomics (higher GDP), comorbidities (obesity, chronic obstructive pulmonary disease, Alzheimer's Disease, depression, 65 years or older mortality %), lifestyle (the prevalence of adult male smoking and consumption of sugar-sweetened beverages), climate, and government measures (less stringent international travel restrictions). Protective factors were related to healthcare (greater healthcare capacity, more equitable access to healthcare, higher number of hospital beds per population), lifestyle (increasing fruits consuming and beans and legumes), climate, and government measures (mask-wearing). However, most of these studies used death data before May 2020, and did not take into account the different course of the epidemic in different countries. In chapter 5, I restricted my study area within a single WHO region—these should be more comparable both in terms of data on predictors and COVID-19 epidemiology. Specifically, I chose the WHO African Region as the study area because Africa has some unique traits compared to other continents—it has the youngest population and the experience in fighting against infectious diseases (Webb, 2013). Meanwhile, it also has a weaker healthcare system and the highest prevalence of comorbidities such as HIV (Lone & Ahmad, 2020; World Health Organisation, 2020c). It was not known whether the relationships between predictors and COVID-19 related infections identified in other continents still held in Africa, e.g. if the young age of the population is a protective factor for COVID-related mortality within Africa. Further, I included a comprehensive list of predictors likely to influence the COVID-19 epidemics, particularly predictors on COVID-19 testing and government interventions.

1.5 Aims and outline

There are three aims in this thesis. The first two aims were to understand predictors of the attributes of each human-infective RNA virus after emergence—discovery (aim 1), geographical extent and disappearance (aim 2). The third aim was to understand predictors of the post-emergence attributes of one single human-infective RNA virus—SARS-CoV-2, including discovery of the first case and the mortality rate.

Aim 1 (Chapter 2 and Chapter 3): To predict the hotspots of discovery of human-infective RNA viruses, based on two types of predictors—predictors reflecting natural virus distribution (e.g. climatic, land use, and biodiversity variables) and predictors reflecting the effort invested in virus discovery (e.g. socio-economic variables). This was done globally (Chapter 2) and then compared across 3 regions with different GDP (United States, China, and Africa) (Chapter 3).

Aim 2 (Chapter 4): To predict the geographical extent and the disappearance of human-infective RNA viruses, based on their biological traits including taxonomy, virus structure, transmission mode, host range, origin, and clinical presentation.

Aim 3 (Chapter 5): To identify the predictors of discovery of the first case and mortality rate of COVID-19 caused by SARS-CoV-2, using the WHO African region as a case study. I aimed to understand how predictors related to demographics, socioeconomics, travel, healthcare, co-morbidities, readiness,

geography, COVID-19 testing, and interventions have caused the different epidemics of COVID-19 in the 47 countries in the WHO African Region.

The outline of the thesis is as follows.

Chapter 2 and Chapter 3: I focused on prediction of hotspots of human-infective RNA virus discovery. In chapter 2, I studied this at the global scale. I used Poisson boosted regression tree (BRT) model to identify predictors with the greatest influence on the virus discovery and predicted the probability of virus discovery in 2010–2019 across the globe. I also conducted stratified analyses (viruses transmissible in humans vs viruses that are strictly zoonotic, vector-borne viruses vs non-vector-borne viruses) to identify important predictors specific to different categories of viruses. In chapter 3, I focused on comparisons of the specific predictors of RNA virus discovery in three different regions—United States, China, and Africa, following a similar process as the global analysis in chapter 2.

Chapter 4: I focused on the prediction of the geographical extent and the disappearance of human-infective RNA viruses. I presented an overview of the geographical extent and the disappearance for each virus. I then investigated how the geographical extent and disappearance were predicted by a group of biological features by fitting the Bernoulli BRT models.

Chapter 5: I investigated the predictors of the timing of the first case and the mortality rate of COVID-19 in the WHO African Region. A set of predictors assigned into nine categories were used, including demographics, socioeconomics, travel, healthcare, co-morbidities, readiness, geography, and

COVID-19 testing and interventions. I applied Cox proportional hazards regression models, generalized linear mixed models, and multinomial logistic regression models as appropriate.

Chapter 6: I discussed the overall implications of findings from the thesis on public health, as well as the future study directions in areas related to this thesis.

Chapter 2 Predictors of global discovery of human-infective RNA viruses

Work in this chapter has been published in PLOS Pathogens.

2.1 Abstract

RNA viruses are a leading cause of human infectious diseases and the prediction of where new RNA viruses are likely to be discovered is a significant public health concern. Here, I geocoded the first peer-reviewed reports of 223 human-infective RNA viruses. Using a boosted regression tree model, I matched these virus data with 33 predictors related to natural virus distribution and research effort to predict the probability of virus discovery across the globe in 2010–2019. Stratified analyses by virus transmissibility and transmission mode were also performed. The historical discovery of human-infective RNA viruses has been concentrated in eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America. The virus discovery can be predicted by a combination of socio-economic, land use, climate, and biodiversity variables. Vector-borne viruses and strictly zoonotic viruses were more associated with climate and biodiversity whereas non-vector-borne and human transmissible viruses were more associated with GDP and urbanization. The areas with the highest predicted probability for virus discovery in 2010–2019 included three new regions including East and

Southeast Asia, India, and Central America, which likely reflected both increasing surveillance and diversity of their virome. My findings can inform priority regions for investment in surveillance systems for new human-infective RNA viruses.

2.2 Introduction

Since the first identification of a virus in humans—yellow fever virus in 1901—viruses have been recognised as a leading cause of human infectious diseases (Woolhouse & Brierley, 2018). Numerous human diseases, from the common cold (Tang et al., 2017) to life-threatening haemorrhagic fevers (Clark et al., 2018), are caused by RNA viruses. RNA viruses such as dengue virus, norovirus, and HIV impose significant burdens on global health and the global economy (Guzman & Isturiz, 2010; Lopman et al., 2016; World Health Organisation, 2018). Despite the striking declines in the incidence and mortality of RNA virus-related diseases in human following the introduction of vaccination, infections due to measles virus, yellow fever virus, and Japanese encephalitis virus continue to endanger human health and cause hundreds to thousands of deaths each year (Barrett, 2018; Moss, 2017; World Health Organisation), particularly in countries with limited resources to launch mass vaccination campaigns.

Human-infective RNA viruses comprise a total of 214 International Committee on Taxonomy of Viruses (ICTV)-recognised species as of July 2017, classified into 55 genera and 22 families (Woolhouse & Brierley, 2018). Many of these—such as rabies virus, dengue virus, and measles virus—have circulated in

humans for thousands of years (Fisher et al., 2018; Guzman & Isturiz, 2010; Moss, 2017), though some—such as HIV-1 and SARS-CoV—have emerged much more recently. Typically, a virus is identified through investigation of the aetiology of a human disease [e.g. yellow fever virus (Reed et al., 1901), measles virus (Goldberger & Anderson, 1911)], although some have been identified during active virus discovery programmes [e.g. Rotavirus C (Bridger et al., 1986), Parechovirus B (Niklasson et al., 2003)]. Viruses such as hepatitis delta virus (Rizzetto et al., 1977) and Highlands J virus (Meehan et al., 2000) were discovered by chance, as incidental findings as part of a disease investigation.

The discovery curve of human viruses, for both RNA viruses and DNA viruses, was described for the first time in 2008 (Woolhouse et al., 2008). Up to nine new human virus species have been detected each year since the 1950s, and this is projected to continue in coming decades (Woolhouse et al., 2008). The factors driving the discovery of human viruses remain to be elucidated, though two previous studies have identified predictors of the emergence of infectious diseases more generally (Allen et al., 2017; Jones et al., 2008). In this chapter, I took a spatiotemporal modelling approach to identify predictors influencing the discovery of RNA viruses in humans. I assume virus discovery is determined by two underlying spatiotemporal patterns: the geographical distribution of viruses in nature, and the process of virus detection—a human activity. Geographical ranges, which vary from worldwide [e.g. Norwalk virus (Lopman et al., 2016), HIV-1 (World Health Organisation, 2018)] to very localised [e.g. Hendra virus (Escaffre et al., 2013), Menangle virus (Philbey et

al., 1998)], are mostly determined by virus natural history, vector distribution (for vector-borne viruses), and non-human host distribution(s) (for zoonotic viruses) (Babayan et al., 2018). In contrast, virus detection reflects scientific resources and research effort (Jones et al., 2008). An uneven distribution of research effort will lead to an uneven distribution of virus discoveries. Geographical ranges and discovery effort are likely to have different drivers (Brierley et al., 2016). Previous studies (Jones et al., 2008; Olival et al., 2017) have attempted to allow for variation in discovery effort, although this is hard to do as no direct and effective measures are available. Here, I took a different approach by identifying predictors of the raw virus discovery data and then interpreted in the discussion whether these effects might relate to virus geographic range or discovery effort or both.

2.3 Materials and Methods

2.3.1 Methods overview

In this study, I followed methods and used code derived from Allen, et al (Allen et al., 2017). I compiled and geocoded the first reports in the peer-reviewed literature of human infection for each RNA virus in my database over a period of 118 years from 1901 to 2018. A Poisson boosted regression tree (BRT) model—a method that handles spatially dependent data well—was fitted to the human-infective RNA virus data with a set of variables thought to be potential predictors. By matching the virus discovery count and all predictors in each 1° resolution grid cell (approximately 110 km at the equator) by decade, I ranked the contribution of each predictor to the predictions. I then used the parameter

estimates from the best fitting BRT model to predict the probability of virus discovery for all grid cells across the globe in 2010–2019 using the values of all predictors in 2015. I also conducted stratified analyses (distinguishing viruses transmissible in humans or strictly zoonotic, and vector-borne or non-vector-borne) to find the predictors of the discovery of specific categories of viruses.

2.3.2 Data set of human-infective RNA viruses and updating

Data on human-infective RNA viruses were derived from an updated version of previously published database (Woolhouse & Brierley, 2017), which contains 214 viruses, with discovery dates between from 1901 to 2017. Search terms, databases searched, and inclusion or exclusion criteria for data collection was provided in the previous paper (Woolhouse & Brierley, 2018). The updated version to 2018 included nine additional human virus species recently recognised by ICTV or newly added to the database: *Nairobi sheep disease orthonairovirus*, *Achimota virus 2*, *Menangle rubulavirus*, *Madariaga virus*, *Pegivirus H*, *Central chimpanzee simian foamy virus*, *Guenon simian foamy virus, Enterovirus H* and *Orthohepevirus C* (Table A.1). The metadata provided information on discovery date, transmissibility, transmission route, and host range (Woolhouse & Brierley, 2018).

I defined “discovery” as the first report of an ICTV-recognised RNA virus species from human(s) in the peer-reviewed literature, and the location of initial human exposure/infection with the virus was taken as the discovery location. When the location was not given from the original paper, the site of the

research laboratory was used as the discovery location (n=3). If neither human exposure/infection location nor research laboratory site were available, the address of the first author was used as the discovery location instead (n=19). In my database, locations of initial human exposure/infection were used for 201 (90%) viruses (Table A.1) and none of these were contracted while travelling. The locations were georeferenced as precisely as possible according to the original literature, ranging from precise coordinates of points to polygon-level data (e.g. city, county, district, state, or country) (Table A.1). Where no point location was provided, location was assigned to a set of grid cells corresponding to the maximum resolution available (e.g. city, district, country). A spatial polygon was created for those locations at administrative level 3 and above. As shown in Table A.1, administrative divisions of different countries vary: level 3 may be city, county, district, municipality; level 2 includes state, province, region, department, district, prefecture, city, and county; level 1 is country. All grid cells falling into these polygons were recorded. Grid cells that have more than ½ area falling in the polygon were also added manually. If no grid cell was identified, the coordinate of the centroid of the polygon was used instead. For locations equal to and lower than administrative level 4, I used the grid cell that the centroid of the location fell into. Table 2.1 summarises types of data within the occurrence database. The majority of occurrence records involved point data (56.0%), while the remainder were recorded at city or state or even country level. Thirty-five viruses were found in multiple locations according to the first published report, and 19 had limited geographical extent (i.e. infecting populations on no more

than two continents). For these, I selected one of these locations based on three criteria in order: (i) the virus being named after the location; (ii) the location was the same as first author or the laboratory; (iii) the location with the highest number of cases. For unspecified locations covering more than one grid cell (Table 2.1), sampling was used in my bootstrap framework as described below.

Table 2.1 Resolution and covered grid cells for global human-infective RNA virus discovery data

	Polygon data			Point data	Total
	Country level	State level	City level		
Virus species counts	3 (1.4%)	44 (19.7%)	50 (22.4%)	126 (56.5%)	223
Gridded cell counts	74	507	53	191*	825

*Grid cell counts here include viruses first detected in multiple points from the literature

2.3.3 Spatial predictors

A set of 33 variables potentially affecting the spatial distribution of RNA virus discovery were collated and used as predictors. Full details of sources, original resolutions, along with the definitions are provided in Table 2.2. The variables were assigned to four groups: climatic, socio-economic, land use, and biodiversity. I expected GDP, GDP growth, and university count to be correlated with discovery effort as they imply more resources that could be invested in virus research (Lipkin, 2013; Rosenberg et al., 2013). Other groups of variables including land use, climate, and biodiversity were more likely to be related to the natural geographic range of the virus (Morse, 1995), i.e. these variables would affect discovery via the intermediate step of emergence.

Table 2.2 List of predictors included in the model for global human-infective RNA virus discovery

Variable of interest	Definition	Group	Resolution	Source
Temperature	1. Annual mean temperature (℃) from 1901 to 2015: calculated from monthly average temperature 2. Annual mean temperature change (℃) from 1901 to 2015: calculated from monthly average temperature 3. Maximum temperature of the warmest month (℃) from 1901 to 2015: calculated from monthly maximum temperature 4. Minimum temperature of the coldest month (℃) from 1901 to 2015: calculated from monthly minimum temperature 5. Annual mean diurnal temperature range (℃) from 1901 to 2015: calculated from monthly diurnal temperature range	Climatic	0.5°	Climatic Research Unit: High-resolution gridded datasets (Harris et al., 2014)
Precipitation	1. Annual total precipitation (mm) from 1901 and 2015: calculated by summing the monthly precipitation 2. Annual total precipitation change (mm) from 1901 and 2015: calculated by summing the monthly precipitation 3. Maximum precipitation (mm) from 1901 to 2015: precipitation of wettest month, calculated from monthly precipitation 4. Minimum precipitation (mm) from 1901 to 2015: precipitation of driest month, calculated from monthly precipitation 5. Annual rainy days (counts) from 1901 to 2015: calculated by summing the monthly rainy-day counts	Climatic	0.5°	Climatic Research Unit: High-resolution gridded datasets (Harris et al., 2014)
Latitude	In angular (ranges from 0° to 90°)	Climatic	-	-
Human population	1. Human population count from 1970 to 2000 (counts, in persons) 2. Population growth: The decadal difference of human population (counts, in persons)	Socio-economic	30"	Global Rural-Urban Mapping project (Center for International Earth Science Information Network - CIESIN - Columbia University, 2017)

Gross domestic product (GDP)	1. Mean GDP (unit: PPP, billion US$2005/yr) for three time steps: 1980, 1990, 2000 and 2010, which are estimated by downscaling actual GDPs by country 2. GDP growth: The decadal difference of GDP (unit: PPP, billion US$2005/yr)	Socio-economic	0.5°	Centre for Global Environmental Research: Global dataset of gridded population and GDP scenarios (Murakami & Yamagata, 2016)
University count	Higher education institutions count offering at least a 4-year professional diploma or a post-graduate degree in each country	Socio-economic	Country level	World Higher Education Database https://www.iau-aiu.net/World-Higher-Education-Database-WHED
Land use	The percentage of 1. Cropland, 2. Pasture, 3. Urban land, 4. Primary land, and 5. Secondary land in each grid cell from 1900 to 2015. Primary land: natural vegetation (either forest or non-forest) that has never been impacted by human activities since 1700; Secondary land is natural vegetation (either forest or non-forest) that is recovering from previous human disturbance. 6. Growth of cropland area, 7. Growth of pasture area, 8. Growth of urbanized land area, 9. Growth of primary land area, and 10. Growth of secondary land area: The percentage of land area change for each category in each grid cell from 1900 to 2015 11. Urbanization of cropland, 12. Urbanization of pasture, 13. Urbanization of primary land, and 14. Urbanization of secondary land: The percentage of land area change from cropland/pasture/primary land/secondary land to urban land in each grid cell from 1900 to 2015 15. The percentage of cultivated and managed vegetation in each grid cell	Land use	0.5°/30"	Harmonized Global Land Use (Chini et al., 2014) EarthEnv (Tuanmu & Jetz, 2014)
Mammal species richness	Mammal species richness for 2015 represented the number of species in a particular class, family or International Union for the Conservation of Nature (IUCN) threatened category.	Biodiversity	30"	Gridded Species Distribution (International Union for Conservation of Nature - IUCN & Center for International Earth Science Information Network - CIESIN - Columbia University, 2015)
Livestock headcount	Domestic animal headcount, summed cattle, buffalo, sheep, goats, pigs	Biodiversity	30"	Gridded Livestock of the World (Robinson et al., 2014)

All predictors and virus locations were matched by 1° spatial grid cell, having rescaled or transformed the data where necessary. For gridded data at the 30" and 0.5° resolution, I have rescaled them to 1° resolution by using the resampling approach in R package “raster”. The “bilinear” interpolation was used to compute values for the re-gridded data. For data at the country level, I first assigned each country to the grid cells it covered, which can be realized by the “over” function in R package “sp”, and then matched variables at the country level with gridded data by coordinates.

My model matched the RNA virus discovery count in each grid cell with historical decadal climatic variables, population, GDP, and land use data (described below), so I extrapolated the data for these variables back to 1901 (Population and GDP only, as climatic variables and land use data have full temporary coverage from 1901 to 2015). I extrapolated the data for population and GDP by using the growth rate at grid cell level or country level. Gridded data for population from 1970 to 2000 were available from Socioeconomic Data and Applications Centre (SEDAC)’s Global Rural-Urban Mapping Project. Gridded population counts after 2000 were estimated from the gridded population data from 2000 to 2020, which were also provided by SEDAC (http://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev10).
To reconcile the population difference from the two databases, I calculated the population growth rate in each grid cell after 2000 based on the gridded data from 2000 to 2020, ignoring the absolute population size. I then estimated the population after 2000 for each grid cell by multiplying the value of gridded population in 2000 (from gridded population data from 1970 to 2000) by the

growth rate I calculated above. Data before 1970 were estimated by population growth at country level, as there was no population data at grid level for this stage. I assumed that the population growth rate for each grid cell in the same country was similar. The total population for each country was obtained from Our World in Data (https://ourworldindata.org/world-population-growth), and the growth rate of population for each country before 1970 was derived from these data. Similarly, by applying the growth rate for the gridded population data for 1970, I calculated the population in each grid cell from 1900 to 1960. Data for GDP before 1980 were estimated by GDP growth at country level. The GDP for each country was obtained from Our World in Data (https://ourworldindata.org/grapher/world-gdp-over-the-last-two-millennia).

2.3.4 Boosted regression tree modelling

By fitting a Poisson BRT model, I estimated the relative risk of RNA virus discovery for each 1° resolution of grid cell across the world as a function of the 33 predictors. BRT is a tree-based machine learning method beginning to be widely used in ecological studies (Redding et al., 2016; Shearer et al., 2018). It applies the technique of boosting to combine many simpler tree models adaptively, and renders improved predictive performance (De'ath, 2007; Elith et al., 2008). Tree-based learning methods are useful tools for modelling non-linear relationships and higher order interactions between variables. In addition, BRT handles spatially dependent data well, as it can capture complex structures within the data that many other modelling methods cannot (Crase et al., 2012). I calculated Moran's I (an index of spatial dependence) to

estimate the ability of the BRT model to account for spatial dependence in the virus data, using package spdep in R v. 3.5.1 (fixed distance weights were generated based on spherical distance, with the cut-off values ranging from one time to thirty times of distance of 1° resolution grid cell at the equator, i.e. 110km to 3300km) (Cliff & Ord, 1981). Unlike the traditional, significance-based approaches, BRT assesses the individual effect of each variable by estimating the relative importance of each variable to the predictions.

The bootstrap resampling approach was applied to account for spatial uncertainty in the location of virus discoveries and generated 95% quantiles. For viruses with imprecise discovery locations, one grid cell was randomly selected each time. For each grid cell with virus discovery, two grid cells with no discovery were randomly selected from all cells throughout the world that were 'virus discovery free' at all time points. So, in each model, 223 grid cells with virus discovery and 446 with no virus discovery were included. I matched the virus data with all predictors (using the same decade for time-varying predictors, e.g. 2010 values of variables were matched with viruses discovered in 2005–2014). The virus count in any given grid cell in each decade followed a Poisson distribution, and the virus discovery count in each grid cell by decade was used as the response variable.

Using bootstrap resampling, I fitted 1000 replicate BRT models and generated relative contribution plots and partial dependence plots with 95% quantiles. The relative contribution, or the influence/weight, of each variable is an indicator of that variable's importance for predicting virus discovery counts.

The relative contributions of all variables of a BRT model sum to 100%, with higher numbers indicating stronger influence on the response. I defined the most influential predictors as those whose relative contribution was greater than the mean level (i.e. 100/(total predictorscounts*100); this study: 100/(33*100)=3.03%) (Shearer et al., 2018). Partial dependence plots are a method of visualizing the relationships between a BRT's predictive variables and its outcome after accounting for the average effects of all other variables. The means of the predictions of all 1000 models were used to predict the probability of virus discovery across the globe in 2010–2019, using 2015 values of the 33 predictors. Using the equation of Poisson probability distribution, I converted the continuous prediction map to a probability map. I used the packages dismo and gbm in R v. 3.5.1 to fit BRT models. Parameters including tree complexity (reflecting the number of nodes in a tree), learning rate (shrinking the contribution of each added tree), and bag fraction (specifying the proportion of data to be selected at each step) were set following Elith et al. (Elith et al., 2008) to make sure each resampling model contained at least 1000 trees. The final parameters of the optimal model had the following values: tree complexity = 5, learning rate = 0.003, bag fraction = 0.5. A cross-validation stagewise function was used to identify the optimal number of trees in each model (Elith et al., 2008). With these parameters, the 1000 replicate BRT models fitted a mean of 1214 trees.

The model's predictive performance was assessed by calculating the cross-validated explained deviance of the bootstrap model (a measure of the goodness-of-fit between the predicted and raw data) (Elith et al., 2008), as well

as by conducting 50 rounds of ten-fold cross-validation. Fifty data sets were randomly selected from the data for 1000 replicated models to conduct 50 rounds of ten-fold cross-validation. As before, 223 discovery present and 446 discovery absent samples were drawn in each of the 50 rounds, and the calculated virus count in each grid cell by decade was used as the response. The dataset of each round was randomly divided into ten subsets. For each round of ten-fold cross-validation, ten different training sets comprising unique combinations of nine subsets were used to fit models, and the remaining one was used to evaluate the predictive performance of the model as a test set. The median infraclass correlation coefficient (ICC) with 95% quantiles was used as the validation statistics. The ICC ranges from 0 and 1, with an ICC of less than 0.40 suggesting a poor predictive ability, 0.40–0.59 suggesting a fair predictive ability, 0.60–0.74 suggesting a good predictive ability, and 0.75–1 suggesting an excellent model (Cicchetti, 1994).

I also performed sensitivity analyses by i) using data from 1980 to 2000 only (as explanatory variables are available without extrapolation only for this period), and ii) removing the 22 discovery reports that were not locations of infected humans (as these are less precise). Model parameters are provided in Table 2.3.

2.3.5 Stratified analysises

Two stratified analyses were conducted to find predictors specific to discoveries of different categories of virus. The first stratified analysis distinguished 131 viruses that were strictly zoonotic (all human infections are

acquired from an infection in a non-human reservoir) and the 92 viruses that could spread within human populations (i.e. are transmissible, directly or indirectly, between humans) (Table A.1), based on previously published data (Woolhouse et al., 2016). A second stratified analysis was performed separately for 93 vector-borne viruses and 130 non-vector-borne viruses (Table A.1). I used the same BRT modelling approach for stratified analyses as I described before, and relative contribution plots and partial dependence plots with 95% quantiles were drawn for each category of virus. Model parameters are provided in Table 2.3. Based on stratified BRT models, predictions of discovery probability for each category of viruses in 2010–2019 were also performed by using 2015 values of the 33 predictors.

All statistical analyses were performed using R software, version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria), and all maps were visualised by using ArcGIS Desktop 10.5.1 (Environmental Systems Research Institute).

Table 2.3 Model parameters for sensitivity analyses and stratified analyses for global human-infective RNA virus discovery

Model	Tree complexity	Learning rate	Bag fraction	No. of trees
Sensitivity analysis (1)	2	0.0015	0.5	1129
Sensitivity analysis (2)	5	0.0030	0.5	1051
Strictly zoonotic	4	0.0020	0.5	1114
Transmissible	2	0.0020	0.5	1430
Vector-borne	2	0.0040	0.5	1147
Non-vector-borne	2	0.0035	0.5	1080

(1) Using data from 1980 to 2000 only; (2) Removing the 22 discovery reports that were not patients' locations.

2.4 Results

The five regions with the highest virus count were eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America [Figure 2.1 (A)]. Strictly zoonotic viruses and vector-borne viruses were mostly discovered from central Africa and north-eastern South America while transmissible viruses and non-vector-borne viruses were mostly discovered from eastern North America and Europe (Figure 2.2). The cumulative discovery count increased slowly before 1950s, and thereafter increased at a constant rate [Figure 2.1 (B)]. There is variation for the rate of detection by geographic region. More viruses have been discovered in North America and Europe, but the numbers have decreased in recent decades. By contrast, an increased number of viruses have been discovered in Asia. Transmissible viruses and non-vector-borne viruses showed a similar temporal pattern with the curve for all human-infective RNA viruses, with an obvious increase in 1950 (Figure 2.2). Strictly zoonotic viruses and vector-borne viruses showed a similar pattern in the early phase, with an obvious increase in 1925, but the numbers of new vector-borne viruses decreased after 1980 (Figure 2.2).

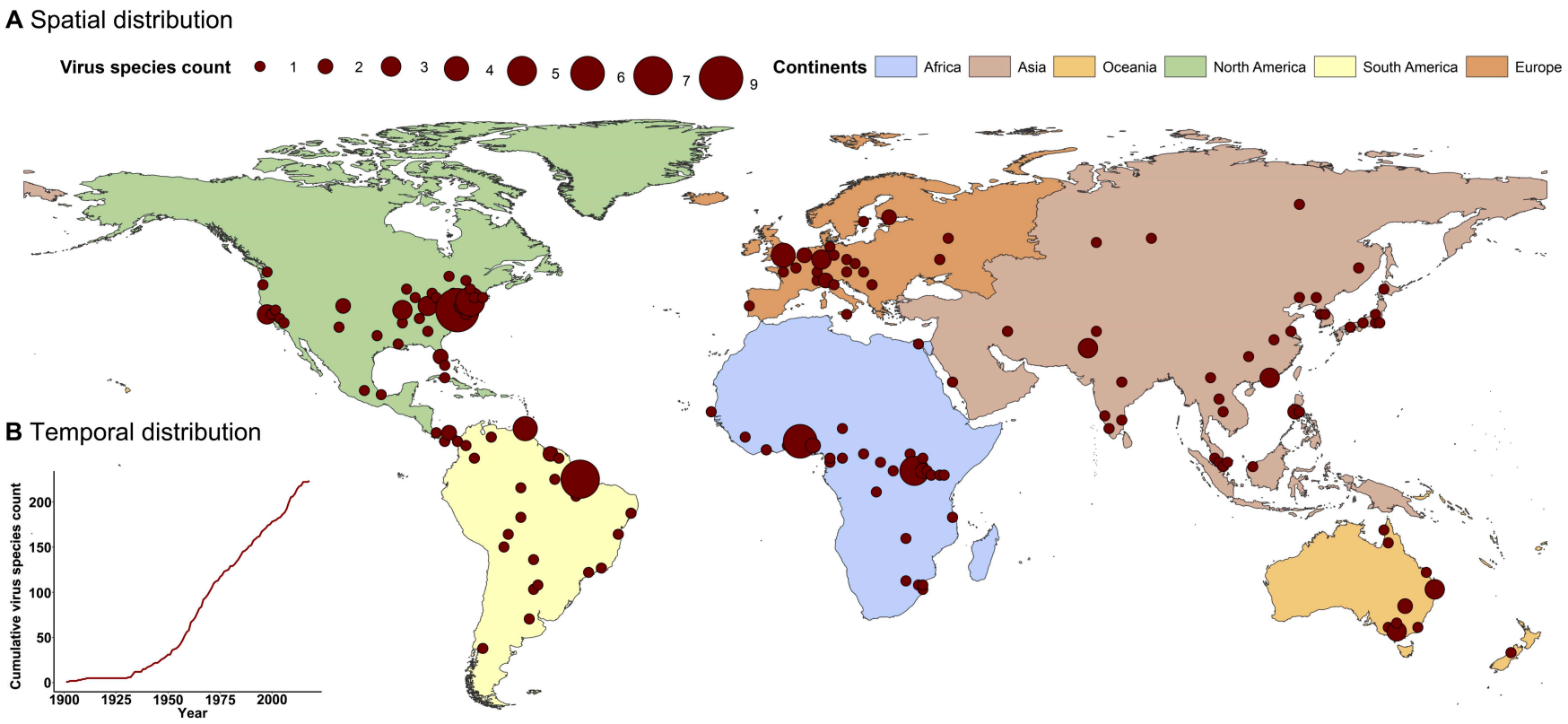


Figure 2.1 Spatiotemporal distribution of global human-infective RNA virus discovery in 1901–2018

A, Spatial distribution. The red spots indicate discovery points or centroids of polygons (administrative regions)—depending on the preciseness of the location provided by the original paper, with the size representing the cumulative virus species count. Centroid is the coordinate of the centre of mass in a spatial object. B, Temporal distribution. The red curve indicates the cumulative virus species discovery count over time.

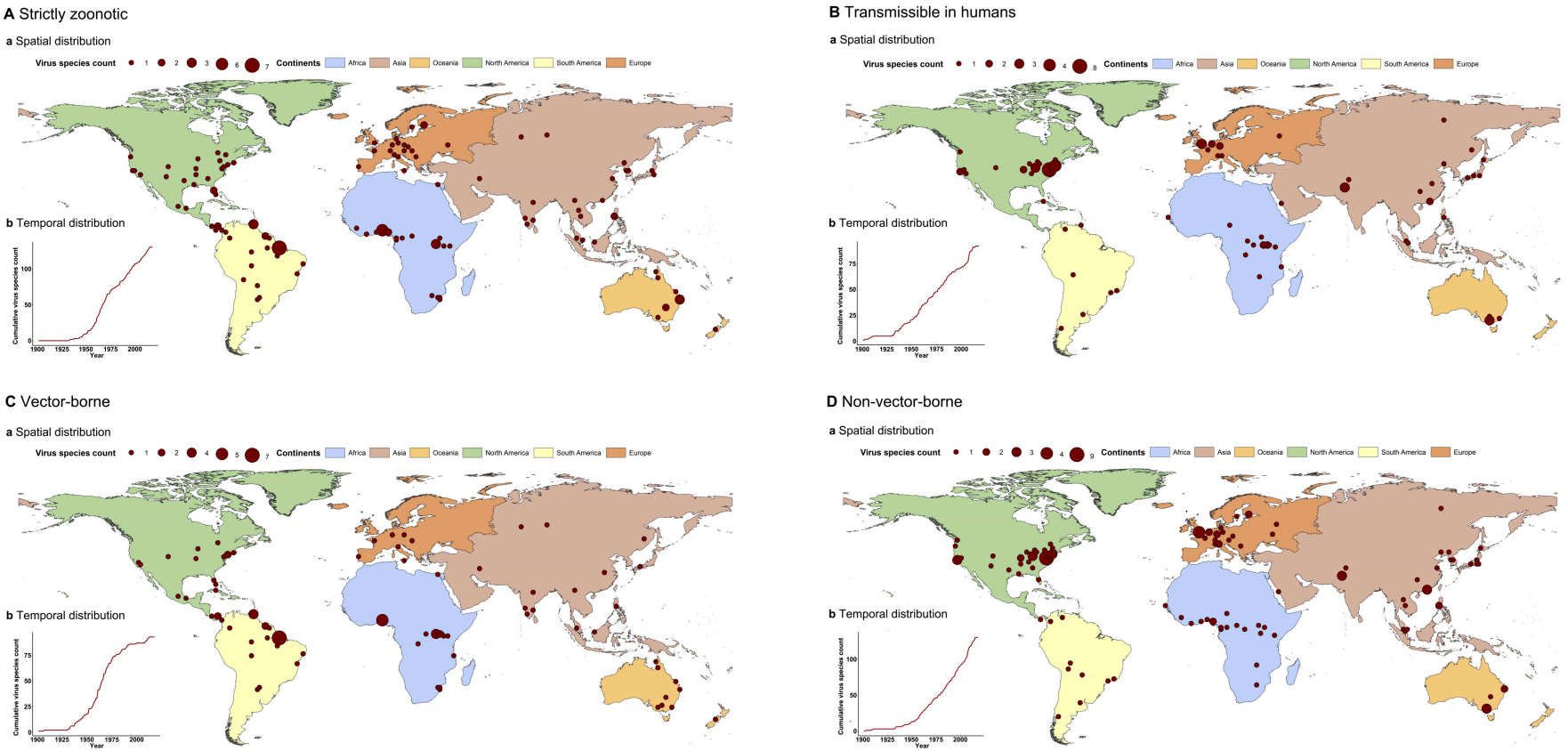


Figure 2.2 Spatiotemporal distribution of global human-infective RNA virus discovery count split by category in 1901–2018

The map was plotted with respect to transmissibility (A, strictly zoonotic; B, transmissible in humans), and transmission mode (C, vector-borne viruses; D, non-vector-borne viruses). In each subplot, the red spots in the map (a) indicate discovery points or centroids of polygons (administrative regions)—depending on the preciseness of the location provided by the original paper, with the size representing the cumulative virus species count. Centroid is the coordinate of the centre of mass in a spatial object. The red curve (b) indicates the cumulative virus species discovery count over time.

Based on the full BRT model involving all 223 viruses, twelve variables had relative contributions greater than the mean (3.03%) (Figure 2.3), including two socio-economic variables (GDP growth: 12.7%, GDP: 9.9%), four variables concerning urbanization [urbanized land: 8.7%, urbanization of secondary land (i.e. the percentage of land area change from secondary land to urban land; secondary land is natural vegetation that is recovering from previous human disturbance, see Table 2.2 for details): 4.8%, growth of urbanized land area: 3.6%, and urbanization of cropland (i.e. the percentage of land area change from cropland to urban land, see Table 2.2 for details): 3.3%], five climatic variables (minimum temperature: 6.3%, precipitation change: 5.0%, latitude: 4.3%, total precipitation: 3.6%, minimum precipitation: 3.5%), and one biodiversity variable (mammal species richness: 5.1%). The partial dependence plots shown in Figure 2.4 showed the relationships between these predictors and virus discovery. For the majority of predictors, the relationship with discovery probability was non-linear, with large effects often seen over a narrow range of values. For example, discovery probability fell sharply if GDP growth was negative, and for very low GDP and low percentage of urbanized land; whereas it rose sharply for high minimum temperature and high mammal richness.

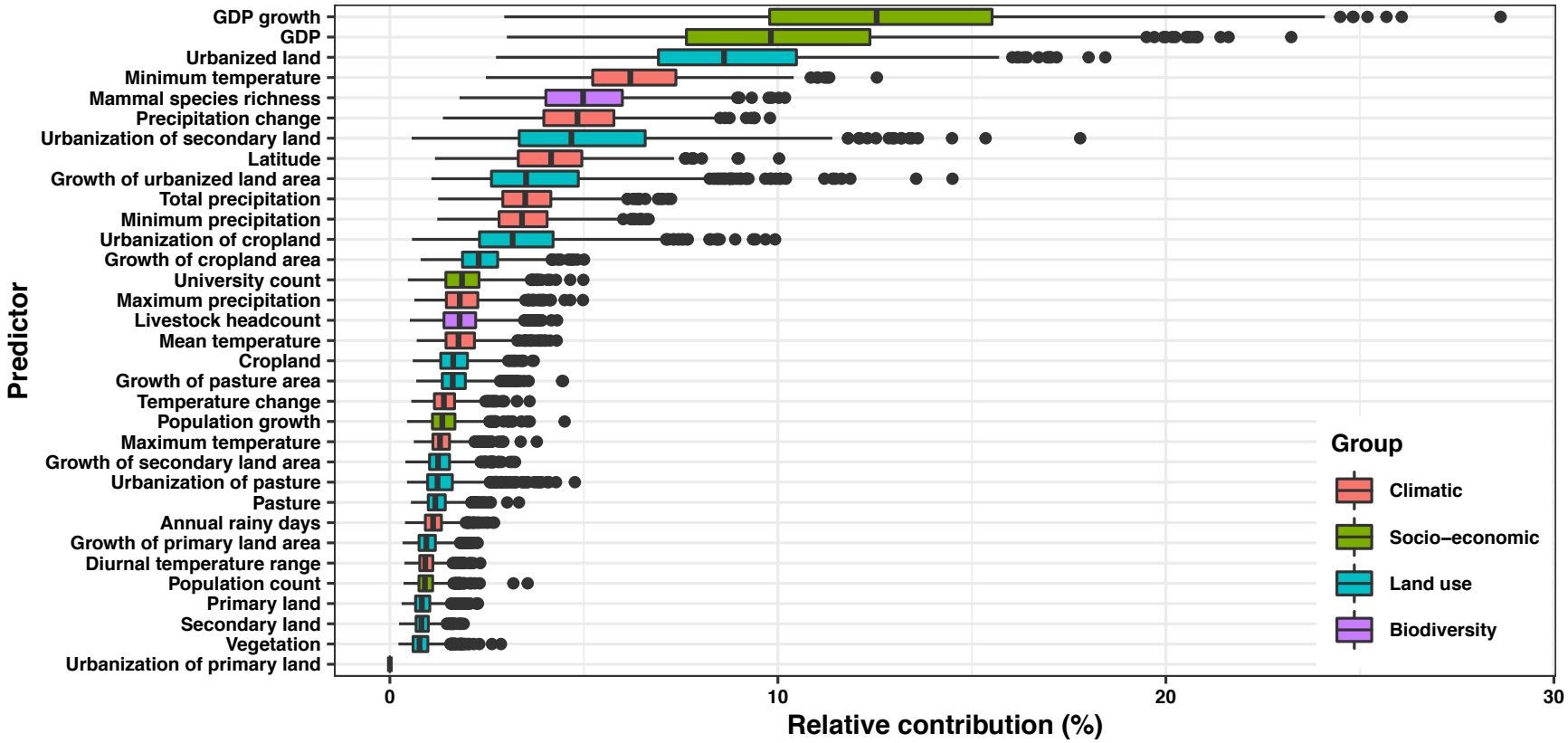


Figure 2.3 Relative contribution of predictors to global human-infective RNA virus discovery in the full model

The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

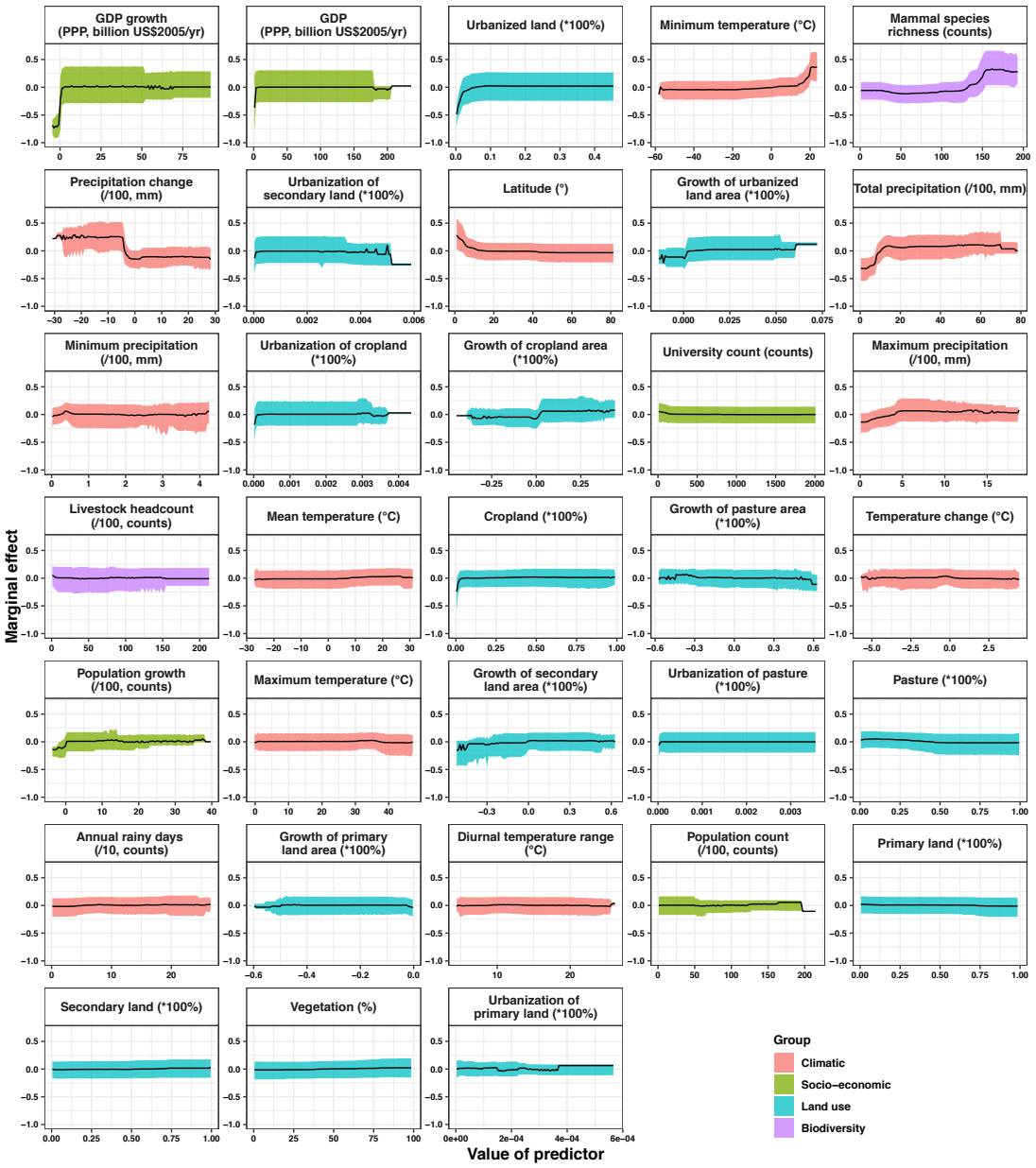


Figure 2.4 Partial dependence plots for all predictors that influence global human-infective RNA virus discovery in the full model

Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

My full BRT model reduced the Moran's I for the raw virus data from a range of 0.04–0.31 to 0.007–0.065 (Figure 2.5), indicating that this modelling method with 33 predictors effectively removed the spatial dependence of the model

residuals. Sensitivity analyses (the analysis using data from 1980 to 2000 and the analysis after removing the 22 viruses with least certain discovery locations) revealed consistent trends with the full model, though with several changes of relative contribution.

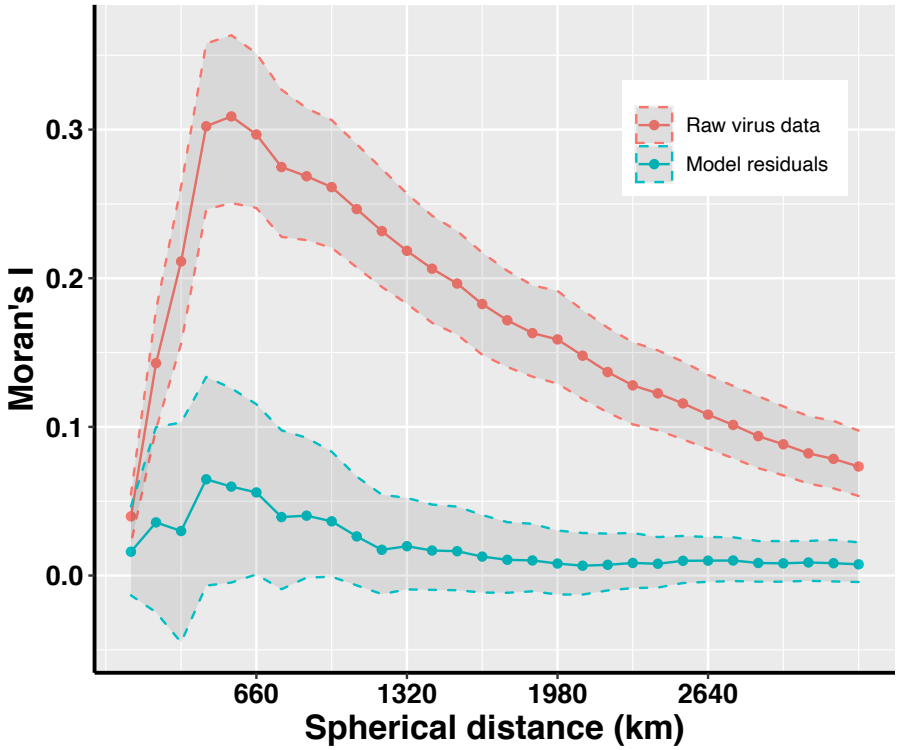


Figure 2.5 Moran's I across different spherical distances for the global raw human-infective RNA virus data and the full model residuals

The solid line and dots represented the median Moran's I value, and the grey area represented its 95% quantiles generated from 1000 samples (red: Raw virus data) or replicate BRT models (blue: Model residuals).

The bootstrap 1000 replicate BRT full model explained a median of 40.9% of deviance in the data (95% quantiles: 36.5%–45.5%). Using 50 runs of 10-fold cross-validation, the model had a median ICC of 0.55 (95% quantiles: 0.50–0.61). The model validation statistics for the stratified model are shown in Table 2.4. Combining these measures, my BRT model predictions ranged from fair to good (Cicchetti, 1994).

Table 2.4 Model validation statistics for stratified analyses for global human-infective RNA virus discovery

Model	% of deviance explained (95% quantiles)	ICC (95% quantiles)
Strictly zoonotic	60.0% (53.9%–65.3%)	0.44 (0.36–0.51)
Transmissible	44.7% (40.1%–49.7%)	0.63 (0.55–0.70)
Vector-borne	63.3% (55.6%–71.7%)	0.43 (0.34–0.52)
Non-vector-borne	40.8% (36.2% –45.3%)	0.67 (0.62–0.73)

ICC, infraclass correlation coefficient

In the transmissibility-stratified BRT model, ten variables had relative contributions greater than 3.03% for discovering strictly zoonotic viruses [Figure 2.6(A), partial dependence plots in Figure 2.7(A)], including four climatic variables (minimum temperature: 13.1%, latitude: 6.2%, precipitation change: 5.3%, total precipitation: 3.6%), three land use variables (urbanized land: 7.7%, urbanization of secondary land: 5.6%, growth of urbanized land area: 5.2%), two socio-economic variables (GDP: 8.3%, GDP growth: 7.9%), and one biodiversity variable (mammal species richness: 5.6%). In contrast, eight variables had relative contributions greater than 3.03% for discovering viruses transmissible in humans [Figure 2.6(B), partial dependence plots in Figure 2.7(B)], including four predictors involving urbanization (urbanized land: 13.6%, urbanization of cropland: 9.3%, urbanization of secondary land: 6.6%, growth of urbanized land area: 3.6%), three socio-economic variables (GDP growth: 14.4%, GDP: 14.0%, population change: 3.6%), and one climatic variable (minimum precipitation: 5.0%).

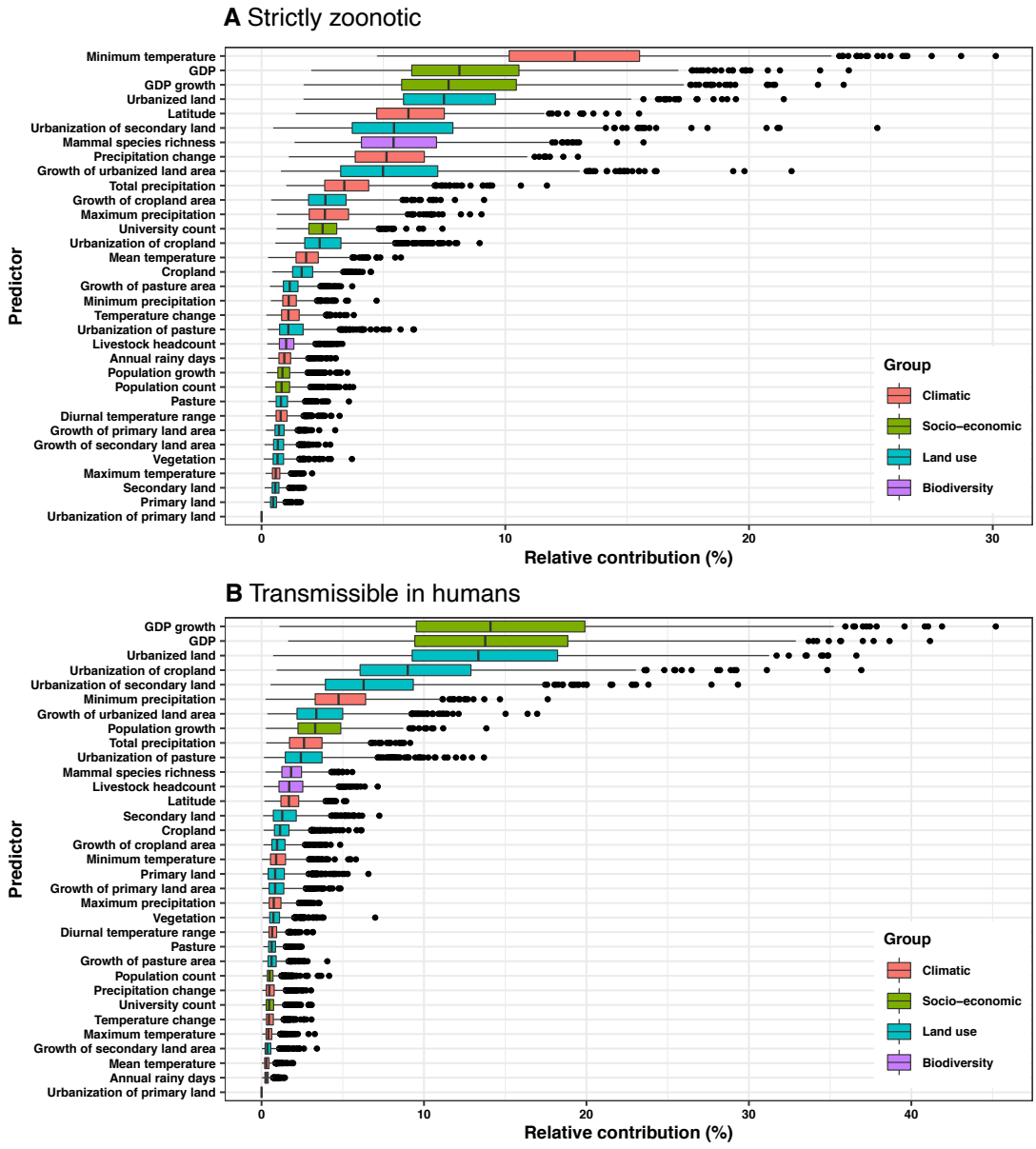


Figure 2.6 Relative contribution of predictors to global human-infective RNA virus discovery in the stratified model by transmissibility

A, Strictly zoonotic; B, Transmissible in humans. The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

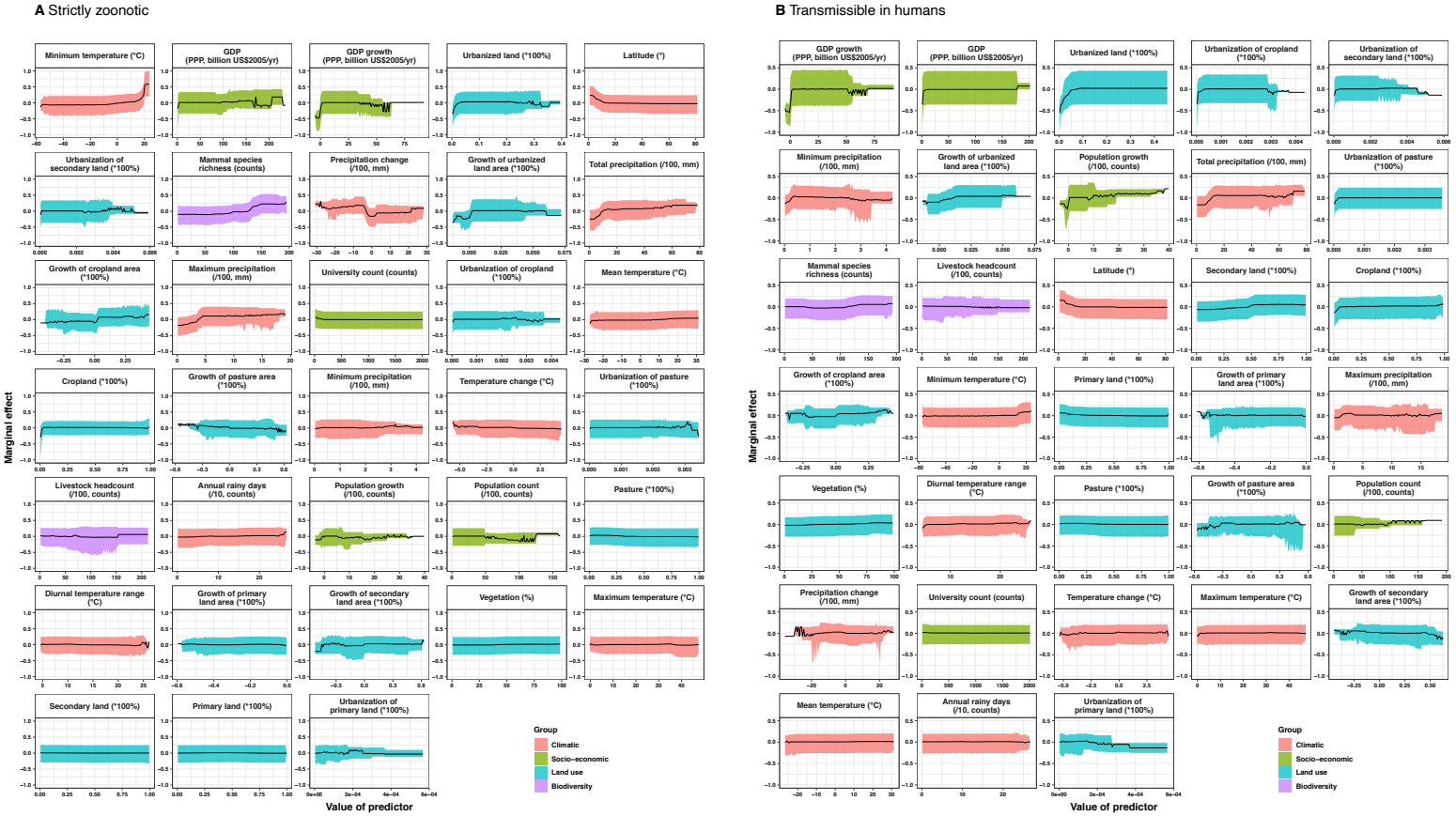


Figure 2.7 Partial dependence plots for all predictors that influence global human-infective RNA virus discovery in the stratified model by transmissibility

A, Strictly zoonotic; B, Transmissible in humans. Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

In the vector-borne-stratified BRT model, thirteen variables had relative contributions greater than 3.03% for discovering vector-borne viruses [Figure 2.8(A), partial dependence plots in Figure 2.9(A)], including five climatic variables (minimum temperature: 17.1%, precipitation change: 7.9%, latitude: 6.2%, total precipitation: 3.8%, maximum precipitation: 3.3%), two socio-economic variables (GDP growth: 7.4%, GDP: 4.4%), one biodiversity variable (mammal species richness, 6.7%), and five land use variables (urbanization of secondary land: 4.8%, urbanized land: 4.1%, growth of cropland area: 3.7%, growth of urbanized land area: 3.6%, growth of pasture area: 3.4%). In contrast, seven variables had relative contributions greater than 3.03% for discovering non-vector-borne viruses [Figure 2.8(B), partial dependence plots in Figure 2.9(B)], including four land use variables (urbanized land: 19.6%, urbanization of secondary land 7.5%, urbanization of cropland: 4.5%, growth of urbanized land area: 3.5%), two socio-economic variables (GDP: 18.7%, GDP growth: 12.4%), and one climatic variable (minimum precipitation: 3.3%).

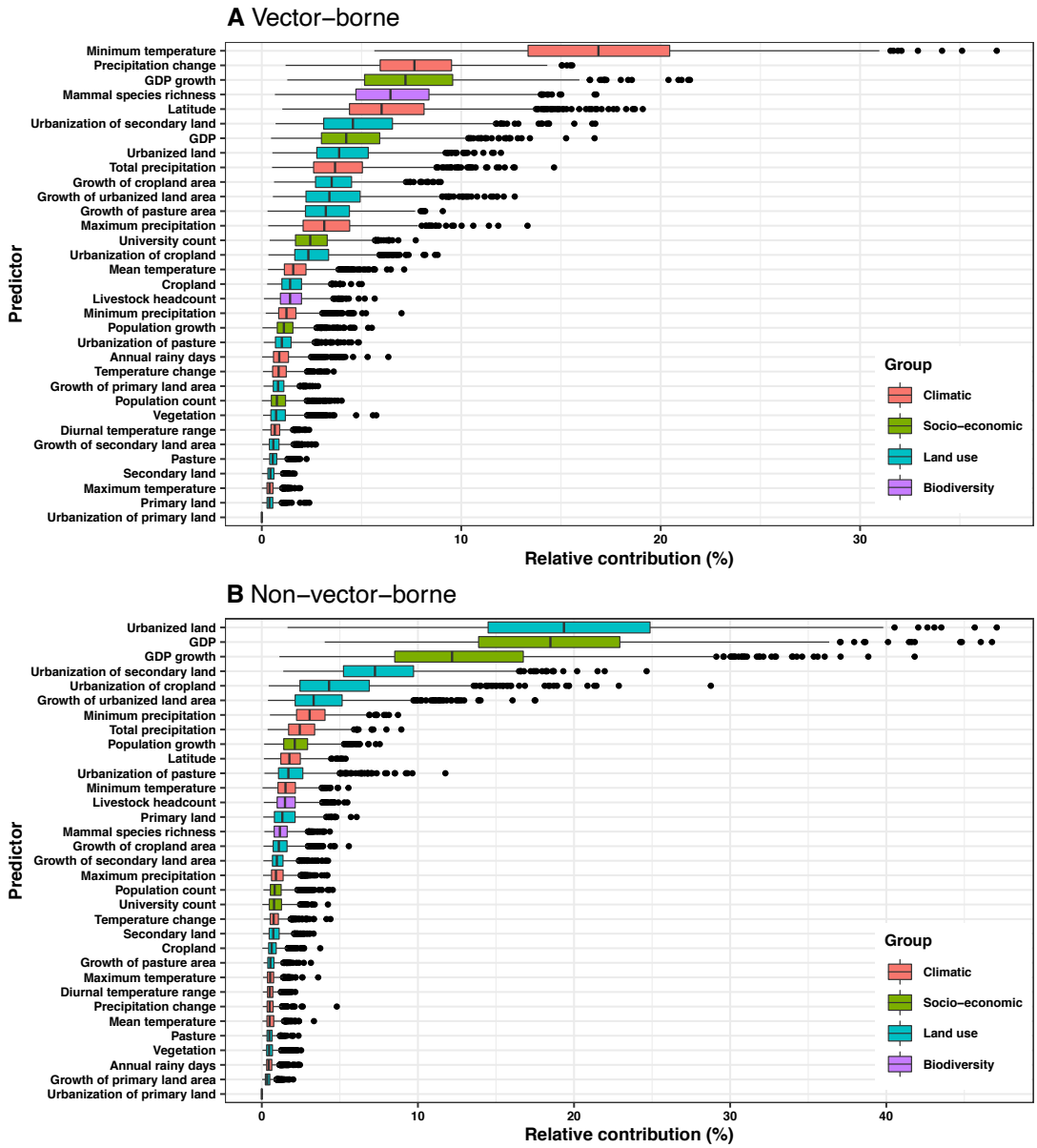


Figure 2.8 Relative contribution of predictors to global human-infective RNA virus discovery in the stratified model by transmission mode

A, Vector-borne; B, Non-vector-borne. The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

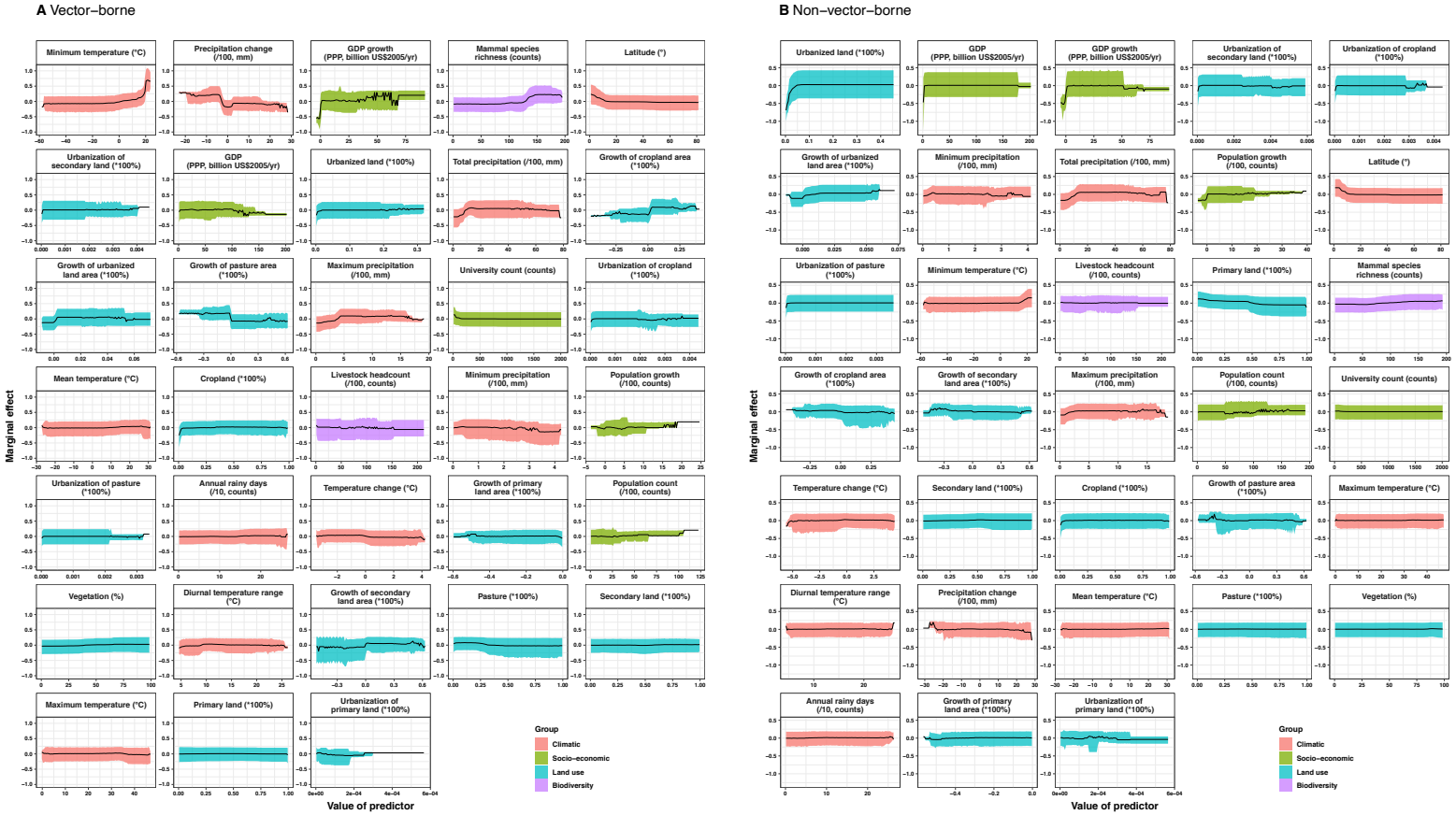


Figure 2.9 Partial dependence plots for all predictors that influence global human-infective RNA virus discovery in the stratified model by transmission mode

A, Vector-borne; B, Non-vector-borne. Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

The summary of the cumulative relative contribution of each group of predictors to human-infective RNA virus discovery in each model is shown in Figure 2.10. In comparison with non-vector-borne and human transmissible viruses, the discovery of vector-borne viruses and strictly zoonotic viruses was better predicted by climatic variables and biodiversity than by socio-economic variables and land use.

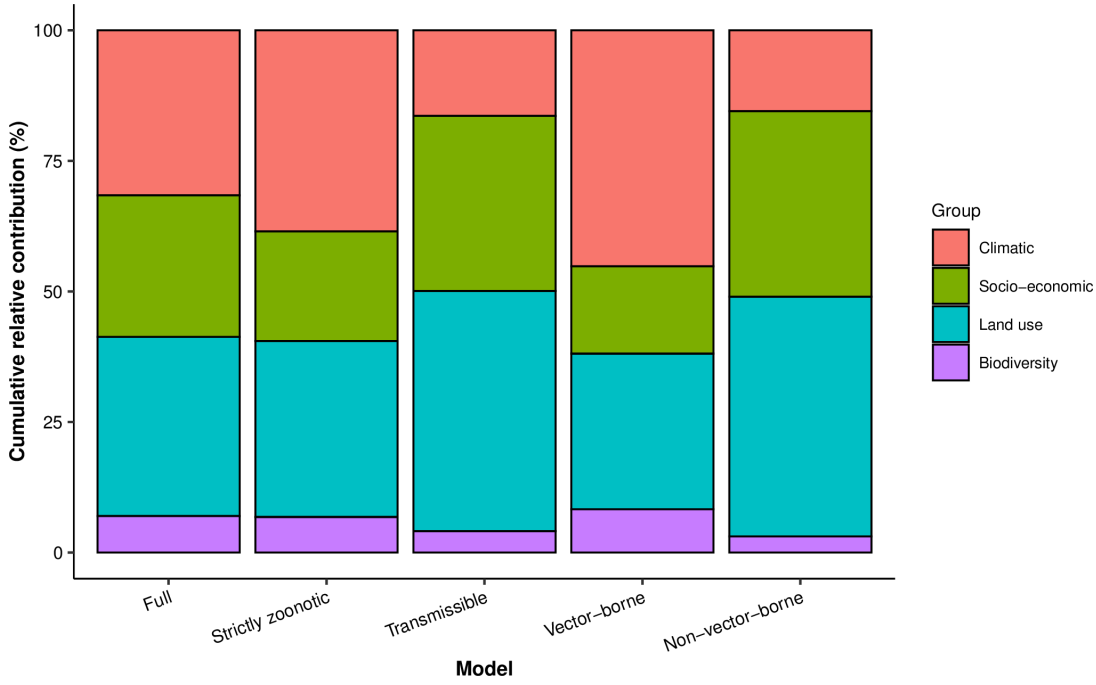


Figure 2.10 Cumulative relative contribution of predictors to global human-infective RNA virus discovery by group in each model

The relative contributions of all predictors sum to 100% in each model, and each colour represents the cumulative relative contribution of all predictors within each group. The relative contribution of different groups to virus discovery varies across each model.

By applying 2015 values of all 33 predictors (Figure A.1) to the fitted full BRT model, I obtained a predicted probability of human-infective RNA virus discovery in 2010–2019 (Figure 2.11). Comparison with Figure 2.1 indicates that virus discoveries remained relatively likely in eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America but,

in addition, I predicted high probabilities of virus discovery across East and Southeast Asia, India, and Central America. All eighteen new virus species since 2010 were discovered in regions of high-risk as predicted by my model (75.0%–99.9% percentiles of predicted probability over the global range), and eleven of them were discovered in very high-risk areas (90.0-99.9% percentiles of predicted probability over the global range). The predictions of discovery for each category of virus are shown in Figure 2.12. Broadly similar patterns as the full prediction model were seen for all four categories: high probabilities of virus discoveries were predicted in East and Southeast Asia, India, and Central America in comparison with the historical distribution (Figure 2.2). However, there was some variation between virus categories: strictly zoonotic viruses were more likely to be discovered in northern South America, central Africa, and Southeast Asia, while transmissible viruses were more likely to be discovered in North America, East Asia, and India (Figure 2.12); and vector-borne viruses were predicted to be more likely to be discovered in northern South America, central Africa, India, and Southeast Asia than non-vector-borne viruses (Figure 2.12).

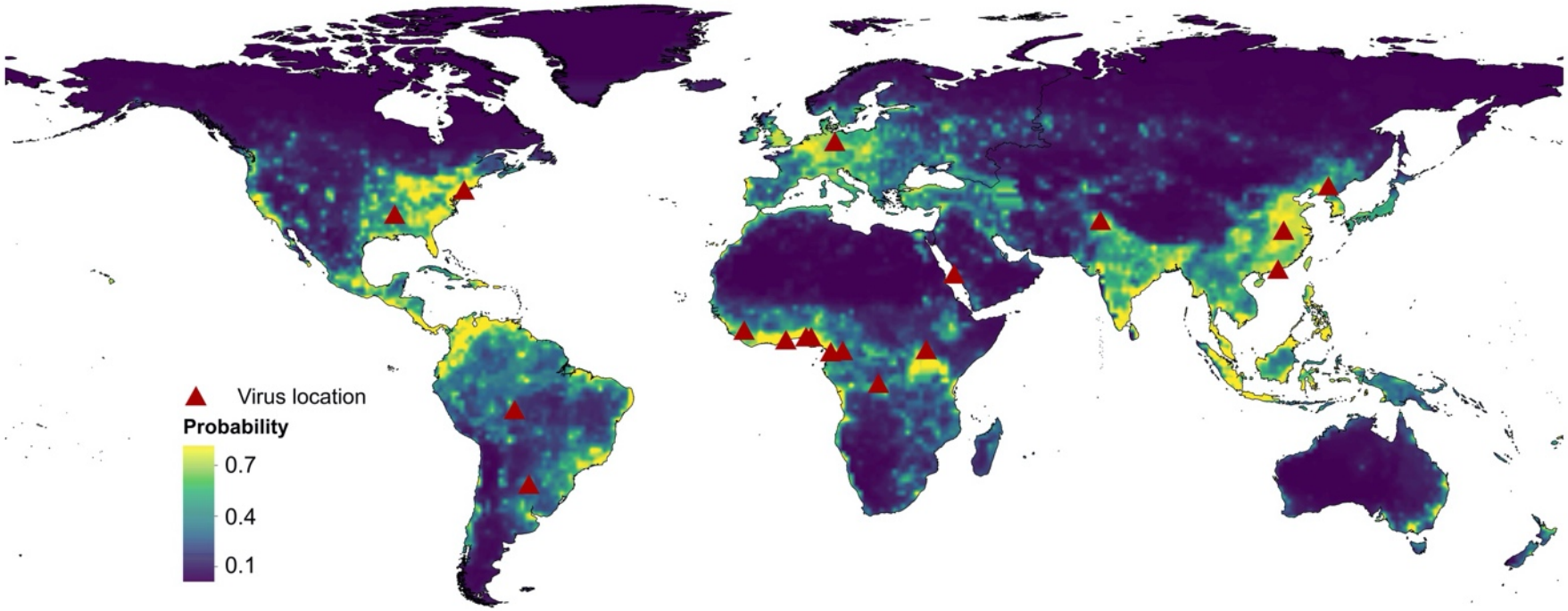


Figure 2.11 Predicted probability of global human-infective RNA virus discovery in 2010–2019

The triangles represented the actual discovery sites from 2010 to 2018, and the background colour represented the predicted discovery probability.

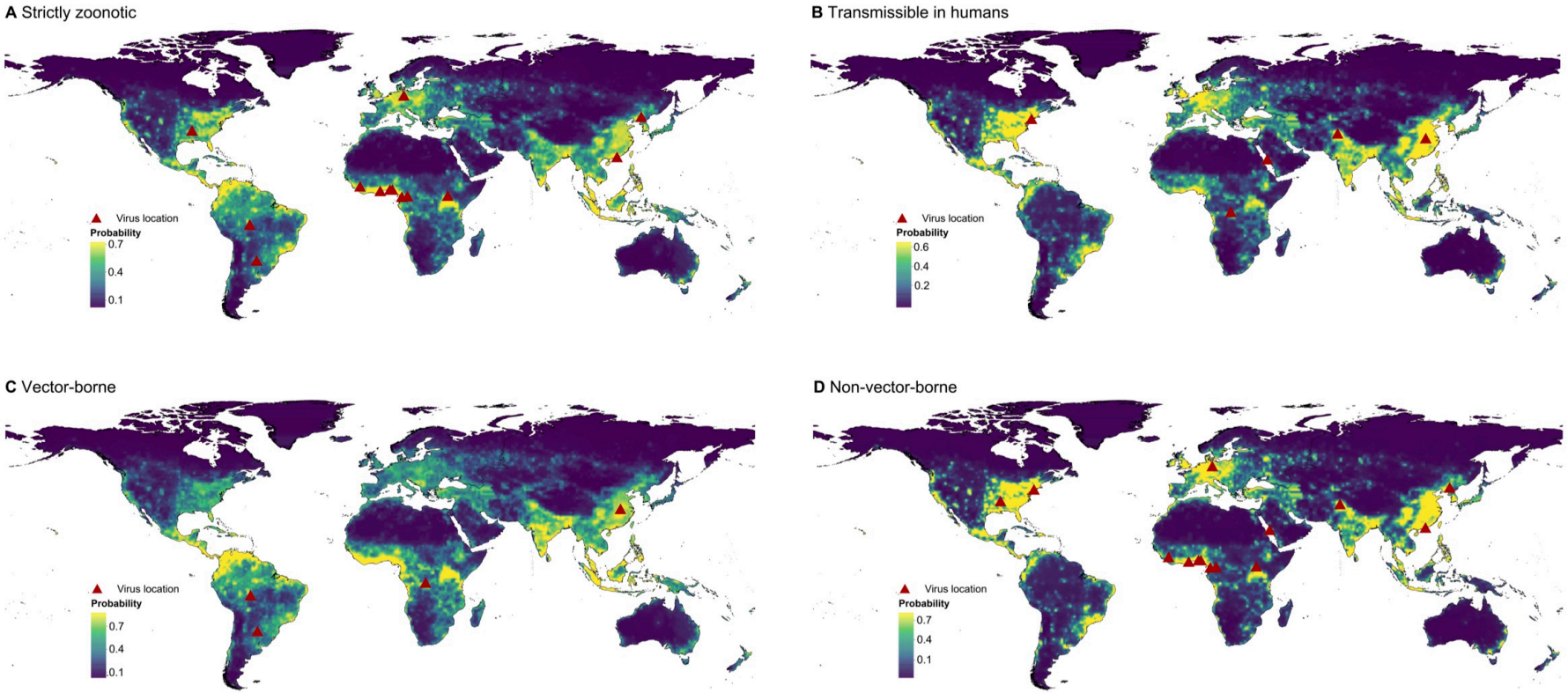


Figure 2.12 Predicted probability of global human-infective RNA virus discovery in 2010–2019 split by category

A Strictly zoonotic; B Transmissible in humans; C Vector-borne viruses; D Non-vector-borne viruses. The triangles represented the actual discovery sites from 2010 to 2018, and the background colour represented the predicted discovery probability.

2.5 Discussion

In this chapter I compiled a large body of information on global spatiotemporal patterns of human-infective RNA virus discovery and developed a spatiotemporal modelling framework to identify predictors of the discovery of new viruses. The maps of human-infective RNA virus discovery indicated five regions with historically high discovery counts: eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America. BRT modelling suggested that virus discovery was well predicted by socio-economic variables (especially GDP and GDP growth), land use variables (especially those related to urbanization), climate variables (including minimum temperature, precipitation change, latitude, minimum precipitation, total precipitation), and biodiversity (especially mammal species richness). The predicted probability map in 2010–2019 identified three new areas across East and Southeast Asia, India, and Central America in addition to the historical high-risk areas.

I focused on the discovery of RNA viruses in human(s) in this chapter, rather than emergence. This was determined by the attribute of the database itself, i.e. the first report of each human-infective RNA virus from the literature review. The discovery location may or may not represent the origin of the virus. For example, HIV-1 is believed to originate from non-human primates in West-central Africa, and is estimated to have transferred to humans in 1920s (Faria et al., 2014), but the first published case from peer-reviewed literature was a

Caucasian and was published by researchers in France (Barre-Sinoussi et al., 1983).

In both the full and the stratified BRT models, GDP and GDP growth were among the top predictors of virus discovery count. This is likely to reflect that richer, more developed areas have more research funding, better access to technologies for virus detection and more effective surveillance systems. In the United States, for example, the Global Virome Project (GVP) launched in 2018 aims to detect and characterize most zoonotic viruses from target zoonotic reservoirs with a cost of $1.2 billion (Carroll et al., 2018). Comparison of Figure 2.1 with Figure A.1 suggested that more viruses have been discovered in developed regions with/without fast GDP growth including North America, Europe, and Australia. I note that more developed countries were more likely to first capture viruses circulating in multiple regions. Over the last 100 years, North America and Europe have witnessed a decreasing fraction of discovered viruses in more recent decades (1985–2018: 32/86=37%) than previously (1901–1984: 78/137=57%), but Asia has accounted for a higher fraction (1901–1984: 16/137=12%; 1985–2018: 22/86=26%). This can be partly explained by the higher GDP and faster GDP growth in Asia in recent decades. In addition, there have also been historical hotspots in individual countries (e.g. Brazil, Nigeria, and Uganda) associated with active virus discovery initiatives such as those supported by the Rockefeller Foundation (RF, the RF supported the discovery of 23 vector-borne viruses since 1951; there was a sharp decline in the discovery of vector-borne viruses when the RF programme ended in the mid-1960s, Figure 2.2C) (Rosenberg et al., 2013).

More viruses are likely to be discovered in the near future in areas with high GDP growth and GDP including most of Asia (except North and Central Asia), Europe and North America.

In contrast to GDP, all other predictors identified in this chapter appear more directly associated with virus geographic distributions, my analysis having the important advantage that their influence was estimated independently of GDP. I note that the relative importance of GDP was less, though still substantial, for strictly zoonotic and vector-borne viruses (two large, overlapping subsets of human-infective RNA viruses—73 out of 93 (78.5%) vector-borne viruses are strictly zoonotic). This likely reflects the fact that most such viruses have geographic ranges restricted by the distributions of their vectors and/or reservoir hosts.

Consistent with this interpretation, predictors related to urbanization—a consistently important category—had greatest influence for human-transmissible and non-vector-borne viruses. This, again, can be explained by the fact that more viruses have been discovered in areas (especially in Asia) which have experienced rapid urbanization in recent decades (especially after 1980 (Chini et al., 2014)). Population density and growth, in contrast, were much less prominent predictors, with particularly little influence on strictly zoonotic and vector-borne viruses. This implies that change in habitat—from natural or rural to urban (Hassell et al., 2017)—has a greater influence on virus discovery (by altering the virus geographic distributions in nature) than human population size or density.

I also found associations between the discovery of RNA viruses and climate: five of the most influential predictors in the full model were minimum temperature, precipitation change, latitude, minimum precipitation, and total precipitation. That warmer and wetter climate (higher minimum temperature, more precipitation and lower latitude) is positively associated with the virus discovery is consistent with previous studies (Allen et al., 2017). Climate variables (especially minimum temperature) were relatively more important predictors of vector-borne and strictly zoonotic virus discovery—both these categories were more often discovered in tropical and sub-tropical regions. Forty two percent (93 out of 223) of human-infective RNA species were vector-borne (Woolhouse & Brierley, 2018) and the distribution and abundance of these viruses was strongly influenced by the impact of climate on vector populations (Jones et al., 2008; Li et al., 2014). That climate was also relatively important for the discovery of strictly zoonotic viruses may be at least partly explained by the fact that 78.5% of vector-borne viruses were strictly zoonotic (Table A.1), although there may also be an association between climate and the distribution of reservoir hosts.

For biodiversity, mammal species richness was shown to make an influential contribution to human-infective RNA virus discovery, again particularly for vector-borne viruses and strictly zoonotic viruses. Most but not all previous studies have indicated that risk of spill-over for a virus from mammal hosts to humans was positively correlated with host species richness (Allen et al., 2017; Jones et al., 2008; Wood et al., 2014) which was consistent with mammals being the main source of zoonotic viruses (Woolhouse et al., 2016) and that

as the mammal species richness increases, so did the richness of the pool of viral zoonoses (Keesing et al., 2010). Where zoonotic viruses are first discovered will be influenced, inter alia, by a range of environmental, ecological, and socioeconomic factors that increase the interaction between humans and mammal reservoirs (Mackey et al., 2014).

My predicted discovery map from the full model, along with two stratified models, identified three areas—East and Southeast Asia, India, and Central America—where more viruses were more likely to be detected in 2010–2019 than have been in the past. Inspection of the historical predicted probabilities of virus discovery in Figure A.2 indicates there has always been and is still fewer discoveries than expected in these regions. This suggested that my model was missing predictors (positive or negative) relevant to these regions. However, as mentioned before, for two predicted high-risk areas—East and Southeast Asia, India—accounted for higher fractions in more recent times. The underlying reason may be that the predictors with the greatest influence on virus discovery, such as GDP and land use variables related to urbanization, have changed substantially over time in these areas (especially China).

This study had several limitations: firstly, as indicated above, my model was missing predictors (positive or negative) relevant to the three newly identified high-risk regions. Second, there is often a lag between virus discovery and publication date, though I used the latter for consistency. Third, there are other potential biases concerning spatiotemporal variation in virus detection methodologies used, and diagnostic accuracy (Woolhouse & Brierley, 2018).

Fourth, I used ICTV species classification following other studies (Shu et al., 2018; Walker et al., 2018), though I note that viral species for each family are defined by independent groups using different criteria, which may lead to over- or under-representation of species entries for certain families in my study compared to their phylogenetic diversity. However, I regard ICTV taxonomy as the most authoritative for comparative analysis. Last, I did not attempt to correct for reporting bias by devising a plausible metric, though previous studies have done so (Allen et al., 2017; Jones et al., 2008). However, I explicitly included predictors that I expect to be correlated with discovery effort, e.g. GDP and university count—these are indirect and likely partial measures of effort.

The strengths of the study include use of a comprehensive data set for human-infective RNA virus discovery, the large set of high-resolution global variables postulated to influence RNA virus discovery, and a more robust model (BRT) combining the strengths of both regression trees and boosting that is capable of solving spatial dependence. I also performed further stratified analyses (distinguishing viruses transmissible in humans or strictly zoonotic, and vector-borne or non-vector-borne) and identified differences between predictors of the discovery of these specific categories of viruses. These results furthered understanding of the spatial distribution of virus discovery for different types, and also demonstrated that such a method can be used to identify such differences between strictly zoonotic and human-transmissible viruses or between vector-borne or non-vector-borne viruses.

In conclusion, the discovery of human-infective RNA viruses showed both spatial and temporal variation, and was a process associated with socio-economic variables, land use, climate, and biodiversity, although the relative importance of these variables differs across different category of RNA viruses. My study helps distinguish the relative contributions of predictors reflecting the natural virus distribution and those reflecting the effort invested in virus discovery to the spatial distribution of first reports of human viruses. New human viruses were more likely to be found in areas with more rapid socio-economic growth. But the underlying geographic distribution of viruses with the potential to infect humans may be somewhat different, reflecting climate, biodiversity and changes in land use. This implies that extra investment in virus discovery in settings that are resource-poor but have other risk factors, such as Africa, central America, and south and southeast Asia, may be warranted.

Chapter 3 Predictors of human-infective RNA virus discovery in the United States, China, and Africa

3.1 Abstract

The variation in the pathogen type as well as the spatial heterogeneity of predictors make the generality of any associations with pathogen discovery debatable. In chapter 2, I confirmed that the global association of a group of predictors differed across different types of RNA viruses, yet there have been no previous comparisons of the specific predictors of RNA virus discovery in different regions. The aim of the chapter was to close the gap by investigating whether predictors of discovery rates within three regions—the United States, China and Africa—differ from one another and from those at the global level. Based on a comprehensive list of human-infective RNA viruses, I collated published data on first discovery of each species in each region. I used a Poisson boosted regression tree (BRT) model to examine the relationship between virus discovery and 33 predictors representing climate, socioeconomics, land use, and biodiversity across each region separately. The discovery probability in three regions in 2010–2019 was mapped using the fitted models and the historical predictors. The numbers of human virus species discovered in the United States, China, and Africa in 1901–2019 were 95, 80 and 107 respectively, with China lagging behind the other two regions.

In each region, discoveries were clustered in hotspots. BRT modelling suggested that in all three regions RNA virus discovery was best predicted by land use and socio-economic variables, followed by climatic variables and biodiversity, though the relative importance of these predictors varied by region. Map of virus discovery probability in 2010–2019 indicated several new hotspots outside historical high-risk areas. Most new virus species since 2010 in each region (6/6 in the United States, 19/19 in China, 12/19 in Africa) were discovered in high-risk areas as predicted by my model. I concluded that the drivers of spatiotemporal variation in virus discovery rates vary in different regions of the world. Within regions virus discovery is driven mainly by land-use and socio-economic variables; climate and biodiversity variables are consistently less important predictors than at a global scale. Potential new discovery hotspots in 2010–2019 are identified. Results from the study could guide active surveillance for new human viruses in local high-risk areas.

3.2 Introduction

RNA viruses are the primary cause for emerging infectious diseases with epidemic potential, given that they have high rate of evolution and high capacity to adapt to new hosts (Woolhouse et al., 2016). In recent decades, infectious diseases caused by severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), Bundibugyo Ebola virus and SARS-CoV-2 present major threats to the health and welfare of humans (Albarino et al., 2013; Ksiazek et al., 2003; Mackay & Arden, 2015; World Health Organisation, 2020a). Detection of formerly

unknown human-infective RNA viruses in the earliest stage after the emergence are essential for controlling the infections they cause. Measures to implement early detection include not only advanced diagnostic techniques (Lipkin & Firth, 2013), but more importantly the idea where to look for them (so-called hotspots) (Morse et al., 2012).

Socio-economic, environmental, and ecological factors related to both virus natural history and research effort have been found to affect the discovery of emerging RNA viruses (Jones et al., 2008; Morse et al., 2012; Rosenberg, 2015; Zhang et al., 2020). However, these factors are highly spatially heterogeneous, making the generality of any associations with discovery debatable. For example, the United States, China, and Africa have experienced different rates of socio-economic, environmental, and ecological changes in the last one hundred years. The United States has always had better resources to discover new viruses than most other countries in the world. For example, the Rockefeller Foundation—a U.S. foundation—supported the discovery of 23 arboviruses in Latin America, Africa, and India in 1951–1969 (Rosenberg, 2015). China has seen urban land coverage more than double and GDP per capita increase by seven times since the 1980s (Ritchie, 2018; Roser, 2013). Nine out of 223 human-infective RNA viruses have been originally discovered in China, and all were discovered after 1982 (Zhang et al., 2020). In contrast, effective surveillance is challenging in less developed regions such as large parts of Africa given resource constraints (Petti et al., 2006).

There have been no previous comparisons of the specific predictors of RNA virus discovery in different regions. In this chapter, I applied a similar methodology from my chapter 2 of global patterns of discovery of human-infective RNA viruses (Zhang et al., 2020) to investigate whether predictors of discovery rates within three regions—the United States, China, and Africa—differ from one another and from those at the global level, using three new virus discovery data sets. Africa, though as a continent as opposed to two single countries (United States and China), I included it because 70% of the least developed countries in the world are in Africa (Department of Economic and Social Affairs, 2021), and this increased the range of predictors I could compare. As in chapter 2, the study unit was at 1° grid cell (details in Methods), and this helped reduce the variabilities in data by including regions at both continent level and country level. I also mapped discovery probability in three regions in 2010–2019 using the fitted models and historical predictors. According to findings from chapter 2 (Zhang et al., 2020), the main predictors of virus discovery at the global scale were GDP-related. This suggests that the patterns of virus discovery I have identified may have been largely driven by research effort rather than the underlying biology. In this chapter, by focusing on more restricted and homogenous regions where the research effort is less variable, I expected to identify predictors more associated with virus biology.

3.3 Materials and Methods

3.3.1 Data sets of human-infective RNA viruses in three regions

With reference to the full list of human-infective RNA virus species (Zhang et al., 2020), I geocoded the first report of each in humans in the United States, China, and Africa separately. The latest version of the full list included 223 human-infective RNA virus species (Table B.1), with *Human torovirus* abolished and a new species—*Heartland banyangvirus*—added by ICTV in 2018 (International Committee on Taxonomy of Viruses, 2018). Data used in this chapter were not subsets of the main data in chapter 2, given that the information on discovery locations and discovery dates for each virus species was re-collected in each specific geographical region.

I followed the same search terms, databases searched, and inclusion or exclusion criteria as the global data set for data collection (Woolhouse & Brierley, 2018). In each region, I decided whether or not each species was ever discovered from human(s) according to peer-reviewed literature. Reference databases included PubMed, Web of Science, Google Scholar, and Scopus. Two Chinese database [i.e. China National Knowledge Infrastructure (CNKI) and Wanfang Data] were also searched when collecting data for China. Reference lists of relevant studies and reviews were also checked manually to find potential earlier discovery papers. The following keywords were used for the retrieval: virus full name or abbreviations or virus synonyms; and human* or person* or case* or patient* or worker* or infection* or disease* or outbreak*

or epidemic*; and region name (Chin* or Taiwan or Hong Kong or Macau; United States or US or USA or America*; Africa* or all African country names). Virus synonyms and abbreviations included early names used in the discovery paper and all subtypes provided by the ICTV 10th report (International Committee on Taxonomy of Viruses, 2018). Evidence which met the following criteria from peer-reviewed literatures were included: (a) Diagnostic methods for RNA virus infection in humans were clearly described, through either viral isolation or serological methods; (b) Specific virus species name or subtypes falling under that species were clearly provided; (c) Both natural infection and iatrogenic or occupational infections were accepted. Evidence which met the following criteria were excluded: (a) Uncertain species due to cross-reactivity with related viruses; (b) Diagnostic methods for virus infection were not specified; (c) Description of clinical symptoms or pathogenicity were not considered as human infection of one certain virus species; (d) Report of '[virus name]-like' or 'potential [virus name] infections', but if '[virus name]-like' was a joint name for a virus species which include several serotypes, then it was accepted; (e) Intentional infections including experimental inoculation or vitro infections; (f) Non-peer-reviewed literature, including media reports, thesis, or unpublished data.

I defined discovery location as where the initial human was exposed to/infected with the virus, as suggested in the first report of human infections from peer-reviewed literature. All locations were geolocated as precisely as possible using methods from the chapter 2 (Zhang et al., 2020). For each region, a polygon was created for those locations at administrative level 3 (county for

the United States; city for China; for Africa, it varied between different countries) and above. Table 3.1 summarises details of data types for virus discovery database in three regions. The majority of discovery locations in the United States and Africa involved point data, while in China the majority involved polygon data at province level. A bootstrap resampling procedure was developed for polygon data covering more than one grid cell (details below). Discovery date of human infection was defined as the publication year in the scientific literature.

Table 3.1 Resolution and covered grid cells for human-infective RNA virus discovery data in the three regions

		Polygon data			Point data	Total
		Country level	State/Province level	City/County level		
United States	**Virus species counts**	NA	14 (14.7%)	11 (11.6%)	70 (73.7%)	95
	Gridded cell counts	NA	189	12	72*	273
China	**Virus species counts**	NA	22 (27.5%)	47 (58.7%)	11 (13.8%)	80
	Gridded cell counts	NA	161	70	12*	243
Africa	**Virus species counts**	7 (6.5)	5 (4.7%)	15 (14.0%)	80 (74.8%)	107
	Gridded cell counts	307	22	17	80	426

*Grid cell counts here include viruses first detected in multiple points from the literature, NA, not applicable

3.3.2 Spatial predictors

As for my global analysis in chapter 2 (Zhang et al., 2020), a suite of 33 global gridded climatic, socio-economic, land use, and biodiversity predictors postulated to affect the spatial distribution of RNA virus discovery were compiled, each at a resolution of 0.5°/30" (except university count having a resolution at country level for Africa and state/province for the United States and China). Data for the United States, China, and Africa were extracted by

restricting the coordinates within each region. The definition, original resolution, and source of each variable were the same as my chapter 2 (Zhang et al., 2020). All predictors were aggregated from their original spatial resolution to 1°×1° resolution grid cell (e.g. as in Figure B.1–Figure B.3 showing distribution of predictors in 2015); data for climatic variables, population, GDP, and land use data without full temporal coverage were extrapolated back to 1901; both following methods from my chapter 2 (Zhang et al., 2020).

3.3.3 Boosted regression trees modelling

I used Poisson boosted regression trees (BRT) model to examine the relationship between discovery of RNA virus and 33 predictors of each 1° resolution of grid cell across each region separately, following codes from my chapter 2 (Zhang et al., 2020) and one previous paper (Allen et al., 2017). Again, the analysis is not the subgroup analysis of the chapter 2, given analyses were redone so that predictors of 'discovery', i.e. first discovery of virus species in each specific region, were compared among three regions. As a tree-based learning method, BRT model can automatically capture complex relationships and interactions between variables, and also can well account for spatial autocorrelation within the data (Crase et al., 2012). I compared Moran's I values of the raw virus data and the model residuals to estimate the ability of the BRT model to account for spatial autocorrelation (Cliff & Ord, 1981). In order to minimise the effect of spatial uncertainty of virus discovery data, I performed 1000 times bootstrap resampling for those discovery locations reported as polygons following methods from my chapter 2.

All BRT models were fitted in R v. 3.6.3, using packages dismo and gbm. BRT models require the user to balance three parameters including tree complexity, learning rate, and bag fraction. I set these parameters as recommended from Elith et al (Elith et al., 2008), and make sure each resampling model contained at least 1000 trees. BRT models identified the final optimal number of trees in each model using a 10-fold cross validation stagewise function (Elith et al., 2008). Table 3.2 summarises the three parameter values of the optimal model as well as the mean optimal number of trees across 1000 replicate models for all three regions.

Table 3.2 Model parameters for human-infective RNA virus discovery in the three regions

Model	Tree complexity	Learning rate	Bag fraction	No. of trees
United States	2	0.0020	0.5	1430
China	2	0.0035	0.5	1473
Africa	2	0.0030	0.5	1446

By fitting 1000 replicate BRT models, the relative contribution plots and partial dependence plots with 95% quantiles were plotted. I defined variables with a relative contribution greater than the mean [100/(33 predictors *100)=3.03%)] as influential predictors in all three regions (Shearer et al., 2018). The partial dependence plots depict the influence of each variable on the response while controlling for the average effects of all the other variables in the model. The map of virus discovery probability across each region in 2010–2019 was derived from the means of the predictions of 1000 replicate models, using values of the 33 predictors in 2015. In order to show discovery hotspots, I converted the prediction map of virus count to a map of probability.

Two statistics were calculated to evaluate the model's predictive performance: a) the deviance of the bootstrap model (Elith et al., 2008), b) infraclass correlation coefficient (ICC) calculated from 50 rounds of ten-fold cross-validation, by following methods from the chapter 2 (Zhang et al., 2020).

Exploratory subgroup analyses (distinguishing viruses firstly discovered in regions and those that had been discovered elsewhere in the world) were performed. I used the same BRT modelling approach as I described above, and relative contribution of each predictor was calculated for each subgroup. I was unable to perform subgroup analysis for China because only 9 human-infective RNA viruses have been firstly discovered in it, and BRT model cannot be fitted to a sample as small as 9.

3.4 Results

The numbers of human virus species discovered in the United States, China, and Africa in 1901–2019 were 95, 80 and 107 respectively (Table C.1). Most first discoveries have been in eastern United States (especially in areas around Maryland, Washington, D.C., and New York), eastern China (developed cities including Beijing, Hong Kong, Shanghai, and Guangzhou), and southern and central Africa (Pretoria and Johannesburg, South Africa; Borno State and Ibadan, Nigeria) (Figure 3.1). A total of 60 virus species were previously reported in all three regions, and 27, 12, 37 species were only found in the United States, China, and Africa respectively (Figure 3.2). In all three regions, smaller proportions of viruses were vector-borne [United States: 23.2% (22/95); China: 21.3% (17/80); Africa: 27.1% (29/107)] and strictly zoonotic

[United States: 30.5% (29/95); China: 16.3% (13/80); Africa: 33.6% (36/107)], compared to larger proportions for both virus types at the global scale [vector-borne: 41.7% (93/223) and strictly zoonotic: 58.7% (131/223)] (Figure 3.2). The 60 shared species were also disproportionally vector-borne [11.7% (7/60)] and strictly zoonotic [7% (4/60), Figure 3.2].

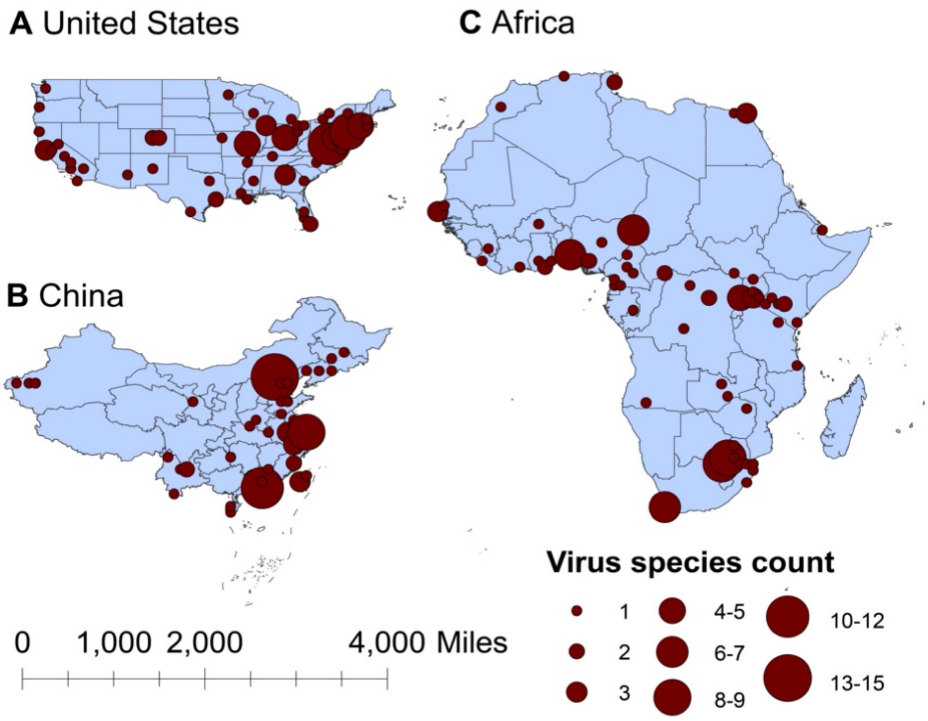


Figure 3.1 Spatial distribution of human-infective RNA virus discovery in the three regions, 1901–2019

A, United States; B, China; C, Africa. Red dots represent discovery points or centroids of polygons, with the size representing the cumulative virus species count.

The discovery curves for the United States and Africa have seen a broadly similar pattern, with China lagging behind these two regions (Figure 3.3). In comparison to the entire world, the median time lag of the virus discovery was 0 [interquartile range (IQR): 2.5], 12 (IQR: 29.5), and 2 (IQR: 10.5) years in the United States, China, and Africa, respectively (Figure 3.4). In China, the time lag was noticeably shorter for viruses discovered after 1975 [before 1975: a median lag of 30.5 (IQR: 30.5) years; after 1975: 2.5 (IQR: 7) years, p value of Wilcoxon rank sum test < 0.001].

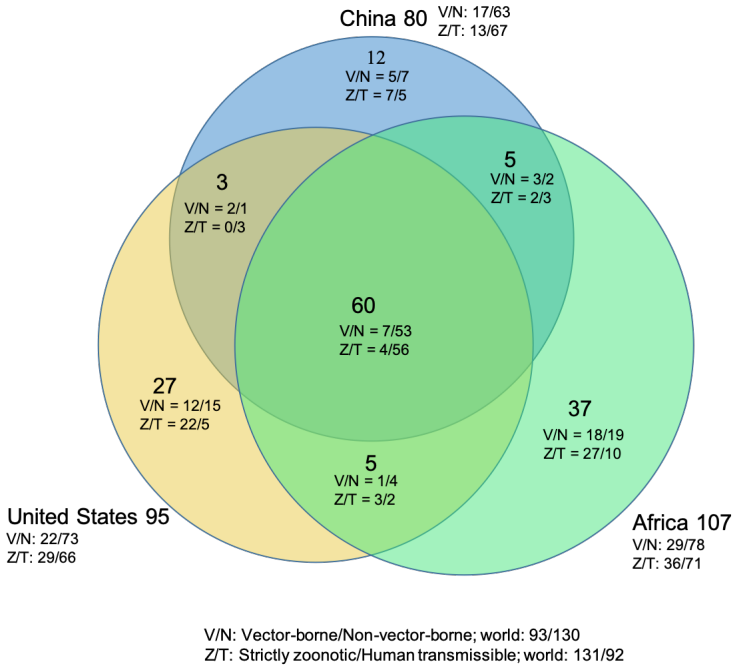


Figure 3.2 Shared human-infective RNA virus species count in the three regions

Under/By the species count the ratios of vector-borne (V) to non-vector-borne (N) viruses and strictly zoonotic (Z) to human transmissible (T) viruses were shown.

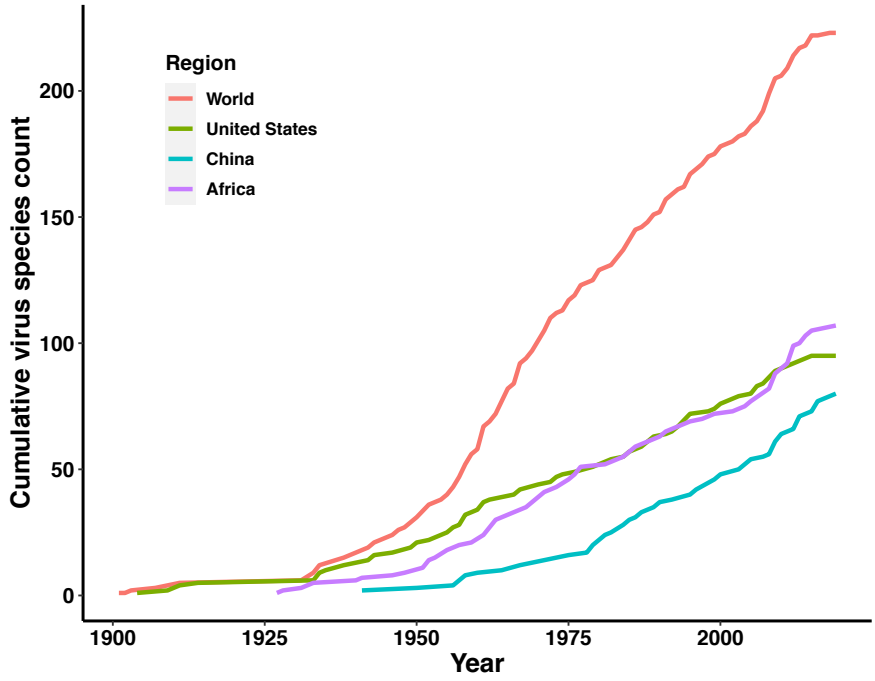


Figure 3.3 Discovery curve of human-infective RNA virus species in the three regions and the world

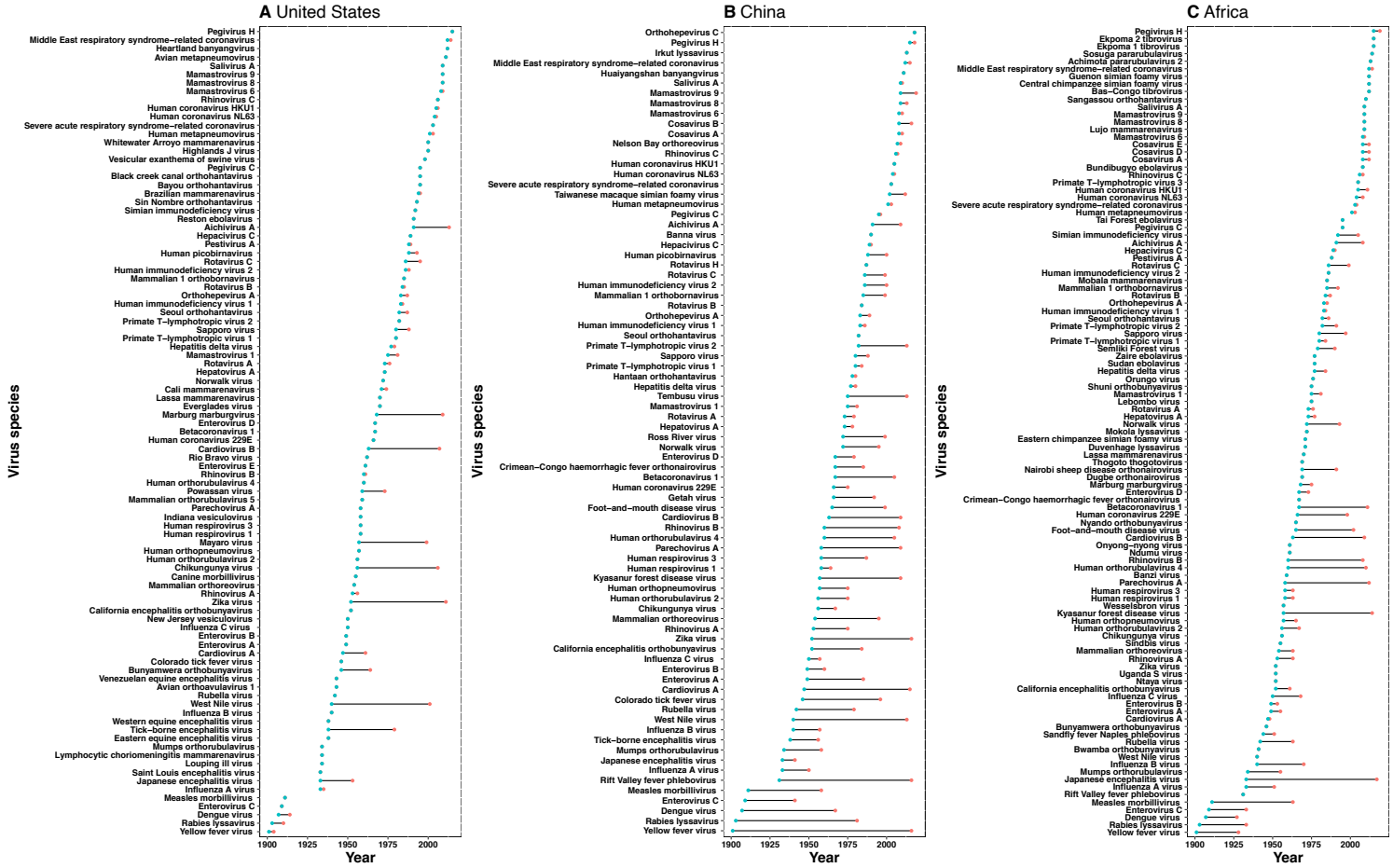


Figure 3.4 Time lag of human-infective RNA virus discovery between the three regions and the world

A, United States; B, China; C, Africa. The blue dots represent the original discovery year of each virus in the world; the red dots represent the discovery year of each virus in three regions; and the segments between them represent the time lag.

In the United States, six variables including three predictors related to land use [urbanized land: relative contribution of 35.8%, urbanization of cropland (i.e. the percentage of land area change from cropland to urban land): 8.0%, growth of urbanized land: 4.1%], two socio-economic variables (GDP growth: 10.0%; GDP: 5.7%), and one climatic variable (diurnal temperature change: 4.9%) were identified as important predictors of discriminating between locations with and without virus discovery (Figure 3.5A). The partial dependence plots shown in Figure 3.6 suggested non-linear relationships between the probability of virus discovery and most predictors. All important predictors presented a positive trend over narrow ranges at lower values.

In China, twelve variables including four socio-economic variables (GDP: 12.7%, university count: 7.5%, GDP growth: 4.6%, population growth: 4.4%), five predictors involving land use [pasture: 8.3%, urbanized land: 8.1%, vegetation: 5.8%, cropland: 5.3%, urbanization of secondary land (the percentage of land area change from secondary land to urban land; secondary land is natural vegetation that is recovering from previous human disturbance): 3.3%], and three climatic variables (maximum precipitation: 4.5%, precipitation change: 3.8%, diurnal temperature range: 3.3%) were identified as important predictors of discriminating between locations with and without virus discovery (Figure 3.5B). GDP, urbanized land, university count, vegetation, GDP growth, maximum precipitation, population growth, and urbanization of secondary land presented a positive trend over narrow ranges at lower levels; pasture, cropland, precipitation change, and diurnal temperature range had non-monotonic/ negative impacts, with highest risks at lower values (Figure 3.7).

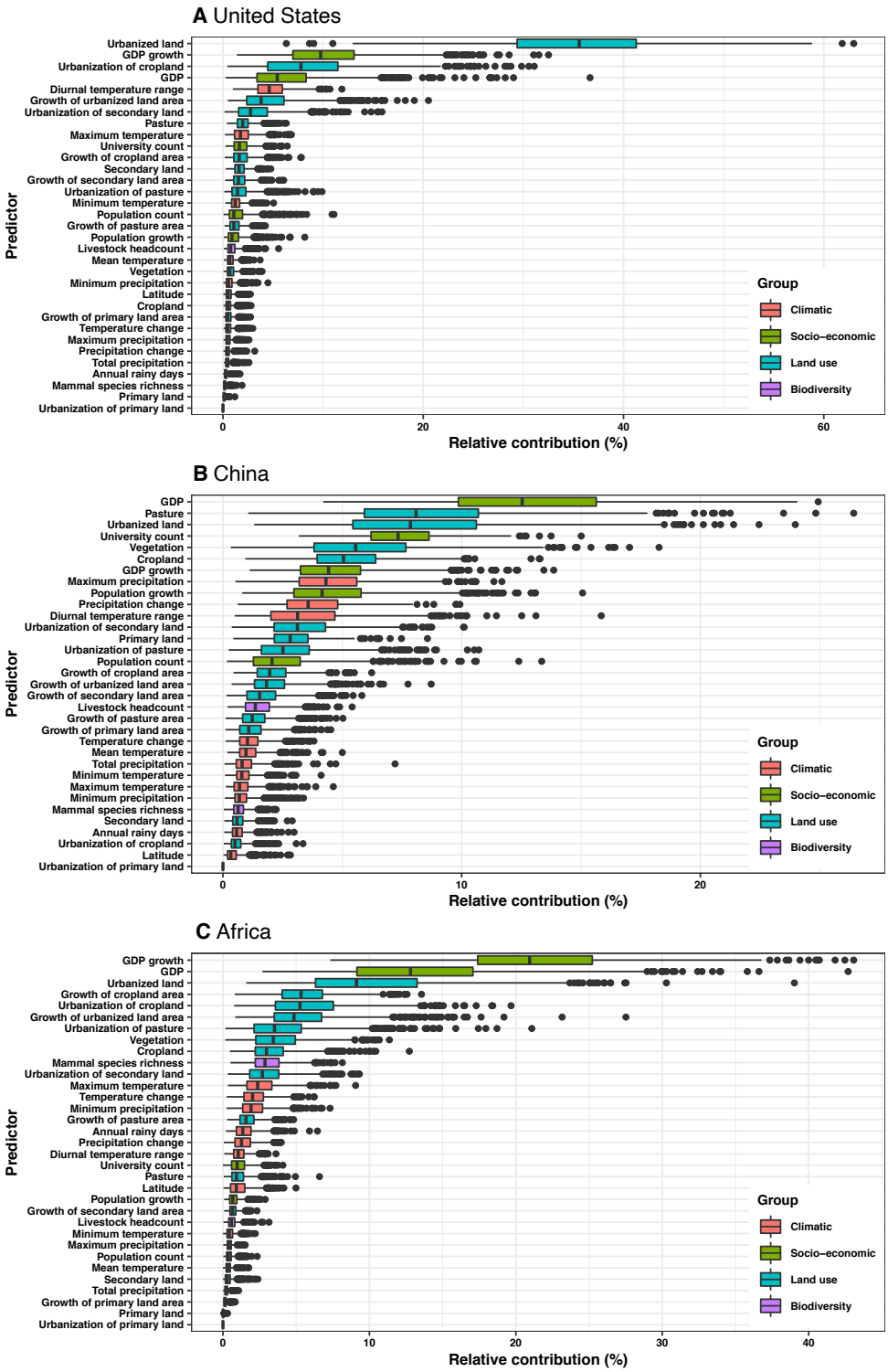


Figure 3.5 Relative contribution of predictors to human-infective RNA virus discovery in the three regions

A, United States; B, China; C, Africa. The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

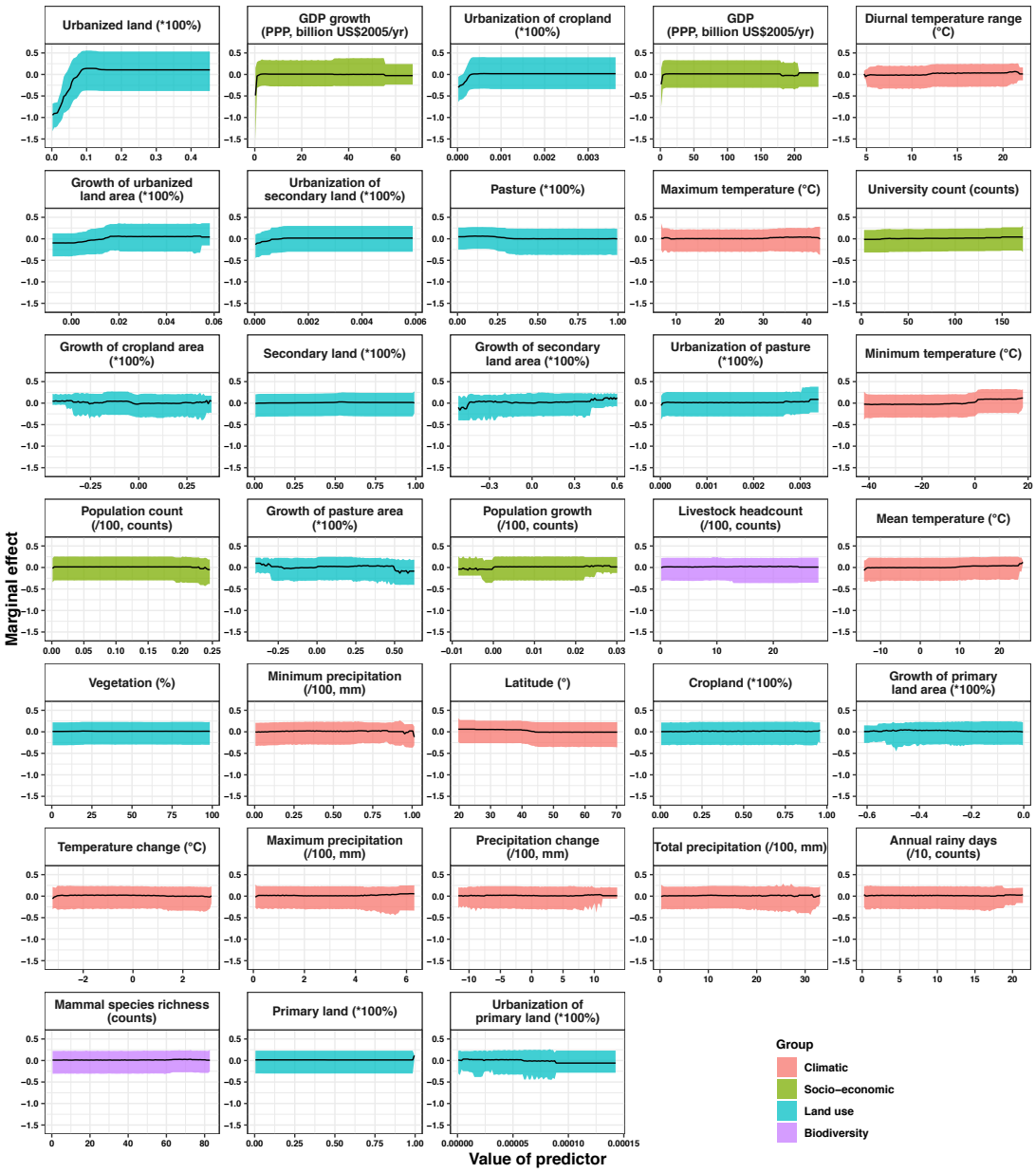


Figure 3.6 Partial dependence plots showing the influence on human-infective RNA virus discovery for all predictors in the Unites States

Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

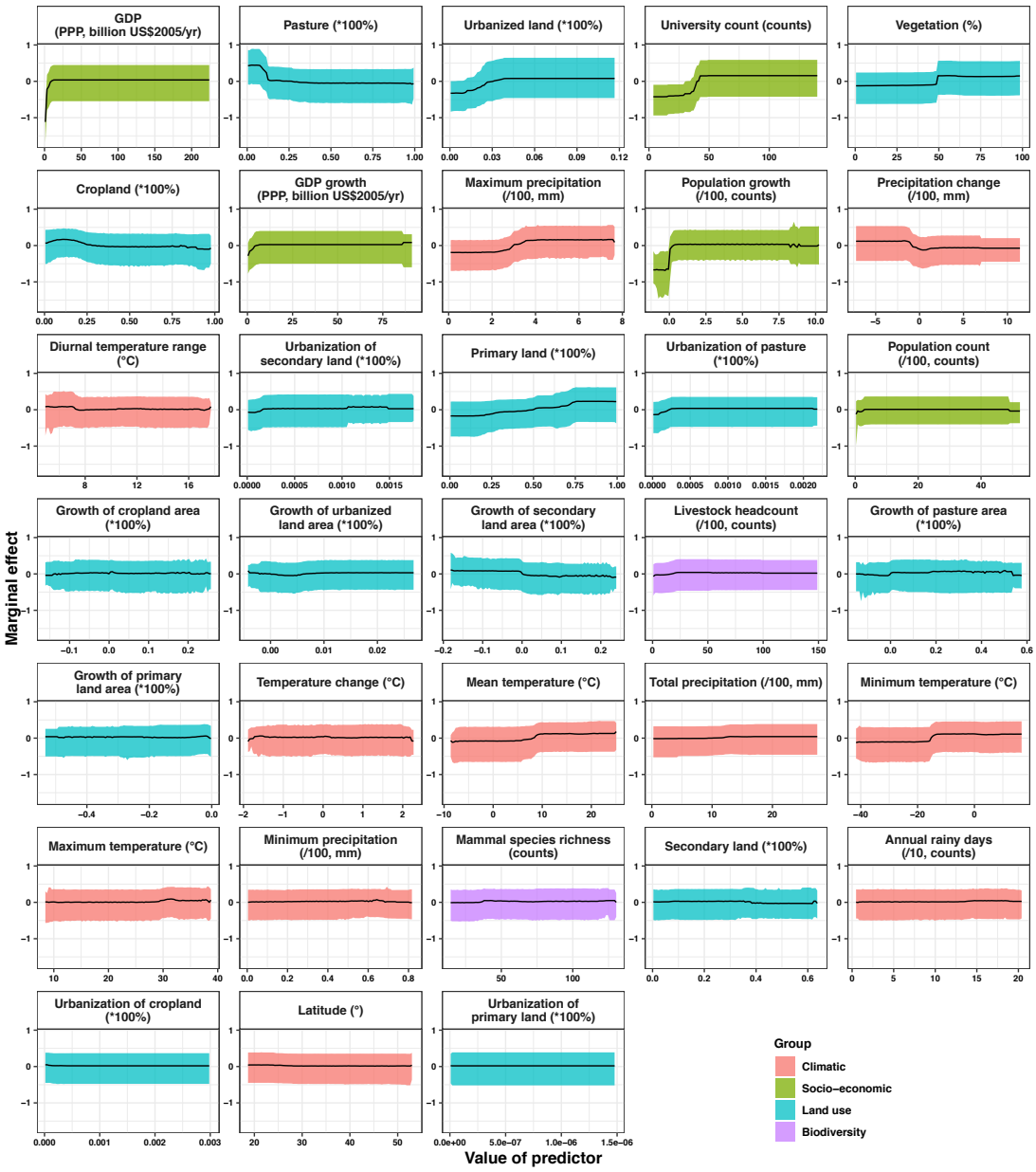


Figure 3.7 Partial dependence plots showing the influence on human-infective RNA virus discovery for all predictors in China

Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

In Africa, ten variables including two socio-economic variables (GDP growth: 21.2%, GDP: 13.0%), seven predictors related to land use (urbanized land: 9.4%, growth of cropland area: 5.6%, urbanization of cropland: 5.5%, growth of urbanized land: 5.1%, urbanization of pasture: 3.8%, vegetation, 3.7%, cropland: 3.2%), and one biodiversity variable (mammal species richness: 3.1%) were identified as important predictors of discriminating between locations with and without virus discovery (Figure 3.5C). All important predictors presented a positive trend over narrow ranges at lower positive values, except mammal species over a large range (Figure 3.8).

My BRT models reduced Moran's I value below 0.15 in all three regions (Figure 3.9), suggesting that BRT models with 33 predictors have adequately accounted for spatial autocorrelations in the raw virus data in all three regions. The model validation statistics for each region are shown in Table 3.3. Combining these measures, my BRT model predictions ranged from fair to good (Cicchetti, 1994).

In all three regions, human-infective RNA virus discovery was best predicted by land use and socio-economic variables, followed by climatic variables and biodiversity (Figure 3.10), whereas virus discovery was more associated with climatic variables and biodiversity at the global level. The comparison between three regions showed that: climatic variables contributed most to the discovery of human-infective RNA viruses in China; land use contributed most to the discovery in the United States; socio-economic variables and biodiversity

contributed most to the discovery in Africa and least to discovery in the United States.

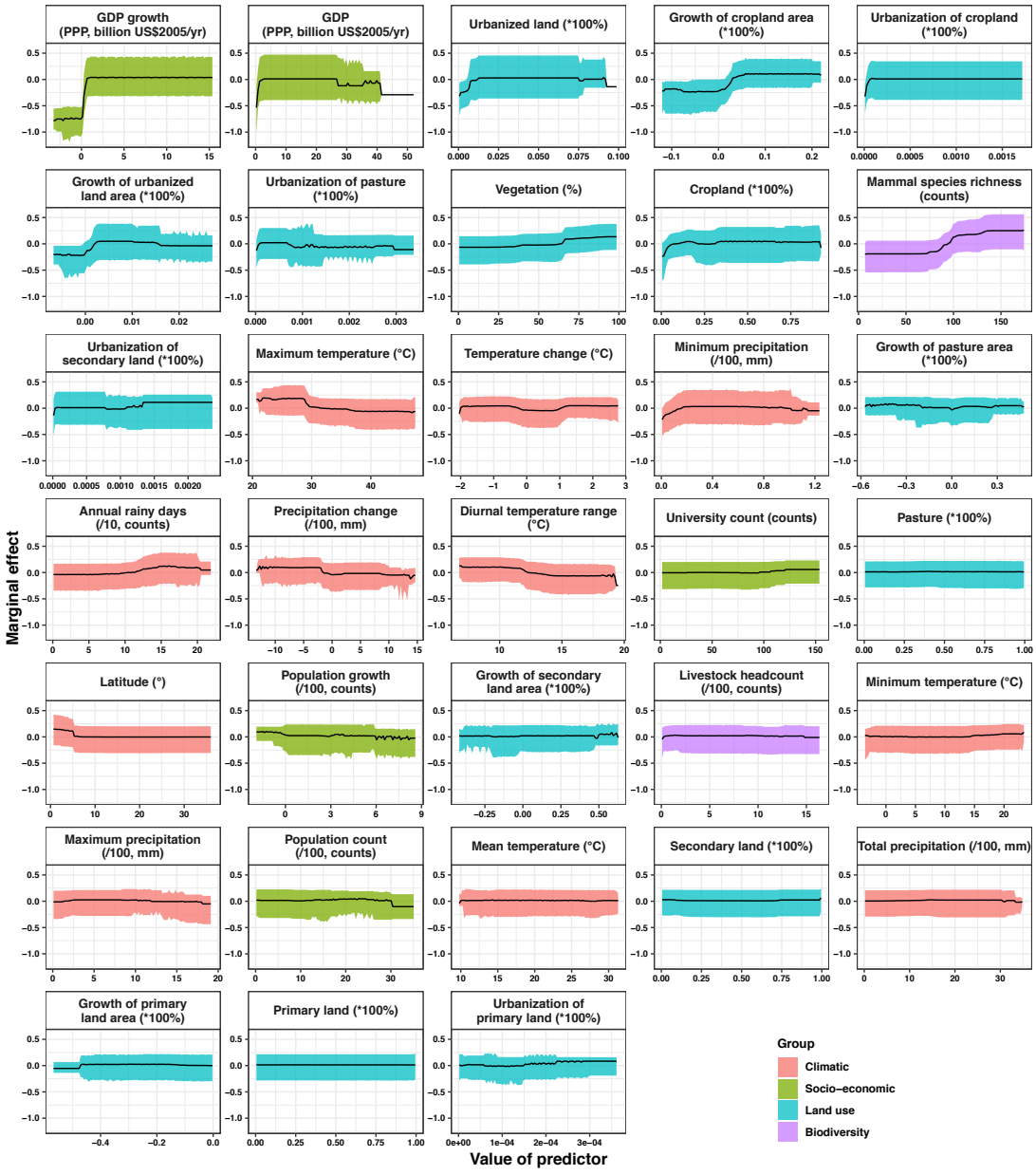


Figure 3.8 Partial dependence plots showing the influence on human-infective RNA virus discovery for all risk factors in Africa

Partial dependence plots show the effect of an individual predictor over its range on the response after factoring out other predictors. Fitted lines represent the median (black) and 95% quantiles (coloured) based on 1000 replicated models. Y axes are centred around the mean without scaling. X axes show the range of sampled values of predictors.

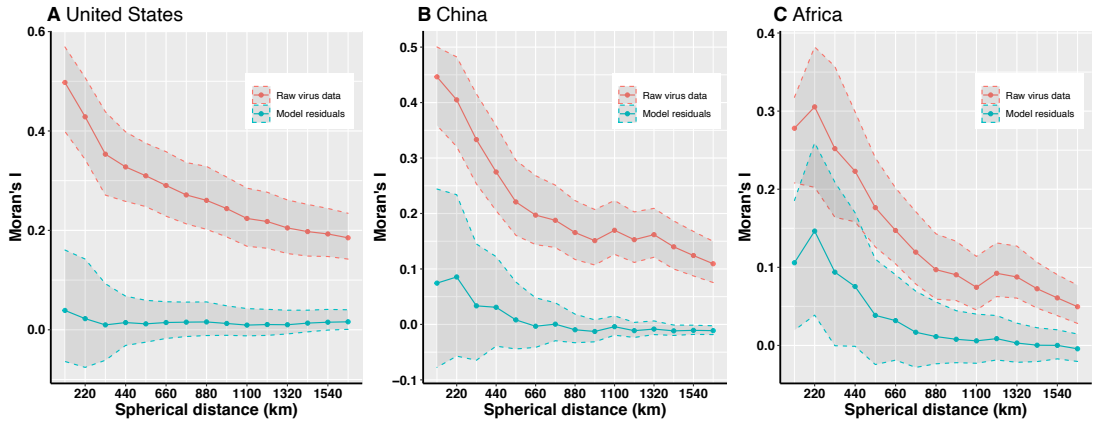


Figure 3.9 Moran's I across different spherical distances for the raw human-infective RNA virus data and the model residuals in the three regions

A, United States; B, China; C, Africa. The solid line and dots represented the median Moran's I value, and the grey area represented its 95% quantiles generated from 1000 samples (Blue: Raw virus data) or replicate BRT models (Red: Model residuals). I used the fixed spherical distance as the neighbourhood weights—as there is no general consensus for selecting cut-off values, I chose spherical distances ranging from one time to fifteen times of distance of 1° grid cell at the equator, i.e. 110km to 1650km, considering the area of three regions. My BRT models reduced Moran's I value from a range of 0.19–0.50 for the raw virus data to 0.009–0.04 for the model residuals in the United States (A), 0.11–0.45 to -0.01–0.09 in China (B), 0.05–0.31 to -0.004–0.15 in Africa (C), suggesting that BRT models with 33 predictors have adequately accounted for spatial autocorrelations in the raw virus data in all three regions.

Table 3.3 Model validation statistics for analyses for human-infective RNA virus discovery in the three regions

Model	% of deviance explained (95% quantiles)	ICC (95% quantiles)
United States	50.5% (44.3%–56.8%)	0.66 (0.60–0.70)
China	42.0% (32.4%–50.8%)	0.52 (0.41–0.60)
Africa	42.4% (34.2%–50.0%)	0.51 (0.44–0.62)

ICC, infraclass correlation coefficient

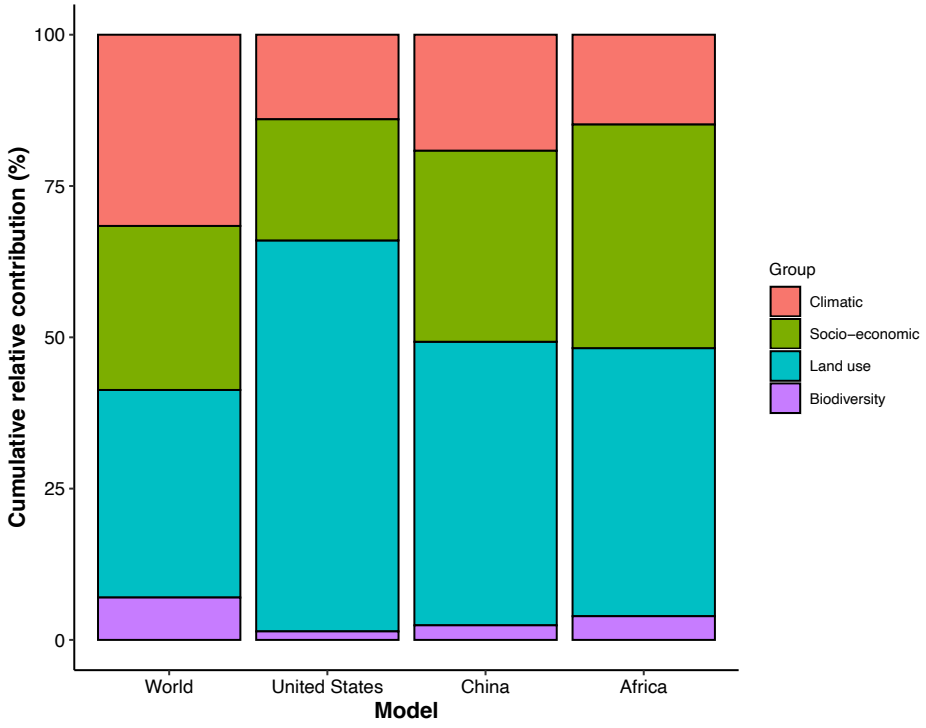


Figure 3.10 Cumulative relative contribution of predictors to human-infective RNA virus discovery by group in each model for different region

The relative contributions of all predictors sum to 100% in each model, and each colour represents the cumulative relative contribution of all predictors within each group.

I mapped human-infective RNA virus discovery probability in 2010–2019 for the three regions, based on the fitted BRT models and values of all 33 predictors in 2015 (Figure B.1–Figure B.3). Outside contemporary risk areas where human-infective RNA viruses were previously discovered in the United States (Figure 3.1A), I predicted high probabilities of virus discovery across southern Michigan, central-Northern Carolina, central Oklahoma, southern Nevada, and north-eastern Utah (Figure 3.11A). Outside contemporary risk areas where human-infective RNA viruses were previously discovered in China (Figure 3.1B), I predicted high probabilities of virus discovery across other eastern China area as well as two western areas including south-central Shaanxi and north-eastern Sichuan (Figure 3.11B). Outside contemporary risk

areas where human-infective RNA viruses were previously discovered in Africa (Figure 3.1C), I predicted high probabilities of virus discovery across northern Morocco, northern Algeria, northern Libya, south-eastern Sudan, central Ethiopia and western Democratic Republic of the Congo (Figure 3.11C). Eighty-four percent (37/44) new virus species since 2010 in each region (6/6 in the United States, 19/19 in China, 12/19 in Africa) were discovered in high-risk areas (85% percentiles of predicted probability across each region) as predicted by my model. Of all the 37 (United States: 6; China: 19; Africa: 12) viruses discovered in high-risk areas in 2010-2019, 13 (United States: 2; China: 7; Africa: 4) viruses were discovered at the potential new hotspots where there have not been any virus discoveries before 2010.

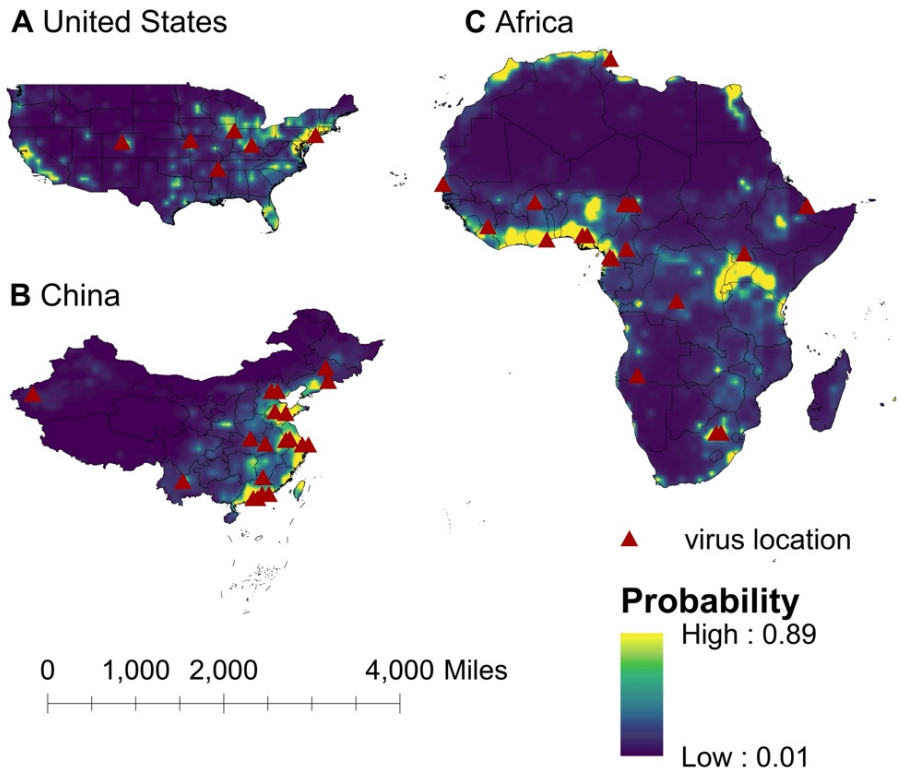


Figure 3.11 Predicted probability of human-infective RNA virus discovery in the three regions in 2010–2019

A, United States; B, China; C, Africa. The triangles represented the actual discovery sites from 2010 to 2019, and the background colour represented the predicted discovery probability.

Based on my subgroup analysis (distinguishing viruses firstly discovered in regions and those that had been discovered elsewhere in the world), discoveries of human-infective RNA viruses firstly discovered from either United States or Africa were better predicted by climatic and biodiversity variables, while discoveries of viruses that had been discovered from elsewhere in the world were better predicted by socio-economic variables (Figure 3.12).

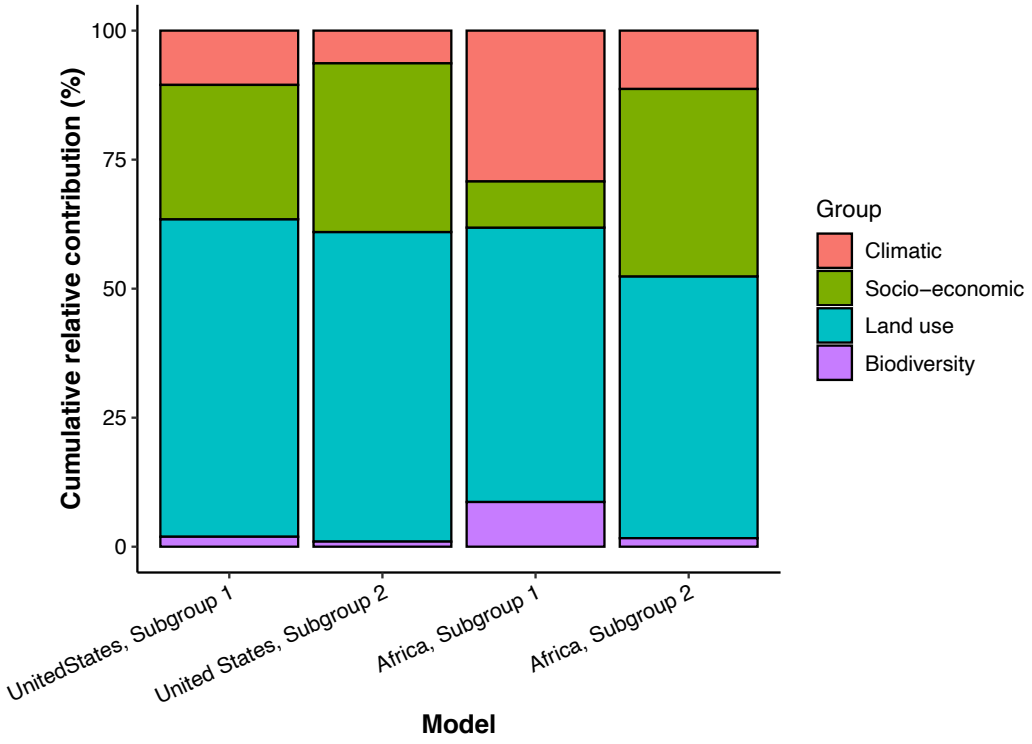


Figure 3.12 Cumulative relative contribution of predictors to human-infective RNA virus discovery by group in each subgroup model

Subgroup 1 represents viruses firstly discovered from the region (United States or Africa); Subgroup 2 represents viruses firstly discovered elsewhere in the world. The relative contributions of all predictors sum to 100% in each model, and each colour represents the cumulative relative contribution of all predictors within each group.

3.5 Discussion

To my knowledge, this analysis represented the first investigation of human-infective RNA virus discovery in three large regions of the world which have experienced distinct socio-economic, ecological, and environmental changes over the last 100 years. In total, 95 human-infective RNA virus species had been found in the United States in 2010–2019; 80 in China; 107 in Africa. The discovery maps of human-infective RNA virus in the three regions indicated areas with historically high discovery counts: eastern and western United States, eastern China, and central and southern Africa. BRT modelling suggested that the relative contribution of 33 predictors to human-infective RNA virus discovery varied across three regions, though climatic and biodiversity variables were consistently less important in all three regions than at a global scale. I mapped the probability of human-infective RNA virus discovery in 2010–2019 which would continue to be high in historical hotspots but, in addition, I identified several new hotspots in central-eastern and southwestern United States, eastern and western China, and northern Africa. These results offered a tool for public health practitioners and policymakers to better understand local patterns of virus discovery and to invest efficiently in surveillance systems at the local level.

In all three regions, GDP and/or GDP growth were identified as important predictors of virus discovery, especially in Africa where GDP and GDP growth were identified as the leading predictors. This is consistent with my analysis in chapter 2 that GDP and GDP growth played a major role in discovering viruses

(Zhang et al., 2020). In general, sufficient economic, human and material resources, the availability of advanced infrastructure and technology, and greater research capabilities in the relative higher-income areas enable the virus discovery (Rosenberg et al., 2013). That this effect applied both within one continent and within single countries such as the United States and China suggested that most virus discoveries were likely passive, i.e. the viruses were detected when they arrived in a location with the resources to detect them. This is plausible because in all regions in my study, human-transmissible viruses accounted for the larger proportion, and my analysis in chapter 2 suggested richer areas were more likely to first capture transmissible viruses (e.g. Influenza virus, Rhinovirus, Rabies lyssavirus, Measles morbillivirus, Mumps orthorubulavirus, Rubella virus, and Norwalk virus) capable of spreading to multiple areas (Zhang et al., 2020). Temporally, in China the rate of discovery increased after economic growth accelerated in the 1980s (Figure 3.3). I note in publications describing first virus discoveries that most historical virus discoveries in Africa received support from the United States and Europe, and this may explain why Africa saw an increased number of virus discoveries after 1950—30 years earlier than China (Figure 3.3). Notably, in China the relative contribution of GDP growth to virus discovery was not as substantial as that in Africa. In contrast, university count was found to be associated with virus discovery, suggesting virus discovery likely being a significant area of research in Chinese universities. My model also suggested socio-economic factors contributed less in the United States than other two regions. The

possible explanation is that the socio-economic level across the entire United States is relatively high and homogenous.

Predictors other than GDP and university count are likely to be linked to virus natural history. In all three regions, the area of urban land and further urbanization made great contribution to virus discovery. This reinforced previous studies that urbanization was linked to the detection of new human pathogens through the denser urban population, increased human-wildlife contact rate, spill-over of human infection from enzootic cycle, and the contamination of the urban environment with microbial agents (Hassell et al., 2017; Olival et al., 2017; Weaver, 2013). In the United States, land use contributed most to virus discovery in comparison to other regions—urbanized land, urbanization of cropland, and growth of urbanized land alone had a relative contribution of 47.9%. However, the leading role of urbanization to virus discovery in the United States remains unclear. It is likely that land use change is driving the emergence of novel viruses, and then the discovery. For example, both index cases of Heartland virus in the United States had a history of exposure to ticks, and is likely to have spilled over to humans as a consequence of human encroachment in the context of increasing urbanization (Mansfield et al., 2017; Savage et al., 2013). The emergence of hantavirus (Sin Nombre orthohantavirus, Bayou orthohantavirus, and Black creek canal orthohantavirus) in the United States is also likely to be associated with land use changes, given that the urban landscapes provide suitable breeding habitat and resources for their natural reservoir rodents (Hassell et al., 2017).

Consistent with the findings in my chapter 2, population growth was identified as a less prominent predictor (Zhang et al., 2020). In China, population growth—though with greater influence than other regions—contributed less than urbanized land and three other land types on virus discoveries. This reinforced my interpretation in chapter 2 that urbanization brings larger changes on human living environment than human population size/growth, and therefore may have influenced virus discovery more greatly.

Climate had less influence on human-infective RNA virus discovery in all three regions in comparison to other predictors, in contrast to virus discovery at a global scale (Zhang et al., 2020). The underlying reason may be that the proportion of vector-borne viruses—whose distribution and abundance is strongly associated with the impact of climate on vector populations (Li et al., 2014)—in all three regions (United States: 23.2%; China: 21.3%; Africa: 27.1%) were less than that in the world (41.7%) (Figure 3.2). Vector-borne viruses tend to have more restricted global ranges, so are less likely to appear in a study of any one region (Zhang et al., 2020).

In addition, a relative smaller proportion of strictly zoonotic viruses in three regions (United States: 30.5%; China: 16.3%; Africa: 33.6%) than that in the world (58.7%) (Figure 3.2) made biodiversity contribute less to virus discovery in the three regions than in the world (Zhang et al., 2020). With exposure to denser mammals played a slightly larger role in virus discovery in Africa than in China and the United States (Figure B.1–Figure B.3).

My discovery probability maps for 2010–2019 in three regions captured most historical hotspots, though several small new areas in central-eastern and southwestern United States, eastern and western China, as well as northern Africa would also make greater contribution to virus discovery (Figure 3.11). My model has a good predictive ability, given 84% (37/44) new virus species in 2010-2019 were discovered in high-risk areas as defined—85% percentiles of discovery probability within each region. Further, 35% (13/37) of those viruses discovered in high-risk areas since 2010 were discovered at the potential new hotspots where there had not been any virus discoveries in the past.

My subgroup analyses suggested in both the United States and Africa, discoveries of viruses firstly discovered in regions were more likely to be associated with climatic and biodiversity variables while discoveries of viruses had been discovered elsewhere in the world were more likely to be associated with socio-economic variables. This is plausible, again because after a novel virus was discovered elsewhere in the world, it is usually areas with a higher socio-economic level firstly capture the virus in the local region.

This study had limitations. First, one common problem for data collected from literature review is the time lag between virus discovery and publication, in which case the virus data are likely to be matched to covariates in later decades. Second, I acknowledge that it is possible I have not identified the earliest report for some well-known viruses such as yellow fever virus, measles virus, especially in the post-vaccination era. Third, I was unable to identify

robust and comprehensive data for all three regions on virus discovery effort, although I interpreted GDP and university count as being an indirect measure of resources available for this activity.

The study added to my chapter 2 (Zhang et al., 2020) in several ways. First, I firstly constructed data sets of human-infective RNA virus discovery reflecting the viral richness in three broad regions of the world. Second, I reduced the heterogeneity of the predictors by focusing on regions, including those predictors reflecting the research effort. Research effort is less variable within restricted regions and therefore has less effect on virus detection. This implies my predicted hotspots stand closer to the virus geographic distribution in nature. Third, the predicted hotspots derived from regional analysis have a higher precision than at a global scale, e.g. specific areas in the United States and China were identified as hotspots from regional analysis, rather than the whole eastern area from the global analysis. This helps target areas for future surveillance.

In conclusion, a heterogeneous pattern of virus discovery-driver relationships was identified across three regions and the globe. Within regions virus discovery was driven more by land-use and socio-economic variables; climate and biodiversity variables were consistently less important predictors than at a global scale. I mapped with good accuracy that in 2010–2019 three regions where human-infective RNA viruses had previously been discovered would continue to be the discovery hotspots, but in addition, several new areas in

each region would make great contribution to virus discovery. Results from the study could guide active surveillance for new human viruses in high-risk areas.

Chapter 4 Predictors of geographical extent and disappearance of human-infective RNA viruses

4.1 Abstract

Understanding the attributes of viruses after their emergence in humans is important for both preventing and controlling the associated outbreaks. However, the factors that determine why some viruses are able to establish sustained infections in humans after their emergence and even spread across the globe, while others are not successful in transmitting and disappear from humans, are not fully understood. In this chapter, I collated information for all 223 human-infective RNA viruses, including their geographical extent and persistence in causing human infections from peer-reviewed literature. Next, I explored what factors predicted sustained human infections using Bernoulli Boosted Regression Tree models. Thirty-seven percent (83) human-infective RNA viral species have wide geographic extents (found in three or more continents). Viral features that predicted these wide geographic extents included transmissibility between humans, a +ssRNA genome, narrow host range [i.e. infecting humans only or humans and other non-human primates (NHP) only], and having a reservoir host in a NHP. Thirty-six percent (81) viruses which at one time were found to infect humans, were then found to have disappeared from humans (no record of infection in the literature for the

past ten years or more). Viruses were more likely to disappear if they were incapable of transmission between humans, have had a localised geographic extent, a dsRNA genome, were non-pathogenic and non-fatal, were discovered through active discovery programmes rather than passive investigation of the aetiology, and were transmitted by vectors and direct contact. Results for both geographical extent and virus disappearance did not change after factoring out reporting effort. I concluded that multiple characteristics determined the geographical extent and disappearance of human-infective RNA viruses; however, transmission ecology and viral genome were consistently the most important predictors. Host range was found to be an important predictor of geographical extent, whereas geographical extent, clinical presentation and the discovery process all contributed to the probability of a virus disappearing, or no longer caused recorded human infections. Understanding the drivers of sustained human transmission in viruses can help public health officials make risk assessments for new viruses both before and after their emergence.

4.2 Introduction

Infections caused by human-infective RNA viruses such as Influenza A virus, HIV-1, and Ebola virus, represent a significant burden on both human health and the economy (Cox & Subbarao, 2000; Hemelaar et al., 2006; Muyembe-Tamfum et al., 2012). Understanding the long-term consequences of viruses after their emergence in humans is of great importance for preventing and controlling associated outbreaks in the future. However, the long-term

consequences after a virus emerges in humans can vary substantially across different virus species. Within the coronavirus family for example, severe acute respiratory syndrome-related coronavirus (SARS-CoV, including the SARS-CoV-2 that has caused the ongoing pandemic), Middle East respiratory syndrome-related coronavirus (MERS-CoV), and the other four human coronavirus—229E, OC43, NL63, and HKU1 have widely circulated in countries across the world after their emergence (de Wit et al., 2016; Gorbalenya et al., 2020; Su et al., 2016). For one less-known coronavirus—alphacoronavirus 1, however, there was only one report of human infections in 2007 from Japan (Terao et al., 2007).

Early detection of viruses with the potential for a worldwide geographic extent after emergence is clearly important for public health. To my knowledge, only one quantitative study has found that viruses with a wide host range and viruses transmitted to humans by direct contact with wild animals predict broader geographic range (C. K. Johnson et al., 2015). Other previous studies, mostly theoretical, have suggested some traits that could estimate whether a virus can cause large-scale spread or not (Morse et al., 2012; Richard et al., 2017; Woolhouse et al., 2016). For example, the level of transmissibility between humans, determined by the basic reproduction number (R_0), is commonly used to evaluate the epidemic potential of a novel pathogen (Woolhouse et al., 2016). Viruses with $R_0 > 1$ such as measles virus (12.5–18), human influenza A virus (1.2–5.2), and Zaire Ebola virus (1.3–3.07) are likely to cause infections with a wide geographic extent (Althaus, 2014; Anderson & May, 1982; Woolhouse et al., 2016). Recent comparative studies linked the

high viral transmissibility in humans to viral traits including viral structure (non-enveloped viruses), transmission routes (non-vector-borne viruses), narrow host range, clinical presentations (low human mortality rate), and viral tissue presence (virus isolated from respiratory tract and central nervous system) (Geoghegan et al., 2016; C. K. Johnson et al., 2015; Walker et al., 2018; Woolhouse et al., 2016). Specifically, these studies supported the evolutionary trade-off between virulence (human mortality rate) and transmissibility which states/assumes virulence constrains transmission between human populations. There were arguments against this, however, e.g. one study suggested viruses incapable of transmission did not show higher virulence, suggesting virulence is likely unselected when humans are dead-end hosts and not contributing to virus spread (Brierley et al., 2019).

Apart from the geographical extent of human-infective RNA viruses, understanding whether they could establish sustained infections in humans or not after emergence is also essential. For example, despite having worldwide extents, human-infective RNA viruses such as the pestivirus A, foot-and-mouth disease virus, and Gareth virus have not been reported to infect humans for more than ten years (Berrios, 2007; Giangaspero & Cominardi, 2006; Li et al., 1992). I have therefore assumed these viruses have disappeared from humans and are less likely to bring significant disease burdens. Although some viruses presumed to have disappeared (e.g. SARS-CoV), do re-emerge in humans (e.g. SARS-CoV-2, the subspecies of SARS-CoV) (Morens & Fauci, 2020). More than 80% of human-infective RNA viruses have zoonotic origins (Woolhouse & Brierley, 2018), and most of them cause single human

infections and humans are dead-end host (Parrish et al., 2008). The possible reasons for disappearance include the limited contact between humans and the zoonotic viruses (e.g. Chapare mammarenavirus, Tai Forest ebolavirus, and Menangle pararubulavirus), host barriers (host tissue specificity, viral entry, receptor binding, innate immune resistance, and anti-viral drugs), and viral evolutional changes (which likely aid unsuccessful adaptation, replication, and transmission in humans). One more reason for viral disappearance is that less research effort has been put into viruses causing no distinct illness in humans (e.g. Cali mammarenavirus, Alphacoronavirus 1) (Buchmeier et al., 1974; Delgado et al., 2008; Le Guenno et al., 1995; Parrish et al., 2008; Philbey et al., 1998; Terao et al., 2007; Trapido & Sanmartin, 1971).

The different patterns of RNA virus infection in humans highlight the importance of revealing factors that determine why some viruses are able to establish worldwide infections in humans and also why some viruses have now disappeared from humans. So in this chapter I aimed to evaluate the geographical extent and disappearance of all known human-infective RNA viruses and determine the corresponding predictors. I hypothesized that 1) transmissible viruses and viruses with lower virulence are more likely to establish wide geographic extents and less likely to disappear; 2) Non-enveloped viruses, Non-vector-borne viruses and viruses with narrow host range are more likely to establish wide geographic extents and less likely to disappear in humans, given they are highly likely to be transmissible and have lower virulence.

4.3 Materials and Methods

4.3.1 Data collection

A full list of human-infective RNA viruses (223 species) was based on an updated version of the database that has been published previously (Woolhouse & Brierley, 2018; Zhang et al., 2020). All viruses were approved by the International Committee on Taxonomy of Viruses (ICTV). Information on the geographical extent (determined by the number of continents that human infections are found in) and disappearance (determined by the year of the last record of human infections, detailed definitions for both outcomes are described below) for each virus were collated from peer-reviewed literature from scientific databases as of 31 December 2019. The scientific databases included PubMed, Web of Science, and Google Scholar. For well-known viruses that have been fully studied, their geographical extents were obtained from published review articles. For those viruses with less research attention, I searched all papers describing infections in humans and recorded the geographical extent of all known cases. For viral disappearance I also used media reports from ProMED-mail (http://www.promedmail.org/). The detailed searching strategy for viral disappearance is shown in Figure 4.1.

Keywords included: [virus name or virus synonyms or abbreviations] AND [human* or person* or case* or patient* or infection* or disease* or outbreak* or epidemic*]. The virus synonyms and abbreviations included early names from the discovery paper and all subtypes provided by the ICTV 10th report

(International Committee on Taxonomy of Viruses, 2018). References of each paper were scanned to detect more potential human infections.

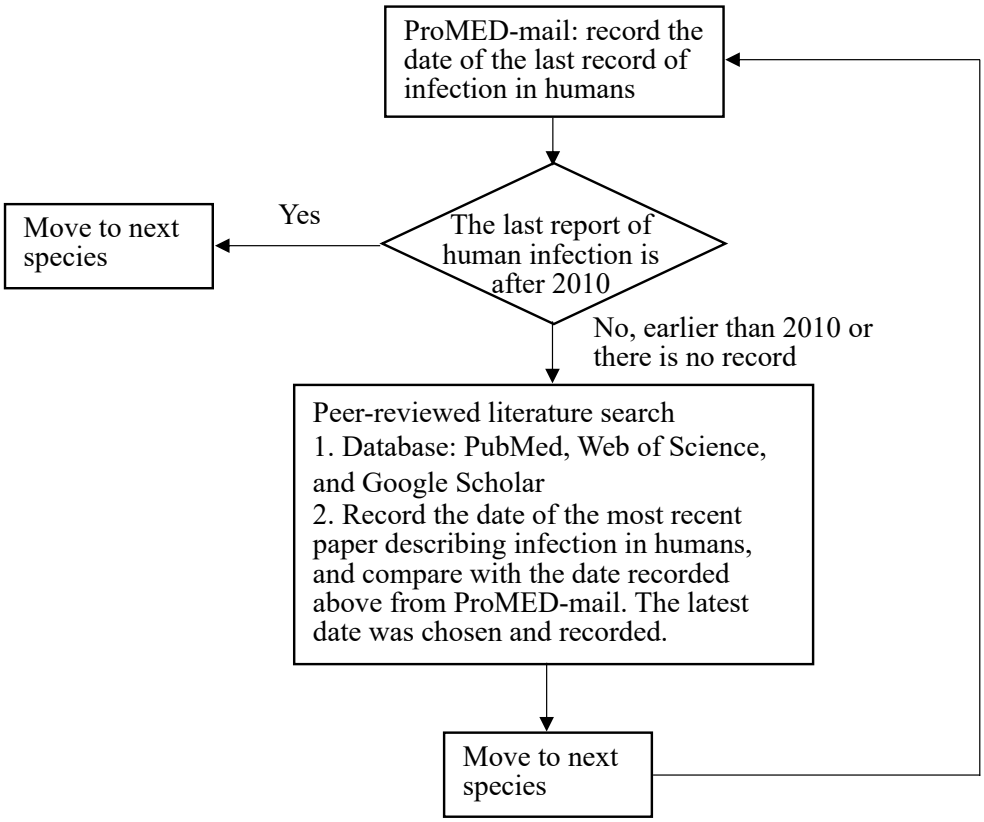


Figure 4.1 Searching strategy for viral disappearance

Information on the last record of human infection was updated as of 31 December 2019.

Table C.1 summarises the geographical extent as well as the state of disappearance for 223 species of human-infective RNA viruses. The median number of continents each virus was found in was one or two, depending on whether the imported cases and laboratory infections are included (Figure 4.2). Therefore, the geographical extent of each virus was classified as either: i) restricted extent: the virus was found to infect humans in two or fewer

continents; or ii) wide extent: the virus was found to infect humans in three or more continents. I also performed a sensitivity analysis by changing the cut-off to one continent (details below). The definition of viral disappearance was that the virus was once found to infect at least one human, but then it disappeared, defined as no known cases in humans found in the literature after 2010 (using both sources of evidence mentioned above).

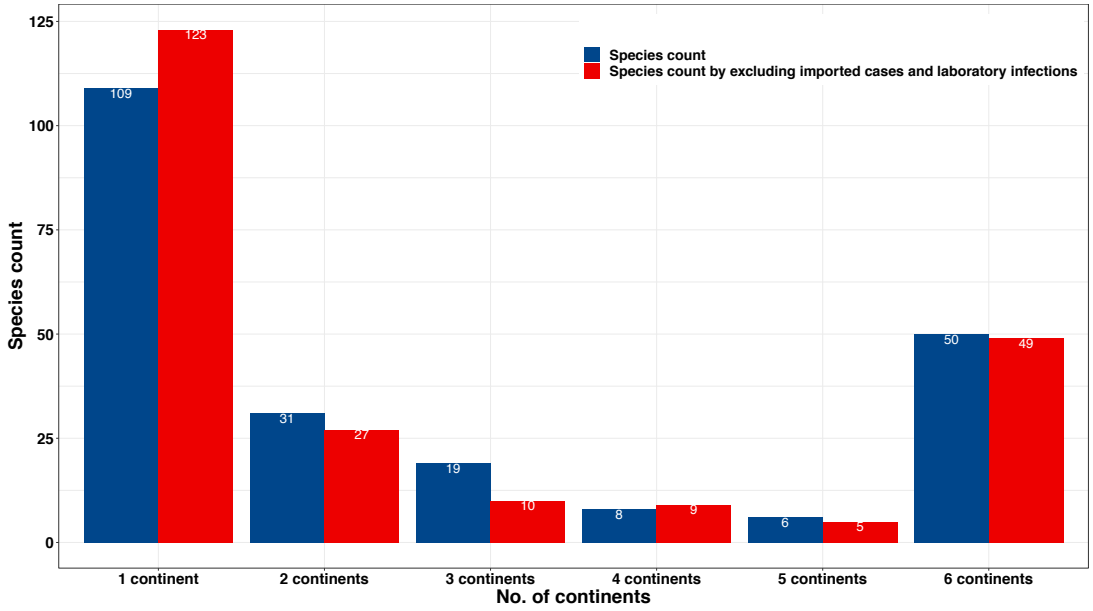


Figure 4.2 Distribution of the number of continents each virus species was found in

The blue bars represent the cumulative number of species found in each continent count. The numbers of continents each virus species was found in were re-calculated after removing imported cases and laboratory infections, and their distributions are shown in red bars. The median number of continents is one (if excluding imported cases and laboratory infections, in red) or two (if including imported cases and laboratory infections, in blue).

To determine what factors may drive both geographic extent and disappearance of human-infective RNA viruses, I considered 22 biological features that reflect key aspects of virus life history and ecology (Brierley et al., 2019; Geoghegan et al., 2016; Guth et al., 2019; C. K. Johnson et al., 2015; Walker et al., 2018; Woolhouse et al., 2016). These factors can be categorised

to six groups—taxonomy, structure, transmission, host range, origin, and clinical presentation (Table 4.1). Information on index case identification and clinical presentation were collected from peer-reviewed literature while all other information were extracted from the published data (Woolhouse & Brierley, 2018). To evaluate the research effort for each virus species, I also searched the Web of Science and recorded the number of publications for each species. Date range is 1 January 1901 to 31 December 2019. Keywords for virus species were the same as I described above. Sensitivity analyses by including the research effort in models for both geographical extent and virus disappearance were performed (details below).

4.3.2 Statistical methods

In the univariable analysis, chi-square test or Fisher's exact probability test was used to compare variables between groups. In the multivariable analysis, I applied Bernoulli Boosted Regression Tree (BRT, see chapter 2.3.4 for detailed descriptions of the model) models to each of the two outcomes—(i) geographical extent and (ii) viral disappearance, a relative new machine learning technique combining boosting and regression trees in a single algorithm (Elith et al., 2008). Packages dismo and gbm in R v. 3.6.3 were used to fit BRT models. The final parameters including tree complexity, learning rate, and bag fraction of the optimal model and the mean number of trees are shown in Table 4.2.

Table 4.1 List of predictors of human-infective RNA virus geographic extent and disappearance

Group	Predictors of interest	Definition and values
Taxonomy	Viral genus (2 predictors)	(i) Genus for each of the 223 viruses (57 genera). (ii) For the two genera with the most human-infective RNA virus species—'Genus *alphavirus*' and 'Genus *flavivirus*'—I included each as a binary predictor in the model: yes (virus species belongs to specified genus) or no (virus species belongs to a different genus).
	Viral family	(i) Family for each of the 223 viruses (23 families). (ii) For the two families with the most human-infective RNA virus species—*flaviviridae* and *picornaviridae*, I only included 'Family *picornaviridae*' as a binary predictor in the model: yes (virus species belongs to specified family) or no (virus species belongs to a different family), because I have included genus *flavivirus*, which is the main genus from family *flaviviridae*.
Structure	Viral envelope	(i) Does the virus have outer envelope—the structure that cover the protective protein capsids of virus? (ii) I included 'Enveloped' as a binary predictor: yes (virus species has an envelope) or no (virus does not have an envelope).
	Genome	(i) Viral genome for each virus. (ii) I included 'Genome' as a categorical predictor: a. positive-sense single-stranded RNA [(+)ssRNA], b. negative-sense single-stranded RNA [(-)ssRNA)], c. single stranded RNA-RT (ssRNA-RT), and d. double stranded RNA (dsRNA).
Transmission	Transmissibility	(i) Whether the virus can be transmitted between humans. (ii) I include 'Transmissibility' as a binary predictor: yes (transmissible) or no (non-transmissible/strictly zoonotic). Non-transmissible/strictly zoonotic viruses have a reproduction number (R_0) of 0 (transmissibility level of 2). Transmissible viruses have a R_0 greater than 0, including viruses with transmissibility level of 3 ($0<R_0\leq1$) and 4 ($R_0>1$).
	Transmission routes (4 predictors)	(i) The route via which the virus is transmitted to humans. (ii) For the four transmission routes—'Vector-borne', 'Inhalation', 'Ingestion', and 'Direct contact', I included each as a binary predictor: yes (virus species is transmitted through specified route) or no (virus species is transmitted through a different route).

Host range	Host range type	(i) The host range for each virus, depending on the type of hosts the virus can infect. (ii) I included 'Host range type' as a categorical predictor: a. Broad (humans and at least one other non-primate host category), b. Narrow (humans only or humans and other primates only), and c. Unknown.
	Host (3 predictors)	(i) The reservoir hosts of each virus. (ii) For the three reservoir hosts—'Non-human primates', 'Other mammals', and 'Birds', I included each as a binary predictor: yes (virus species has the specified animal as the host) or no (virus species do not have the specified animal as the host).
Origin	Index case identification	(i) The way the index case(s) was/were detected, according to the description from the first report of human infection. (ii) I included 'Index case identification' as a binary predictor: surveillance-based (identification through active discovery programmes) or non-surveillance-based (identification through passive investigation of the aetiology).
Clinical presentation	Pathogenicity	(i) Whether the virus can cause diseases or not. (ii) I included 'Pathogenicity' as a binary predictor: yes (pathogenic) or no (not pathogenic).
	Symptoms after infection (5 predictors)	(i) The main symptoms after infection. (ii) For the five symptoms that viruses can cause—'Blood symptoms', 'Viral fever', 'Neurologic infection', 'Respiratory infection', and 'Digestive infection', I included each as a binary predictor: yes (virus species can cause the specified symptom) or no (virus species can not cause the symptom).
	Fatality	(i) Whether the virus is fatal or not. (ii) I included 'Fatality' as a binary predictor: yes (fatal) or no (not fatal).

Next, I developed a bootstrapping procedure to generate a robust estimate of model performance. First, I randomly extract 1000 bootstrapped samples from my 223 records of virus geographical extent/viral disappearance and their predictors, with each sample having 223 records (meaning I allow repetition of records in each sample). By fitting 1000 replicate BRT models using the 1000

bootstrapped samples, the relative contribution plots and partial dependence plots with 95% quantiles were obtained. The most influential predictors were defined as those whose relative contribution was greater than the mean (i.e. 100/(total predictors count*100); this study: 100/(22*100)=4.55%) (Shearer et al., 2018). Apart from the 22 biological features presented in Table 4.1, the geographical extent (wide vs. restricted) was also considered as a predictor of viral disappearance. So a relative contribution greater than 100/(23*100)=4.35% was regarded as influential predictors of the lack of human viral infection in the last 10 years.

For the geographical extent, I performed two sensitivity analyses by i) using two or more continents as wide extent and one continent as restricted extent, and ii) using three or more continents as wide extent but excluding the continent if in it there were only imported cases or lab infections. Relative contribution plots were plotted by using the same BRT modelling strategy as I described above. The final parameters for both models are shown in Table 4.2.

Table 4.2 Model parameters for human-infective RNA virus geographic extent and disappearance

Model	Tree complexity	Learning rate	Bag fraction	No. of trees
Geographical extent				
Main model	4	0.010	0.5	1331
Sensitivity analysis (i)	4	0.010	0.5	1927
Sensitivity analysis (ii)	4	0.010	0.5	1419
Sensitivity analysis (iii)	4	0.010	0.5	1132
Disappearance				
Main model	5	0.008	0.5	1511
Sensitivity analysis (1)	5	0.008	0.5	1695

(i) using 2 or more continents as wide geographical extent and 1 continent as restricted geographical extent; (ii) using 3 or more continents as wide geographical extent but excluding the continent if in it there are only imported cases or lab infection; (iii) and (1) including the number of publications per species.

To factor out research effort, one further sensitivity analysis was performed for both geographical extent and virus disappearance by including the number of publications per species in the models. The number of publications per species entered the model as a four-level categorical predictor, using the lower quartile, median, and upper quartile as the cut-off values. I included the number of publications per species in model for viral disappearance—though geographical extent can also partly reflect the research effort—given that the BRT model can deal with the interaction between predictors (Elith et al., 2008). Similarly, relative contribution plots were plotted by using the same BRT modelling strategy as I described above. The final parameters for the two models are shown in Table 4.2.

I calculated two different statistics to evaluate the model's predictive performance: i) the cross-validated explained deviance (a measure of the goodness-of-fit between the predicted and raw data) of the bootstrapped model (Elith et al., 2008), ii) the area under the receiver operating characteristic curve (AUC) calculated from 50 rounds of ten-fold cross-validation (Mandrekar, 2010), following the similar procedure of calculating ICC in chapter 2.3.4. The receiver operator curves (ROCs) were drawn and the median of AUC was calculated. The AUC ranges from 0 and 1, with an AUC of 0.5 suggesting predictive ability no better than random, 0.7 to 0.8 suggesting acceptable predictive ability, 0.8 to 0.9 suggesting excellent predictive ability, and greater than 0.9 suggesting outstanding predictive ability (Mandrekar, 2010).

All statistical analyses were performed using the R software (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria). P values <0.05 were considered statistically significant.

4.4 Results

4.4.1 Geographical extent of human-infective RNA viruses

A total of 83 viruses (37%) were found to have wide geographical extents, with human infections occurring in three or more continents, while the remaining 140 (63%) were found to be currently restricted to just one or two continents (Figure 4.2, Table C.1). The geographical extent of human-infective RNA viruses varied between different viral biological features (Figure 4.3). In the univariable analysis, I found that human-infective RNA viruses with wide geographic extent were more likely to be in non-*alphavirus* genus, be in *picornaviridae* family, lack an outer envelope, have a +(ss)RNA genome, be transmissible between humans, be transmitted through ingestion or direct contact but not through vector-born transmission, to have a narrow host range (restricted in human or human and NHPs), to not have other mammalian hosts, to be detected from symptomatic cases, to be pathogenic, to cause digestive or respiratory infections, but not cause viral fever, and finally, to be fatal ($p<0.05$).

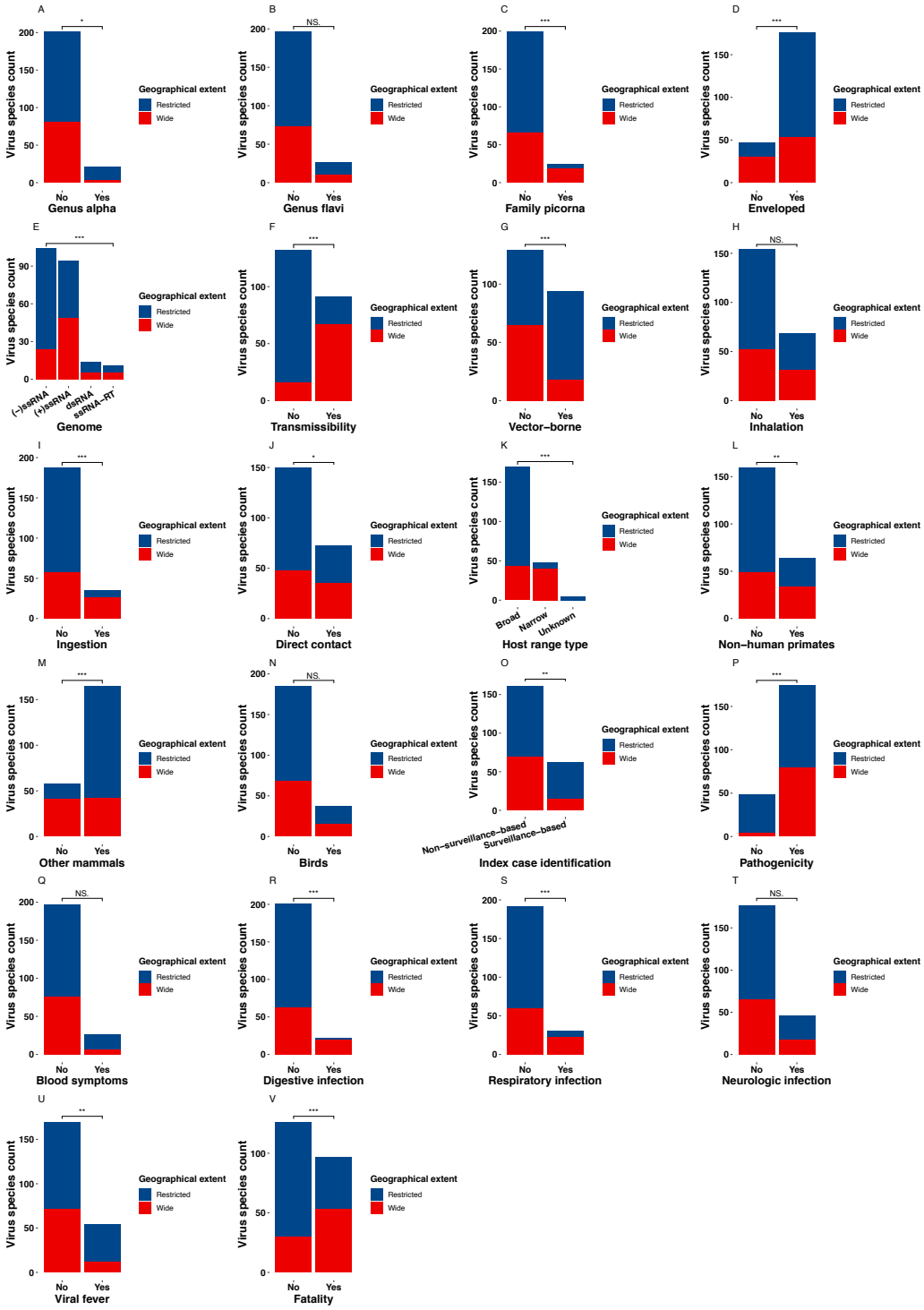


Figure 4.3 Number of human-infective RNA viruses per group with respect to geographical extent

Proportions were compared by using chi-square test or Fisher's exact probability test. "***" <=0.001, "**" <=0.01, "*" <0.05, NS. Not significant.

In the BRT model for geographical extent of viruses, I found that 4 predictors had relative contributions greater than the mean (4.55%) based on the 1000 replicate models (Figure 4.4), including transmissibility (32.6%), genome (9.0%), host range type (8.8%), and NHPs host (6.2%). The partial dependence plots provide detailed descriptions of these relationships (Figure 4.5). Human-to-human transmissible viruses were more likely to have a wide geographic extent, with human infections found across more than three continents in the world compared to non-transmissible viruses. The genome type of the virus was also important for predicting virus geographical extent, with +(ss)RNA viruses having a higher probability of being widely distributed between humans than other virus genome types. Host range was also predictive of virus geographical extents, with surprisingly viruses with a narrow host range more likely to be associated with a wide geographical extent of human infections. Viruses having NHPs as hosts showed a significantly higher probability of being widely distributed between humans compared to viruses without NHPs hosts. All 22 predictors explained a median of 55.8% (95%CI: 44.3%–68.3%) of the deviance in the data based on the 1000 bootstrap BRT models. Using 50 runs of 10-fold cross-validation, the model had a median AUC of 0.94 (95% quantiles: 0.90–0.97) (Figure 4.6 (A)), which represented an outstanding predictive performance (Mandrekar, 2010).

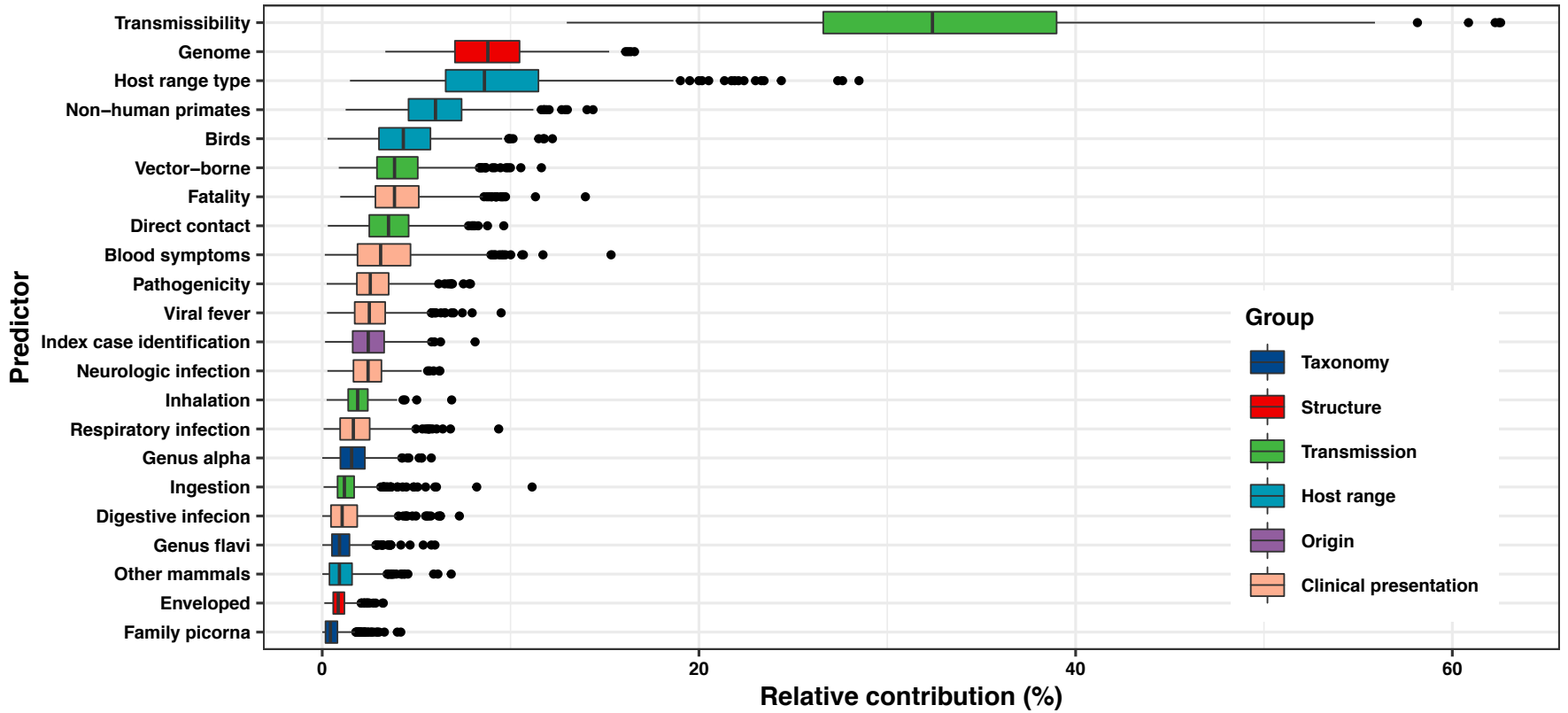


Figure 4.4 Relative contribution of predictors to human-infective RNA virus geographical extent

The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

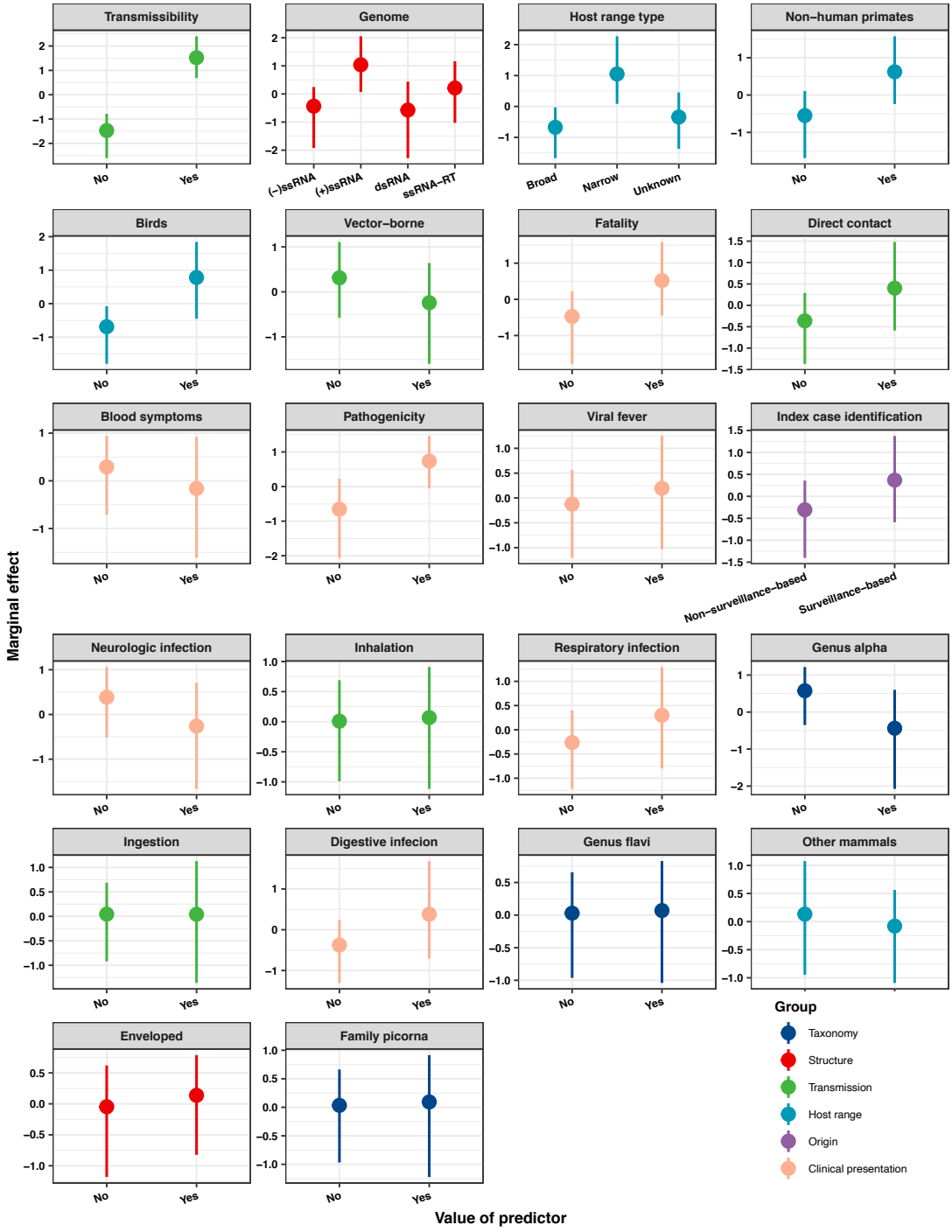


Figure 4.5 Partial dependence plots showing the influence on human-infective RNA virus geographical extent

Partial dependence plots show the effect of an individual predictor over its values on the response after factoring out other predictors. The dots represent the median and the whiskers represent 95% quantiles based on 1000 replicated models.

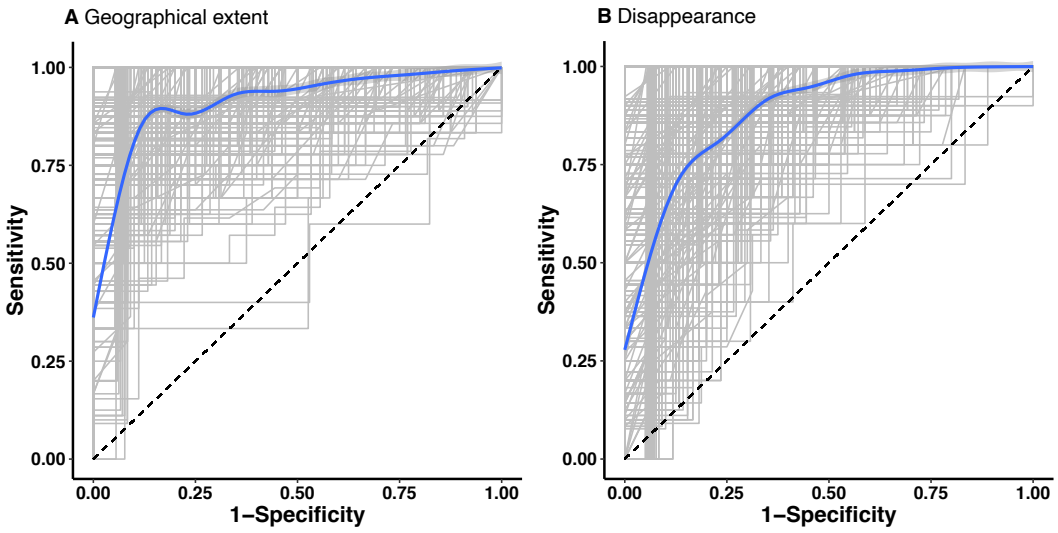


Figure 4.6 Receiver operating characteristic curves (ROCs) of the 50 rounds of 10-fold cross-validation for boosted regression tree models

A, Geographical extent; B, Disappearance. The blue curve denotes the averaged curve of ROC over 50 rounds of ten-fold cross-validation and the grey curves denotes each ROC from the 50 rounds of ten-fold cross-validation. Y axis denotes the sensitivity/true positive rate, and X axis denotes the 1–specificity/false positive rate. Dashed black line indicates no predictive power.

For the geographical extent, sensitivity analyses by i) using two or more continents as wide extent and one continent as restricted extent, and ii) using three or more continents as wide extent but excluding the continent if in it there were only imported cases or lab infections have revealed several changes of relative contribution, but the first three important predictors kept consistent with the main model (Figure 4.7).

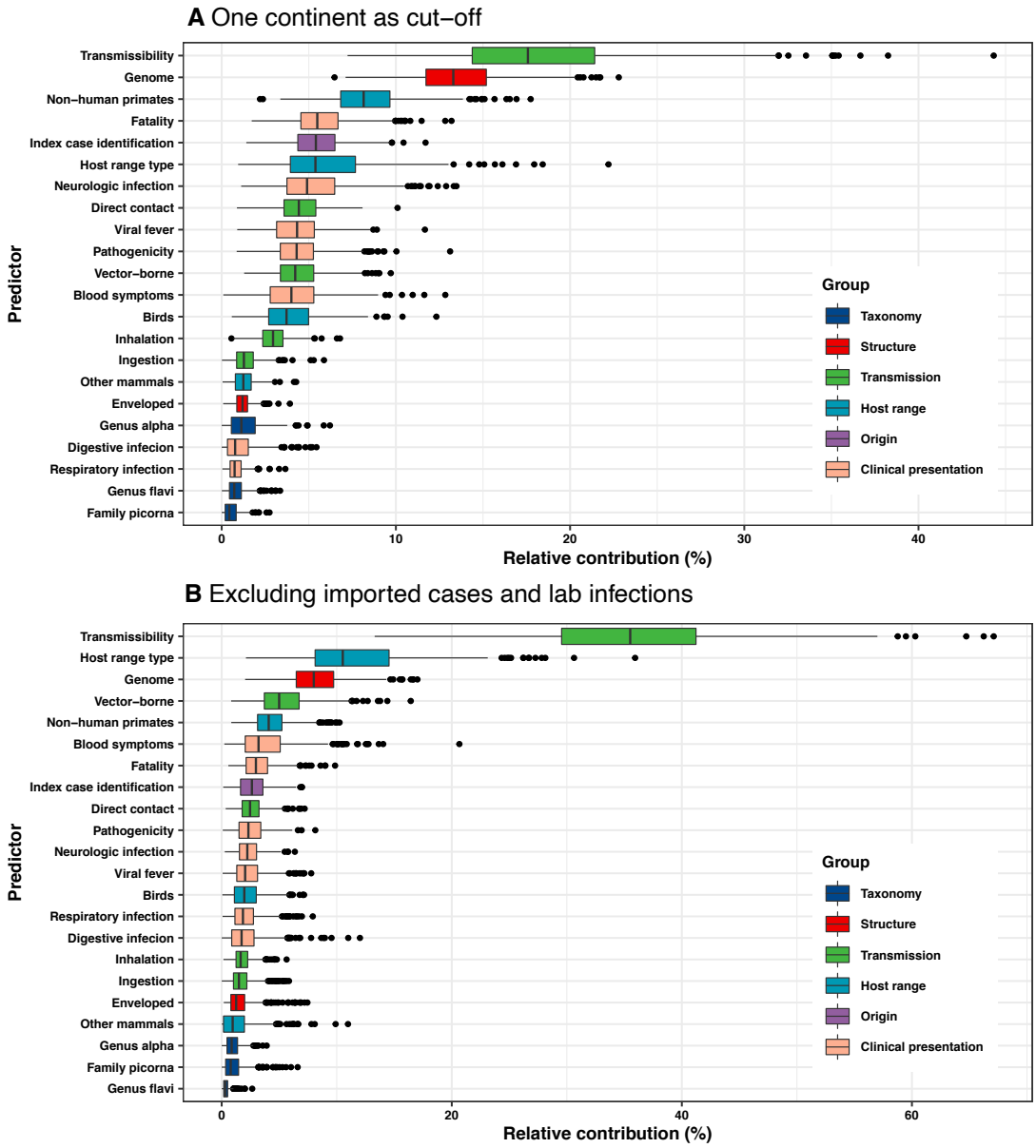


Figure 4.7 Relative contribution of predictors to human-infective RNA virus geographical extent with different definitions

A, using two or more continents as wide distribution and one continent as restricted geographical extent; B, using three or more continents as wide geographical extent but excluding the continent if in it there were only imported cases or lab infections. The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

Sensitivity analysis by including research effort in the model showed that the number of publications per species was the second important predictor of

geographical extent (Figure 4.8), and the explained deviance in the data increased from a median of 55.8% in the main model to 63.7% based on the 1000 bootstrap BRT models. After factoring out the effect of research effort, the relative contributions of other important predictors were broadly consistent with the main model (Figure 4.4).

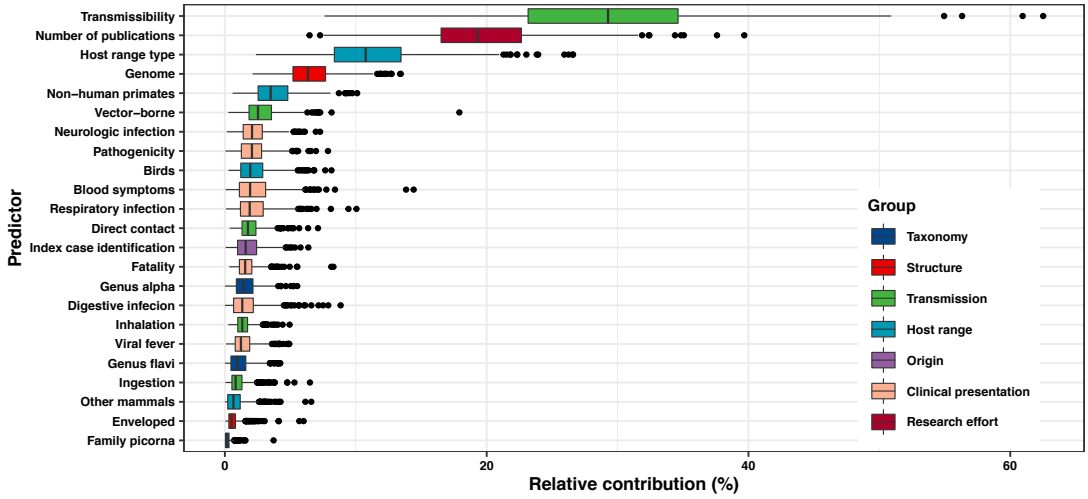


Figure 4.8 Relative contribution of predictors to human-infective RNA virus geographical extent by including the research effort

The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

4.4.2 Disappearance of human-infective RNA viruses

A total of 81 viruses (36%) were found to have disappeared, as measured by a lack of detected human infections over the last 10 years, while the remaining 142 (64%) had evidence of human infections during this period (Table C.1). Figure 4.9 shows the 'survival years' of 223 RNA viruses in humans (as measured from the first documented case to the last documented case); the

median was 25 years. The viral family *peribunyaviridae* had the largest proportion (65%) of viruses that disappeared from causing recorded human infections, followed by *rhabdoviridae* (61%), *reoviridae* (54%), and *nairoviridae* (50%). In the univariable analysis, viral disappearance was found to be determined by several key biological features (Figure 4.10), with viruses having been found to not infect humans in the last 10 years to be more likely to be non-transmissible, transmitted by vectors and not through ingestion, with a broad host range (capable of infecting other mammals except NHPs), able to infect other mammals, detected through surveillance, non-pathogenic, not able to cause digestive, or respiratory infections, non-fatal, and with a restricted geographic extent ($p<0.05$).

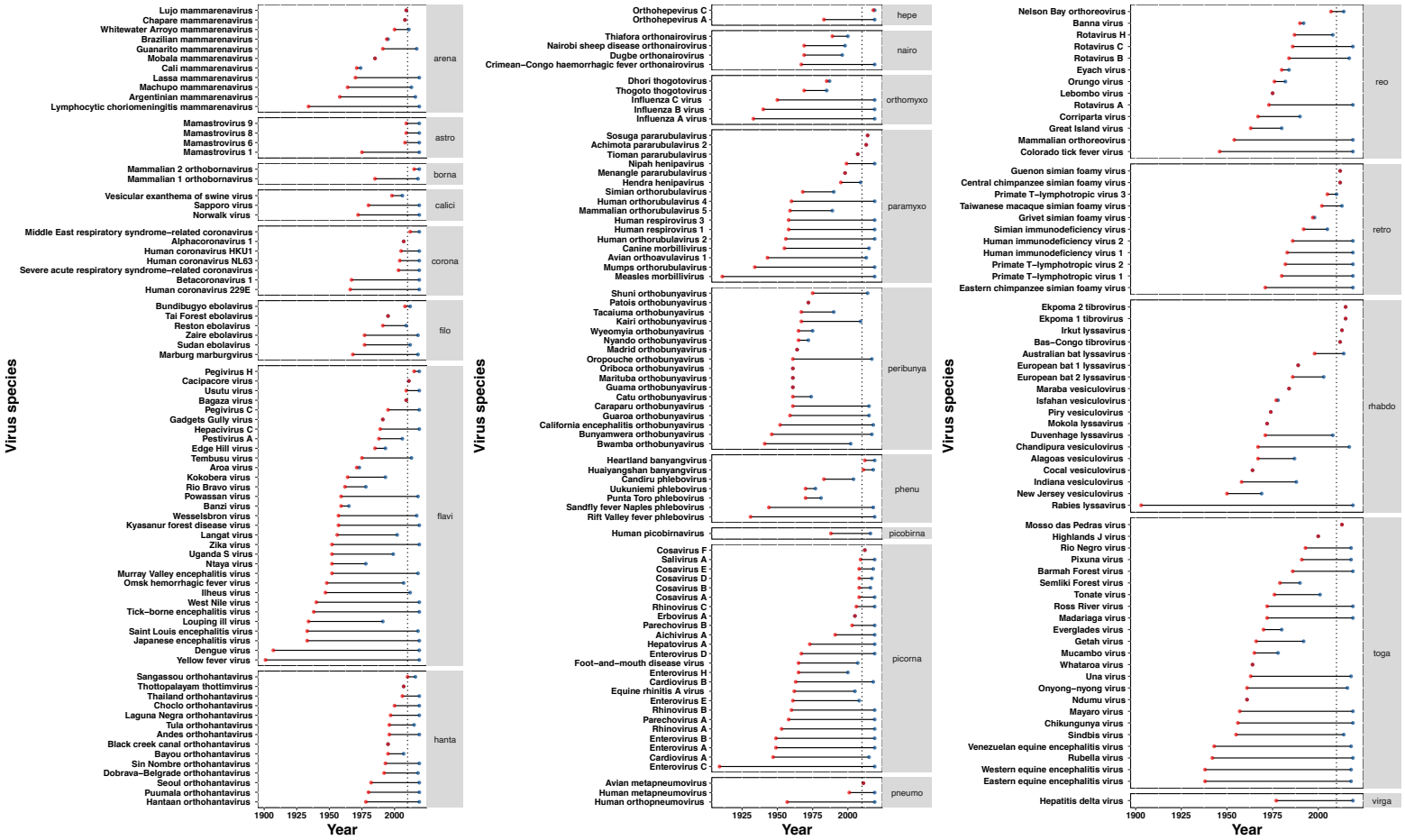


Figure 4.9 Survival years of each human-infective RNA virus species in humans

The red dots denote the discovery date of each virus in humans, and the blue dots denote the year of the last report from humans. The vertical dash lines denote the year of 2010.

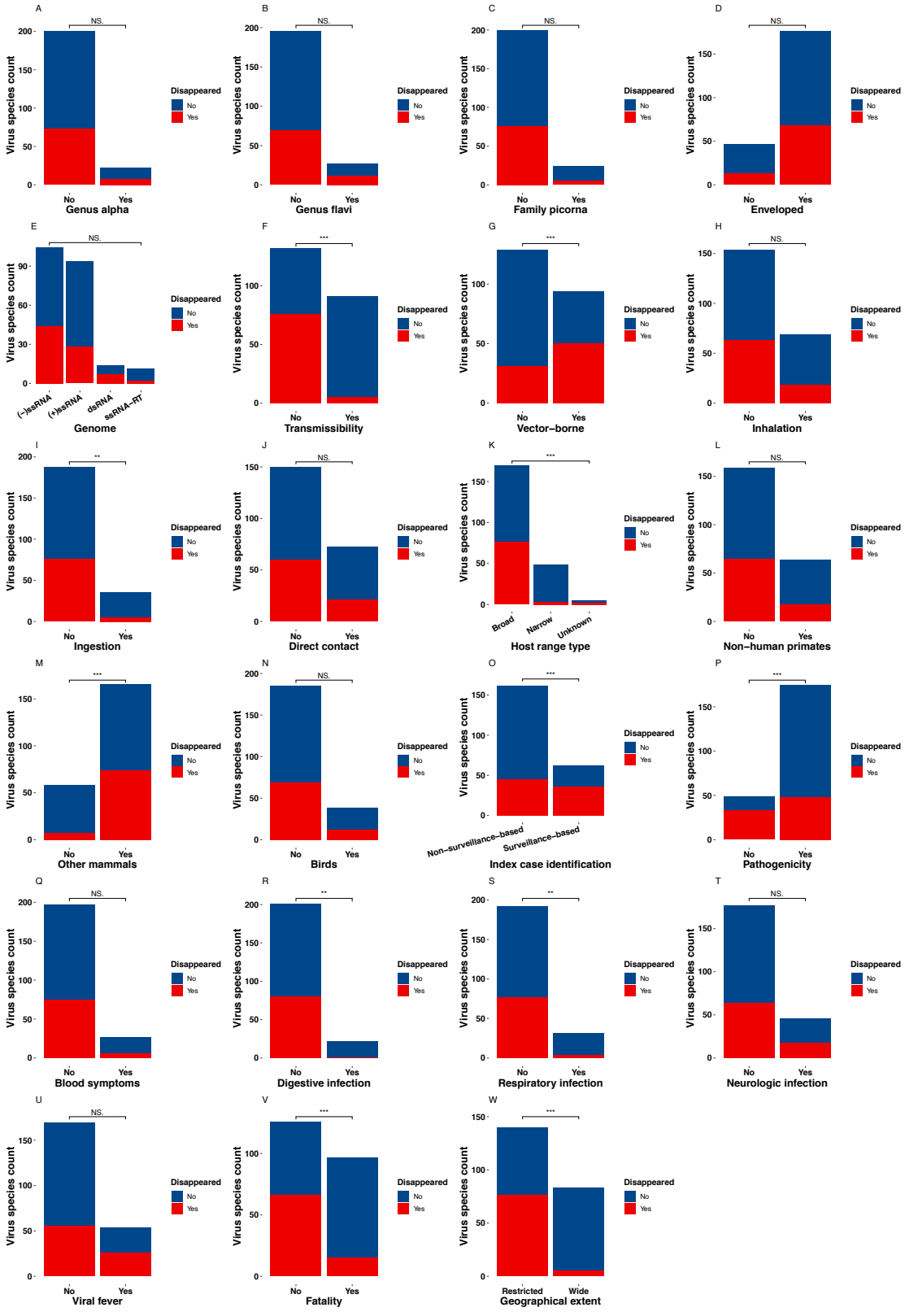


Figure 4.10 Number of human-infective RNA viruses per group with respect to viral disappearance

Proportions were compared by using chi-square test or Fisher's exact probability test. "***" <=0.001, "**" <=0.01, "*" <0.05, NS. Not significant.

In the BRT model for the disappearance of viral infections in humans over the last 10 years, I found that 8 predictors had relative contributions greater than the mean (4.35%) based on the 1000 replicate models (Figure 4.11), including transmissibility (20.9%), geographic extent (12.2%), genome (8.8%), fatality (7.4%), pathogenicity (6.0%), index case identification (5.6%), vector-borne (5.0%), and direct contact (4.4%). The partial dependence plots shown inFigure 4.12 Figure 4.12 provide detailed descriptions of these relationships. Non-transmissible viruses were more likely to have disappeared from humans than transmissible viruses. Geographical extent was also predictive of virus disappearance in my model. Virus with a restricted extent was associated with an increased likelihood of disappearance between humans. Genome was another factor for virus disappearance, with dsRNA viruses having a higher probability of disappearing between humans than other virus genome types. Non-fatal and non-pathogenic viruses were associated with a higher likelihood of disappearance than fatal and pathogenic viruses. Further, viruses discovered through surveillance showed a significantly higher probability of disappearance among humans compared to viruses discovered from symptomatic cases. Vector-borne viruses and viruses transmitted via direct contact were also more likely to have disappeared from humans. All 23 predictors explained a median of 42.6% (95%CI: 32.7%–55.7%) of deviance in the data based on the 1000 bootstrap BRT models. Using 50 runs of 10-fold cross-validation, the model had a median AUC of 0.91 (95% quantiles: 0.87–0.94) (Figure 4.6 (B)), which also represented an outstanding predictive performance (Mandrekar, 2010).

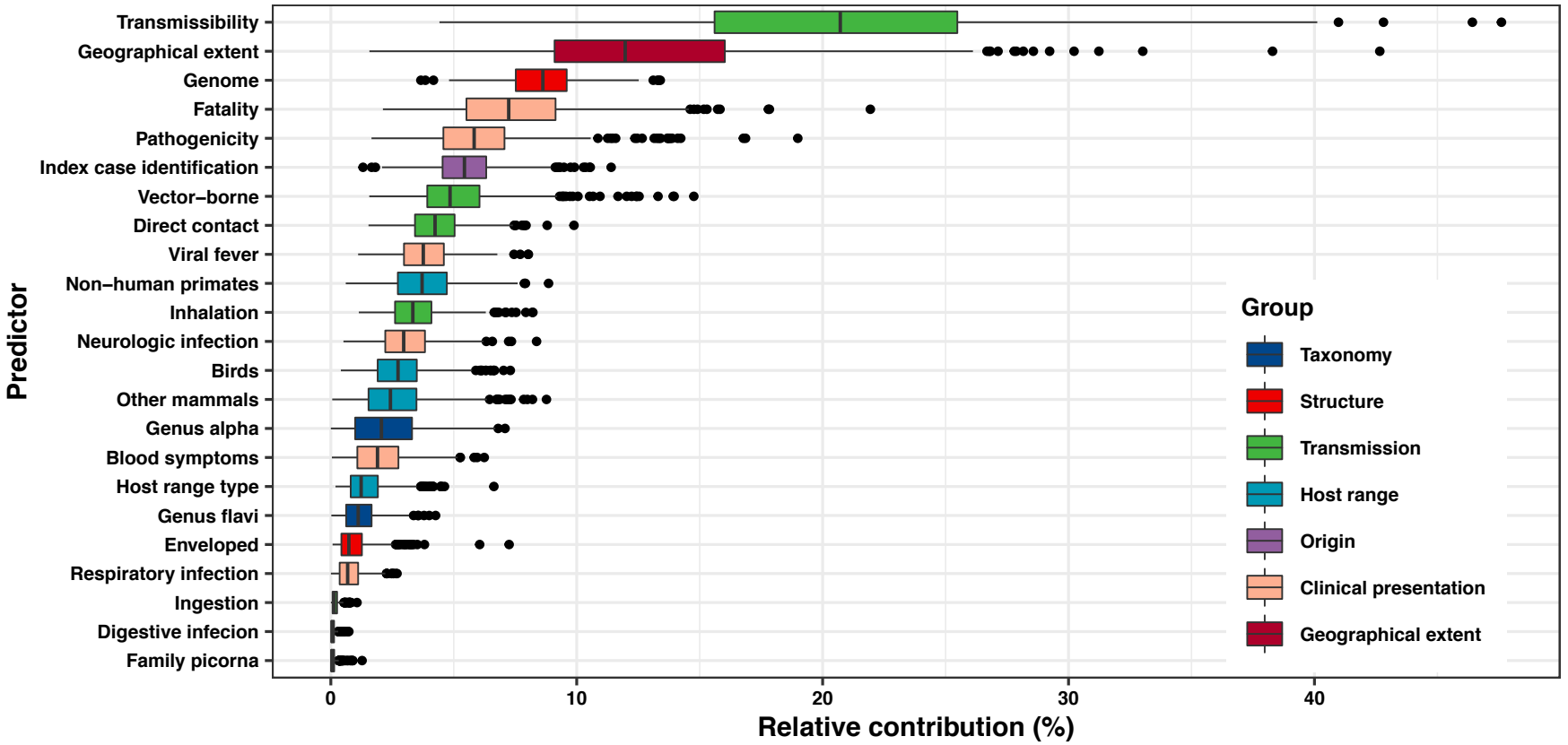


Figure 4.11 Relative contribution of predictors to human-infective RNA virus disappearance

The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers

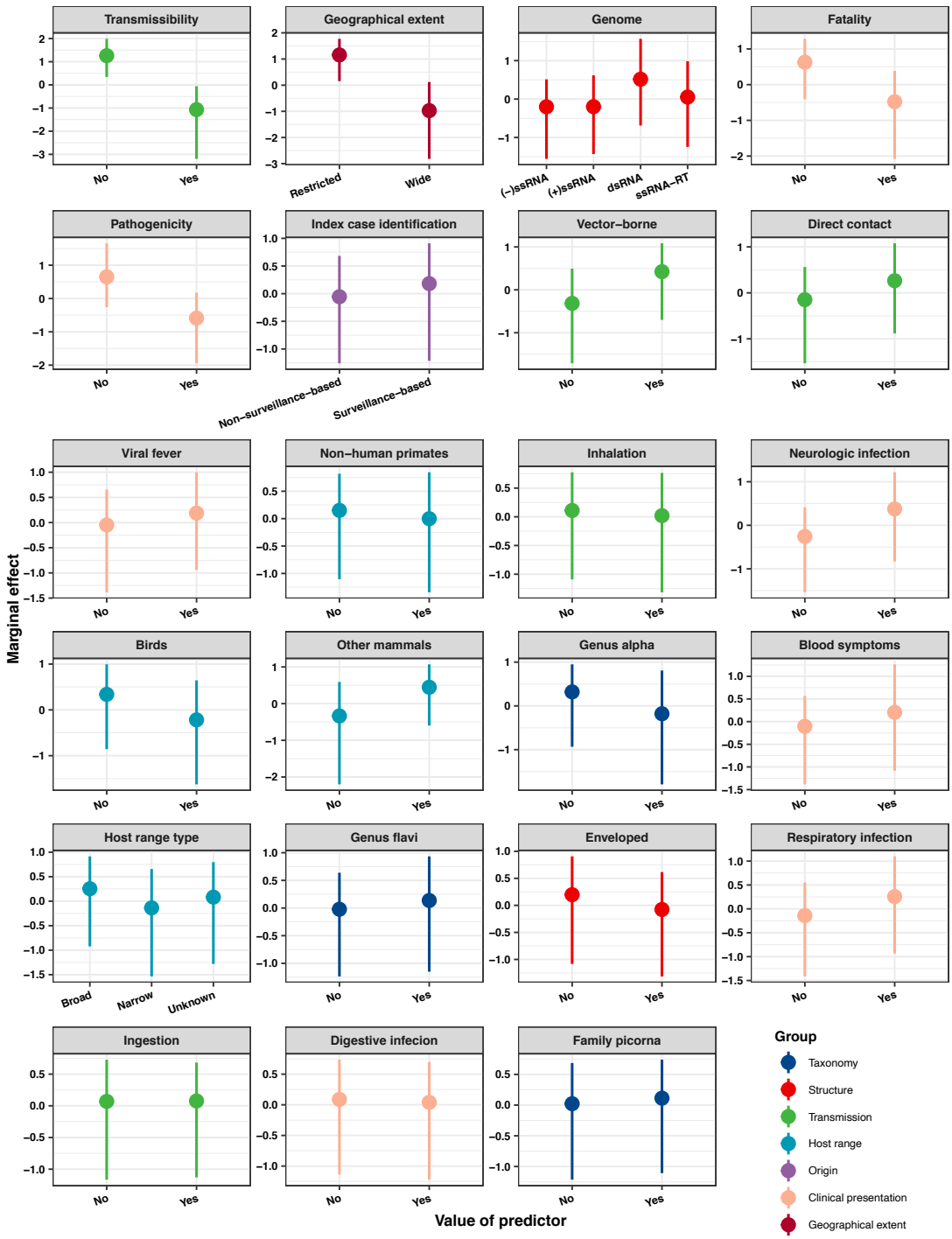


Figure 4.12 Partial dependence plots showing the influence on human-infective RNA virus disappearance

Partial dependence plots show the effect of an individual predictor over its values on the response after factoring out other predictors. The dots represent the median and the whiskers represent 95% quantiles based on 1000 replicated models.

Sensitivity analysis by including research effort in the model showed that the number of publications per species was the most important predictor of viral disappearance (Figure 4.13), and the explained deviance in the data increased from a median of 42.6% to 46.5% based on the 1000 bootstrap BRT models. After factoring out the effect of research effort, the relative contributions of other important predictors were broadly consistent with the main model (Figure 4.11).

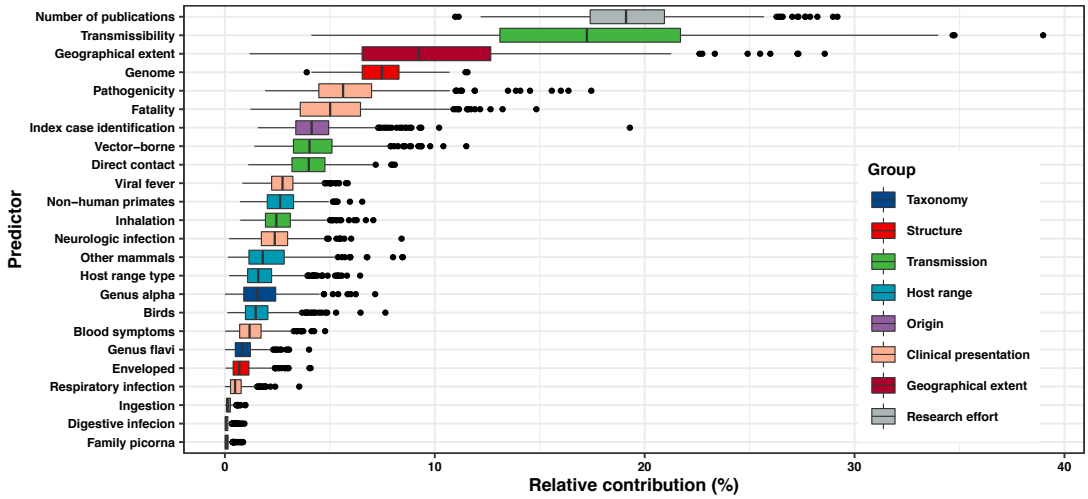


Figure 4.13 Relative contribution of predictors to human-infective RNA virus disappearance by including the research effort

The boxplots show the median (black bar) and interquartile range (box) of the relative contribution across 1000 replicate models, with whiskers indicating minimum and maximum and black dots indicating outliers.

4.5 Discussion

In this chapter, I collated the information for all known human-infective RNA viruses with respect to their geographical extent and disappearance, and explored the associated biological predictors including viral taxonomy, structure/genome, transmission, host range, origin, and clinical presentation. I

found that 37% of human-infective RNA viruses had wide geographical extent (able to infect humans on 3+ continents), and viruses were more likely to have a wide geographical extent if they were transmissible, had a genome type of +(ss)RNA, had a narrow host range, and had NHPs as the host. For viral disappearance, I found that 36% of human-infective RNA viruses have disappeared, i.e. no reported human infection for at least the past 10 years, and the top eight predictors included transmissibility, extent type, genome, fatality, pathogenicity, index case identification, vector-borne, and direct contact. Viruses were more likely to disappear from humans if they were non-transmissible, had a localised extent, had a genome type of dsRNA, were non-fatal and non-pathogenic, have been discovered through surveillance, were vector-borne, and were transmitted by direct contact. I acknowledged that some of these predictors are not independent of each other, and I discussed here where such correlations might exist.

Transmissibility was the most prominent factor for geographical extent. This is not surprising given that human transmissible viruses are more likely to spread, cross geographical borders and be introduced to humans in new areas. Transmissibility is one important parameter for calculating R_0. R_0 is defined as the number of secondary cases generated by an index case in a previously unexposed population in absence of control measures in disease transmission (Woolhouse et al., 2016). We can calculate R_0 as the transmission rate (determined by the transmissibility of the infectious agent and the contact rates) multiplied by the average infectious period. So, theoretically, the larger the value of R_0, the more secondary cases generated by an index case and so the

more likely it is to generate a worldwide epidemic. For example, the R_0 of SARS-CoV and SARS-CoV-2 in the early phase was 2 to 3 (Lipsitch et al., 2003; Petersen et al., 2020; Riley et al., 2003), and this led to massive transmissions between humans in the world. In contrast, MERS-CoV has a R_0 of 0.69 and therefore it has never caused a global epidemic.

Transmissibility was also the most important predictor of viral disappearance, with non-transmissible viruses having a higher probability to disappear from humans. Human infections can still occur in viruses that are not capable of human to human transmission (e.g. Tai Forest ebolavirus, Hantaan orthohantavirus), due to spill-over events of zoonotic viruses when humans living or working in proximity with habitats of infected non-human hosts (Lee et al., 1978; Miranda et al., 1991). Of the 81 viruses that were classified as disappeared in my data (with most having a localised geographic extent), 51 have had less than five documented spill-over events since their discovery of causing infection in humans (Table C.1), and of note, most of these viruses were either not found to cause illness or only to cause mild non-specific symptoms. I inferred the disappearance of these non-transmissible is either a true disappearance due to low frequency of contacts between humans and animal hosts, reduced laboratory infections thanks to better-equipped laboratories, well-trained staff and safer procedures, or a false disappearance due to underreporting (less research effort), as these uncommon viruses are associated with less disease burdens. The observation that vector-borne viruses were more likely to disappear from humans was predicted, given that 78.5% of vector-borne viruses were strictly zoonotic and not capable of

documented human to human transmission (Zhang et al., 2020). It has been suggested that vector-borne viruses were less likely to be transmissible between humans due to the evolutionary trade-offs for infecting both invertebrate vectors and vertebrate hosts, the low viral loads in human populations, and antagonistic pleiotropy of vector-borne viruses in humans (Geoghegan et al., 2016; Walker et al., 2018).

It was surprising that viruses which have a narrow host range have a higher probability of a wide geographical extent in humans than viruses with a broad host range, and viruses which also have non-human primate hosts were more likely to have wide geographic extents. This can be explained partly by the fact that 84% (41/49) of viruses with a narrow host range are transmissible between humans, while 70% (119/169) of viruses with a broad host range are non-transmissible/strictly zoonotic (Woolhouse & Brierley, 2018). My results supported previous theories on host range-transmissibility trade-offs, i.e. zoonotic viruses (i.e. viruses broad host range) emerging from hosts with a higher host phylogenetic distance from humans was associated with a reduced human transmissibility due to viral adaptation trade-offs (Guth et al., 2019; Walker et al., 2018). In addition, zoonoses emerging from more distantly related hosts have been found to have higher mortality rates in human populations and therefore may be less likely to be transmitted to new susceptible populations given a reduced number of mobile active human cases (Brierley et al., 2019; Guth et al., 2019). The distribution of viruses with a broad host range were also likely to be restricted by the distribution of their animal hosts. For example, in genus *hantavirus*, *Seoul orthohantavirus* is the only

hantaviral pathogen having a wide geographical extent while other *hantavirus* including *Hantaan orthohantavirus*, *Dobrava-Belgrade orthohantavirus*, and *Bayou orthohantavirus* etc. were restricted to the old world or to the new world, mainly constrained by the distribution of their rodent reservoirs (Milholland et al., 2018).

Genome structure was also one of the top predictors of viral geographic extents, with (+)ssRNA viruses having a wider extent than other genome types. While, few studies have investigated the associations between the virus genome and epidemic potential, it has been suggested that double-stranded genome (including dsDNA and dsRNA) were more likely to be transmissible than single-stranded genome (including ssDNA, (+)ssRNA, and (-)ssRNA), though the underlying mechanism is unclear (Walker et al., 2018). In contrast, my BRT model suggested dsRNA viruses had the lowest probability of infecting humans across a wide geographical extent, though my data found that dsRNA viruses were the most likely (57%) to contain human transmissible viruses. If excluding the dsRNA viruses, the probability of having a wide geographic extent for the remaining three genome types were consistent with the proportions of transmissible viruses, i.e. (+)ssRNA > (ss)RNA-RT > (-)ssRNA viruses. Of note, most transmissible viruses with a (+)ssRNA genome (80%) have a transmission level of 4 (viruses that are capable of epidemic spread in humans) compared to only 41% of viruses with a (-)ssRNA genome. It has been hypothesised that (+)ssRNA viruses are more transmissible than (-)ssRNA viruses due to a simpler replication process of

(+)ssRNA viruses, while (-)ssRNA must be converted to positive-sense RNA by an RNA polymerase before translation (Ferrero et al., 2018).

I also found an association between genome and virus disappearance, with viruses with dsRNA genome being more likely to disappear from human populations. It is possible that dsDNA viruses have a slower rate of evolution than ssRNA viruses and therefore a lower likelihood of adaptation to new hosts (Aris-Brosou et al., 2019). In addition, I found that both non-fatal and non-pathogenic viruses had a higher probability of disappearance from human populations. I suggested that underreporting of these types of non-pathogenic infections may be driving this result. I also noted that 66% of viruses that can cause fatalities were capable of human-to-human transmission, whereas 23% of non-fatal viruses were transmissible. However, these results conflicted with previous analyses that human transmissibility was negatively associated with virus virulence (high fatality rate) (Brierley et al., 2019; Geoghegan et al., 2016). The possible reason is that I have included many rare viruses which cause asymptomatic or mild symptoms. That whether fatality and pathogenicity are related to viral disappearance remains to be validated by further research.

My analysis revealed that the discovery process that found the first index cases, here described as viral origin, was also an important predictor in determining whether a virus is likely to disappear from causing recorded human infections. Specifically, viruses that were discovered through active surveillance, which were also mostly from asymptomatic cases and/or sporadic human infections from spill-over events, were more likely to

disappear compared to viruses discovered when investigating the aetiology. Again, this may be driven by less research effort, and/or true less frequent human infection may lead to the disappearance of these viruses (Berrios, 2007; Rezza et al., 2017).

While I have characterised an important set of biological predictors that contributed to the geographic extent and probability of disappearance for human-infective RNA viruses, this analysis has limitations. First, I have only considered viruses at the level of species, but variation for geographical extent may occur within individual species as well. For example, genotype I and II of *Orthohepevirus A* are the primary cause of hepatitis E endemic in developing regions of Asia, Africa, and South America; zoonotic spill-over of genotype III and IV has been found in cases of hepatitis E in developed regions, whereas genotype VII was only seen in one case from the Middle East (Dalton et al., 2008; Lee et al., 2016). Second, I only included virus species recognised by ICTV—and as such, the result of the study cannot be extrapolated to those that have not been included by ICTV. Third, I only considered biological factors in this study as I wanted to identify predictors specific to each virus. However, it is clear that socio-economic factors including urbanization (via the by-product effects such as overcrowd migrants, lack of sanitation facilities, alterations of the ecological system) and increased international travel (via transportation of passengers, vectors, and animals that carrying pathogens) can also aid the spread of viruses (Lipkin, 2013; Morse, 1995; Neiderud, 2015). I included publications per species to correct for the research effort, but this is unlikely to capture the influence of all socio-economic factors. Fourth, the same problem

of research effort exists in predictors such as transmission, host range, and clinical presentation, given these data were also collected from the peer-reviewed literature (Woolhouse & Brierley, 2017). For example, host range may be categorised incorrectly for some viruses because it is difficult for scientists to identify and report all their animal hosts. Fifth, I included index case identification as a fixed variable, but it is likely to evolve with time (time-varying), e.g. virus species discovered in recent decades have been more likely detected by active discovery programmes. Therefore, the true effect of index case identification on both outcomes is likely to vary over time.

The strength of the analysis is that the result has important implications for the management of emerging viruses at the earliest stages. My models are capable of predicting the geographical extent and viral disappearance of emerging human-infective RNA viruses by using their biological features. The early knowledge on which viruses are likely to have a wide geographical extent or to continuously infect human would help public health practitioners make control measures according to the risk assessment.

In conclusion, among human-infective RNA viruses, most have restricted geographic ranges (63%) and were still causing detectable human infection in the last ten years (64%). Transmission ecology and genome type were consistently the most important predictors of both geographical extent and viral disappearance. Host range was found to be an important predictor of geographical extent, though less important for disappearance. Geographical extent, discovery process and clinical presentation all contributed to the

probability of a virus disappearing. My analysis therefore reveals that multiple biological features determine geographical extent and disappearance of human-infective RNA viruses. This conclusion can be used to inform risk assessments for new viruses after their emergence.

Chapter 5 Predictors of COVID-19 epidemics in countries of the World Health Organisation African Region

Work in this chapter has been published in Nature Medicine.

5.1 Abstract

WHO African Region countries have experienced very different COVID-19 epidemics. This study aimed to identify predictors of the timing of the first COVID-19 case and the per capita mortality rate during the first and second pandemic wave in the region, and to test for any impact of countermeasures. I performed a region-wide, country-based observational study. Data on COVID-19 cases and deaths for all 47 countries in the WHO African Region were obtained from the WHO COVID-19 Dashboard. A set of predictors classified to nine categories were collected and used as explanatory variables. I applied Cox proportional hazards regression models, generalized linear mixed models, and multinomial logistic regression models as appropriate. Predictors of an earlier first case were a more urban population, high international connectivity and more land borders, and better COVID-19 test capacity. Predictors of a high per capita mortality rate during the first wave were a more urban population, higher pre-pandemic international connectivity,

and higher prevalence of HIV. The stringency and timing of government restrictions on behaviour were not associated with a lower per capita mortality rate in the first wave. A more urban population and a higher infectious disease resilience score were associated with more stringent restrictions and/or a higher per capita mortality rate in the first wave. The predictor set for the second wave was similar to the first. Second wave per capita mortality could be predicted from that of the first wave. These results were not altered when measures of national testing effort were included in the models. I concluded that COVID-19 in Africa arrived earlier and caused greater mortality in countries with more international travel and a more urban population. Mortality was exacerbated by high HIV prevalence; this could be a direct or indirect effect, or both. Countries rated as better prepared and having more resilient health systems were worst affected by the disease, the imposition of restrictions or both, making any benefit of more stringent countermeasures difficult to detect. The COVID-19 pandemic highlights unanticipated vulnerabilities to infectious disease in Africa that should be taken into account in future pandemic preparedness planning.

5.2 Introduction

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) presents an urgent threat to both global health and the global economy. By early March 2021, more than 119 million cases worldwide have been reported, with more than 2.6 million deaths as of early March 2021 (World Health Organisation, 2021c). In response to the

emerging threat posed by this disease, unprecedented public health interventions including social distancing, contact tracing, and large-scale lockdowns of the population have been implemented (Hsiang et al., 2020). Despite this, the burden of the disease has continued to rise, but with substantial variation among countries and regions and with countries in many regions around the world experiencing multiple waves (World Health Organisation, 2021c). As of 14 March 2021, the WHO African Region had reported over 2.9 million cases of infection and more than 74,000 deaths (World Health Organisation, 2021c). A third wave is currently in progress.

Gaining an understanding of variation in the progression of the pandemic in different countries will aid the response to future pandemics. Current evidence from high- and middle-income countries suggests that demographics (e.g. proportion of population aged 65 years or older), co-morbidities, healthcare resources, and stringency of response are important risk factors for COVID-19 related infections (Hsiang et al., 2020; Ji et al., 2020; Kandel et al., 2020; Zheng et al., 2020). It had been suggested that Africa would be more susceptible to SARS-CoV-2, given the higher prevalence of pre-existing conditions including tuberculosis, malaria, AIDS, diabetes, undernourishment, and other communicable and non-communicable co-morbidities, as well as lower accessibility to healthcare (Brauer et al., 2020; Walker et al., 2020). Using the data for COVID-19 cases and deaths from WHO COVID-19 Dashboard, this study aimed to identify predictors of the timing of the first case and the per capita mortality rate in the first and second COVID-19 pandemic

waves in the WHO African Region, and to evaluate the effectiveness of intervention measures on COVID-19 related deaths.

5.3 Materials and Methods

5.3.1 Study design and study area

I performed a region-wide, country-based observational study that included all 47 Member States of the WHO African Region (Figure 5.1)**.** The WHO African Region has a total population of 1,019,922,000, with the median age varying from 15.0 in Niger to 34.6 in Mauritius (World Health Organisation, 2020c). About 50% of the population in the WHO African Region lack access to essential medicines (Kirigia et al., 2009). Globally, 22 out of the 25 countries regarded as most vulnerable to infectious diseases are in sub-Saharan Africa (Moore et al., 2016).

I extracted data for daily cases and deaths for each country in the WHO African Region and calculated the following three outcomes: timing of the first case and per capita mortality rates in the first and second waves. Predictors relating to demographics, socioeconomics, travel, healthcare, co-morbidities, readiness, and geography were extracted from public data sources. The ratio of total COVID-19 mortality to reported COVID-19 mortality was obtained from the Institute for Health Metrics and Evaluation (Institute for Health Metrics and Evaluation, 2021). The COVID-19 test data quality and the government response data were collected by the Tackling Infections to Benefit Africa (TIBA) Pandemic Response Unit. COVID-19 testing policy data were taken from the

Oxford COVID-19 Government Response Tracker (OxCGRT). Total numbers of tests per capita was collected by Africa CDC (Salyer et al., 2021). Statistical models were fitted to evaluate the relationships among the three outcomes and predictors. I also run a secondary analysis for an outcome combining per capita mortality in the first wave and stringency index.

The start date of the analysis was set as 25 February 2020 when the first case was reported from the WHO African Region (in Algeria). I collated values of predictor variables as close to this date as possible.

5.3.2 Outcomes

My first outcome—the timing of the first case—refers to the day on which the first official laboratory-confirmed COVID-19 case/cases was/were reported to the WHO (**Error! Reference source not found.**A), largely based on case definitions defined by the WHO (World Health Organisation, 2020e).

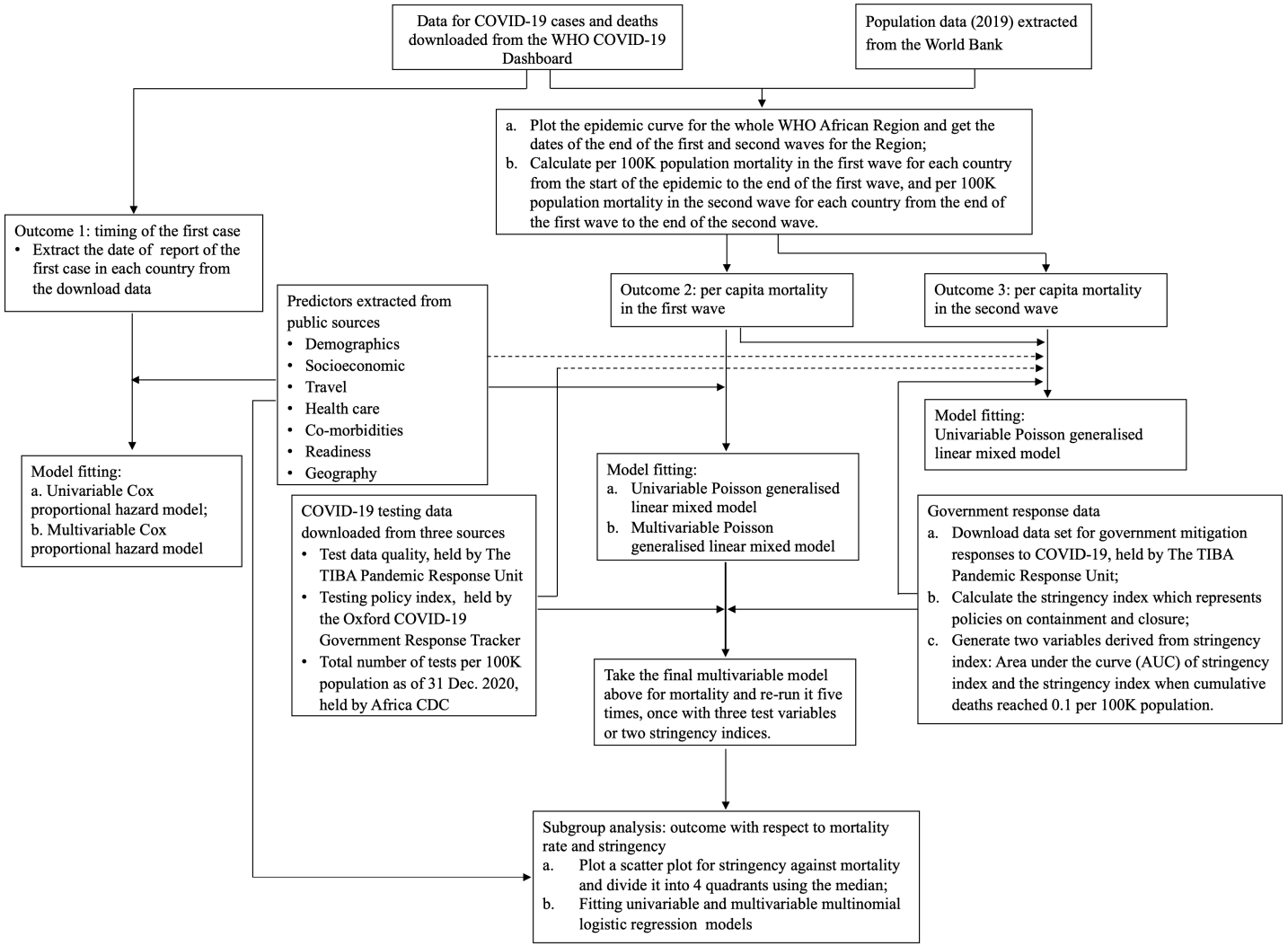
My other outcomes are the total deaths per 100K population (per capita mortality rate) during the first and second waves. According to international guidelines for certificate and coding of COVID-19 as cause of death (World Health Organisation, 2020d), a death due to COVID-19 is defined for surveillance purposes as a death resulting from a clinically compatible illness, in a probable or confirmed COVID-19 case, unless there is a clear alternative cause of death that cannot be related to COVID disease (e.g. trauma). There should be no period of complete recovery from COVID-19 between illness and death. A death due to COVID-19 may not be attributed to another disease (e.g.

cancer) and should be counted independently of preexisting conditions that are suspected of triggering a severe course of COVID-19.

The pandemic curve for daily new deaths for the whole WHO African Region was plotted by using 21-day kernel smoothing (**Error! Reference source not found.**B). Kernel smoothing is a common non-parametric method for revealing trends in curves. Here, I used Nadaraya–Watson estimator (can be seen as a weighted average, and a higher weight was assigned to daily new deaths closer to the target date) to fit a local estimate on each target date using daily new deaths within a 21-day window (10 days before and after the target date) (Hazelton, 2014) and smooth pandemic curves for daily new deaths was plotted based on these estimates. I chose the date with the first minimum daily new deaths (31 October 2020) as the end of the first wave and the date of the second minimum daily new deaths (14 March 2021) as the end of the second wave, and I calculated per capita mortality rate in each wave for each country.

Data on COVID-19 cases and deaths for all 47 Member States in the WHO African Region were taken from the WHO COVID-19 Dashboard (World Health Organisation, 2021d). The data include daily new cases, cumulative cases, daily new deaths and cumulative deaths.

Figure 5.1 Flow diagram for analysing predicators of COVID-19 epidemics in the WHO African Region



5.3.3 Predictors

A set of predictors considered likely to affect the timing of the first case and the per capita mortality rate were collected and included as explanatory variables. The definition, time range, details of missing data, and data sources are reported in Table 5.1. Predictors were classified in nine categories: demographics, socioeconomics, travel, healthcare, co-morbidities, readiness, geography, COVID-19 testing and interventions. Demographic and socioeconomic variables may predict both vulnerability to severe disease (e.g. by age) and transmission potential (e.g. urban versus rural populations) (Hradsky & Komarek, 2021; Lawal, 2021). Healthcare, readiness, and COVID-19 testing variables may predict the capability to detect and/or treat cases (Chaudhry et al., 2020; Okeahalam et al., 2020). Travel and the number of shared borders are likely to affect the imported cases from neighbouring countries (Chinazzi et al., 2020). Comorbidities are related to vulnerability to dying from infection (Hashim et al., 2020). Latitude is related to climate, which may affect transmission rates (Tzampoglou & Loukidis, 2020).

Data on COVID-19 testing were obtained from three sources. Testing effort was extracted from a recent report of the COVID-19 pandemic in Africa up to the end of December 2020 (Salyer et al., 2021). The predictor variable was total number of tests divided by per 100K population. Testing policy index data were collected by the OxCGRT, which records government policy on access to testing. The ordinal scores are shown in Table 5.1, and I calculated days with testing policy index above 2 during the first wave (25/02/2020 to

31/10/2020). Testing policy index at the start of the second wave on 1st November 2020 was used as a baseline predictor of per capita mortality in the second wave. A test data quality index was generated by the TIBA Pandemic Response Unit (see D.1 Supplementary Methods), and was placed into four categories (no data, basic, satisfactory, and good; Table 5.1). Estimated ratios of total COVID-19 mortality to reported COVID-19 mortality were obtained from the Institute for Health Metrics and Evaluation (Table 5.1) (Institute for Health Metrics and Evaluation, 2021).

Government response data were collected by the TIBA Pandemic Response Unit, and the stringency index representing policies on containment and closure was calculated using a method developed by OxCGRT (D.1 Supplementary Methods) (Hale et al., 2020). Two variables related to the stringency index were generated: area under the curve (AUC) of stringency index scores and stringency index score when cumulative mortality reached 0.1 per 100K population (D.1 Supplementary Methods).

Google mobility data (https://www.google.com/covid19/mobility/) available for 25 WHO African Region Member States were used to validate the data for the stringency index (D.1 Supplementary Methods).

Table 5.1 Predictors of interest for COVID-19 epidemics in the WHO African Region

Category	Definition from data source	Time range	Source
Demographics	1. Population, total Population, total is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship-except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. The values shown are midyear estimates.	2018, except that Eritrea was using data of 2011	The World Bank https://data.worldbank.org/indicator
	2. Population density (people per sq. km of land area) Midyear population divided by land area in square kilometres. Definition of population is as shown above. Land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases the definition of inland water bodies includes major rivers and lakes.	2018, except that Eritrea was using data of 2011. Data of South Sudan was added manually, using the above total population data and land area from the government (http://www.goss-online.org/about.html).	
	3. Urban population (% of total population) People living in urban areas as defined by national statistical offices. The data are collected and smoothed by United Nations Population Division.	2018, except that Eritrea was using data of 2011	
	4. Population ages 65 and above (% of total population) Population ages 65 and above as a percentage of the total population. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship.		
	5. Sex ratio (Male/Female) Ratio of male population count to female population count		
Socioeconomic	1. GDP per capita (current US$) Gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars.	2018, except that Eritrea was using data of 2011	The World Bank https://data.worldbank.org/indicator
	2. Human development index Human development index is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living.	2018	United Nations http://hdr.undp.org/en/data
Travel	1. Number of international airports Total number of international airports within each country, extracted from a repository of air traffic flow.	2019	The World Bank https://datacatalog.worldbank.org/dataset/global-airports
	2. Volume of international air travel Total seats from the most recent year (2019) for all airports with international air travel within each country.		

Healthcare	1. Current health expenditure (% of GDP) Level of current health expenditure expressed as a percentage of GDP. Estimates of current health expenditures include healthcare goods and services consumed during each year. This indicator does not include capital health expenditures such as buildings, machinery, IT and stocks of vaccines for emergency or outbreaks.	2017	The World Bank https://data.worldbank.org/indicator
	2. Infectious disease resilience index A composite index that helps identify countries that are potentially most vulnerable to poorly controlled infectious disease outbreaks because of a confluence of factors ranging across multiple domains, including political, economic, public health, medical, demographic, and disease dynamics. 0 indicating the country most vulnerable to infectious disease outbreaks and 1 indicating the most resilient country.	Data published in 2016	Rand Corporation https://www.rand.org/pubs/research_reports/RR1605.html
Co-morbidities	1. DALY rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases Age-standardized DALY (Disability-Adjusted Life Year) rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases. DALYs are used to measure total burden of disease - both from years of life lost and years lived with a disability. One DALY equals one lost year of healthy life.	2017	Our world in data https://ourworldindata.org/burden-of-disease
	2. DALY rates per 100,000 individuals from non-communicable diseases Age-standardized DALY rates per 100,000 individuals from non-communicable diseases (NCDs). DALYs are used to measure total burden of disease - both from years of life lost and years lived with a disability. One DALY equals one lost year of healthy life.	2017	
	3. Prevalence of HIV, total (% of population ages 15-49) Prevalence of HIV refers to the percentage of people ages 15-49 who are infected with HIV.	2018, two countries (Seychelles, São Tomé and Príncipe) were missing.	The World Bank https://data.worldbank.org/indicator
	4. Diabetes prevalence (% of population ages 20 to 79) Diabetes prevalence refers to the percentage of people ages 20-79 who have type 1 or type 2 diabetes.	2019	
Readiness	1. COVID-19 test capacity The ability to test for potential COVID-19 cases. There is no report case of COVID-19 in the WHO African Region when the data was released.	Based on newsletter released on 25 Feb. 2020	WHO Region office for Africa https://www.afro.who.int/health-topics/coronavirus-covid-19
	2. COVID-19 readiness status Levels of readiness are assessed on the basis of information from an extensive survey that is given to WHO country offices in 47 countries in Africa. The survey questions are broadly grouped under nine response pillars: logistics; coordination; risk communication and community engagement; laboratory capacity; Points of Entry (PoE); surveillance; infection prevention and control; Rapid Response Teams (RRT); and case management. Readiness status have been classified into three levels: limited, moderate and adequate.	Based on newsletter released on 13 Feb. 2020; Three countries (Cape Verde, Mauritius, Seychelles) were missing.	

Geography	1. Number of borders Number of countries with shared boundary.	2020	Data and Maps for ArcGIS (formerly Esri Data & Maps, https://www.arcgis.com/home/group.html?id=24838c2d95e14dd18c25e9bad55a7f82#overview)
	2. Latitude The absolute degrees from equator.		
COVID-19 testing	1. Days with testing policy index >= 2 2. Testing policy index on 1st November 2020 The Oxford COVID-19 Government Response Tracker (OxCGRT) systematically collects publicly available information on several different indicators of response that governments have taken to respond to the COVID-19 pandemic, including school closures, travel restrictions etc. One of the indicators is testing policy, which record government policy on who has access to testing. The ordinal scores are shown below. The data is daily- based data and I calculated days with testing policy index above 2 between 25/02/2020 and 31/10/2020 as a predictor of per capita mortality in the first wave. Testing policy index at the start of the second wave on 1st November 2020 was also collected as a baseline predictor of per capita mortality in the second wave. *0 – No testing policy* *1 – Only those who both (a) have symptoms AND (b) meet specific criteria (e.g. key workers, admitted to hospital, came into contact with a known case, returned from overseas)* *2 – testing of anyone showing COVID-19 symptoms* *3 – open public testing (eg "drive through" testing available to asymptomatic people)* *No data*	2020	OxCGRT https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
	3. Test data quality The TIBA Pandemic Response Unit has created a new data set recording COVID-19 testing data as of 31 October 2020 for WHO African Region, and qualified the data quality into four categories. *No data – viz. no data from official sources* *Basic data – data that is published irregularly, or that does not distinguish between people tested and tests conducted* *Satisfactory data – data that has none of the shortcomings of "basic data"* *Good data – data that goes beyond daily numbers of tests or people tested. It may provide information on the positivity rate of tests, and on which tests were routine surveillance or linked to contact tracing for example. It may distinguish in which laboratory tests have been conducted, or provide additional information on the gender or age or regional origin of people tested.*	2020	GitLab https://git.ecdf.ed.ac.uk/epigroup/covid-19/tiba_testing_data/blob/master/TIBA_PRU_Testing_Data.xlsx

	4. Total number of tests per 100K population as of 31 Dec. 2020 Africa CDC collected data on total number of tests per capita for each African country (https://africacdc.org/covid-19/). I obtained data as of 31 December 2020 from one paper published on Lancet on 24 March 2021. For nine countries (Algeria, Benin, Comoros, Eritrea, Liberia, Mauritius, Sao Tome and Principe, Seychelles, and Sierra Leone), the testing data were incomplete; For United Republic of Tanzania, testing information was not reported for more than 2 weeks.	2020	Lancet https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00632-2/fulltext
Interventions	1. Area under the curve (AUC) of stringency index 2. Stringency index when cumulative deaths reached 0.1 per 10K population. Stringency index represents policies on containment and closure, calculated from the data set for government mitigation responses to COVID-19 for WHO African Region, held by the TIBA Pandemic Response Unit. Stringency index is the average normalised strictness values of 12 subcategories of measures, excluding the governance and socio-economic measures and surveillance and testing from public health measures.	2020	GitLab https://git.ecdf.ed.ac.uk/epigroup/covid-19/tiba_pru_measures/blob/master/TIBA_PRU_measures.xlsx

For some variables I can't get the latest data for all countries, i.e. the date for some variables may be not consistent, but I set a criterion that the date should be after 2010, otherwise the value was set as missing.

5.3.4 Statistical Methods

All 47 Member States were included in the model for the timing of the first case, but the number of Member States included in the model for per capita mortality in two waves depended on the completeness of the data. The epidemic curves for both daily cases and deaths in each country within the WHO African Region were plotted to evaluate the completeness of the data. The government of United Republic of Tanzania stopped reporting COVID-19 cases/deaths since 8 May and, therefore, was excluded (World Health Organisation, 2021d).

For predictors, the most recent available data were used—and no earlier than 2010. If one predictor has missing values, one column of binary indicator was added showing which country has missing data and which has not, and both

the raw data and the indicator were included in the model. All predictors used had data available for at least 90% of countries.

Spearman's rank correlation was used to test for a correlation between predictors. Predictors with a correlation coefficient greater than 0.6 were not included in the same multivariable model.

Cox proportional hazards regression models were used to determine hazard ratios (HRs) and 95% confidence intervals (CIs) for individual predictors of timing of the first case. A univariable model was fitted first. Only predictors quantified on or before the start date were included in this analysis. COVID-19 test capacity, COVID-19 readiness status, and the number of borders entered the model as binary variables where 'no', 'limited and moderate', and 'no border' were set as the reference levels, respectively. For COVID-19 readiness status, I combined ''limited and 'moderate' into one single level—'limited and moderate', because few countries were at the 'limited' level (2 countries) (Figure D.1Q). Three countries (Cape Verde, Mauritius, Seychelles) with unknown COVID-19 readiness status were also included in the 'limited and moderate' level. Other variables entered the model as continuous variables, and all continuous variables were standardised before entering the model, by subtracting the mean and dividing by the standardized deviation. Variables with *P* values less than 0.2 were considered for inclusion in a multivariable model. If multiple variables with *P* values less than 0.2 were highly correlated (correlation coefficient greater than 0.6), only one variable was selected each time to enter the multivariable model. The multivariable model with the lowest

corrected Akaike information criterion (AICc) was taken as the best model (Emiliano et al., 2014), but models with +2 AICc score were also retained.

I used a generalized linear mixed model (GLMM) with a Poisson error distribution to identify predictors of per capita mortality rate in the first wave. I used the reported deaths times the ratio of total COVID-19 mortality to reported COVID-19 mortality (Figure D.1T) as the outcome, population size as an offset and country as a random effect. The risk ratios (RRs) and 95% CIs were calculated. Five countries (the United Republic of Tanzania having incomplete data, Burundi, Eritrea, and Seychelles being clear outliers and Seychelles and São Tomé and Príncipe having missing data for HIV prevalence) were excluded (also for the multinomial logistic model below for outcome with respect to per capita mortality in the first wave and stringency. Days with testing policy index >=2 entered the model as a binary variable (using median as the cut-off) where 'below median' was set as the reference level. Three countries (Guinea Bissau, Equatorial Guinea, and Comoros) with missing days with testing policy index were included in the 'below median' level. I treated test data quality as binary, combining no data and basic data to the lower level (reference level), and satisfactory data and good data to the higher level. Univariable models and the best multivariable model were fitted using the same approach as for the timing of the first case. I then added the two stringency scores (AUC of stringency index in Figure D.1Y and stringency index when cumulative deaths reached 0.1 per 100,000 population in Figure D.1Z) to the best multivariable model for per capita mortality in the first wave and checked for significantly improved model fit (lower AICc). I first estimated

the correlations between the two stringency scores and the set of selected predictors in the best multivariable model, using the Spearman rank correlation test. Then, I took the best multivariable model and re-ran it by adding each stringency score. Again, only stringency scores with correlation coefficients less than 0.6 with the set of selected predictors were included in the multivariable model. I repeated this exercise for the three testing variables, i.e. adding days with testing policy index >= 2 (Figure D.1U), test data quality (Figure D.1W), and tests per capita (Figure D.1X) to the best multivariable model for per capita mortality in the first wave and test if the result kept consistent after adjusting for COVID-19 testing.

I carried out a secondary analysis using the original set of predictors of COVID-19 mortality in the first wave to predict an outcome combining per capita mortality in the first wave and stringency index (Table 5.1). In this analysis, countries were placed into four groups based on the medians of total per capita mortality in the first wave and of the AUC of stringency index (high stringency/high mortality, high stringency/low mortality, low stringency/high mortality, and low stringency/low mortality). Multinomial logistic regression was used to estimate the relationship between these outcomes and the set of predictors, and the odds ratios (ORs) and 95% CIs were calculated. Univariable models and the best multivariable model were fitted using the same approach as for the first wave mortality rate. Low stringency/low mortality was set as the reference level. COVID-19 readiness status and number of borders were excluded from the model because no country in the low/low level

had adequate COVID-19 readiness status, and there was no island nation in the high/high level.

For the second wave mortality rate analysis, I fitted only the univariable model using the same approach as for first wave mortality rate. I dropped predictors related to travel and readiness, given that these pre-pandemic predictors cannot represent the baseline level at the start of the second wave. I added per capita mortality in the first wave (**Error! Reference source not found.**C) and testing policy index on 1 November 2020 (Figure D.1V) as two new predictors. Testing policy index on 1 November 2020 entered the model as a binary predictor where 'below 2' was set as the reference level. AUC of stringency in the first wave (Figure D.1Y), test data quality in the first wave (Figure D.1 W), and tests per capita as of 31 Dec. 2020 (Figure D.1X) were considered as predictors of second wave mortality rate, respectively.

R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria, 2020) was used in all statistical analyses. R packages used for model fitting included survival, lme4, nnet and mgcv. A two-sided *P* value <0.05 was regarded as statistically significant. The raw African shapefile used in the study was obtained from Data and Maps for ArcGIS (formerly Esri Data & Maps, https://www.arcgis.com/home/group.html?id=24838c2d95e14dd18c25e9bad55a7f82#overview).

5.4 Results

On 25 February 2020, Algeria was the first country in the WHO African Region to report COVID-19 cases (Error! Reference source not found.). Thirty-one out of 47 countries reported their first cases in the 2 weeks from 12 March to 26 March 2020. Lesotho was the last of the 47 countries to report its first case, on 14 May 2020. There was no apparent relationship between the timing of the first COVID-19 case and the first death (**Error! Reference source not found.**A).

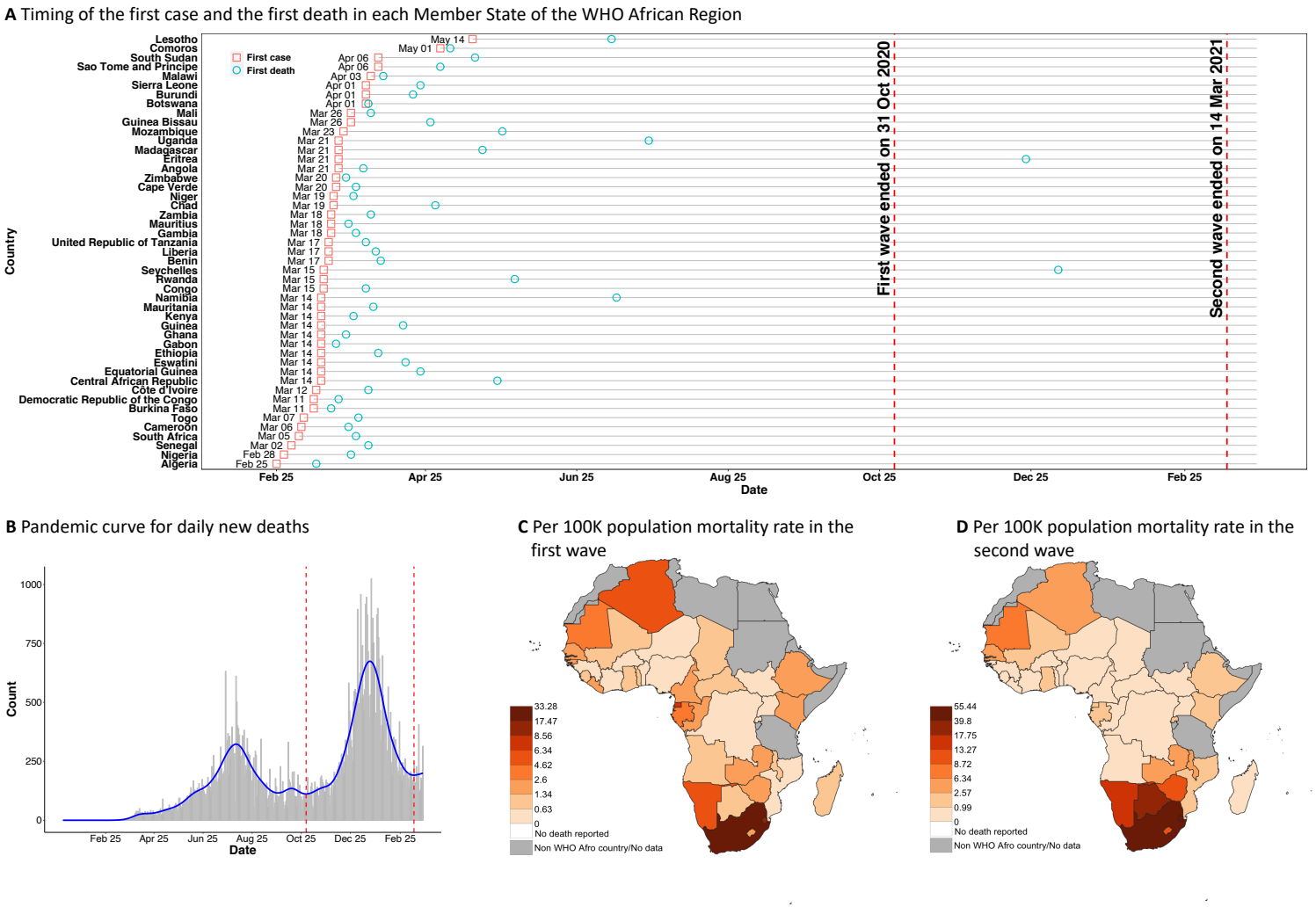
The 47 Member States reported a total of 29,635 COVID-19 deaths in the first wave and 44,850 deaths in the second wave. However, Tanzania discontinued reporting of COVID-19-related deaths from 8 May 2020, and Burundi, Eritrea and Seychelles were outliers (0.009, 0 and 0 per 100,000 population first wave mortality rates, respectively). São Tomé and Príncipe, as well as Seychelles, had missing data on the prevalence of HIV. These five countries were, therefore, excluded from the mortality rate analyses, giving a sample size of 42. Daily new deaths in the whole WHO African Region peaked on 5 August 2020 in the first wave and on 18 January 2021 in the second wave (**Error! Reference source not found.**B), lagging 16 and 7 d behind the peak of daily new cases in the first and second waves, respectively. The WHO African Region as a whole experienced a higher second wave peak than the first wave: 323 deaths (on 5 August 2020), and 675 deaths (on 18 January 2021), respectively. In the first wave, the highest mortality per 100,000 population was reported from South Africa (33.3), followed by Cape Verde (17.5) and Eswatini

(8.6 (**Error! Reference source not found.**C). In the second wave, the highest mortality per 100,000 population was also reported from South Africa (55.4), followed by Eswatini (39.8) and Botswana (17.7). Twenty countries had higher or similar mortality rates in the second wave than in the first wave, whereas 23 countries had lower mortality rates in the second wave than in the first wave (Figure 5.3).

I included 47 countries and 15 predictors (Figure D.1A-K, P-S) in the Cox regression model for timing of the first case. Spearman's correlation identified five pairs of predictors with correlation coefficients greater than 0.6 (Figure 5.4). The univariable Cox regression model identified total population size, number of international airports, volume of international air travel, COVID-19 test capacity, and COVID-19 readiness status as risk factors for earlier detection of the first case, and current health expenditure (% of GDP) as protective factor (Figure 5.5, Table D.1). In the multivariable model, the percentage of urban population (hazard ratio (HR) = 1.40, 95% confidence interval (CI) 1.01–1.95), number of international airports (HR = 1.48, 95% CI 1.02–2.14), volume of international air travel (HR = 1.52, 95% CI 1.10–2.11), COVID-19 test capacity (HR = 3.86, 95% CI 1.83–8.15) and number of borders (HR = 2.87, 95% CI 1.12–7.32) were identified as risk factors for earlier detection of the first case (Figure 5.5, Table D.1).

Figure 5.2 COVID-19 pandemic in the WHO African Region

A, Timeline of the first case and first death; B, Pandemic curve for daily new deaths; C, Map of per capita mortality rates in the first wave and D, in the second wave.



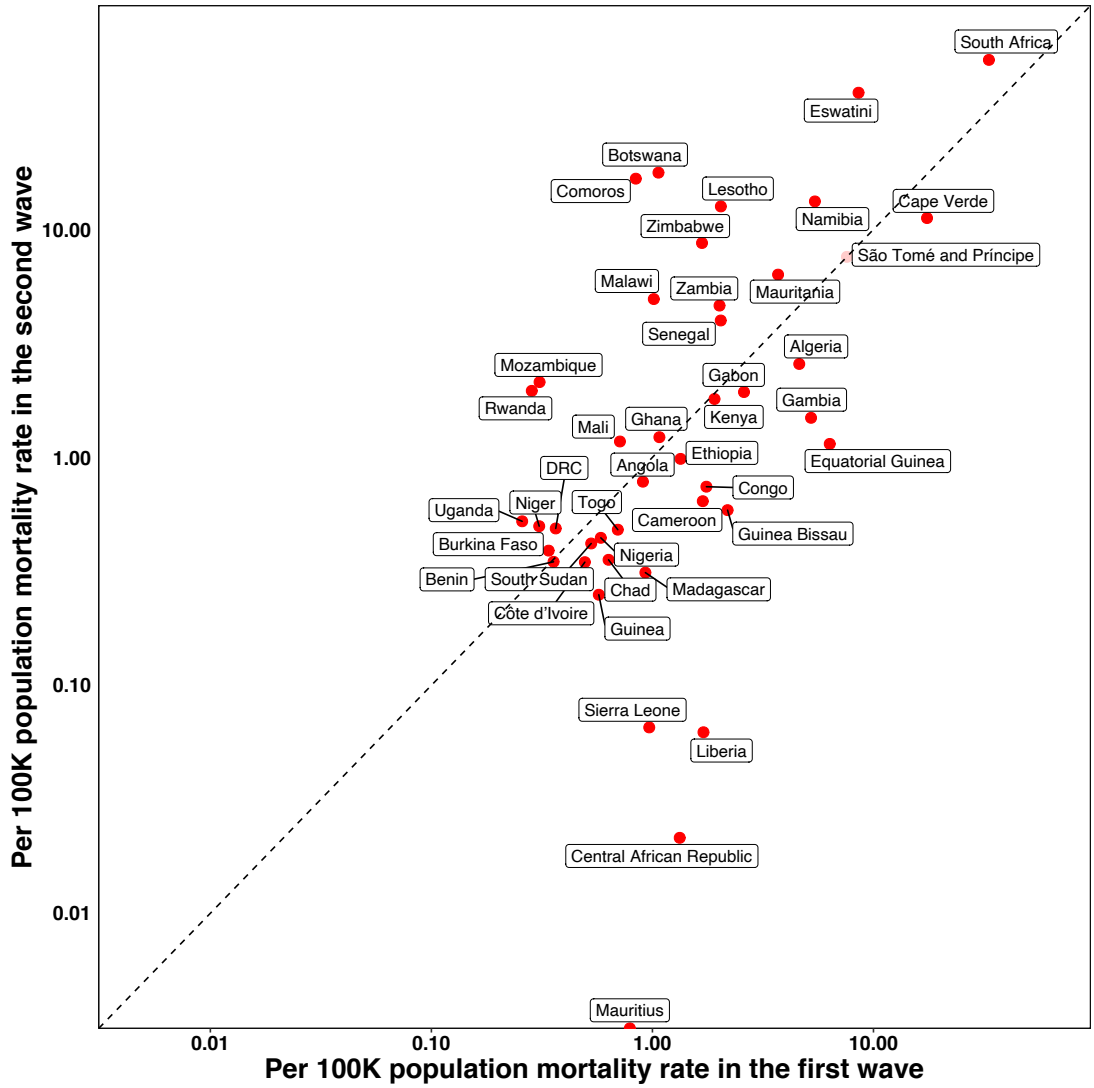


Figure 5.3 Scatter plot of per capita mortality rates in the first and second COVID-19 waves

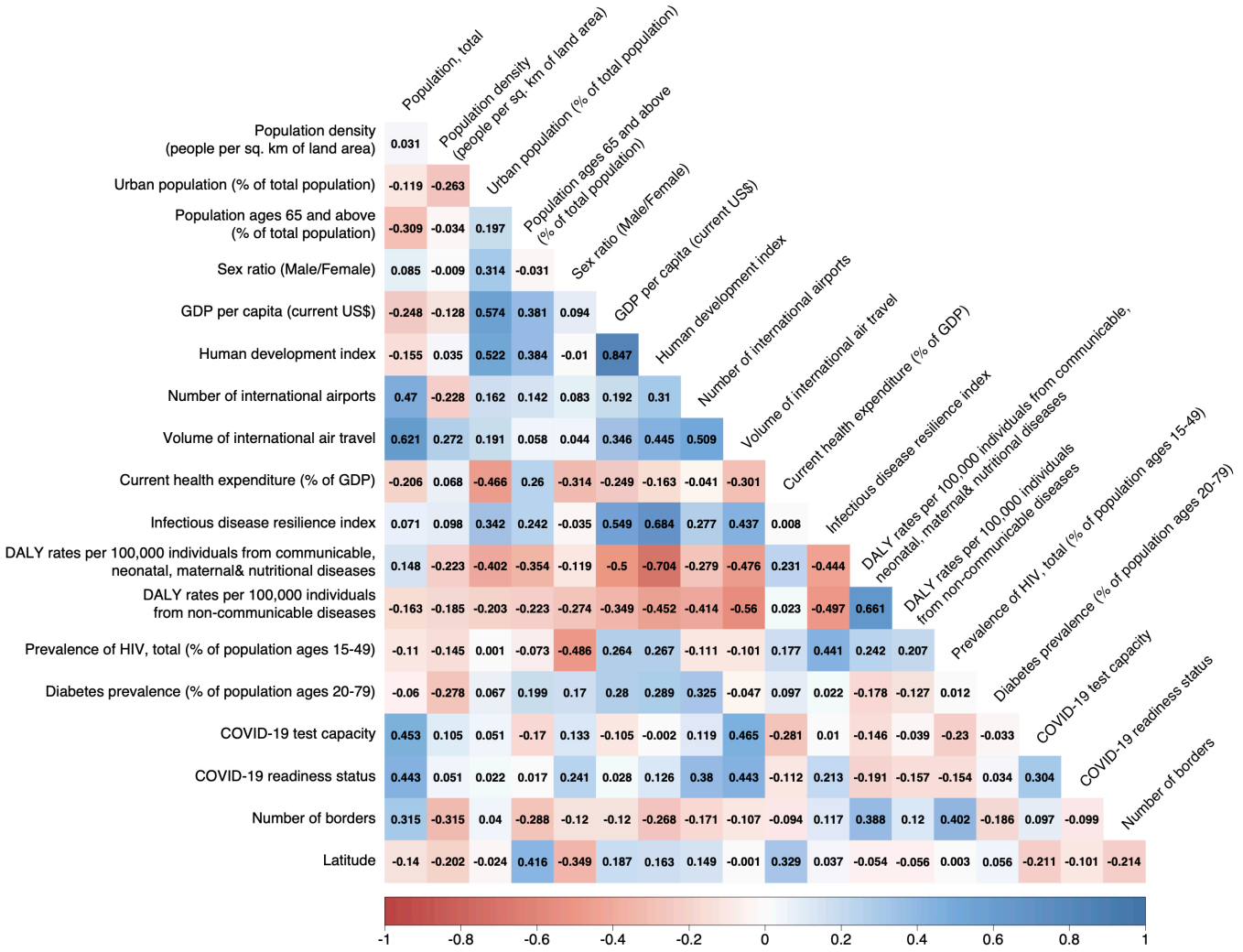
Axes on log10 scale with points falling on the axes denoting zero deaths. The dashed line indicates identical levels of mortality rates in two waves. Tanzania is not shown due to incomplete data. DRC, Democratic Republic of the Congo.

I included 42 countries and 18 predictors (Figure D.1B-S) in the GLMM models for per capita mortality in the first wave. In the univariable analyses, the percentage of urban population, GDP per capita, human development index, volume of international air travel, infectious disease resilience index, prevalence of HIV, and latitude were risk factors (Figure 5.6, Table D.2). The correlation between the time to first case and per capita mortality was not

significant (p=0.22). In the multivariable GLMM model, the percentage of urban population (risk ratio (RR) = 1.61, 95% CI 1.25–2.06), volume of international air travel (RR = 1.31, 95% CI 1.04–1.66) and prevalence of HIV (RR = 1.40, 95% CI 1.10–1.78) were risk factors for mortality rate in the first wave (Figure 5.6, Table D.3). Percentage of urban population was included in all models within +2-corrected Akaike information criterion (AICc) scores (Methods); volume of international air travel and HIV prevalence were included in most but not all.

Figure 5.4 Correlation matrix for predictors in the first wave of COVID-19

Positive correlations are displayed in blue and negative correlations in red colour. Colour intensity is proportional to the correlation coefficients.



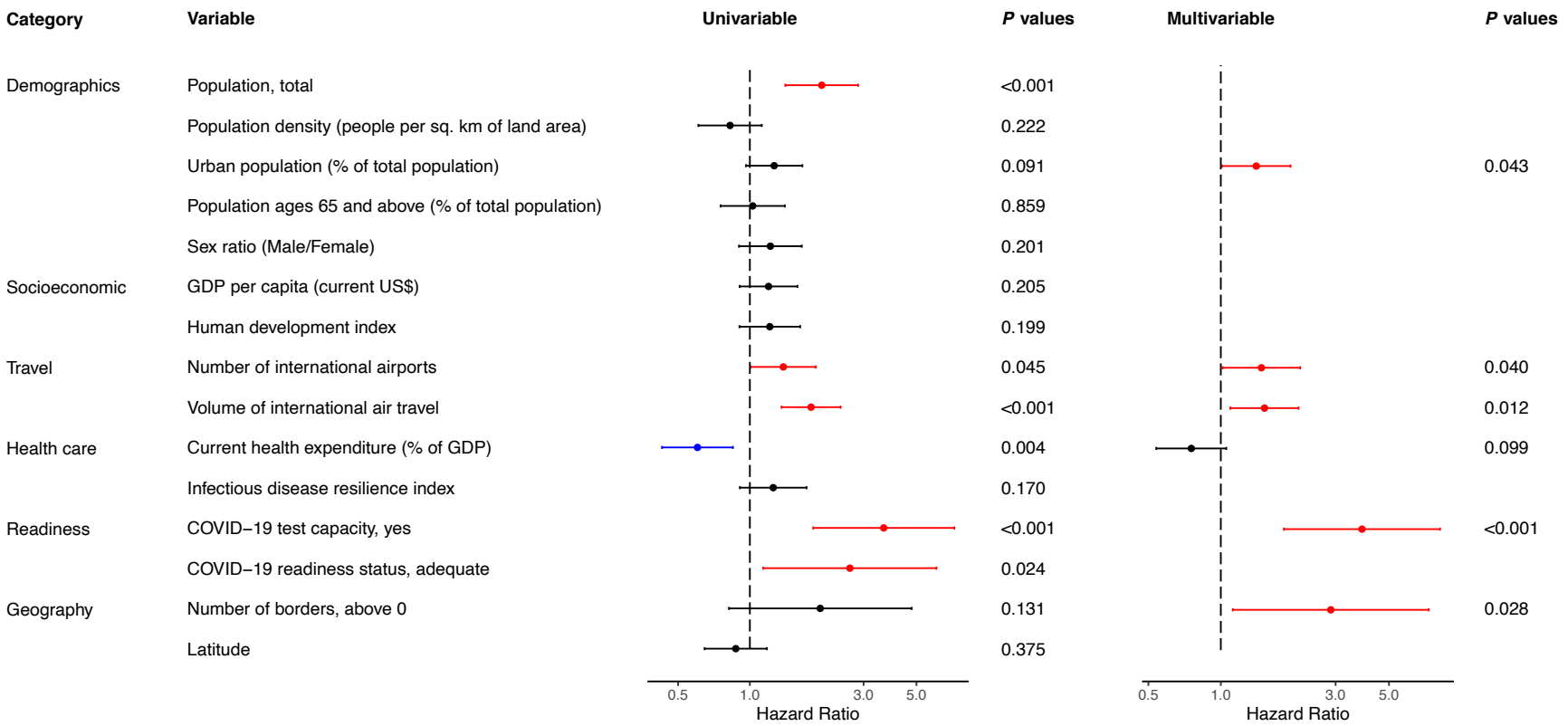


Figure 5.5 Hazard ratios and 95% confidence intervals for predictors of timing of the first COVID-19 case in univariable and multivariable cox regression model

Error bars are shown. Statistically significant risk factors are in red; protective factors are in blue.

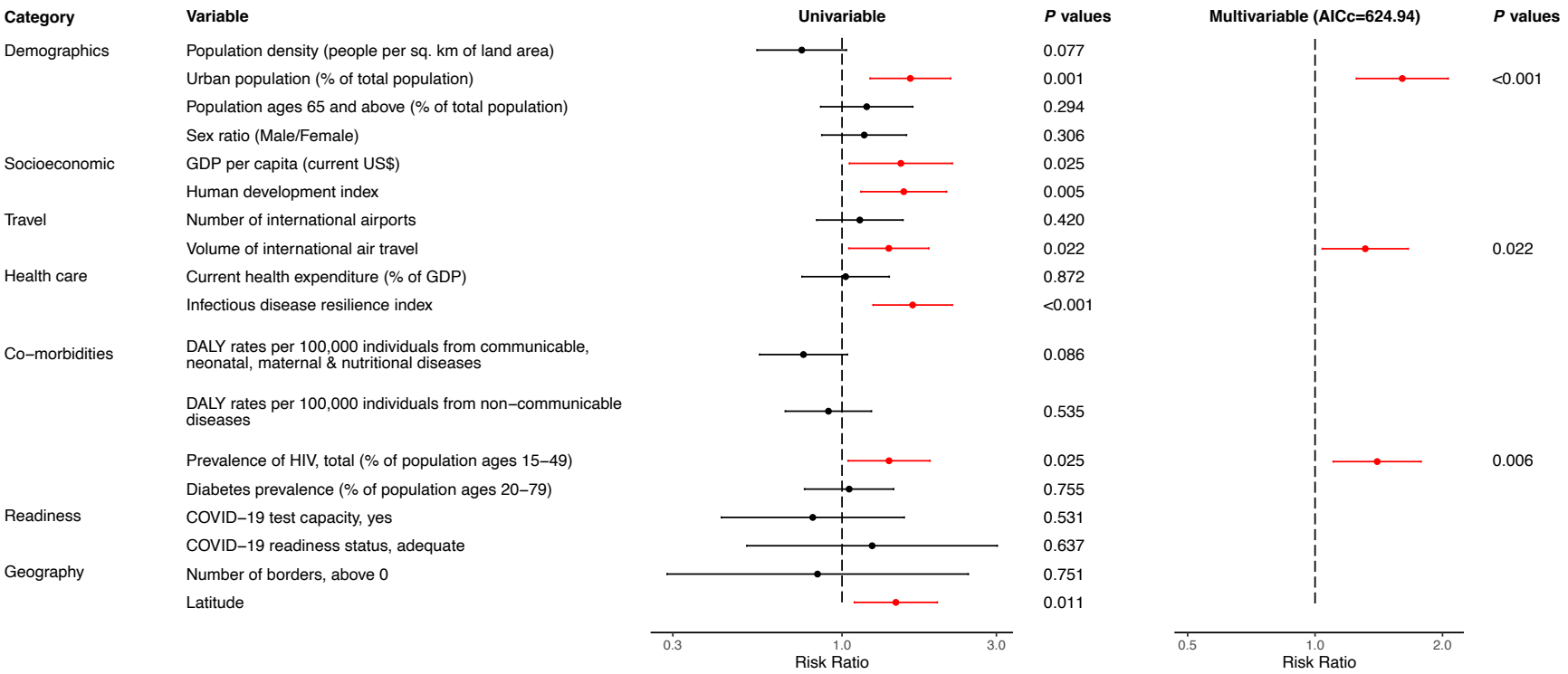


Figure 5.6 Risk ratios and 95% confidence intervals for predictors of per capita COVID-19 mortality in the first wave in univariable and multivariable Poisson generalized linear mixed model

Error bars are shown. Statistically significant risk factors are in red.

None of the predictors in the best multivariable model was correlated with the any of the COVID-19 testing variables (correlation coefficients < 0.6) (Figure 5.7). I then re-ran the best multivariable GLMM models with each additional testing variable (Figure D.1U, Figure D.1W-X). No test variable was associated with the per capita mortality rate and reduced the AICc, and there were no changes in the RRs estimated by the best multivariable model (Figure 5.8, Table D.3).

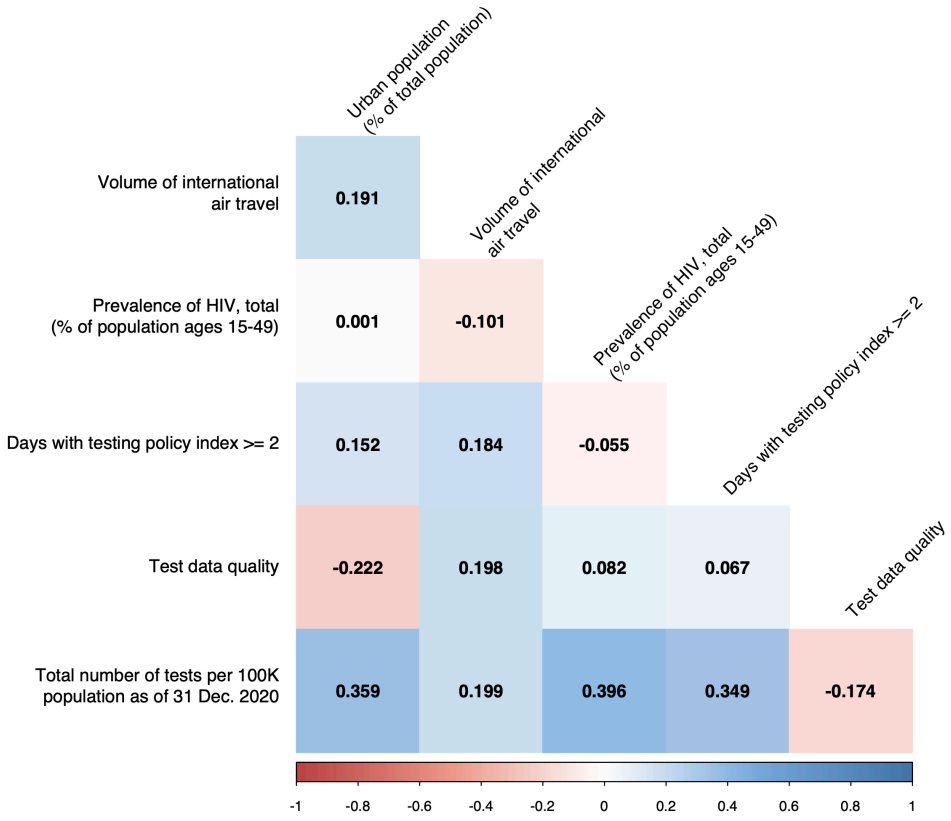


Figure 5.7 Correlation matrix for significant predictors in multivariable model for per capita COVID-19 mortality in the first wave and three test variables

Positive correlations are displayed in blue and negative correlations in red colour. Colour intensity is proportional to the correlation coefficients.

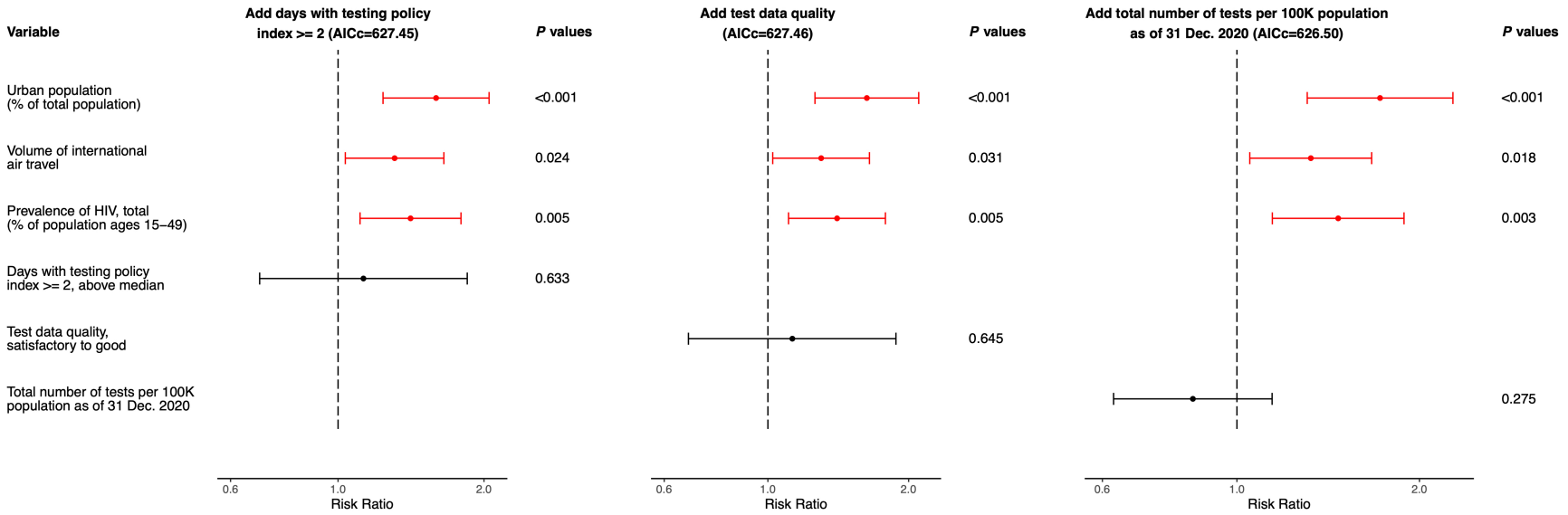


Figure 5.8 Risk ratios and 95% confidence intervals for three test variables for per capita COVID-19 mortality in the first wave in multivariable Poisson generalized linear mixed model

Error bars are shown. Statistically significant risk factors are in red.

There was a good consistency between the stringency index and percent change of residential mobility as indicated by the Google mobility data. After controlling for temporal and random effects, the stringency index was non-linearly associated with the residential mobility (P<0.0001), with an effective degree of freedom of 8.66. The R^2 of the model is 0.77 and the explained deviance is 77.5% (see Supplementary Methods).

None of the predictors in the best multivariable model was correlated with the two stringency scores (correlation coefficients < 0.6) (Figure 5.9). Again, I then re-ran the best multivariable GLMMs, once with each stringency score (Figure D.1Y, Z). No stringency score was associated with the per capita mortality rate, and none reduced the AICc (Figure 5.10, Table D.3). I explored other thresholds of cumulative per capita mortality all produced consistent results.

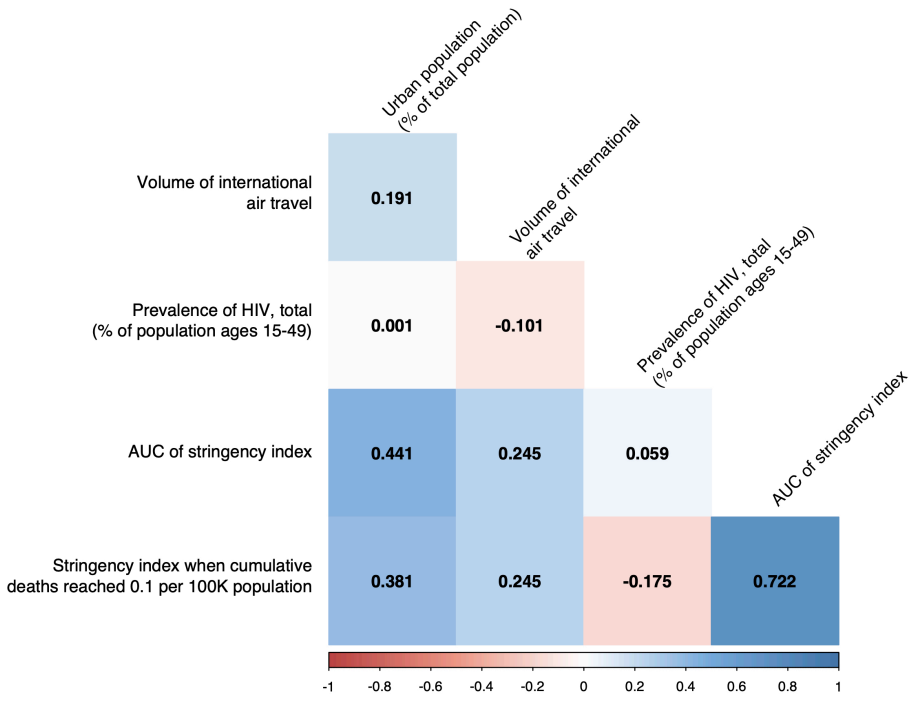


Figure 5.9 Correlation matrix for significant predictors in multivariable model for per capita COVID-19 mortality in the first wave and two stringency indices

Positive correlations are displayed in blue and negative correlations in red colour. Colour intensity is proportional to the correlation coefficients.

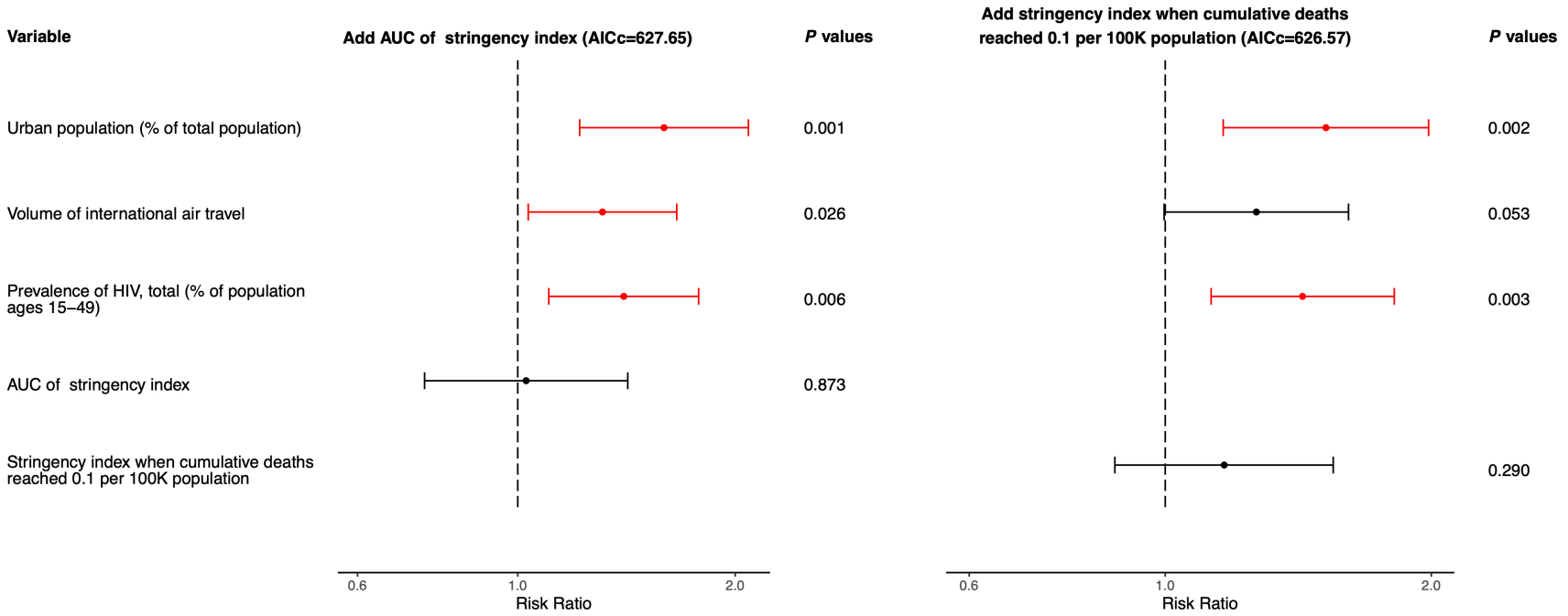


Figure 5.10 Risk ratios and 95% confidence intervals for two stringency indices for per capita COVID-19 mortality in the first wave in multivariable Poisson generalized linear mixed model

Error bars are shown. Statistically significant risk factors are in red.

There were 11, 10, 10, and 11 countries in the categories of high (AUC of stringency index)/high (per capita mortality), high/low, low/high, and low/low, respectively (Figure 5.11A). In the univariable multinomial logistic model, the percentage of urban population, infectious disease resilience index, and human development index were risk factors for one or more categories relative to low/low (Figure 5.12A, Table D.4). In the multivariable multinomial logistic model, the percentage of urban population and infectious disease resilience index were risk factors for high/high, low/high and/or high/low relative to low/low (Figure 5.12B, Table D.5). As above, I also added the three COVID-19 testing predictors into the best multivariable multinomial logistic model and the results remained consistent (Table D.5).

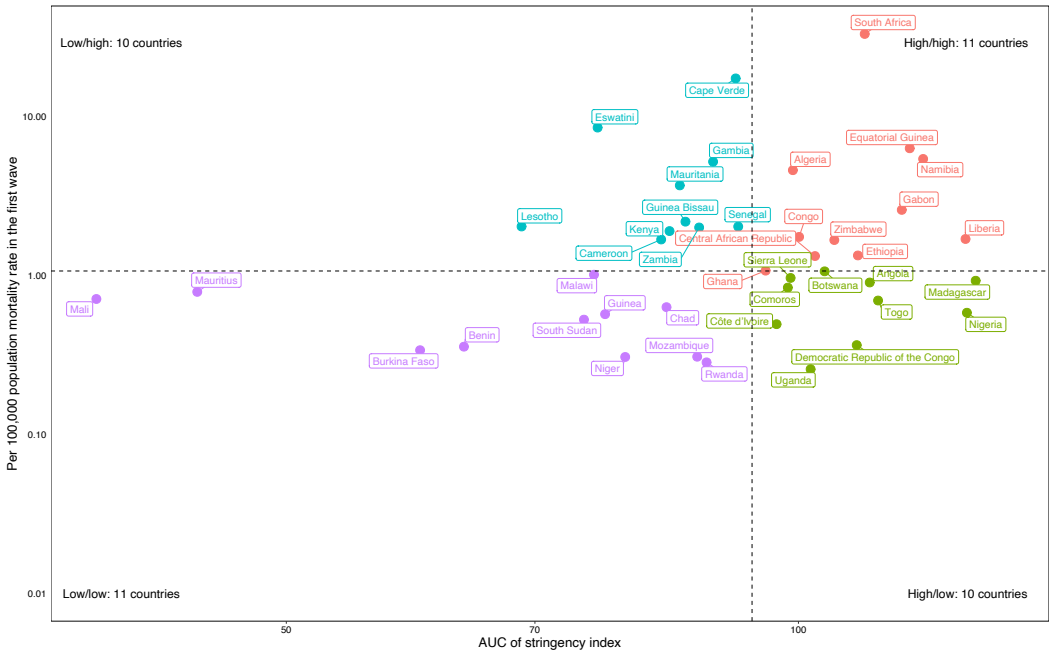
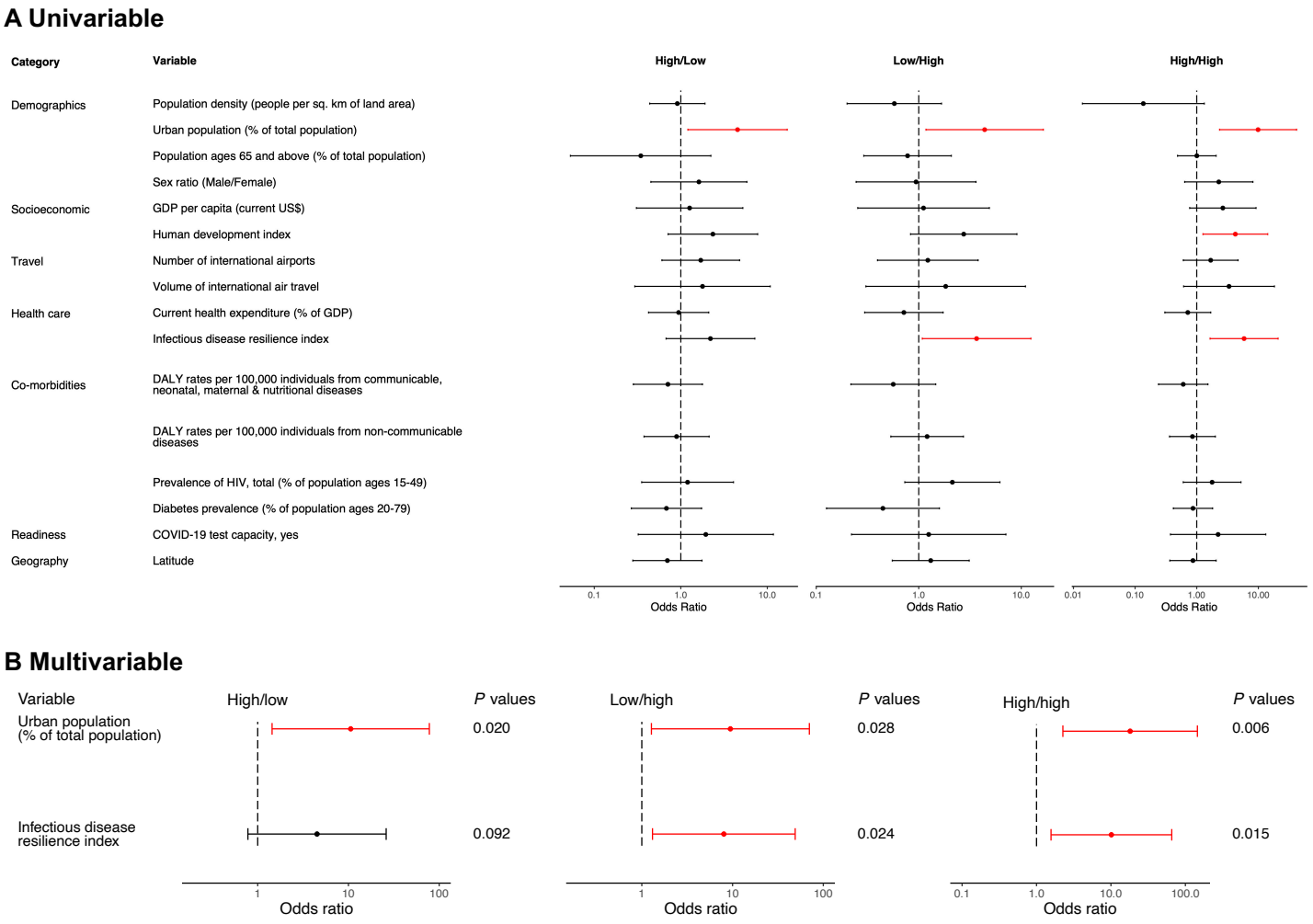


Figure 5.11 Scatter plot of outcome with respect to AUC of stringency index and per capita COVID-19 mortality rate in the first wave

Vertical axis has log10 scale and points falling on the horizontal axis denotes zero deaths.

Figure 5.12 Odds ratios and 95% confidence intervals for predictors of outcome with respect to AUC of stringency index and COVID-19 mortality rate in the first wave in the multinomial logistic regression model

A, Univariable; B, Multivariable. Covid-19 readiness status was excluded from the model because there is no country with adequate Covid-19 readiness status in the reference low/low level and putting it in the model will generate super wide 95%Cis. Error bars are shown. Statistically significant risk factors are in red; protective factors are in blue.



I included 42 countries and 19 predictors (Figure D.1B-G, Figure D.1J-O, R, S, V-Y, and **Error! Reference source not found.**C) in the univariable GLMM model for per capita mortality in the second wave. Consistent with the results for the univariable analysis of the first wave, human development index, infectious disease resilience index, prevalence of HIV, and latitude were risk factors for per capita mortality in the second wave (Figure 5.13, Table D.6). Per capita mortality rate in the first wave was also a risk factor. DALYs per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases was identified as protective factors.

5.5 Discussion

In this chapter, I have identified statistical predictors of the timing of the first case and the per capita mortality rates during the first and second COVID-19 pandemic waves for countries in the WHO African Region. The percentage of urban population, number of international airports, volume of pre-pandemic international air travel, COVID-19 test capacity and number of borders were predictors of the earlier detection of the first case. The percentage of urban population, volume of pre-pandemic international air travel and prevalence of HIV were risk factors for per capita mortality rate in the first pandemic wave. Stringency and timing of government restrictions were not associated with the mortality rate, but countries with higher proportions of urban population and higher infectious disease resilience scores were at increased risk of an adverse outcome, defined as either high AUC of stringency index and/or high per capita mortality. Predictors of per capita mortality rates in the two waves

were broadly consistent, and per capita mortality rate in the first wave was predictive of per capita mortality rate in the second wave.

The association between laboratory capacity to test for COVID-19 cases (evaluated before the detection of COVID-19 in the WHO African Region) and earlier detection of first COVID-19 cases was expected. This result highlights the importance and urgency of ensuring adequate preparedness, especially in the earliest stages of a pandemic, noting that COVID-19 was first detected in Africa over 7 weeks after it was first detected in China (Zhu et al., 2020).

We found that countries with more international airports and a greater volume of pre-pandemic international air travel detected their first COVID-19 cases earlier, and island nations detected their first COVID-19 cases later. Flight connectivity to China was found to be a risk factor for earlier detection of COVID-19, irrespective of their preparedness status as measured by the Global Health Security and Joint External Evaluation scores (Haider et al., 2020), but genome sequencing data suggest that early cases in Africa were mainly imported from Europe and not China (Lu et al., 2021; Ngoi et al., 2021).

Pre-pandemic volume of international air travel also predicts per capita mortality during the first wave. I interpret this as indicating that wider seeding of an epidemic before travel restrictions were imposed (as they were in all countries in my study) resulted in a larger epidemic.

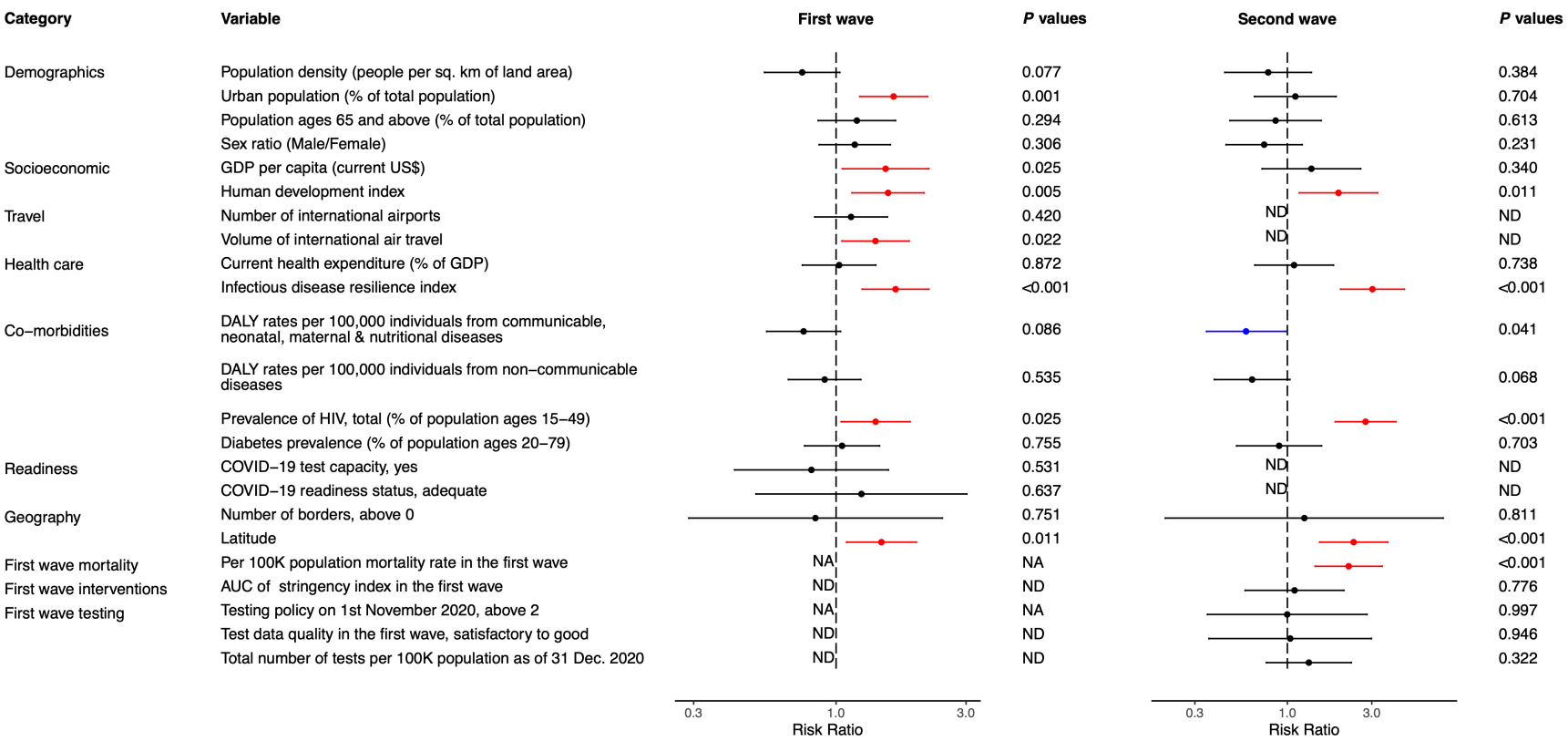


Figure 5.13 Risk ratios and 95% confidence intervals for predictors of per capita COVID-19 mortality in the first wave and second wave in univariable Poisson generalized linear mixed model

Error bars are shown. Statistically significant risk factors are in red; protective factors are in blue. NA, not applicable. ND, not done.

A more urban population predicts both earlier detection of COVID-19 and a higher first wave mortality rate. Urban environments are recognised as risk factors for the transmission of respiratory pathogens in general (Norwegian Institute of Public Health, 2020). Other studies found an association between a more urban population and the number of COVID-19 cases (Li et al., 2020) and that countries with higher socio-economic development Belgium, United Kingdom, and Italy, have higher COVID-19 mortality rates (Chaudhry et al., 2020; Hashim et al., 2020). Countries with a more urban population and greater socio-economic development might have lower COVID-19 case fatality rates (CFRs) (Asfahan et al., 2020; Li et al., 2020). However, my study focused on per capita mortality as CFR is heavily influenced by COVID-19 testing capability, which is highly heterogeneous across countries (Lawal, 2021; Rice et al., 2021; Salyer et al., 2021).

I also found that a higher prevalence of HIV was associated with a higher mortality rate in the first pandemic wave. HIV has been associated with severe COVID-19 during the pandemic; a large population-based study in South Africa found that HIV doubled (HR=2.14) the risk of COVID-19 mortality (Boulle et al., 2020). A meta-analysis of 22 studies worldwide also found that HIV-positive status was associated with an increased risk of COVID-19 mortality (Ssentongo et al., 2021). The underlying reasons might include a high prevalence of comorbidities in patients with HIV and severe COVID-19 and persistent immune suppression in severe COVID-19 (Boulle et al., 2020). In my study, statistical models replacing HIV with other common comorbidities—tuberculosis (which is strongly correlated with HIV), chronic obstructive

pulmonary disease, hypertensive heart disease and obesity—fitted the data less well, although it is possible that HIV status acts as a marker for a basket of these and other comorbidities. Alternatively, any link could be wholly or partially indirect if HIV prevalence is correlated with behavioural, lifestyle or socioeconomic variables not included in my analysis.

I found that stringency and timing of government restrictions were not associated with the mortality rate in the first pandemic wave. Some studies have found that measures including internal 'lockdown' and rapid border closures are not associated with COVID-19 mortality (Chaudhry et al., 2020; Leffler et al., 2020), whereas others found that rapid implementation of restrictions reduced COVID-19 mortality (Fountoulakis et al., 2020). There is a complex cause-and-effect relationship between restrictions and mortality rate, and my results should not be interpreted as demonstrating that restrictions are ineffective, only that any effect is difficult to detect by a retrospective statistical analysis (Salyer et al., 2021; Woolhouse, 2011). This is expected if countries that imposed more stringent restrictions more quickly did so in response to the observed or anticipated severity of their epidemic, and if differences in stringency, at best, only partially mitigated the outcome.

As the response to the pandemic is likely to be damaging in its own right (for example, through negative effects on human well-being, the economy, education and work), an alternative approach is to consider stringency as an outcome variable. The preferred outcome is a low per capita mortality rate and fewer restrictions as measured by the stringency index. Taking this approach,

I found that countries were more likely to achieve a good outcome if they had a less urban population and low infectious disease resilience. Infectious disease resilience is a composite index that considers multiple factors ranging across multiple domains, including political, economic, public health, medical, demographic and disease dynamics (Table 5.1). It is positively correlated to GDP per capita, the human development index, volume of international travel, and prevalence of HIV, and negatively correlated to DALY rates from both communicable diseases and non-communicable diseases (Figure 5.4). This result contradicts speculation that poor countries with a low resilience would be most affected by COVID-19 [see also ref. (Haider et al., 2020)]. In Africa, more urbanized countries and those considered *more* resilient to infectious diseases suffered more from both the direct and indirect impacts of the pandemic. However, this was also true in the United States, which had one of the highest Joint External Evaluation scores (JEE, one score evaluating national capacity to prevent, detect and respond to public health) and preparedness ranking (pre COVID) globally (Stowell & Garfield, 2021).

Similar results for the first and second waves suggest that there were no major shifts in the epidemiology of COVID-19 over the study period, implying no systematic differences in vulnerabilities to the two waves. There was no relationship between stringency of measures taken during the first wave and the severity of the second wave. This indicates that, regardless of the stringency and effectiveness of the government response, intrinsic differences among countries have a substantial effect on the course of national epidemics.

This study has some limitations. It is an observational study of country-level data and cannot demonstrate a direct, causal link between predictors and outcome. Effects due to unmeasured confounders may influence the results and interpretation. Statistical power is limited by sample size, so the final multivariable models include only those predictors with the strongest effects; others might have effect sizes too small to be retained in the models. Given the enormous number of combinations of predictors that could be considered, it is possible that the best fitting models were not identified. Data quality has been raised as an issue (Rice et al., 2021). Some, possibly substantial, under-ascertainment of COVID-19 deaths is likely in Africa, as elsewhere (Institute for Health Metrics and Evaluation, 2021), and could affect my findings if the degree of under-ascertainment was correlated with predictors included in my analysis. I directly addressed this issue by including in my analyses independent estimates of under-reporting of COVID-19 deaths generated by the Institute for Health Metrics and Evaluation (Institute for Health Metrics and Evaluation, 2021). These estimates range up to approximately 75% of COVID-19 deaths unreported (in Burkina Faso, Nigeria, and the Democratic Republic of the Congo). The WHO definition of a COVID-19 death does not require a positive test result, but it is possible that ascertainment is influenced by testing capacity. However, my main results are robust to inclusion of indicators of testing effort in my statistical models, although I note that test volume data were not collected over exactly the same time period.

The stringency variable is a composite index of government policies, reflecting that many countries implemented measures as a package. Not all policies are

expected to have equal impact, and a wide range of combinations of measures was implemented across the region. I validated the stringency index by comparison with Google mobility data. I found a strong association, indicating that the index is related to real-world behaviour by at least a subset of the population. However, the association weakened over time, as has been reported elsewhere (Nouvellet et al., 2021).

My study had several strengths. I considered countries from a single WHO region; these should be more comparable both in terms of data on predictors and COVID-19 epidemiology. I restricted my analysis to outcome variables judged to be most reliably estimated—date of first case and mortality—while correcting for under-reporting/under-ascertainment. The evident plausibility of the results of my date of first case analysis improves confidence that the predictor and outcome data are fitted for purpose.

In conclusion, I have identified risk factors associated with poor direct and indirect outcomes of the first two waves of the COVID-19 pandemic in the WHO African Region countries. My key finding is that countries that were assumed to be better prepared and better equipped to respond to the pandemic were also the most vulnerable to it. These data should be taken into account in future pandemic preparedness planning for WHO African Region countries.

Chapter 6 Concluding remarks

6.1 Summary of findings

In this thesis, I have identified predictors of three attributes of human-infective RNA viruses after emergence: discovery, geographical extent and disappearance. Predictions of hotspots for discovering novel human-infective RNA viruses, the geographical extent the viruses can reach, and the ability to establish sustained infections in humans before or at the initial emergence stage are crucial for controlling the infections they cause. Predictors of post-emergence attributes for one single human-infective RNA virus—SARS-CoV-2 in the WHO African Region were also explored. The main findings are summarised as follows.

1. Globally, GDP growth, GDP, and urbanization were identified as the top predictors of virus discovery. The predicted discovery hotspots included both historical hotspots (i.e. eastern North America, Europe, central Africa, eastern Australia, and north-eastern South America), and new hotspots (i.e. East and Southeast Asia, India, and Central America). Stratified analyses based on transmission mode and transmissibility in humans suggested discovery of vector-borne viruses and strictly zoonotic viruses was more correlated with climatic variables and biodiversity, whereas discovery of non-vector-borne viruses and human-transmissible viruses was strongly correlated with GDP and urbanization.

2. The numbers of human virus species discovered in the United States, China, and Africa up to 2019 were 95, 80 and 107 respectively, with China lagging behind the other two regions. Most virus discovery firsts have been in eastern United States, eastern China, and central and southern Africa. GDP and land use continued to be the top predictors of RNA virus discovery in three different regions—United States, China, and Africa, but climate and biodiversity variables were consistently less important predictors than at a global scale. A map of virus discovery probability in 2010–2019 indicated several new hotspots in central-eastern and southwestern United States, eastern and western China, and northern Africa.

3. Thirty-seven percent (83/223) human-infective RNA virus species have wide geographic distributions extending over three or more continents. Viral features that predicted these wide extents included transmissibility between humans, a +ssRNA genome, narrow host range [i.e. infecting only humans or non-human primates (NHPs)], and a reservoir host in NHPs. For 36% (81/223) viruses, there were no recorded human cases in the past ten years or more. Viruses were more likely to disappear if they were incapable of transmission between humans, had a localised distribution, had a dsRNA genome, were non-pathogenic and non-fatal, were discovered through active discovery programmes rather than passive investigation of the aetiology, and were transmitted by vectors and direct contact.

4. The first countries within the WHO African region to record a case of COVID-19 were more likely to have a high proportion of urban population, high volume

of international air travel and more land borders, and better COVID-19 test capacity. Predictors of a high per capita mortality rate during the first wave were a higher proportion of urban population, higher volume of pre-pandemic international air travel and higher prevalence of HIV. The stringency and timing of government restrictions on behaviour were not associated with a lower per capita mortality rate in the first wave. A higher proportion of urban population and a higher infectious disease resilience score were associated with more stringent restrictions and/or a higher per capita mortality rate in the first wave. The predictor set for the second wave was similar, and mortality per capita in the first wave predicted that in the second wave. These results were not altered when measures of national testing effort were included in the models.

6.2 Public health significance

The findings in this thesis confirmed that the process of virus discovery is related to a series of socio-economic, land use, climate, and biodiversity variables. The new hotspots identified in the thesis for human virus discovery including East and Southeast Asia, India, and Central America are mostly located in developing areas which are generally resource-poor. Attention needs to be directed toward establishing or improving early warning systems for novel viruses in these regions. Discovery process could directly rely on advanced diagnostic technology and the effort to discover novel viruses, or indirectly be accelerated by rapidly locating regions where novel viruses have emerged. Therefore, in addition to the improvement of testing capacity of novel viruses, more intensive monitoring should also focus on the human-animal

interface with a high probability of spill-over, especially in discovery hotspots identified in this thesis where there is a high biodiversity, high temperatures, and a greater level of urbanization.

I also confirmed that the predictors of virus discovery vary between different virus types. Discovery of non-vector-borne viruses and human-transmissible viruses was more strongly correlated with GDP and urbanization than discovery of vector-borne viruses and strictly zoonotic viruses. Predictions showed that transmissible viruses (as opposed to strictly zoonotic viruses) are more likely to be discovered in North America, East Asia, and India, and non-vector-borne viruses (as opposed to vector-borne viruses) are predicted to be more likely to be discovered in North America, Europe, and East Asia. In combination with the findings in chapter 4 that transmissible viruses and/or non-vector-borne viruses are more likely to establish sustained infections in humans with wide geographical extents, North America, Europe and particularly East Asia and India, should be regarded as the priority areas for viral surveillance.

The discovery of human-infective RNA viruses in China lagged behind the United States and Africa, although after 1975 the discovery has increased at a similar rate as the other two regions. In addition to the global hotspots for virus discovery, eastern China and southern Africa were identified as the regional hotspots, and extra surveillance of novel viruses is needed in these regions, especially given the recent discovery of SARS-CoV-2 in Wuhan of China.

Viral features predicting the geographical extents and disappearance of emerging RNA viruses were identified. If the novel virus is transmissible, non-vector-borne, has a (+)ssRNA genome, and a narrow host range (i.e. infecting only humans or NHPs) then there is cause for concern (though viruses from distantly related animals are also harmful). At the earliest emergence of a novel virus, the models can be used to predict its geographical extents and probability of disappearance. The key is to collect as many viral features for the novel virus to use as predictors in the model. The problem is many features including the host range, transmission mode, and taxonomy are mostly unknown at the early stage of emergence. However, the recent advancement of sequencing technology has made this possible. For example, the genome of SARS-CoV-2 and its transmission modes were identified within several weeks after the initial outbreak (Lu et al., 2020).

The proportion of urban population, volume of international air travel, COVID-19 test capacity and number of borders were predictors of the earlier detection of the first COVID-19 case. This result highlights the importance of ensuring adequate preparedness in the earliest stages of a pandemic, including strengthening laboratory capacity for diagnosing and managing cases, screening of incoming passengers at entry points and/or closing borders, and enhancing surveillance measures in urban areas. In addition to the proportion of urban population and volume of international air travel, mortality due to COVID-19 was also exacerbated by high HIV prevalence. More attention should be paid to SARS-CoV-2 co-infections with HIV, particularly in African countries with a higher prevalence of HIV. Negative impacts of COVID-19 (high

mortality and/or high stringency) were greatest in countries with a high proportion of urban populations and a higher infectious disease resilience score, contradicting speculation that poor countries with a low resilience would be more affected by COVID-19. This suggests an urgency to understand what has constituted the current 'vulnerability' to infectious disease in Africa (Moore et al., 2016). Different infectious diseases vary massively in terms of the impact on population health and economy. I suggest generating several vulnerability indices by classifying infectious diseases into different groups, using attributes of pathogens including transmission routes, sources of infections, transmission level. COVID-19 and AIDS that have no known animals as the active source of infections have a higher morbidity/mortality in developed African countries (Ritchie et al., 2020; Roser & Ritchie, 2018), and contradicts the current 'vulnerability' map (Moore et al., 2016). In contrast, Ebola virus disease that are transmitted to people from fruit bats before spreading between humans via direct contact have a high morbidity in less developed regions (World Health Organisation, 2021a). Vector-borne viruses such yellow fever and Zika virus infect humans in central Africa where there is a proper climatic condition for their vectors (Noorbakhsh et al., 2019; Shearer et al., 2018), regardless of the socio-economic level.

6.3 Future study directions

Throughout the thesis, I considered viruses at the species level and did not take into account the variation between subspecies for discovery, geographical extent or disappearance. Manifestations after emergence can be

different between subspecies. One good example is the influenza A virus. Worldwide there are approximately 14 human-infective Influenza A subtypes. They were firstly discovered in the UK [H1N1 in 1933 (Smith et al., 1933)], Singapore [H2N2 in 1957 during the Asia flu pandemic (Assaad et al., 1980)], United States [H7N7 in 1967 (DeLay et al., 1967; Lang et al., 1981) and H7N2 in 2002 (CDC, 2004)], China [three from Hong Kong including H3N2 in 1968 during the Hong Kong flu pandemic (Coleman et al., 1968), H5N1 in 1997 (de Jong et al., 1997), and H9N2 in 1999 (Peiris et al., 1999); three from mainland China including H1N2 in 1992 (Guo et al., 1992), H7N9 (Gao et al., 2013), and H10N8 (To et al., 2014) in 2013; H6N1 from Taiwan in 2013 (Yuan et al., 2013)], Canada ([H7N3 in 2004 (Tweed et al., 2004)], Egypt [H10N7 in 2004 (Pan American Health Organisation, 2004)], and Japan [H5N2 in 2005 (Ogata et al., 2008)]. Of these, only H1N1, H2N2 and H3N2 have led to major pandemics of influenza in humans since 1918 (WHO Regional Office for South-East Asia, 2009), whereas human infections of some viruses such as H10N7 and H9N2 have been reported in restricted areas [Egypt and Australia for H10N7 (Arzey et al., 2012); Egypt, China, and Bangladesh for H9N2 (Hagag et al., 2019)]. Another example is the two SARS coronaviruses—SARS-CoV and SARS-CoV-2—I mentioned several times in the thesis. SARS-CoV was eventually eradicated from humans as of January 2004 by measures of infection control, whereas SARS-CoV-2 has caused the ongoing pandemic with sustained human-to-human infection after emergence in 2019 and does not show any signs of disappearing as of mid-2021 (Petersen et al., 2020). Therefore, to fill the gap additional efforts are needed to widen the database to include

subspecies. Updated analysis with the new database with subspecies will help fully understand the impact of human-infective RNA viruses on population health.

The study in this thesis is retrospective, and in chapter 2 to chapter 5 I did not have complete data for some predictors of each outcome of interest. In chapter 2 and 3, I was able to match RNA virus discovery data with historical climatic variables, population, GDP, and land use data by time, but not with biodiversity, because only static data on biodiversity is available. In chapter 4, a few potential predictors such as the general viral load in patients and the availability of effective vaccination are likely related to geographical extent and disappearance of viruses. In chapter 5, stringency index and tests per capita are two important predictors of COVID-19 mortality. Data on both predictors were collected manually from multiple online resources. However, the data provided from the various sources was not always consistent. For example, COVID-19 testing referred in some instances to the number of tests performed while in other instances it referred to the number of individuals tested. Therefore, future effort to fill these data gaps is crucial.

Another important gap is a universal method to calculate the ascertainment bias for novel pathogens. Ascertainment bias reflects the effort researchers are making to detect new viruses, which can play a major role in discovering a new virus. The spatial distribution of any hotspots will be strongly affected by the reporting effort, with the larger numbers of viruses identified in developed areas with more funding schemes and research projects, as well as locations

where there is more intensive surveillance. Within my model I used GDP and the number of universities as proxies for discovery effort, however, a more accurate index considering research funding (for example) on emerging viruses could be considered. I reviewed and tested previous strategies researchers have used to adjust for ascertainment bias, including frequency of the country listed as the address for authors in scientific papers, and frequency of publications for each pathogen from scientific databases, but the results were not encouraging. For example, Jones et al (Jones et al., 2008) explained the reporting effort by calculating the number of papers published by each country (denoted by the address for every author) in the Journal of Infectious Diseases (JID) since 1973. I tested the efficacy of the method by plotting the relationship between published human-infective RNA virus count and total number of papers from journals which published all human-infective RNA viruses in Web of Science (as of 21 Feb 2018). Both the total number of papers (Figure 6.1A) and total number of papers on viruses (Figure 6.1B) are of little relevance to the published human virus count in my database, though the number of papers has a positive relationship with the number of papers on viruses (Figure 6.1C). I also noted that papers in JID (highlighted in blue in Figure 6.1) may not be able to fully explain the reporting efforts for newly discovered viruses. Olival et al (Olival et al., 2017) improved the method of Jones et al (Jones et al., 2008) by searching the number of publications for each of 586 virus species they have studied using a keyword search by virus name in PubMed and Web of Science. I found the results using this method were similar to that of Jones et al (Jones et al., 2008). Allen et al (Allen et al.,

2017) derived a different index for reporting bias, based on the spatial distribution of place names in peer-reviewed biomedical literature. The disadvantage of this method is it may not represent the reporting effort, because many place names are not related to zoonotic viruses. Other papers have also used occurrence records for a similar class of observations as a surrogate for background sampling effort (Phillips et al., 2009). For example, in my case, I may model the distribution of RNA viruses by utilizing occurrence records of DNA viruses. However, this strategy is not applicable as I was unable to get access to a database on DNA viruses. In future research, more reliable data and methods can be explored and compared.

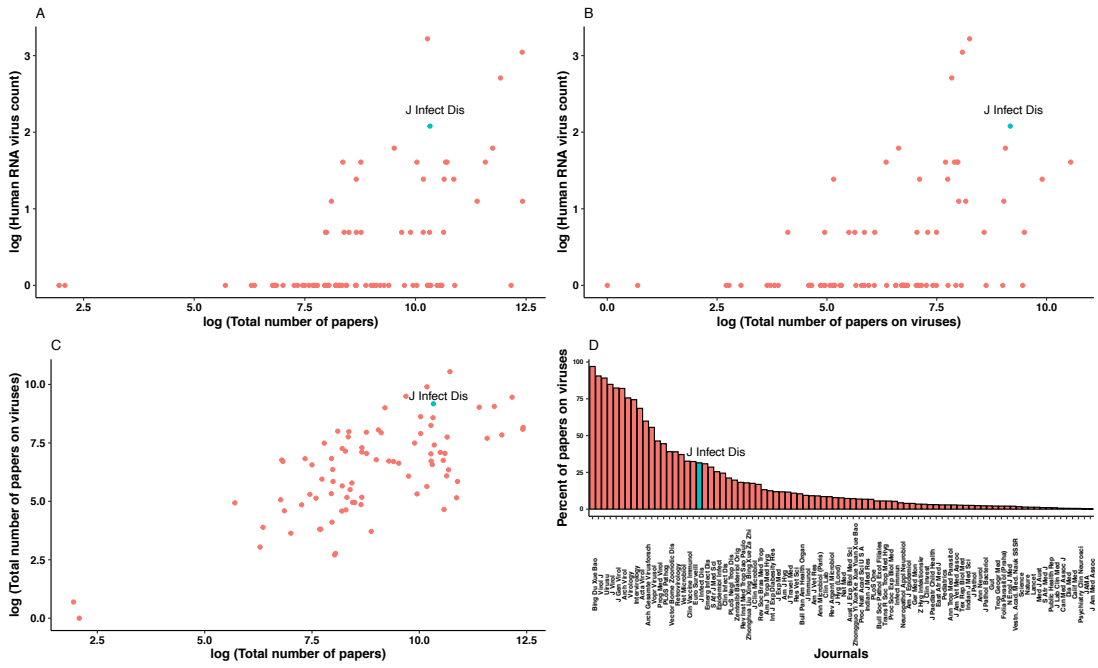


Figure 6.1 Relationship between published human-infective RNA virus count and total number of papers from the journals which published all human-infective RNA viruses in Web of Science

A, total number of papers vs. published human virus count; B, total number of papers on viruses vs. published human virus count; C, total number of papers vs. total number of papers on viruses; D, Percent of papers on viruses in each journal. J Infect Dis (JID) is highlighted in blue.

In chapter 2, I have explored predictors of discovering viruses with different transmissibility and mode of transmission. Future work could focus on other important characteristics, such as clinical presentation, mortality, and host types. Understanding variation of predictors and hotspots for viruses with different characteristics will not only help guide allocation of resources, but also may reveal some unknown mechanisms for virus emergence or discovery. For example, if viruses with high virulence tend to be discovered in a specific area, investigations might reveal certain animal hosts in this area that harbour viruses with high virulence.

In chapter 4, I have only investigated the relationships between viral features and two attributes of human-infective RNA viruses after emergence. However, attributes of the same viruses in their non-human hosts are also important, given these consequences are shaped by viruses' interactions with their hosts. In chapter 4, I found that viruses with a narrow host ranges (i.e. infecting only humans or NHPs) and with a reservoir host in NHPs are more likely to have a wide extent and less likely to disappear, though previous studies suggested NHPS play a peripheral role in the epidemics of human viruses with transmission level 4 (Woolhouse et al., 2016). However, it is worthwhile to investigate which NHP species likely harbour a higher proportion of viruses capable of establish sustained infection in humans with/without wide geographical extent.

6.4. Insights for future research and surveillance on emerging viruses

The hotspot modelling reported in this thesis informs surveillance strategies for the next virus with an epidemic potential. Because zoonotic viruses are responsible for most historical endemics and epidemic diseases, an ambitious aim is to construct a comprehensive data set of unknown viruses with epidemic potential. The subjects of surveillance include the specific animals likely harbour high-risk viruses, humans having a high contacting rate with animals, and the animal-human interfaces. In fact, several projects have started working on the unknown zoonotic viruses, such as the Global Virome project (GVP), the PREDICT project, and the Vietnam Initiative on Zoonotic Infections (VIZIONS) (Carroll et al., 2018; Morse et al., 2012; Rabaa et al., 2015). For example, the VIZIONS project focused on identifying diseases of unknown origin and associated pathogens by using both hospital-based surveillance data and samples from high-risk cohorts (zoonotic infection) in Vietnam (Rabaa et al., 2015). More projects like this should be done in other hotspots in Southern and South-eastern Asia, Africa, and Central and Southern America. Given the prior knowledge on which groups of viruses pose the great risk and where there is the highest risk of spill-over, as well as the advancement of the genome-sequencing techniques (Morse et al., 2012; Olival et al., 2017; Parrish et al., 2008), a nearly full list of zoonotic viruses can be gradually made.

To obtain a clearer picture of zoonotic viruses, scientists also need to try to find more advanced technology to reveal features that facilitate the spill-over

and spread. After years of research to trace the potential animal host of SARS-CoV-1, we know more about coronaviruses in animals, but we were still not prepared for the COVID-19 (Cui et al., 2019; Morens & Fauci, 2020). Changing our behaviour to avoid spill-over is indeed important but understanding what traits of viruses and animals enable the spill-over and spread would get us better prevent the future pandemics. My analyses of geographical extent and viral disappearance can be used for risk assessment at the earliest stage of discovery, but the information necessary for prediction of them is often not available. Viral genomic data might be a good alternative in future studies.

6.5 Conclusions

The results of the studies described in this thesis showed that there are variations in predictors of discovery both between different virus types and geographical regions, and identified high-risk regions for virus discovery beyond their historical extent. The studies also provided proof-of-principle for the prediction of attributes such as mortality, geographical extent, and disappearance for new human-infective RNA viruses. These results help identify priority regions for investment in surveillance systems for new human viruses, and to make risk assessments once they have emerged.

References

Albarino, C. G., Shoemaker, T., Khristova, M. L., Wamala, J. F., Muyembe, J. J., Balinandi, S., . . . Stroher, U. (2013). Genomic analysis of filoviruses associated with four viral hemorrhagic fever outbreaks in Uganda and the Democratic Republic of the Congo in 2012. *Virology*, *442*(2), 97-100.

Allen, T., Murray, K. A., Zambrana-Torrelio, C., Morse, S. S., Rondinini, C., Di Marco, M., . . . Daszak, P. (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nat Commun*, *8*(1), 1124.

Althaus, C. L. (2014). Estimating the Reproduction Number of Ebola Virus (EBOV) During the 2014 Outbreak in West Africa. *PLoS Curr*, *6*.

Anderson, R. M., & May, R. M. (1982). Directly transmitted infections diseases: control by vaccination. *Science*, *215*(4536), 1053-1060.

Aris-Brosou, S., Parent, L., & Ibeh, N. (2019). Viral Long-Term Evolutionary Strategies Favor Stability over Proliferation. *Viruses*, *11*(8).

Arsalan, M., Mubin, O., Alnajjar, F., & Alsinglawi, B. (2020). COVID-19 Global Risk: Expectation vs. Reality. *Int J Environ Res Public Health*, *17*(15).

Arzey, G. G., Kirkland, P. D., Arzey, K. E., Frost, M., Maywood, P., Conaty, S., . . . Selleck, P. (2012). Influenza virus A (H10N7) in chickens and poultry abattoir workers, Australia. *Emerg Infect Dis*, *18*(5), 814-816.

Asfahan, S., Shahul, A., Chawla, G., Dutt, N., Niwas, R., & Gupta, N. (2020). Early trends of socio-economic and health indicators influencing case fatality rate of COVID-19 pandemic. *Monaldi Arch Chest Dis*, *90*(3).

Assaad, F. A., Bres, P., Chu, C. M., Dowdle, W. R., Kendal, A. P., Kilbourne, E. D., . . . Yakhno, M. (1980). A revision of the system of nomenclature for influenza viruses: a WHO memorandum. *Bull World Health Organ*, *58*(4), 585-591.

Babayan, S. A., Orton, R. J., & Streicker, D. G. (2018). Predicting reservoir hosts and arthropod vectors from evolutionary signatures in RNA virus genomes. *Science*, *362*(6414), 577-580.

Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., . . . Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, *220*(4599), 868-871.

Barrett, A. D. T. (2018). The reemergence of yellow fever. *Science*, *361*(6405), 847-848.

Baylis, M. (2017). Potential impact of climate change on emerging vector-borne and other infections in the UK. *Environ Health*, *16*(Suppl 1), 112.

Bergstrom, C. T., McElhany, P., & Real, L. A. (1999). Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. *Proc Natl Acad Sci U S A*, *96*(9), 5095-5100.

Berrios, E. P. (2007). [Foot and mouth disease in human beings. A human case in Chile]. *Rev Chilena Infectol*, *24*(2), 160-163.

Boulle, A., Davies, M. A., Hussey, H., Ismail, M., Morden, E., Vundle, Z., . . . Tamuhla, T. (2020). Risk factors for COVID-19 death in a population cohort study from the Western Cape Province, South Africa. *Clin Infect Dis*.

Brauer, M., Zhao, J. T., Bennitt, F. B., & Stanaway, J. D. (2020). Global Access to Handwashing: Implications for COVID-19 Control in Low-Income Countries. *Environ Health Perspect*, *128*(5), 57005.

Bridger, J. C., Pedley, S., & McCrae, M. A. (1986). Group C rotaviruses in humans. *J Clin Microbiol*, *23*(4), 760-763.

Brierley, L., Pedersen, A. B., & Woolhouse, M. E. J. (2019). Tissue tropism and transmission ecology predict virulence of human RNA viruses. *PLoS Biol*, *17*(11), e3000206.

Brierley, L., Vonhof, M. J., Olival, K. J., Daszak, P., & Jones, K. E. (2016). Quantifying Global Drivers of Zoonotic Bat Viruses: A Process-Based Perspective. *Am Nat*, *187*(2), E53-64.

Buchmeier, M., Adam, E., & Rawls, W. E. (1974). Serological evidence of infection by Pichinde virus among laboratory workers. *Infect Immun*, *9*(5), 821-823.

Carrillo-Larco, R. M., & Castillo-Cara, M. (2020). Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach. *Wellcome Open Res*, *5*, 56.

Carroll, D., Daszak, P., Wolfe, N. D., Gao, G. F., Morel, C. M., Morzaria, S., . . . Mazet, J. A. K. (2018). The Global Virome Project. *Science*, *359*(6378), 872-874.

CDC. (2004). Update: influenza activity--United States, 2003-04 season. *MMWR Morb Mortal Wkly Rep*, *53*(13), 284-287.

Center for International Earth Science Information Network - CIESIN - Columbia University. (2017). *Global Population Count Grid Time Series Estimates* NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4CC0XNV

Chakraborti, S., Maiti, A., Pramanik, S., Sannigrahi, S., Pilla, F., Banerjee, A., & Das, D. N. (2020). Evaluating the plausible application of advanced machine learnings in exploring determinant factors of present pandemic: A case for continent specific COVID-19 analysis. *Sci Total Environ*, 142723.

Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. *Chaos Solitons Fractals*, *135*, 109850.

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., & Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*, *25*, 100464.

Chen, J. (2020). Pathogenicity and transmissibility of 2019-nCoV-A quick overview and comparison with other emerging viruses. *Microbes Infect*, *22*(2), 69-71.

Chen, J., & Lu, H. (2016). Yellow fever in China is still an imported disease. *Biosci Trends*, *10*(2), 158-162.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., . . . Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395-400.

Chini, L. P., Hurtt, G. C., & Frolking, S. (2014). Harmonized Global Land Use for Years 1500 -2100, V1. In: ORNL Distributed Active Archive Center.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290.

Clark, L. E., Mahmutovic, S., Raymond, D. D., Dilanyan, T., Koma, T., Manning, J. T., . . . Abraham, J. (2018). Vaccine-elicited receptor-binding site antibodies neutralize two New World hemorrhagic fever arenaviruses. *Nat Commun*, *9*(1), 1884.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models and applications.* Pion Limited.

Coleman, M. T., Dowdle, W. R., Pereira, H. G., Schild, G. C., & Chang, W. K. (1968). The Hong Kong-68 influenza A2 variant. *Lancet*, *2*(7583), 1384-1386.

Cox, N. J., & Subbarao, K. (2000). Global epidemiology of influenza: past and present. *Annu Rev Med*, *51*, 407-421.

Crase, B., Liedloff, A. C., & Wintle, B. A. (2012). A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography*, *35*(10), 879-888.

Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta Biomed*, *91*(1), 157-160.

Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol*, *17*(3), 181-192.

Dalton, H. R., Bendall, R., Ijaz, S., & Banks, M. (2008). Hepatitis E: an emerging infection in developed countries. *Lancet Infect Dis*, *8*(11), 698-709.

Day, T. (2002). The evolution of virulence in vector-borne and directly transmitted parasites. *Theor Popul Biol*, *62*(2), 199-213.

de Jong, J. C., Claas, E. C., Osterhaus, A. D., Webster, R. G., & Lim, W. L. (1997). A pandemic warning? *Nature*, *389*(6651), 554.

de Roode, J. C., Yates, A. J., & Altizer, S. (2008). Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. *Proc Natl Acad Sci U S A*, *105*(21), 7489-7494.

de Wit, E., van Doremalen, N., Falzarano, D., & Munster, V. J. (2016). SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol*, *14*(8), 523-534.

De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*, *88*(1), 243-251.

DeLay, P. D., Casey, H. L., & Tubiash, H. S. (1967). Comparative study of fowl plague virus and a virus isolated from man. *Public Health Rep*, *82*(7), 615-620.

Delgado, S., Erickson, B. R., Agudo, R., Blair, P. J., Vallejo, E., Albarino, C. G., . . . Nichol, S. T. (2008). Chapare virus, a newly discovered

arenavirus isolated from a fatal hemorrhagic fever case in Bolivia. *PLoS Pathog*, *4*(4), e1000047.

Department of Economic and Social Affairs. (2021). *List of Least Developed Countries*. Retrieved 7 April 2021 from https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/ldc_list.pdf.

Dolan, P. T., Whitfield, Z. J., & Andino, R. (2018). Mapping the Evolutionary Potential of RNA Viruses. *Cell Host Microbe*, *23*(4), 435-446.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *J Anim Ecol*, *77*(4), 802-813.

Emiliano, P. C., Vivanco, M. J. F., & de Menezes, F. S. (2014). Information criteria: How do they behave in different models? *Computational Statistics & Data Analysis*, *69*, 141-153.

Ergönül, Ö., Akyol, M., Tanrıöver, C., Tiemeier, H., Petersen, E., Petrosillo, N., & Gönen, M. (2021). National case fatality rates of the COVID-19 pandemic. *Clin Microbiol Infect*, *27*(1), 118-124.

Escaffre, O., Borisevich, V., & Rockx, B. (2013). Pathogenesis of Hendra and Nipah virus infection in humans. *J Infect Dev Ctries*, *7*(4), 308-311.

Ewald, P. W. (1983). Host-Parasite Relations, Vectors, and the Evolution of Disease Severity. *Annual Review of Ecology and Systematics*, *14*(1), 465-485.

Faria, N. R., Rambaut, A., Suchard, M. A., Baele, G., Bedford, T., Ward, M. J., . . . Lemey, P. (2014). HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*, *346*(6205), 56-61.

Ferrero, D., Ferrer-Orta, C., & Verdaguer, N. (2018). Viral RNA-Dependent RNA Polymerases: A Structural Overview. *Subcell Biochem*, *88*, 39-71.

Fisher, C. R., Streicker, D. G., & Schnell, M. J. (2018). The spread and evolution of rabies virus: conquering new frontiers. *Nat Rev Microbiol*, *16*(4), 241-255.

Fountoulakis, K. N., Fountoulakis, N. K., Koupidis, S. A., & Prezerakos, P. E. (2020). Factors determining different death rates because of the COVID-19 outbreak among countries. *J Public Health (Oxf)*, *42*(4), 681-687.

Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., . . . Hahn, B. H. (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature*, *397*(6718), 436-441.

Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., . . . Shu, Y. (2013). Human infection with a novel avian-origin influenza A (H7N9) virus. *N Engl J Med*, *368*(20), 1888-1897.

Garon, J. R., Cochi, S. L., & Orenstein, W. A. (2015). The Challenge of Global Poliomyelitis Eradication. *Infectious Disease Clinics*, *29*(4), 651-665.

Gatherer, D. (2009). The 2009 H1N1 influenza outbreak in its historical context. *J Clin Virol*, *45*(3), 174-178.

Geddes, A. M. (2006). The history of smallpox. *Clinics in Dermatology*, *24*(3), 152-157.

Geoghegan, J. L., Senior, A. M., Di Giallonardo, F., & Holmes, E. C. (2016). Virological factors that increase the transmissibility of emerging human viruses. *Proc Natl Acad Sci U S A*, *113*(15), 4170-4175.

Giangaspero, M., & Cominardi, P. F. (2006). Parsonage-Turner syndrome associated with anti-bovine viral diarrhoea virus antibodies. *Vet Ital*, *42*(3), 255-259, 249-253.

Goldberger, J., & Anderson, J. F. (1911). The nature of the virus of measles. *Journal of the American Medical Association*, *LVII*(12), 971-972.

Gorbalenya, A. E., Baker, S. C., Baric, R. S., de Groot, R. J., Drosten, C., Gulyaeva, A. A., . . . Coronaviridae Study Group of the International Committee on Taxonomy of, V. (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, *5*(4), 536-544.

Gubler, D. J., Vasilakis, N., & Musso, D. (2017). History and Emergence of Zika Virus. *J Infect Dis*, *216*(suppl_10), S860-S867.

Guo, Y. J., Xu, X. Y., & Cox, N. J. (1992). Human influenza A (H1N2) viruses isolated from China. *J Gen Virol*, *73 (Pt 2)*, 383-387.

Guth, S., Visher, E., Boots, M., & Brook, C. E. (2019). Host phylogenetic distance drives trends in virus virulence and transmissibility across the animal-human interface. *Philos Trans R Soc Lond B Biol Sci*, *374*(1782), 20190296.

Guzman, A., & Isturiz, R. E. (2010). Update on the global spread of dengue. *Int J Antimicrob Agents*, *36 Suppl 1*, S40-42.

Hagag, N. M., Erfan, A. M., El-Husseiny, M., Shalaby, A. G., Saif, M. A., Tawakol, M. M., . . . Naguib, M. M. (2019). Isolation of a Novel Reassortant Highly Pathogenic Avian Influenza (H5N2) Virus in Egypt. *Viruses*, *11*(6).

Haider, N., Yavlinsky, A., Chang, Y. M., Hasan, M. N., Benfield, C., Osman, A. Y., . . . Kock, R. (2020). The Global Health Security index and Joint External Evaluation score for health preparedness are not correlated with countries' COVID-19 detection response time and mortality outcome. *Epidemiol Infect*, *148*, e210.

Hale, T., Webster, S., Petherick, A., Phillips, T., & Kira, B. (2020). *Oxford COVID-19 government response tracker*. Blavatnik School of Government. Retrieved 16 Nov 2020 from https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623-642.

Hashim, M. J., Alsuwaidi, A. R., & Khan, G. (2020). Population Risk Factors for COVID-19 Mortality in 93 Countries. *J Epidemiol Glob Health*, *10*(3), 204-208.

Hassell, J. M., Begon, M., Ward, M. J., & Fevre, E. M. (2017). Urbanization and Disease Emergence: Dynamics at the Wildlife-Livestock-Human Interface. *Trends Ecol Evol*, *32*(1), 55-67.

Hazelton, M. L. (2014). Kernel Smoothing. In N. Balakrishnan, T. Colton, B. Everit, W. Piegorsch, F. Ruggeri, & J. L. Teugels (Eds.), *Wiley StatsRef:*

Statistics Reference Online https://doi.org/https://doi.org/10.1002/9781118445112.stat06538

Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., . . . Ghys, P. D. (2018). Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis. *Lancet Infect Dis*, 1-13.

Hemelaar, J., Gouws, E., Ghys, P. D., & Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *Aids*, *20*(16), W13-23.

Hradsky, O., & Komarek, A. (2021). Demographic and public health characteristics explain large part of variability in COVID-19 mortality across countries. *Eur J Public Health*.

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., . . . Wu, T. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*.

Huremović, D. (2019). Brief History of Pandemics (Pandemics Throughout History). In D. Huremović (Ed.), *Psychiatry of Pandemics: A Mental Health Response to Infection Outbreak* (pp. 7-35). Springer International Publishing. https://doi.org/10.1007/978-3-030-15346-5_2

Institute for Health Metrics and Evaluation. (2021). *Estimation of total mortality due to COVID-19*. Retrieved 16 June 2021 from http://www.healthdata.org/special-analysis/estimation-excess-mortality-due-covid-19-and-scalars-reported-covid-19-deaths

International Committee on Taxonomy of Viruses. (2018). *Virus Taxonomy: The ICTV Report on Virus Classification and Taxon Nomenclature*. Retrieved 1 April 2019 from https://talk.ictvonline.org/ictv-reports/ictv_online_report/

International Union for Conservation of Nature - IUCN, & Center for International Earth Science Information Network - CIESIN - Columbia University. (2015). *Gridded Species Distribution: Global Mammal Richness Grids, 2015 Release* NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4N014G5

Ji, Y., Ma, Z., Peppelenbosch, M. P., & Pan, Q. (2020). Potential association between COVID-19 mortality and health-care resource availability. *Lancet Glob Health*, *8*(4), e480.

Johnson, C. K., Hitchens, P. L., Smiley Evans, T., Goldstein, T., Thomas, K., Clements, A., . . . Mazet, J. K. (2015). Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci Rep*, *5*, 14830.

Johnson, P. T., de Roode, J. C., & Fenton, A. (2015). Why infectious disease research needs community ecology. *Science*, *349*(6252), 1259504.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990-993.

Kandel, N., Chungong, S., Omaar, A., & Xing, J. (2020). Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries. *Lancet*, *395*(10229), 1047-1053.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., . . . Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, *468*(7324), 647-652.

Kile, J. C., Ren, R., Liu, L., Greene, C. M., Roguski, K., Iuliano, A. D., . . . Davis, C. T. (2017). Update: Increase in Human Infections with Novel Asian Lineage Avian Influenza A(H7N9) Viruses During the Fifth Epidemic - China, October 1, 2016-August 7, 2017. *MMWR Morb Mortal Wkly Rep*, *66*(35), 928-932.

Kirigia, J. M., Sambo, H. B., Sambo, L. G., & Barry, S. P. (2009). Economic burden of diabetes mellitus in the WHO African region. *BMC Int Health Hum Rights*, *9*, 6.

Kranjac, A. W., & Kranjac, D. (2020). Decomposing Differences in Coronavirus disease 2019-related Case-Fatality Rates across Seventeen Nations. *Pathog Glob Health*, 1-8.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., . . . Anderson, L. J. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med*, *348*(20), 1953-1966.

Kucharski, A. J., Klepac, P., Conlan, A. J. K., Kissler, S. M., Tang, M. L., Fry, H., . . . group, C. C.-w. (2020). Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Infect Dis*, *20*(10), 1151-1160.

Lang, G., Gagnon, A., & Geraci, J. R. (1981). Isolation of an influenza A virus from seals. *Arch Virol*, *68*(3-4), 189-195.

Lawal, Y. (2021). Africa's low COVID-19 mortality rate: A paradox? *Int J Infect Dis*, *102*, 118-122.

Le Guenno, B., Formenty, P., Wyers, M., Gounon, P., Walker, F., & Boesch, C. (1995). Isolation and partial characterisation of a new strain of Ebola virus. *Lancet*, *345*(8960), 1271-1274.

Lee, G. H., Tan, B. H., Teo, E. C., Lim, S. G., Dan, Y. Y., Wee, A., . . . Teo, C. G. (2016). Chronic Infection With Camelid Hepatitis E Virus in a Liver Transplant Recipient Who Regularly Consumes Camel Meat and Milk. *Gastroenterology*, *150*(2), 355-357.e353.

Lee, H. W., Lee, P. W., & Johnson, K. M. (1978). Isolation of the etiologic agent of Korean Hemorrhagic fever. *J Infect Dis*, *137*(3), 298-308.

Leffler, C. T., Ing, E., Lykins, J. D., Hogan, M. C., McKeown, C. A., & Grzybowski, A. (2020). Association of Country-wide Coronavirus Mortality with Demographics, Testing, Lockdowns, and Public Wearing of Masks. *Am J Trop Med Hyg*, *103*(6), 2400-2411.

Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., . . . Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, *438*(7068), 575-576.

Li, L. M., Grassly, N. C., & Fraser, C. (2014). Genomic analysis of emerging pathogens: methods, application and future trends. *Genome Biol*, *15*(11), 541.

Li, M., Zhang, Z., Cao, W., Liu, Y., Du, B., Chen, C., . . . Wang, X. (2020). Identifying novel factors associated with COVID-19 transmission and fatality using the machine learning approach. *Sci Total Environ*, 142810.

Li, X. D., Qiu, F. X., Yang, H., Rao, Y. N., & Calisher, C. H. (1992). Isolation of Getah virus from mosquitos collected on Hainan Island, China, and results of a serosurvey. *Southeast Asian J Trop Med Public Health*, *23*(4), 730-734.

Lipkin, W. I. (2013). The changing face of pathogen discovery and surveillance. *Nat Rev Microbiol*, *11*(2), 133-141.

Lipkin, W. I., & Firth, C. (2013). Viral surveillance and discovery. *Curr Opin Virol*, *3*(2), 199-204.

Lipsitch, M., Cohen, T., Cooper, B., Robins, J. M., Ma, S., James, L., . . . Murray, M. (2003). Transmission dynamics and control of severe acute respiratory syndrome. *Science*, *300*(5627), 1966-1970.

Lipsitch, M., & Moxon, E. R. (1997). Virulence and transmissibility of pathogens: what is the relationship? *Trends Microbiol*, *5*(1), 31-37.

Liu, W., Sun, F. J., Tong, Y. G., Zhang, S. Q., & Cao, W. C. (2016). Rift Valley fever virus imported into China from Angola. *Lancet Infect Dis*, *16*(11), 1226.

Lone, S. A., & Ahmad, A. (2020). COVID-19 pandemic - an African perspective. *Emerg Microbes Infect*, *9*(1), 1300-1308.

Longdon, B., Hadfield, J. D., Day, J. P., Smith, S. C., McGonigle, J. E., Cogni, R., . . . Jiggins, F. M. (2015). The causes and consequences of changes in virulence following pathogen host shifts. *PLoS Pathog*, *11*(3), e1004728.

Lopman, B. A., Steele, D., Kirkwood, C. D., & Parashar, U. D. (2016). The Vast and Varied Global Burden of Norovirus: Prospects for Prevention and Control. *PLoS Med*, *13*(4), e1001999.

Lu, L., Lycett, S., Ashworth, J., Mutapi, F., & Woolhouse, M. (2021). What are SARS-CoV-2 genomes from the WHO Africa region member states telling us? *BMJ Glob Health*, *6*(1).

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., . . . Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*, *395*(10224), 565-574.

Mackay, I. M., & Arden, K. E. (2015). MERS coronavirus: diagnostics, epidemiology and transmission. *Virol J*, *12*, 222.

Mackey, T. K., Liang, B. A., Cuomo, R., Hafen, R., Brouwer, K. C., & Lee, D. E. (2014). Emerging and reemerging neglected tropical diseases: a review of key characteristics, risk factors, and the policy and innovation environment. *Clin Microbiol Rev*, *27*(4), 949-979.

Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*, *5*(9), 1315-1316.

Mansfield, K. L., Jizhou, L., Phipps, L. P., & Johnson, N. (2017). Emerging Tick-Borne Viruses in the Twenty-First Century. *Front Cell Infect Microbiol*, *7*, 298.

Medeiros de Figueiredo, A., Daponte, A., Moreira Marculino de Figueiredo, D. C., Gil-García, E., & Kalache, A. (2020). [Case fatality rate of COVID-19: absence of epidemiological pattern]. *Gac Sanit*.

Meehan, P. J., Wells, D. L., Paul, W., Buff, E., Lewis, A., Muth, D., . . . Tsai, T. F. (2000). Epidemiological features of and public health response to a St. Louis encephalitis epidemic in Florida, 1990-1. *Epidemiol Infect*, *125*(1), 181-188.

Milholland, M. T., Castro-Arellano, I., Suzan, G., Garcia-Pena, G. E., Lee, T. E., Jr., Rohde, R. E., . . . Mills, J. N. (2018). Global Diversity and Distribution of Hantaviruses and Their Hosts. *Ecohealth*, *15*(1), 163-208.

Miranda, M. E., White, M. E., Dayrit, M. M., Hayes, C. G., Ksiazek, T. G., & Burans, J. P. (1991). Seroepidemiological study of filovirus related to Ebola in the Philippines. *Lancet*, *337*(8738), 425-426.

Mollentze, N., Babayan, S. A., & Streicker, D. G. (2021). Identifying and prioritizing potential human-infecting viruses from their genome sequences. *PLoS Biol*, *19*(9), e3001390.

Mollentze, N., & Streicker, D. G. (2020). Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. *Proc Natl Acad Sci U S A*, *117*(17), 9423-9430.

Moore, M., Gelfeld, B., & Okunogbe, A. (2016). *Identifying future disease hot spots: Infectious Disease Vulnerability Index*. Retrieved 16 Nov 2020 from www.rand.org/pubs/research_reports/RR1605.html

Morens, D. M., & Fauci, A. S. (2020). Emerging Pandemic Diseases: How We Got to COVID-19. *Cell*, *182*(5), 1077-1092.

Morse, S. S. (1995). Factors in the emergence of infectious diseases. *Emerg Infect Dis*, *1*(1), 7-15.

Morse, S. S., Mazet, J. A., Woolhouse, M., Parrish, C. R., Carroll, D., Karesh, W. B., . . . Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet*, *380*(9857), 1956-1965.

Moss, W. J. (2017). Measles. *Lancet*, *390*(10111), 2490-2502.

Moya, A., Holmes, E. C., & Gonzalez-Candelas, F. (2004). The population genetics and evolutionary epidemiology of RNA viruses. *Nat Rev Microbiol*, *2*(4), 279-288.

Murakami, D., & Yamagata, Y. (2016). *Estimation of gridded population and GDP scenarios with spatially explicit statistical downscaling*.

Muyembe-Tamfum, J. J., Mulangu, S., Masumu, J., Kayembe, J. M., Kemp, A., & Paweska, J. T. (2012). Ebola virus outbreaks in Africa: past and present. *Onderstepoort J Vet Res*, *79*(2), 451.

Neiderud, C. J. (2015). How urbanization affects the epidemiology of emerging infectious diseases. *Infect Ecol Epidemiol*, *5*, 27060.

Ngoi, J. M., Quashie, P. K., Morang'a, C. M., Bonney, J. H., Amuzu, D. S., Kumordjie, S., . . . Awandare, G. A. (2021). Genomic analysis of SARS-CoV-2 reveals local viral evolution in Ghana. *Exp Biol Med (Maywood)*, *246*(8), 960-970.

Niklasson, B., Heller, K. E., Schonecker, B., Bildsoe, M., Daniels, T., Hampe, C. S., . . . Lernmark, A. (2003). Development of type 1 diabetes in wild bank voles associated with islet autoantibodies and the novel ljungan virus. *Int J Exp Diabesity Res*, *4*(1), 35-44.

Noorbakhsh, F., Abdolmohammadi, K., Fatahi, Y., Dalili, H., Rasoolinejad, M., Rezaei, F., . . . Nicknam, M. H. (2019). Zika Virus Infection, Basic and Clinical Aspects: A Review Article. *Iran J Public Health*, *48*(1), 20-31.

Norwegian Institute of Public Health. (2020). *Urbanization and preparedness for outbreaks with high-impact respiratory pathogens*. Retrieved 28 February 2021 from https://apps.who.int/gpmb/assets/thematic_papers_2020/tp_2020_4.pdf

Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K. E. C., Baguelin, M., Bhatt, S., . . . Donnelly, C. A. (2021). Reduction in mobility and COVID-19 transmission. *Nat Commun*, *12*(1), 1090.

Ogata, T., Yamazaki, Y., Okabe, N., Nakamura, Y., Tashiro, M., Nagata, N., . . . Kawada, Y. (2008). Human H5N2 avian influenza infection in Japan and the factors associated with high H5N2-neutralizing antibody titer. *J Epidemiol*, *18*(4), 160-166.

Okeahalam, C., Williams, V., & Otwombe, K. (2020). Factors associated with COVID-19 infections and mortality in Africa: a cross-sectional study using publicly available data. *BMJ Open*, *10*(11), e042750.

Olival, K. J., Hosseini, P. R., Zambrana-Torrelio, C., Ross, N., Bogich, T. L., & Daszak, P. (2017). Host and viral traits predict zoonotic spillover from mammals. *Nature*, *546*(7660), 646-650.

Olivieri, A., Palù, G., & Sebastiani, G. (2021). COVID-19 cumulative incidence, intensive care, and mortality in Italian regions compared to selected European countries. *Int J Infect Dis*, *102*, 363-368.

Pan American Health Organisation. (2004). Avian Influenza Virus A (H10N7) Circulating among Humans in Egypt.

Pan, J., St Pierre, J. M., Pickering, T. A., Demirjian, N. L., Fields, B. K. K., Desai, B., & Gholamrezanezhad, A. (2020). Coronavirus Disease 2019 (COVID-19): A Modeling Study of Factors Driving Variation in Case Fatality Rate by Country. *Int J Environ Res Public Health*, *17*(21).

Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E. C., Burke, D. S., Calisher, C. H., . . . Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol Mol Biol Rev*, *72*(3), 457-470.

Patz, J. A., Daszak, P., Tabor, G. M., Aguirre, A. A., Pearl, M., Epstein, J., . . . Disease, E. (2004). Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. *Environ Health Perspect*, *112*(10), 1092-1098.

Peiris, J. S., Lai, S. T., Poon, L. L., Guan, Y., Yam, L. Y., Lim, W., . . . group, S. s. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*, *361*(9366), 1319-1325.

Peiris, M., Yuen, K. Y., Leung, C. W., Chan, K. H., Ip, P. L., Lai, R. W., . . . Shortridge, K. F. (1999). Human infection with influenza H9N2. *Lancet*, *354*(9182), 916-917.

Petersen, E., Koopmans, M., Go, U., Hamer, D. H., Petrosillo, N., Castelli, F., . . . Simonsen, L. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *Lancet Infect Dis*, *20*(9), e238-e244.

Petti, C. A., Polage, C. R., Quinn, T. C., Ronald, A. R., & Sande, M. A. (2006). Laboratory Medicine in Africa: A Barrier to Effective Health Care. *Clinical Infectious Diseases*, *42*(3), 377-382.

Philbey, A. W., Kirkland, P. D., Ross, A. D., Davis, R. J., Gleeson, A. B., Love, R. J., . . . Hyatt, A. D. (1998). An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats. *Emerg Infect Dis*, *4*(2), 269-271.

Phillips, S. J., Dudik, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl*, *19*(1), 181-197.

Piret, J., & Boivin, G. (2020). Pandemics Throughout History. *Front Microbiol*, *11*, 631736.

Pulliam, J. R., & Dushoff, J. (2009). Ability to replicate in the cytoplasm predicts zoonotic transmission of livestock viruses. *J Infect Dis*, *199*(4), 565-568.

Rabaa, M. A., Tue, N. T., Phuc, T. M., Carrique-Mas, J., Saylors, K., Cotten, M., . . . Baker, S. (2015). The Vietnam Initiative on Zoonotic Infections (VIZIONS): A Strategic Approach to Studying Emerging Zoonotic Infectious Diseases. *Ecohealth*, *12*(4), 726-735.

Redding, D. W., Moses, L. M., Cunningham, A. A., Wood, J., & Jones, K. E. (2016). Environmental-mechanistic modelling of the impact of global change on human zoonotic disease emergence: a case study of Lassa fever. *Methods in Ecology and Evolution*, *7*(6), 646-655.

Reed, W., Carroll, J. S., & Agramonte, A. (1901). The etiology of yellow fever: An additional note. *Journal of the American Medical Association*, *XXXVI*(7), 431-440.

Rezza, G., Chen, R., & Weaver, S. C. (2017). O'nyong-nyong fever: a neglected mosquito-borne viral disease. *Pathog Glob Health*, *111*(6), 271-275.

Rice, B. L., Annapragada, A., Baker, R. E., Bruijning, M., Dotse-Gborgbortsi, W., Mensah, K., . . . Metcalf, C. J. E. (2021). Variation in SARS-CoV-2 outbreaks across sub-Saharan Africa. *Nature Medicine*, *27*(3), 447-453.

Richard, M., Knauf, S., Lawrence, P., Mather, A. E., Munster, V. J., Muller, M. A., . . . Kuiken, T. (2017). Factors determining human-to-human transmissibility of zoonotic pathogens via contact. *Curr Opin Virol*, *22*, 7-12.

Riley, S., Fraser, C., Donnelly, C. A., Ghani, A. C., Abu-Raddad, L. J., Hedley, A. J., . . . Anderson, R. M. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science*, *300*(5627), 1961-1966.

Ritchie, H. (2018). *"Urbanization". Published online at OurWorldInData.org.* https://ourworldindata.org/urbanization

Ritchie, H., Ortiz-Ospina, E., Beltekian, D., Mathieu, E., Hasell, J., Macdonald, B., . . . Roser, M. (2020). *Coronavirus Pandemic (COVID-19). Published online at OurWorldInData.org.* Retrieved 16 May 2021 from https://ourworldindata.org/coronavirus

Rizzetto, M., Canese, M. G., Arico, S., Crivelli, O., Trepo, C., Bonino, F., & Verme, G. (1977). Immunofluorescence detection of new antigen-antibody system (delta/anti-delta) associated to hepatitis B virus in liver and in serum of HBsAg carriers. *Gut*, *18*(12), 997-1003.

Robinson, T. P., Wint, G. R., Conchedda, G., Van Boeckel, T. P., Ercoli, V., Palamara, E., . . . Gilbert, M. (2014). Mapping the global distribution of livestock. *PLoS One*, *9*(5), e96084.

Rosenberg, R. (2015). Detecting the emergence of novel, zoonotic viruses pathogenic to humans. *Cell Mol Life Sci*, *72*(6), 1115-1125.

Rosenberg, R., Johansson, M. A., Powers, A. M., & Miller, B. R. (2013). Search strategy has influenced the discovery rate of human viruses. *Proceedings of the National Academy of Sciences*, 201307243.

Roser, M. (2013). *"Economic Growth". Published online at OurWorldInData.org.* https://ourworldindata.org/economic-growth

Roser, M., & Ritchie, H. (2018). *HIV / AIDS. Published online at OurWorldInData.org.* Retrieved 16 May 2021 from https://ourworldindata.org/hiv-aids

Salyer, S. J., Maeda, J., Sembuche, S., Kebede, Y., Tshangela, A., Moussif, M., . . . Nkengasong, J. (2021). The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. *Lancet*.

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustain Cities Soc*, *62*, 102418.

Savage, H. M., Godsey, M. S., Lambert, A., Panella, N. A., Burkhalter, K. L., Harmon, J. R., . . . Nicholson, W. L. (2013). First detection of heartland virus (Bunyaviridae: Phlebovirus) from field collected arthropods. *Am J Trop Med Hyg*, *89*(3), 445-452.

Sharma, S. K., & Mohan, A. (2006). Multidrug-Resistant Tuberculosis: A Menace That Threatens To Destabilize Tuberculosis Control. *Chest*, *130*(1), 261-272.

Shearer, F. M., Longbottom, J., Browne, A. J., Pigott, D. M., Brady, O. J., Kraemer, M. U. G., . . . Golding, N. (2018). Existing and potential infection risk zones of yellow fever worldwide: a modelling analysis. *Lancet Global Health*, *6*(3), E270-E278.

Shu, X., Zang, X., Liu, X., Yang, J., & Wang, J. (2018). Predicting MicroRNA Mediated Gene Regulation between Human and Viruses. *Cells*, *7*(8).

Smith, W., Andrewes, C. H., & Laidlaw, P. P. (1933). A virus obtained from influenza patients. *Lancet*, *222*(5732), 66-68.

Sornette, D., Mearns, E., Schatz, M., Wu, K., & Darcet, D. (2020). Interpreting, analysing and modelling COVID-19 mortality data. *Nonlinear Dyn*, 1-26.

Ssentongo, P., Heilbrunn, E. S., Ssentongo, A. E., Advani, S., Chinchilli, V. M., Nunez, J. J., & Du, P. (2021). Epidemiology and outcomes of COVID-19 in HIV-infected individuals: a systematic review and meta-analysis. *Sci Rep*, *11*(1), 6283.

Stamm, L. V. (2015). Yaws: 110 Years After Castellani's Discovery of Treponema pallidum subspecies pertenue. *The American Journal of Tropical Medicine and Hygiene*, *93*(1), 4-6.

Stowell, D., & Garfield, R. (2021). How can we strengthen the Joint External Evaluation? *BMJ Glob Health*, *6*(5).

Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C. K., Zhou, J., . . . Gao, G. F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol*, *24*(6), 490-502.

Sun, H., Jia, F. J., & Huang, B. C. (2016). Research progress and epidemic situation of the Zika Virus [Article in Chinese]. *Chin J Diagnostics (Electronic Edition)*, *04*(01), 66-69.

Swei, A., Couper, L. I., Coffey, L. L., Kapan, D., & Bennett, S. (2020). Patterns, Drivers, and Challenges of Vector-Borne Disease Emergence. *Vector Borne Zoonotic Dis*, *20*(3), 159-170.

Tang, J. W., Lam, T. T., Zaraket, H., Lipkin, W. I., Drews, S. J., Hatchette, T. F., . . . investigators, I. (2017). Global epidemiology of non-influenza RNA respiratory viruses: data gaps and a growing need for surveillance. *Lancet Infect Dis*, *17*(10), e320-e326.

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. (2001). Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci*, *356*(1411), 983-989.

Terao, Y., Takagi, H., Phan, T. G., Okitsu, S., & Ushijima, H. (2007). Identification of antibody against porcine coronavirus in human milk. *Clin Lab*, *53*(3-4), 129-130.

To, K. K., Tsang, A. K., Chan, J. F., Cheng, V. C., Chen, H., & Yuen, K. Y. (2014). Emergence in China of human disease due to avian influenza A(H10N8)--cause for concern? *J Infect*, *68*(3), 205-215.

Trapido, H., & Sanmartin, C. (1971). Pichinde virus, a new virus of the Tacaribe group from Colombia. *Am J Trop Med Hyg*, *20*(4), 631-641.

Tuanmu, M.-N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, *23*(9), 1031-1045.

Tweed, S. A., Skowronski, D. M., David, S. T., Larder, A., Petric, M., Lees, W., . . . Mak, A. (2004). Human illness from avian influenza H7N3, British Columbia. *Emerg Infect Dis*, *10*(12), 2196-2199.

Tzampoglou, P., & Loukidis, D. (2020). Investigation of the Importance of Climatic Factors in COVID-19 Worldwide Intensity. *Int J Environ Res Public Health*, *17*(21).

Walker, J. W., Han, B. A., Ott, I. M., & Drake, J. M. (2018). Transmissibility of emerging viral zoonoses. *PLoS One*, *13*(11), e0206926.

Walker, P. G. T., Whittaker, C., Watson, O. J., Baguelin, M., Winskill, P., Hamlet, A., . . . Ghani, A. C. (2020). The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science*.

Wang, N., Li, S. Y., Yang, X. L., Huang, H. M., Zhang, Y. J., Guo, H., . . . Shi, Z. L. (2018). Serological Evidence of Bat SARS-Related Coronavirus Infection in Humans, China. *Virol Sin*, *33*(1), 104-107.

Weaver, S. C. (2013). Urbanization and geographic expansion of zoonotic arboviral diseases: mechanisms and potential strategies for prevention. *Trends Microbiol*, *21*(8), 360-363.

Webb, J. L. A., Jr. (2013). Historical epidemiology and infectious disease processes in Africa. *The journal of African History*, *54*(1), 3-10.

WHO Regional Office for South-East Asia. (2009). *Pandemic H1N1 2009*. Retrieved 7 April 2021 from https://apps.who.int/iris/handle/10665/205605

Wolfe, N. D., Dunavan, C. P., & Diamond, J. (2007). Origins of major human infectious diseases. *Nature*, *447*(7142), 279-283.

Wood, C. L., Lafferty, K. D., DeLeo, G., Young, H. S., Hudson, P. J., & Kuris, A. M. (2014). Does biodiversity protect humans against infectious disease? *Ecology*, *95*(4), 817-832.

Woolhouse, M. (2011). How to make predictions about future infectious disease risks. *Philos Trans R Soc Lond B Biol Sci*, *366*(1573), 2045-2054.

Woolhouse, M., & Brierley, L. (2017). *Epidemiological characteristics of human-infective RNA viruses, 1900-2017 [dataset]*. http://dx.doi.org/10.7488/ds/2138

Woolhouse, M., & Brierley, L. (2018). Epidemiological characteristics of human-infective RNA viruses. *Sci Data*, *5*, 180017.

Woolhouse, M., Brierley, L., McCaffery, C., & Lycett, S. (2016). Assessing the Epidemic Potential of RNA and DNA Viruses. *Emerg Infect Dis*, *22*(12), 2037-2044.

Woolhouse, M., & Gaunt, E. (2007). Ecological origins of novel human pathogens. *Crit Rev Microbiol*, *33*(4), 231-242.

Woolhouse, M., Howey, R., Gaunt, E., Reilly, L., Chase-Topping, M., & Savill, N. (2008). Temporal trends in the discovery of human viruses. *Proc Biol Sci*, *275*(1647), 2111-2115.

World Health Organisation. (2015). *Japanese encephalitis report 2015. Geneva: World Health Organization*. Retrieved 1 December 2018 from https://www.who.int/en/news-room/fact-sheets/detail/japanese-encephalitis

World Health Organisation. (2018). *HIV/AIDS report 2018. Geneva: World Health Organization*. Retrieved 19 July 2018 from https://www.who.int/en/news-room/fact-sheets/detail/hiv-aids

World Health Organisation. (2020a). *Coronavirus disease (COVID-2019) situation reports*. Retrieved 26 April 2020 from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation430

World Health Organisation. (2020b). *Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO, 2003-2020*. Retrieved 7 April 2021 from https://www.who.int/influenza/human_animal_interface/2020_MAY_tableH5N1.pdf

World Health Organisation. (2020c). *Global Health Observatory data repository*. Retrieved 12 June 2020 from http:// www.who.int/whosis/data/

World Health Organisation. (2020d). *International guidelines for certificate and classification (coding) of COVID-19 as cause of death*. Retrieved 16 December 2020 from https://www.who.int/classifications/icd/Guidelines_Cause_of_Death_COVID-19.pdf

World Health Organisation. (2020e). *WHO COVID-19: Case Definitions*. Retrieved 16 December 2020 from https://www.who.int/publications/i

World Health Organisation. (2021a). *Ebola virus disease*. Retrieved 16 June 2021 from https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease

World Health Organisation. (2021b). *Nipah virus disease - India*. Retrieved 17 October 2021 from https://www.who.int/emergencies/disease-outbreak-news/item/nipah-virus-disease---india

World Health Organisation. (2021c). *Weekly epidemiological update on COVID-19*. Retrieved 01 June 2021 from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

World Health Organisation. (2021d). *WHO Coronavirus Disease (COVID-19) Dashboard*. Retrieved 01 June 2021 from https://covid19.who.int/table

Yan, Y. S., Zheng, Z. S., Chen, G., Zheng, J., Yan, P. P., & Shao, Y. M. (2000). Confirmation of the first HIV-2 case in China [Article in Chinese]. *Journal of Chinese AIDS&STD Prevention and Control*, *06*(01), 16-18.

Yuan, J., Zhang, L., Kan, X., Jiang, L., Yang, J., Guo, Z., & Ren, Q. (2013). Origin and molecular characteristics of a novel 2013 avian influenza A(H6N1) virus causing human infection in Taiwan. *Clin Infect Dis*, *57*(9), 1367-1368.

Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D., & Fouchier, R. A. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*, *367*(19), 1814-1820.

Zhang, F., Chase-Topping, M., Guo, C. G., van Bunnik, B. A. D., Brierley, L., & Woolhouse, M. E. J. (2020). Global discovery of human-infective RNA viruses: A modelling analysis. *PLoS Pathog*, *16*(11), e1009079.

Zheng, Z., Peng, F., Xu, B., Zhao, J., Liu, H., Peng, J., . . . Tang, W. (2020). Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect*.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., . . . Research, T. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*, *382*(8), 727-733.

Appendix A Supplementary material for chapter 2: Predictors of global discovery of human-infective RNA viruses

A.1 Supplementary Figures

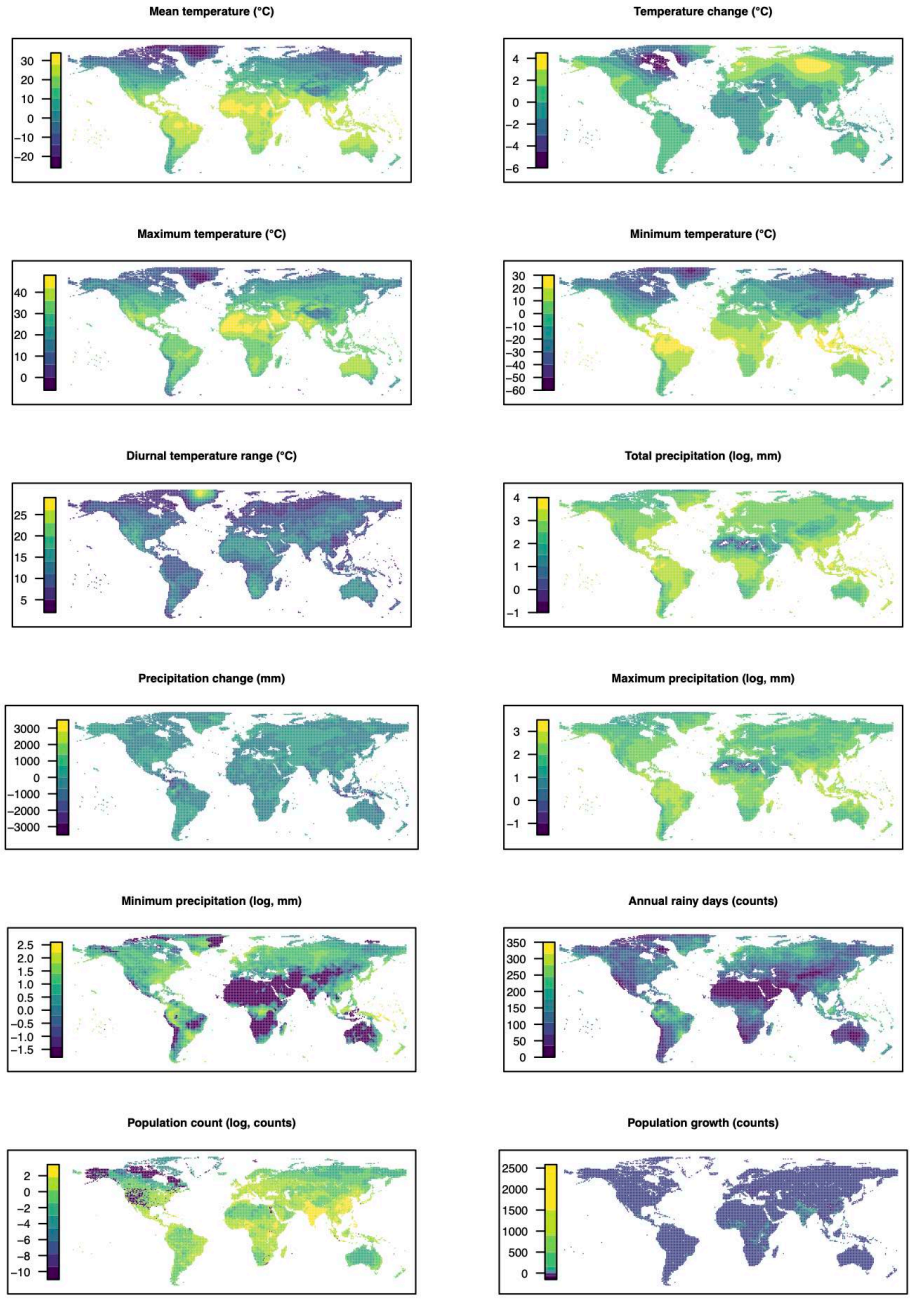


Figure A.1 Global distribution maps of 32 predictors in 2015

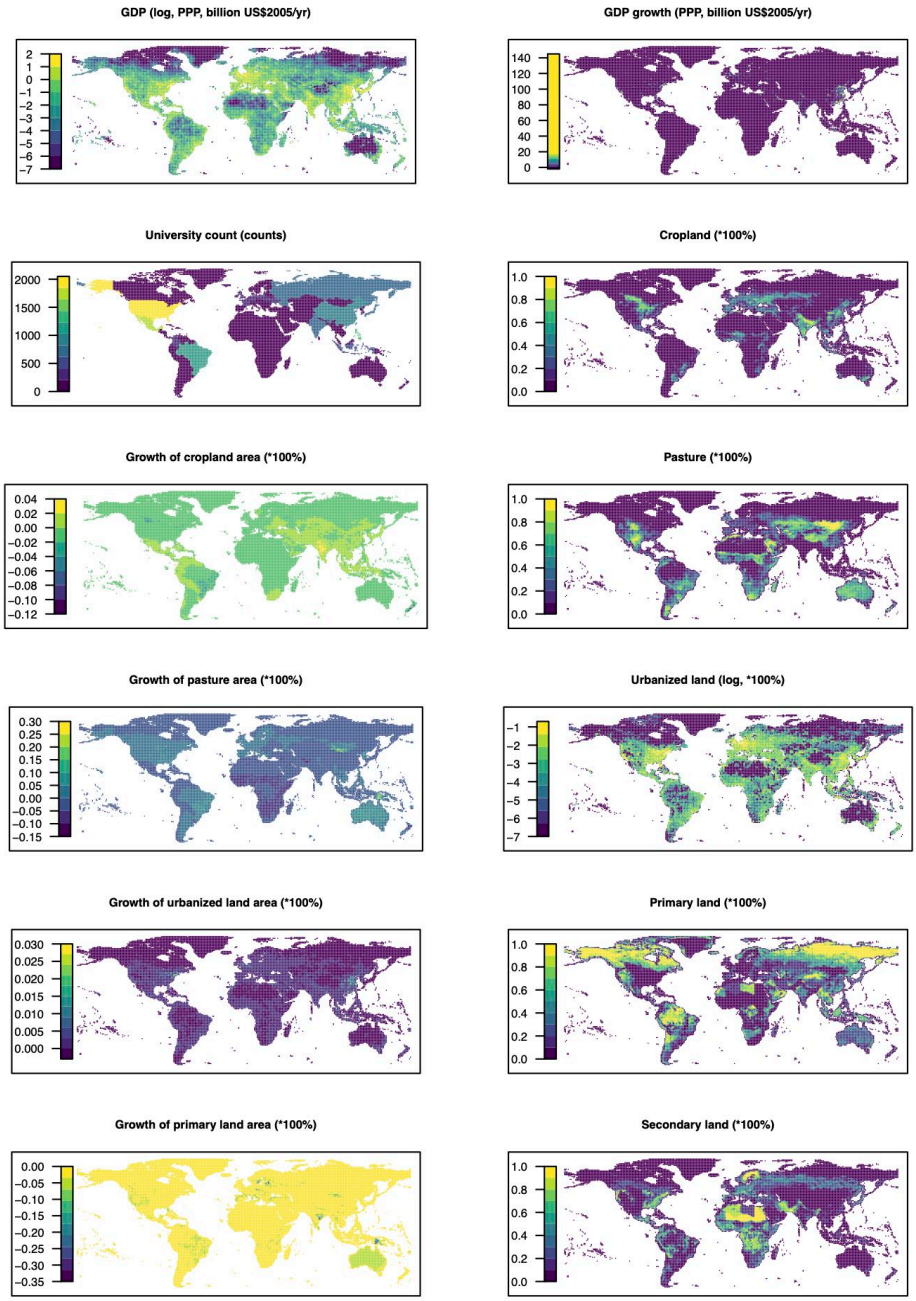


Figure A.1 (continued 1)

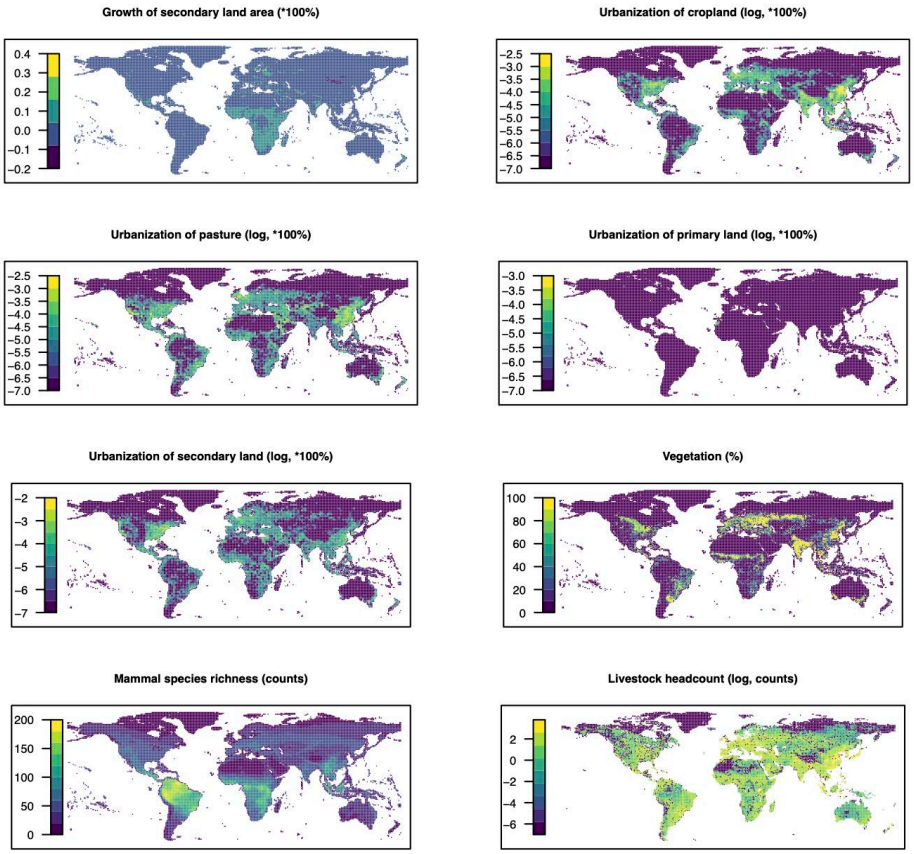


Figure A.1 (continued 2)

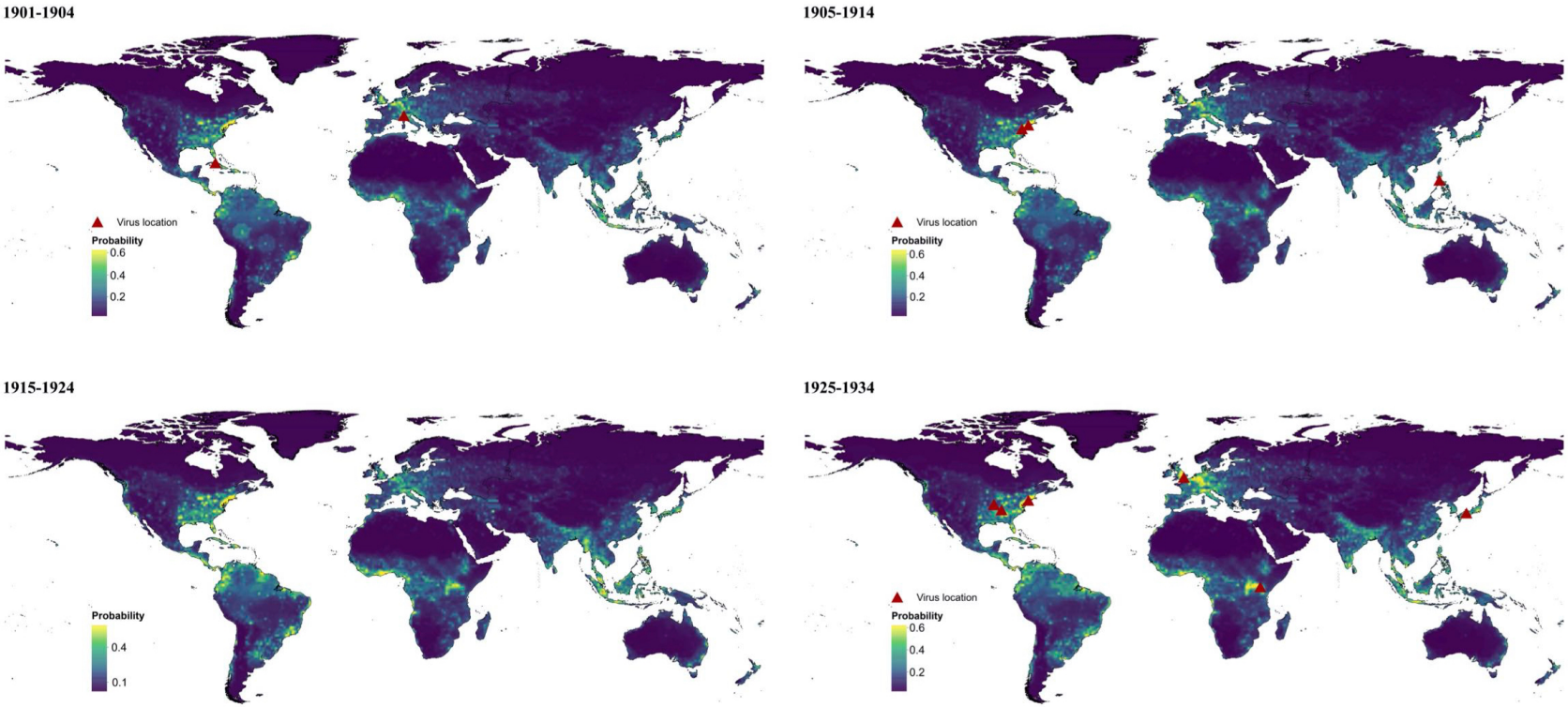


Figure A.2 Historical predicted probability of human-infective RNA virus discovery in the globe by decade (except the first period with four years)

The triangles represented the actual discovery sites in each decade, and the background colour represented the predicted discovery probability.

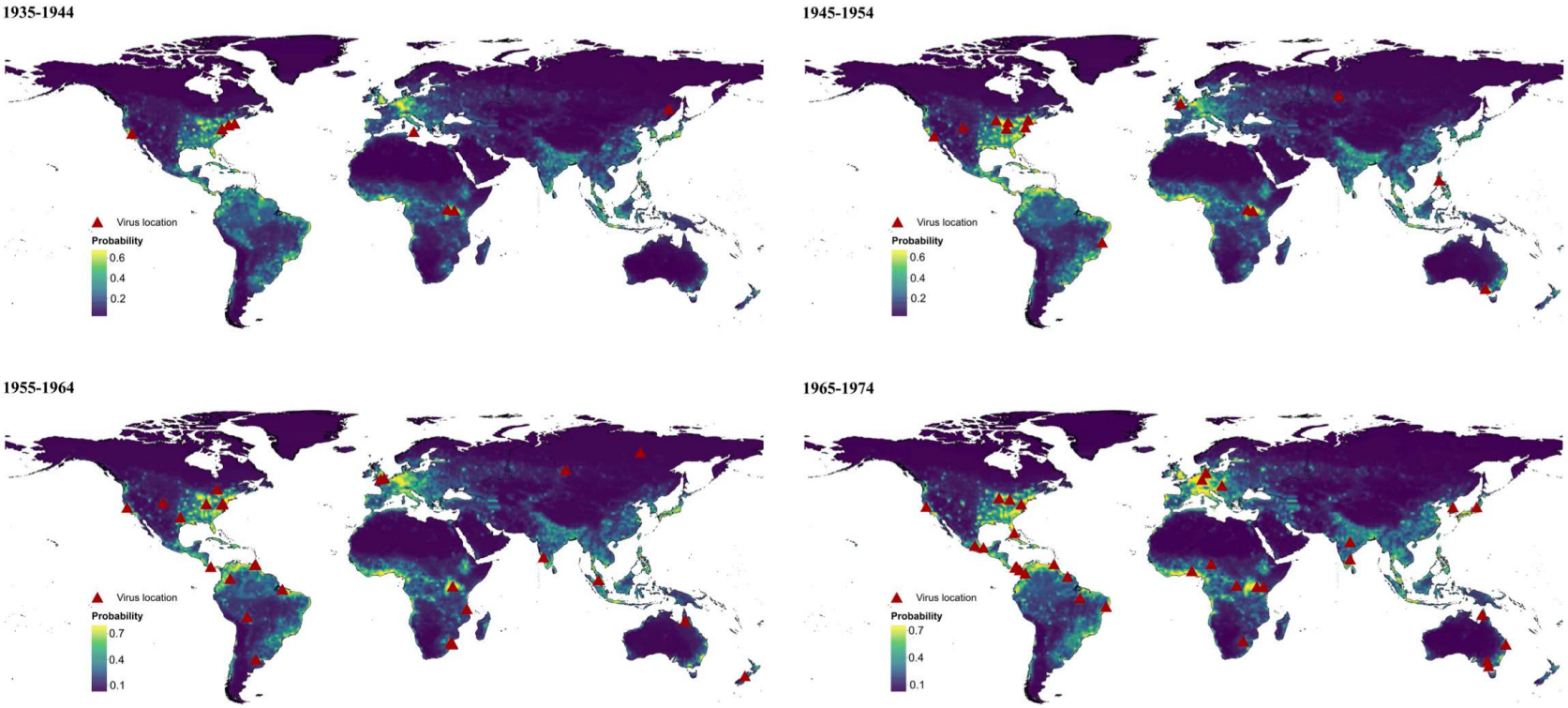


Figure A.2 (continued 1)

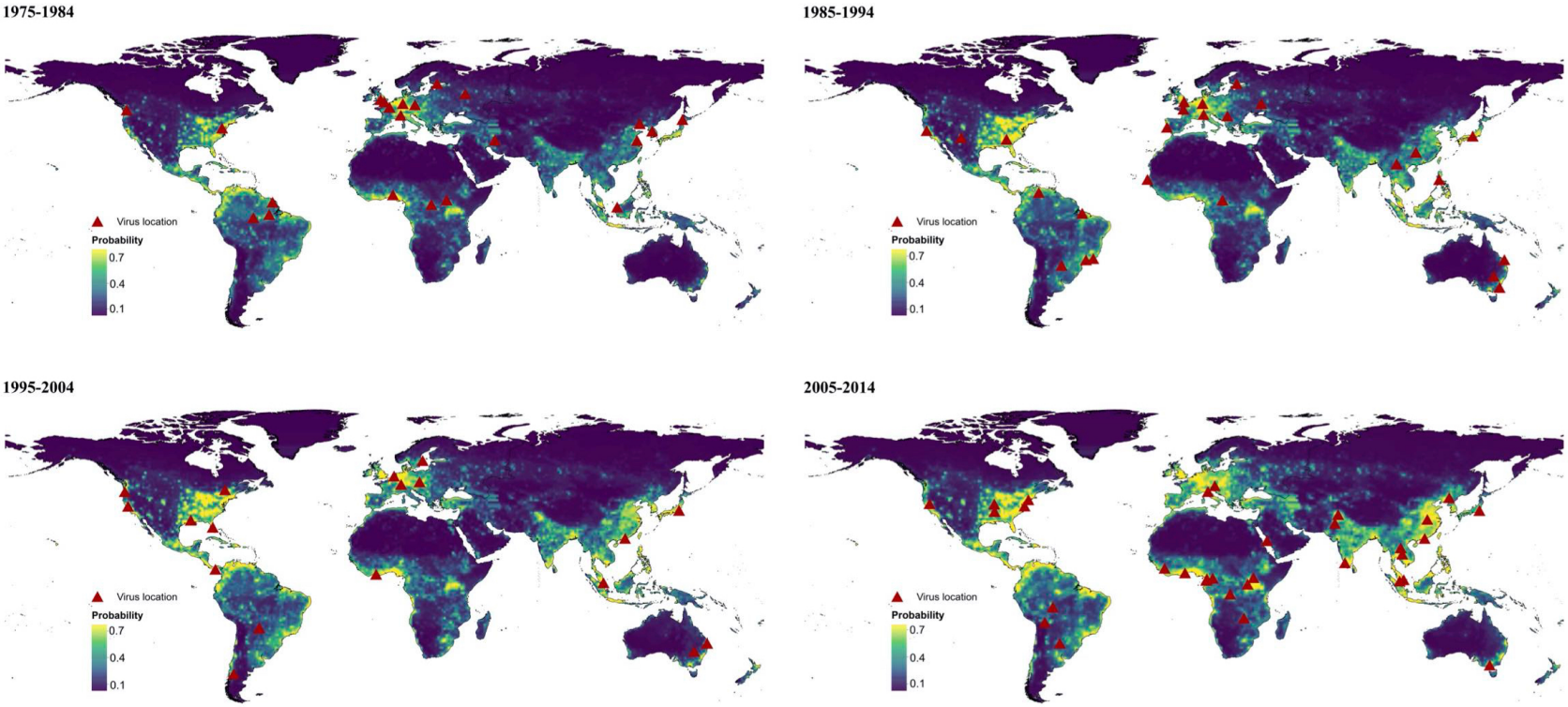


Figure A.2 (continued 2)

A.2 Supplementary Tables

Table A.1 Summary of the global human-infective RNA virus discovery database

Species	Genus	Family	Year	Geographical information						Transmission level	Vector-borne
				Origin	Lat	Long	Type	Level	Location of patient		
Argentinian mammarenavirus [1]	*mammarenavirus*	*arena*	1958	Junín, Buenos Aires Province, Argentina	-34.59	-60.95	city	-	Y	Transmissible in humans	N
Brazilian mammarenavirus [2]	*mammarenavirus*	*arena*	1994	Jardim Sabiá Distrcit, São Paulo city, Brazil	-23.55	-46.63	district	-	Y	Transmissible in humans	N
Cali mammarenavirus [3]	*mammarenavirus*	*arena*	1971	Guatapé, Antioquia Department, Colombia	6.23	-75.16	municipality	3	Y	Strictly zoonotic	N
Chapare mammarenavirus [4]	*mammarenavirus*	*arena*	2008	Samuzabeti town, Chapare Province, Bolivia	-17.41	-66.17	town	-	Y	Strictly zoonotic	N
Guanarito mammarenavirus [5]	*mammarenavirus*	*arena*	1991	Guanarito, Portuguesa State, Venezuela	8.51	-69.00	city	3	Y	Transmissible in humans	N
Lassa mammarenavirus [6]	*mammarenavirus*	*arena*	1970	Lassa, Borno State, Nigeria	10.69	13.27	town	-	Y	Transmissible in humans	N
Lujo mammarenavirus [7]	*mammarenavirus*	*arena*	2009	Lusaka city, Lusaka District, Zambia	-15.39	28.32	city	3	Y	Transmissible in humans	N
Lymphocytic choriomeningitis mammarenavirus [8]	*mammarenavirus*	*arena*	1934	St. Louis county, Missouri, USA	38.73	-90.38	county	3	Y	Transmissible in humans[†]	N
Machupo mammarenavirus [9]	*mammarenavirus*	*arena*	1964	Beni Department, Bolivia	-14.38	-65.1	department	2	Y	Transmissible in humans	N
Mobala mammarenavirus [10]	*mammarenavirus*	*arena*	1985	Bouboui & Gomoka village, Boali town, Central Africa	4.89	18.14	village	-	Y	Strictly zoonotic	N
Whitewater Arroyo mammarenavirus [11]	*mammarenavirus*	*arena*	2000	Alameda County, California, USA	37.6	-121.72	county	3	Y	Strictly zoonotic	N
Mamastrovirus 1 [12]	*mamastrovirus*	*astro*	1975	London, UK	51.51	-0.13	city	3	N	Transmissible in humans	N
Mamastrovirus 6 [13]	*mamastrovirus*	*astro*	2008	Melbourne, Victoria, Australia	-37.81	144.96	city	-	Y	Transmissible in humans	N

Mamastrovirus 8 [14]	*mamastrovirus*	*astro*	2009	St. Louis city, USA	38.63	-90.2	city	-	Y	Transmissible in humans	N
Mamastrovirus 9 [15]	*mamastrovirus*	*astro*	2009	Accomack and Northampton Counties, Virginia, USA	37.71	-75.81	county	3	Y	Transmissible in humans	N
Mammalian 1 orthobornavirus [16]	*orthobornavirus*	*borna*	1985	Gießen, Gießen District, Hesse State, Germany	50.58	8.68	town	-	Y	Strictly zoonotic	N
Mammalian 2 orthobornavirus [17]	*orthobornavirus*	*borna*	2015	Saxony-Anhalt State, Germany	51.95	11.69	state	2	Y	Strictly zoonotic	N
Norwalk virus [18]	*norovirus*	*calici*	1972	Norwalk, Ohio, USA	41.24	-82.62	city	-	Y	Transmissible in humans	N
Sapporo virus [19]	*sapovirus*	*calici*	1980	Sapporo, Hokkaido Prefecture, Japan	43.06	141.35	city	3	Y	Transmissible in humans	N
Vesicular exanthema of swine virus [20]	*vesivirus*	*calici*	1998	Corvallis, Oregon, USA	44.56	-123.26	city	-	N	Strictly zoonotic	N
Alphacoronavirus 1 [21]	*alphacoronavirus*	*corona*	2007	Chiba City, Chiba Prefecture, Japan	35.61	140.11	city	3	Y	Strictly zoonotic	N
Human coronavirus 229E [22]	*alphacoronavirus*	*corona*	1966	Chicago, USA	41.88	-87.63	city	-	Y	Transmissible in humans	N
Human coronavirus NL63 [23]	*alphacoronavirus*	*corona*	2004	Rotterdam, Netherlands	51.92	4.48	city	3	Y	Transmissible in humans	N
Betacoronavirus 1 [24]	*betacoronavirus*	*corona*	1967	Bethesda, Montgomery County, Maryland, USA	38.98	-77.09	CDP	-	Y	Transmissible in humans	N
Human coronavirus HKU1 [25]	*betacoronavirus*	*corona*	2005	Hong Kong, China	22.4	114.11	city	2	Y	Transmissible in humans	N
Middle East respiratory syndrome-related coronavirus [26]	*betacoronavirus*	*corona*	2012	Jeddah, Makkah region, Saudi Arabia	21.29	39.24	city	-	Y	Transmissible in humans	N
Severe acute respiratory syndrome-related coronavirus [27]	*betacoronavirus*	*corona*	2003	Hong Kong, China	22.4	114.11	city	2	Y	Transmissible in humans	N
Human torovirus [28]	*torovirus*	*corona*	1984	Birmingham, UK	52.49	-1.89	city	-	Y	Transmissible in humans	N
Bundibugyo ebolavirus [29]	*ebolavirus*	*filo*	2008	Bundibugyo & Kikyo town, Bwamba county, Uganda	0.71	30.06	town	-	Y	Transmissible in humans	N
Reston ebolavirus [30]	*ebolavirus*	*filo*	1991	Manila, Philippines	14.6	120.98	city	3	Y	Strictly zoonotic	N

Sudan ebolavirus [31]	*ebolavirus*	*filo*	1977	Maridi Town, Maridi County, South Sudan	4.91	29.45	town	-	Y	Transmissible in humans	N
Tai Forest ebolavirus [32]	*ebolavirus*	*filo*	1995	Abidjan, Abidjan department, Côte d'Ivoire	5.36	-4.01	city	-	Y	Strictly zoonotic	N
Zaire ebolavirus [33]	*ebolavirus*	*filo*	1977	Yambuku village, DRC	2.83	22.22	village	-	Y	Transmissible in humans	N
Marburg Marburgvirus [34]	*marburgvirus*	*filo*	1968	Marburg, Germany	50.8	8.77	town	-	Y	Transmissible in humans	N
Aroa virus [35]	*flavivirus*	*flavi*	1971	La Arenosa village, Panama City, Panama	9.04	-79.95	village	-	Y	Strictly zoonotic	Y
Bagaza virus [36]	*flavivirus*	*flavi*	2009	Kerala state, India	10.85	76.27	state	2	Y	Strictly zoonotic	Y
Banzi virus [37]	*flavivirus*	*flavi*	1959	Maponde's Kraal (Usutu river), South Africa	-26.52	31.67	village	-	Y	Strictly zoonotic	Y
Cacipacore virus [38]	*flavivirus*	*flavi*	2011	Theobroma city, State of Rondônia, Brazil	-10.16	-62.36	city	3	Y	Strictly zoonotic	Y
Dengue virus [39]	*flavivirus*	*flavi*	1907	Fort William McKinley, Philippines (now Fort Bonifacio, located in Taguig City, Metro Manila, Philippines)	14.56	121.07	area	-	Y	Transmissible in humans	Y
Edge Hill virus [40]	*flavivirus*	*flavi*	1985	New South Wales, Australia	-31.25	146.92	state	2	Y	Strictly zoonotic	Y
Gadgets Gully virus [41]	*flavivirus*	*flavi*	1991	Heron Island, Great Barrier reef, Australia	-23.44	151.91	area	-	Y	Strictly zoonotic	Y
Ilheus virus [42]	*flavivirus*	*flavi*	1947	Ilheus city, State of Bahia, Brazil	-14.79	-39.05	city	3	Y	Strictly zoonotic	Y
Japanese encephalitis virus [43]	*flavivirus*	*flavi*	1935	Okayama, Okayama Prefecture, Japan	34.66	133.92	city	3	Y	Transmissible in humans[†]	Y
Kokobera virus [40]	*flavivirus*	*flavi*	1964	Mitchell River, Australia; Lockhart River, Australia (Cape York Peninsula)	-16.17	142.84	area	-	Y	Strictly zoonotic	Y
Kyasanur forest disease virus [44]	*flavivirus*	*flavi*	1957	Shimoga District, State of Karnataka, India	13.93	75.57	district	3	Y	Strictly zoonotic	Y
Langat virus [45]	*flavivirus*	*flavi*	1956	Kuala Lumpur, Malaysia	3.14	101.69	city	2	Y	Strictly zoonotic	Y

Louping ill virus [46]	*flavivirus*	*flavi*	1934	New York, New York State, USA	40.71	-74.01	city	-	Y	Strictly zoonotic	Y
Murray Valley encephalitis virus [47]	*flavivirus*	*flavi*	1952	Murray Valley, Northern Victoria, Australia	-36.9	142.96	area	-	Y	Strictly zoonotic	Y
Ntaya virus [48]	*flavivirus*	*flavi*	1952	Bwamba county, Uganda	0.75	30.02	county	3	Y	Strictly zoonotic	Y
Omsk hemorrhagic fever virus [49]	*flavivirus*	*flavi*	1948	Omsk Oblast, Russia	54.99	73.32	federal subject	2	Y	Strictly zoonotic	Y
Powassan virus [50]	*flavivirus*	*flavi*	1959	Powassan, Ontario, Canada	46.08	-79.37	municipality	-	Y	Strictly zoonotic	Y
Rio Bravo virus [51]	*flavivirus*	*flavi*	1962	Dallas city, Texas, USA	32.78	-96.8	city	-	Y	Strictly zoonotic	N
Saint Louis encephalitis virus [52]	*flavivirus*	*flavi*	1933	St. Louis City, USA	38.63	-90.2	city	-	Y	Strictly zoonotic	Y
Tembusu virus [53]	*flavivirus*	*flavi*	1975	Kampong Tijirak village, Kuching Town, Malaysia	1.55	110.36	village	-	Y	Strictly zoonotic	Y
Tick-borne encephalitis virus [54]	*flavivirus*	*flavi*	1938	Khabarovsk Krai, Russia	48.52	135.1	federal subject	2	Y	Transmissible in humans †	Y
Uganda S virus [55]	*flavivirus*	*flavi*	1952	Bwamba county, Bundibugyo District, Uganda	0.75	30.02	county	3	Y	Strictly zoonotic	Y
Usutu virus [56]	*flavivirus*	*flavi*	2009	Emilia Romagna region, Italy	44.6	11.22	region	2	Y	Strictly zoonotic	Y
Wesselsbron virus [57]	*flavivirus*	*flavi*	1957	Lake Simbu region, KwaZulu-Natal Province, South Africa	-27.36	32.32	area	-	Y	Strictly zoonotic	Y
West Nile virus [58]	*flavivirus*	*flavi*	1940	Omogo, West Nile district, Uganda	0.42	33.21	village	-	Y	Transmissible in humans†	Y
Yellow fever virus [59]	*flavivirus*	*flavi*	1901	Quemados town, Cuba	22.79	-80.25	town	-	N*	Transmissible in humans	Y
Zika virus [60]	*flavivirus*	*flavi*	1952	Zika forest, Uganda	0.12	32.53	area	-	Y	Transmissible in humans	Y
Hepacivirus C [61]	*hepacivirus*	*flavi*	1989	Emeryville, California, USA	37.83	-122.29	city	-	N	Transmissible in humans	N
Pegivirus C [62]	*pegivirus*	*flavi*	1995	Yokohama, Kanagawa Prefecture, Japan	35.69	139.69	city	3	Y	Transmissible in humans	N
Pegivirus H [63]	*pegivirus*	*flavi*	2015	New York, New York State, USA	40.71	-74.01	city	-	Y	Transmissible in humans	N

Pestivirus A [64]	*pestivirus*	*flavi*	1988	Milan, Province of Milan, Lombardy region, Italy	45.46	9.19	city	-	Y	Strictly zoonotic	N
Andes orthohantavirus [65]	*orthohantavirus*	*hanta*	1996	El Bolsón, Río Negro Province, Argentina	-41.96	-71.54	town	-	Y	Transmissible in humans	N
Bayou orthohantavirus [66]	*orthohantavirus*	*hanta*	1995	Louisiana state, USA	30.98	-91.96	state	2	Y	Strictly zoonotic	N
Black creek canal orthohantavirus [67]	*orthohantavirus*	*hanta*	1995	Miami-Dade County, Florida, USA	25.76	-80.33	county	3	Y	Strictly zoonotic	N
Choclo orthohantavirus [68]	*orthohantavirus*	*hanta*	2000	Las Tablas Town, Las Tablas District, Panama	7.77	-80.27	town	-	Y	Strictly zoonotic	N
Dobrava-Belgrade orthohantavirus [69]	*orthohantavirus*	*hanta*	1992	Belgrade, Yugoslavia (now Belgrade, Serbia)	44.79	20.45	city	2	Y	Strictly zoonotic	N
Hantaan orthohantavirus [70]	*orthohantavirus*	*hanta*	1978	Seoul, Korea	37.57	126.98	city	2	Y	Strictly zoonotic	N
Laguna Negra orthohantavirus [71]	*orthohantavirus*	*hanta*	1997	Chaco region, Paraguay	-20.09	-59.47	region	2	Y	Strictly zoonotic	N
Puumala orthohantavirus [72]	*orthohantavirus*	*hanta*	1980	Helsinki, Uusimaa region, Finland	60.17	24.94	city	-	Y	Strictly zoonotic	N
Sangassou orthohantavirus [73]	*orthohantavirus*	*hanta*	2010	Sangassou village, Macenta district, Forest Guinea	8.24	-9.32	village	-	Y	Strictly zoonotic	N
Seoul orthohantavirus [74]	*orthohantavirus*	*hanta*	1982	Jiangsu, China	33.14	119.79	province	2	Y	Strictly zoonotic	N
Sin Nombre orthohantavirus [75]	*orthohantavirus*	*hanta*	1993	New Mexico, USA	34.52	-105.87	state	2	Y	Strictly zoonotic	N
Thailand orthohantavirus [76]	*orthohantavirus*	*hanta*	2006	Surin province, Thailand	14.88	103.49	province	2	Y	Strictly zoonotic	N
Thottapalayam orthohantavirus [77]	*orthohantavirus*	*hanta*	2007	Nong khai Province, Thailand	17.88	102.74	province	2	Y	Strictly zoonotic	N
Tula orthohantavirus [78]	*orthohantavirus*	*hanta*	1996	Moravia, Czech Republic	49.2	16.61	region	2	Y	Strictly zoonotic	N
Orthohepevirus A [79]	*orthohepevirus*	*hepe*	1983	Moscow Oblast, Russia (now Moscow city)	55.34	38.29	federal subject	2	N	Transmissible in humans	N
Orthohepevirus C [80]	*orthohepevirus*	*hepe*	2018	Hong Kong, China	22.4	114.11	city	2	Y	Strictly zoonotic	N
Crimean-Congo haemorrhagic fever orthonairovirus [81]	*orthonairovirus*	*nairo*	1967	Kisangani, Tshopo province, DRC	0.53	25.19	city	-	Y	Transmissible in humans	N

Dugbe orthonairovirus [82]	*orthonairovirus*	*nairo*	1969	Ibadan, Nigeria	7.35	3.88	city	-	Y	Strictly zoonotic	Y
Nairobi sheep disease orthonairovirus [83]	*orthonairovirus*	*nairo*	1969	Vellore, Vellore District, Tamil Nadu state, India	12.92	79.13	city	-	Y	Strictly zoonotic	Y
Thiafora orthonairovirus [84]	*orthonairovirus*	*nairo*	1989	Saulges village, Western France	47.98	-0.41	village	-	Y	Strictly zoonotic	Y
Influenza A virus [85]	*Alphainfluenzavirus*	*orthomyxo*	1933	Mill hill, London, UK	51.62	-0.22	city	-	N	Transmissible in humans	N
Influenza B virus [86]	*Betainfluenzavirus*	*orthomyxo*	1940	Irvington village, Greenburgh town, New York, USA	41.03	-73.87	village	-	Y	Transmissible in humans	N
Influenza C virus [87]	*Gammainfluenzavirus*	*orthomyxo*	1950	Ann Arbor city, Michigan state, USA	42.28	-83.74	city	-	Y	Transmissible in humans	N
Dhori thogotovirus [88]	*thogotovirus*	*orthomyxo*	1985	Évora District, Beja District and Portalegre District, Portugal	39.4	-8.22	district	2	Y	Strictly zoonotic	Y
Thogoto thogotovirus [82]	*thogotovirus*	*orthomyxo*	1969	Ibadan, Nigeria	7.35	3.88	city	-	Y	Strictly zoonotic	Y
Avian avulavirus 1 [89]	*avulavirus*	*paramyxo*	1943	Washington, D. C., USA	38.91	-77.04	city	-	Y	Strictly zoonotic	N
Hendra henipavirus [90]	*henipavirus*	*paramyxo*	1995	Hendra, Brisbane, Queensland, Australia	-27.42	153.07	suburb	-	Y	Strictly zoonotic	N
Nipah henipavirus [91]	*henipavirus*	*paramyxo*	1999	Singapore	1.36	103.87	country	1	Y	Transmissible in humans	N
Canine morbillivirus [92]	*morbillivirus*	*paramyxo*	1955	Buffalo, New York, USA	42.89	-78.88	city	-	Y	Strictly zoonotic	N
Measles morbillivirus [93]	*morbillivirus*	*paramyxo*	1911	Washington, D. C., USA	38.91	-77.04	city	-	N	Transmissible in humans	N
Human respirovirus 1 [94]	*respirovirus*	*paramyxo*	1958	Washington, D. C., USA	38.91	-77.04	city	-	Y	Transmissible in humans	N
Human respirovirus 3 [94]	*respirovirus*	*paramyxo*	1958	Washington, D. C., USA	38.91	-77.04	city	-	Y	Transmissible in humans	N
Achimota rubulavirus 2 [95]	*rubulavirus*	*paramyxo*	2013	Volta, Ghana	6.05	0.37	region	2	Y	Strictly zoonotic	N
Human rubulavirus 2 [96]	*rubulavirus*	*paramyxo*	1956	Cincinnati, Ohio, USA	39.1	-84.51	city	-	N	Transmissible in humans	N
Human rubulavirus 4 [97]	*rubulavirus*	*paramyxo*	1960	Bethesda, Montgomery County, Maryland, USA	38.98	-77.09	CDP	-	N	Transmissible in humans	N

Mammalian rubulavirus 5 [98]	*rubulavirus*	*paramyxo*	1959	Stanford, Santa Clara County, California, USA	37.42	-122.17	CDP	-	N	Strictly zoonotic	N
Menangle rubulavirus [99]	*rubulavirus*	*paramyxo*	1998	New South Wales, Australia	-31.25	146.92	state	2	Y	Strictly zoonotic	N
Mumps rubulavirus [100]	*rubulavirus*	*paramyxo*	1934	Nashville, Tennessee, USA	36.16	-86.78	city	-	Y	Transmissible in humans	N
Simian rubulavirus [101]	*rubulavirus*	*paramyxo*	1968	Gosen-shi city, Niigata-ken Prefecture, Japan	37.74	139.18	city	3	Y	Strictly zoonotic	N
Sosuga rubulavirus [102]	*rubulavirus*	*paramyxo*	2014	Area between South Sudan and Uganda	3.76	32.82	area	-	Y	Strictly zoonotic	N
Tioman rubulavirus [103]	*rubulavirus*	*paramyxo*	2007	Tioman Island, Rompin District, Pahang, Malaysia	2.79	104.17	island	-	Y	Strictly zoonotic	N
Bunyamwera orthobunyavirus [104]	*orthobunyavirus*	*peribunya*	1946	Bwamba county, Bundibugyo District, Uganda	0.75	30.02	county	3	Y	Strictly zoonotic	Y
Bwamba orthobunyavirus [105]	*orthobunyavirus*	*peribunya*	1941	Bwamba county, Bundibugyo District, Uganda	0.75	30.02	county	3	Y	Transmissible in humans	Y
California encephalitis orthobunyavirus [106]	*orthobunyavirus*	*peribunya*	1952	Kern county, California, USA	35.49	-118.86	county	3	Y	Strictly zoonotic	Y
Caraparu orthobunyavirus [107]	*orthobunyavirus*	*peribunya*	1961	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Catu orthobunyavirus [107]	*orthobunyavirus*	*peribunya*	1961	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Guama orthobunyavirus [107]	*orthobunyavirus*	*peribunya*	1961	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Guaroa orthobunyavirus [108]	*orthobunyavirus*	*peribunya*	1959	Guaroa, Meta Department, Colombia	3.71	-73.24	municipality	3	Y	Strictly zoonotic	Y
Kairi orthobunyavirus [109]	*orthobunyavirus*	*peribunya*	1967	Brokopondo District, Suriname	4.77	-55.05	district	2	Y	Strictly zoonotic	Y
Madrid orthobunyavirus [110]	*orthobunyavirus*	*peribunya*	1964	Almirante Town, Changuinola District, Panama	9.3	-82.42	town	-	Y	Strictly zoonotic	Y
Marituba orthobunyavirus [107]	*orthobunyavirus*	*peribunya*	1961	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Nyando orthobunyavirus [111]	*orthobunyavirus*	*peribunya*	1965	Kano Plains, near Kisumu, Kisumu County, Kenya	-0.12	34.97	village	-	Y	Strictly zoonotic	Y
Oriboca orthobunyavirus [107]	*orthobunyavirus*	*peribunya*	1961	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y

Oropouche orthobunyavirus [112]	*orthobunyavirus*	*peribunya*	1961	Oropouche, Trinidad and Tobago	10.22	-61.52	island	-	Y	Transmissible in humans	Y
Patois orthobunyavirus [113]	*orthobunyavirus*	*peribunya*	1972	Sontecomapan, Catemaco Municipality, Mexico	18.45	-95.1	village&city	-	Y	Strictly zoonotic	Y
Shuni orthobunyavirus [114]	*orthobunyavirus*	*peribunya*	1975	Ibadan, Nigeria	7.38	3.95	city	-	Y	Strictly zoonotic	Y
Tacaiuma orthobunyavirus [109]	*orthobunyavirus*	*peribunya*	1967	Brokopondo District, Suriname	4.77	-55.05	district	2	Y	Strictly zoonotic	Y
Wyeomyia orthobunyavirus [115]	*orthobunyavirus*	*peribunya*	1965	Darién Province, Panama	7.87	-77.84	province	2	Y	Strictly zoonotic	Y
Candiru phlebovirus [116]	*phlebovirus*	*phenu*	1983	Alenquer city, state of Pará, Brazil	-1.94	-54.73	city	3	Y	Strictly zoonotic	Y
Punta Toro phlebovirus [117]	*phlebovirus*	*phenu*	1970	Jungle area of Panama	9.5	-79.4	area	-	N*	Strictly zoonotic	Y
Rift Valley fever phlebovirus [118]	*phlebovirus*	*phenu*	1931	Rift Valley, Rift Valley Province, Kenya	-0.28	36.07	province	2	Y	Transmissible in humans†	Y
Sandfly fever Naples phlebovirus [119]	*phlebovirus*	*phenu*	1944	Sicily (region level), Italy	37.6	14.02	region	2	Y	Strictly zoonotic	Y
SFTS phlebovirus [120]	*phlebovirus*	*phenu*	2011	Huaiyangshan, China	31.37	115.39	natural region	-	Y	Transmissible in humans	Y
Uukuniemi phlebovirus [121]	*phlebovirus*	*phenu*	1970	Nitra region, Central Slovakia	47.87	18.19	region	2	Y	Strictly zoonotic	Y
Human picobirnavirus [122]	*picobirnavirus*	*picobirna*	1988	Rio de Janerio city, State of Rio de Janerio, Brazil	-22.91	-43.17	city	3	Y	Transmissible in humans	N
Equine rhinitis A virus [123]	*aphthovirus*	*picorna*	1962	Beckenham, Kent, UK	51.41	-0.03	village	-	Y	Strictly zoonotic	N
Foot-and-mouth disease virus [124]	*aphthovirus*	*picorna*	1965	Lübeck, Schleswig-Holstein State, Germany	53.87	10.69	city	3	Y	Strictly zoonotic	N
Cardiovirus A [125]	*cardiovirus*	*picorna*	1947	Manila, Philippines	14.6	120.98	city	3	Y	Strictly zoonotic	N
Cardiovirus B [126]	*cardiovirus*	*picorna*	1963	Town of Vilyuysk, Sakha Republic, Russsia	63.75	121.62	river	-	Y	Transmissible in humans	N
Cosavirus A [127]	*cosavirus*	*picorna*	2008	Mutiple locations in Pakistan	30.38	69.35	mutiple locations#	-	Y	Transmissible in humans	N
Cosavirus B [127]	*cosavirus*	*picorna*	2008	Mutiple locations in Pakistan	30.38	69.35	mutiple locations#	-	Y	Transmissible in humans	N

Cosavirus D [127]	*cosavirus*	*picorna*	2008	Mutiple locations in Pakistan	30.38	69.35	mutiple locations[#]	-	Y	Transmissible in humans	N
Cosavirus E [128]	*cosavirus*	*picorna*	2008	Melbourne, Victoria, Australia	-37.81	144.96	city	-	Y	Transmissible in humans	N
Cosavirus F [129]	*cosavirus*	*picorna*	2012	Islamabad, Pakistan	33.68	73.05	city	-	Y	Transmissible in humans	N
Enterovirus A [130]	*enterovirus*	*picorna*	1949	New York State, USA	43.3	-74.22	state	2	Y	Transmissible in humans	N
Enterovirus B [130]	*enterovirus*	*picorna*	1949	Wilmington, Delaware, USA	39.74	-75.54	city	-	Y	Transmissible in humans	N
Enterovirus C [131]	*enterovirus*	*picorna*	1909	New York, New York State, USA	40.71	-74.01	city	-	Y	Transmissible in humans	N
Enterovirus D [132]	*enterovirus*	*picorna*	1967	Berkeley, California, USA	37.87	-122.27	city	-	N	Transmissible in humans	N
Enterovirus E [133]	*enterovirus*	*picorna*	1961	Denver, Colorado, USA	39.74	-104.99	city	-	Y	Strictly zoonotic	N
Enterovirus H [134]	*enterovirus*	*picorna*	1965	Korea	37.66	127.98	country	1	Y	Strictly zoonotic	N
Rhinovirus A [135]	*enterovirus*	*picorna*	1953	Salisbury, UK	51.07	-1.79	city	3	N*	Transmissible in humans	N
Rhinovirus B [136]	*enterovirus*	*picorna*	1960	Salisbury, UK	51.07	-1.79	city	-	N	Transmissible in humans	N
Rhinovirus C [137]	*enterovirus*	*picorna*	2006	New York, New York State, USA	40.71	-74.01	city	-	Y	Transmissible in humans	N
Erbovirus A [138]	*erbovirus*	*picorna*	2005	Styria, Austria	47.36	14.47	state	2	Y	Strictly zoonotic	N
Hepatovirus A [139]	*hepatovirus*	*picorna*	1973	Bethesda, Montgomery County, Maryland, USA	38.98	-77.09	CDP	-	N	Transmissible in humans	N
Aichivirus A [140]	*kobuvirus*	*picorna*	1991	Aichi Prefecture, Japan	35.18	136.91	prefecture	2	Y	Transmissible in humans	N
Parechovirus A [141]	*parechovirus*	*picorna*	1958	Cincinnati, Ohio, USA	39.1	-84.51	city	-	Y	Transmissible in humans	N
Parechovirus B [142]	*parechovirus*	*picorna*	2003	Stockholm, Sweden	59.33	18.07	city	3	Y	Strictly zoonotic	N
Salivirus A [143]	*salivirus*	*picorna*	2009	Northern California, USA	38.84	-120.9	state	2	Y	Transmissible in humans	N

Avian metapneumovirus [144]	*metapneumovirus*	*pneumo*	2011	Memphis, Tennessee, USA	35.15	-90.05	city	-	Y	Strictly zoonotic	N
Human metapneumovirus [145]	*metapneumovirus*	*pneumo*	2001	Rotterdam, Netherlands	51.92	4.48	city	3	Y	Transmissible in humans	N
Human orthopneumovirus [146]	*orthopneumovirus*	*pneumo*	1957	Baltimore, Maryland, USA	39.29	-76.61	city	-	N	Transmissible in humans	N
Colorado tick fever virus [147]	*coltivirus*	*reo*	1946	Denver, Colorado, USA	39.74	-104.99	city	-	N	Transmissible in humans[†]	Y
Eyach virus [148]	*coltivirus*	*reo*	1980	Prague, Czech Republic	50.08	14.44	city	2	Y	Strictly zoonotic	Y
Corriparta virus [149]	*orbivirus*	*reo*	1967	Aurukun, North Queensland, Australia	-13.36	141.73	town	-	Y	Strictly zoonotic	Y
Great Island virus [150]	*orbivirus*	*reo*	1963	Kemerovo District, Kemerovo Oblast, Russia	55.35	86.06	district	3	Y	Strictly zoonotic	Y
Lebombo virus [114]	*orbivirus*	*reo*	1975	Ibadan, Nigeria	7.38	3.95	city	-	Y	Strictly zoonotic	Y
Orungo virus [114]	*orbivirus*	*reo*	1976	Ibadan, Nigeria	7.38	3.95	city	-	Y	Strictly zoonotic	Y
Mammalian orthoreovirus [151]	*orthoreovirus*	*reo*	1954	Cincinnati, Ohio, USA	39.1	-84.51	city	-	Y	Transmissible in humans	N
Nelson Bay orthoreovirus [152]	*orthoreovirus*	*reo*	2007	Melaka state, Malaysia	2.19	102.25	state	2	Y	Transmissible in humans	N
Rotavirus A [153]	*rotavirus*	*reo*	1973	Parkville, Melbourne, Victoria, Australia	-37.81	144.96	village	-	Y	Transmissible in humans	N
Rotavirus B [154]	*rotavirus*	*reo*	1984	Jinzhou, China	41.1	121.13	city	3	Y	Transmissible in humans	N
Rotavirus C [155]	*rotavirus*	*reo*	1986	London, United Kingdom	51.51	-0.13	city	3	Y	Transmissible in humans	N
Rotavirus H [156]	*rotavirus*	*reo*	1987	Huaihua city, Hunan Province, China	27.55	109.96	city	3	Y	Transmissible in humans	N
Banna virus [157]	*seadornavirus*	*reo*	1990	Xishuangbanna, Yunnan Province, China	22.01	100.8	city	3	Y	Strictly zoonotic	Y
Primate T-lymphotropic virus 1 [158]	*deltaretrovirus*	*retro*	1980	Bethesda, Montgomery County, Maryland, USA	38.98	-77.09	CDP	-	Y	Transmissible in humans	N
Primate T-lymphotropic virus 2 [159]	*deltaretrovirus*	*retro*	1982	Seattle, Washington, USA	47.61	-122.33	city	-	Y	Transmissible in humans	N

Primate T-lymphotropic virus 3 [160]	*deltaretrovirus*	*retro*	2005	Remote settlement in the Océan department, South Province, Cameroon	2.5	10.5	area[#]	-	Y	Strictly zoonotic	N
Human immunodeficiency virus 1 [161]	*lentivirus*	*retro*	1983	Paris, France	48.86	2.35	city	3	N	Transmissible in humans	N
Human immunodeficiency virus 2 [162]	*lentivirus*	*retro*	1986	Dakar city, Dakar region, Senegal	14.72	-17.47	city	-	Y	Transmissible in humans	N
Simian immunodeficiency virus [163]	*lentivirus*	*retro*	1992	Atlanta, Georgia, USA	33.75	-84.39	city	-	N	Strictly zoonotic	N
Central chimpanzee simian foamy virus [164]	*simiispumavirus*	*retro*	2012	Near Dja Nature Reserves, Southern Cameroon	4.5	13.5	approximate location[#]	-	Y	Strictly zoonotic	N
Eastern chimpanzee simian foamy virus [165]	*simiispumavirus*	*retro*	1971	Kenya	-0.02	37.91	country	1	Y	Strictly zoonotic	N
Grivet simian foamy virus [166]	*simiispumavirus*	*retro*	1997	Freiburg, Baden-Württemberg, Germany	48.00	7.84	city	3	N	Strictly zoonotic	N
Guenon simian foamy virus [164]	*simiispumavirus*	*retro*	2012	Near lolodrof, Southern Cameroon	3.23	10.73	approximate location[#]	-	Y	Strictly zoonotic	N
Taiwanese macaque simian foamy virus [167]	*simiispumavirus*	*retro*	2002	Ottawa, Canada	45.42	-75.7	city	3	Y	Strictly zoonotic	N
Australian bat lyssavirus [168]	*lyssavirus*	*rhabdo*	1998	Brisbane, Queensland, Australia	-27.47	153.03	city	-	Y	Strictly zoonotic	N
Duvenhage lyssavirus [169]	*lyssavirus*	*rhabdo*	1971	Pretoria, Tshwane City, Gauteng province, South Africa	-25.75	28.23	city	-	Y	Strictly zoonotic	N
European bat 1 lyssavirus [170]	*lyssavirus*	*rhabdo*	1989	Belgorod, Belgorod Oblast, Russia	50.6	36.6	City	3	Y	Strictly zoonotic	N
European bat 2 lyssavirus [171]	*lyssavirus*	*rhabdo*	1986	Helsinki, Uusimaa region, Finland	60.17	24.94	city	-	Y	Strictly zoonotic	N
Irkut lyssavirus [172]	*lyssavirus*	*rhabdo*	2013	Tonghua county, Tonghua City, Jilin Province, China	41.68	125.76	county	-	Y	Strictly zoonotic	N
Mokola lyssavirus [173]	*lyssavirus*	*rhabdo*	1972	Ibadan, Nigeria	7.38	3.95	city	-	Y	Strictly zoonotic	N
Rabies lyssavirus [174]	*lyssavirus*	*rhabdo*	1903	Pavia, province of Pavia, region of Lombardy, Italy	45.18	9.16	city	-	N	Transmissible in humans[†]	N

Bas-Congo tibrovirus [175]	*tibrovirus*	*rhabdo*	2012	Mangala village, Boma Bungu Health Zone, DRC	-4.04	21.76	village	-	Y	Transmissible in humans	Y
Ekpoma 1 tibrovirus [176]	*tibrovirus*	*rhabdo*	2015	Irrua, Edo State, Nigeria	6.74	6.22	village	-	Y	Strictly zoonotic	N
Ekpoma 2 tibrovirus [176]	*tibrovirus*	*rhabdo*	2015	Irrua, Edo State, Nigeria	6.74	6.22	village	-	Y	Strictly zoonotic	N
Alagoas vesiculovirus [177]	*vesiculovirus*	*rhabdo*	1967	State of Alagoas, Brazil	-9.57	-36.78	state	2	Y	Strictly zoonotic	Y
Chandipura vesiculovirus [178]	*vesiculovirus*	*rhabdo*	1967	Nagpur city, Nagpur district, Maharashtra state, India	21.15	79.09	city	-	Y	Strictly zoonotic	Y
Cocal vesiculovirus [179]	*vesiculovirus*	*rhabdo*	1964	Nariva swamp, Trinidad and Tobago	10.43	-61.06	island	-	Y	Strictly zoonotic	Y
Indiana vesiculovirus [180]	*vesiculovirus*	*rhabdo*	1958	Beltsville, Prince George's County, Maryland, USA	39.05	-76.9	CDP	-	Y	Strictly zoonotic	Y
Isfahan vesiculovirus [181]	*vesiculovirus*	*rhabdo*	1977	Dormian village, Isfahan Province, Iran	33.28	52.36	village	-	Y	Strictly zoonotic	Y
Maraba vesiculovirus [182]	*vesiculovirus*	*rhabdo*	1984	Serra Norte area, State of Pará, Brazil	-6.04	-50.18	area	-	Y	Strictly zoonotic	Y
New Jersey vesiculovirus [183]	*vesiculovirus*	*rhabdo*	1950	Madison, Wisconsin, USA	43.07	-89.4	city	-	Y	Strictly zoonotic	Y
Piry vesiculovirus [184]	*vesiculovirus*	*rhabdo*	1974	Marabá city, State of Pará, Brazil	-5.38	-49.13	city	3	Y	Strictly zoonotic	Y
Barmah Forest virus [185]	*alphavirus*	*toga*	1986	South coast of New South Wales, Australia	-36.68	149.66	state	2	Y	Transmissible in humans	Y
Chikungunya virus [186]	*alphavirus*	*toga*	1956	Newala district, Tanzania	-10.64	39.24	district	3	Y	Transmissible in humans	Y
Eastern equine encephalitis virus [187]	*alphavirus*	*toga*	1938	Southwestern Massachusetts, USA	42.20	-71.10	state	2	Y	Strictly zoonotic	Y
Everglades virus [188]	*alphavirus*	*toga*	1970	Homestead, Florida, USA	25.47	-80.48	city	-	Y	Strictly zoonotic	Y
Getah virus [189]	*alphavirus*	*toga*	1966	Brisbane, Queensland, Australia	-27.47	153.025	city	-	Y	Strictly zoonotic	Y
Highlands J virus [190]	*alphavirus*	*toga*	2000	Florida, USA	27.66	-81.52	state	2	Y	Strictly zoonotic	Y
Madariaga virus [191]	*alphavirus*	*toga*	1972	Port of Spain, Trinidad	10.67	-61.52	city	2	Y	Strictly zoonotic	Y

Mayaro virus [192]	*alphavirus*	*toga*	1957	Mayaro County, Trinidad and Tobago	10.28	-61.03	county	-	Y	Strictly zoonotic	Y
Mosso das Pedras virus [193]	*alphavirus*	*toga*	2013	Chaco province, Argentina	-27.43	-59.02	state	2	Y	Strictly zoonotic	Y
Mucambo virus [194]	*alphavirus*	*toga*	1965	Mexico city, Mexico	19.43	-99.13	city	2	Y	Strictly zoonotic	Y
Ndumu virus [195]	*alphavirus*	*toga*	1961	Ndumu, KwaZulu-Natal Province, South Africa	-26.93	32.26	town/city	-	Y	Strictly zoonotic	Y
Onyong-nyong virus [196]	*alphavirus*	*toga*	1961	Entebbe, Wakiso District, Uganda	0.05	32.46	city	3	Y	Transmissible in humans	Y
Pixuna virus [197]	*alphavirus*	*toga*	1991	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Rio Negro virus [198]	*alphavirus*	*toga*	1993	General Belgrano Island, Formosa, Argentina	-26.19	-58.18	island	-	Y	Strictly zoonotic	Y
Ross River virus [199]	*alphavirus*	*toga*	1972	Edward River, New South Wales, Australia	-35.4	144.25	river	-	Y	Transmissible in humans	Y
Semliki Forest virus [200]	*alphavirus*	*toga*	1979	Gießen, Gießen District, Hesse State, Germany	50.58	8.68	town	-	Y	Transmissible in humans	Y
Sindbis virus [201]	*alphavirus*	*toga*	1955	Cairo city, Cairo Governorate, Egypt	30.04	31.24	city	-	Y	Strictly zoonotic	Y
Tonate virus [202]	*alphavirus*	*toga*	1976	French Guyana, now French Guiana	3.93	-53.13	Region	2	Y	Strictly zoonotic	Y
Una virus [203]	*alphavirus*	*toga*	1963	Belém, State of Pará, Brazil	-1.39	-48.42	city	3	Y	Strictly zoonotic	Y
Venezuelan equine encephalitis virus [204]	*alphavirus*	*toga*	1943	New York, New York State, USA	40.71	-74.01	city	-	Y	Transmissible in humans	Y
Western equine encephalitis virus [187]	*alphavirus*	*toga*	1938	Fresno, California, USA	36.75	-119.77	city	-	Y	Transmissible in humans[†]	Y
Whataroa virus [205]	*alphavirus*	*toga*	1964	Greymouth town, Grey District, New Zealand	-42.45	171.21	town	-	Y	Strictly zoonotic	Y
Rubella virus [206]	*rubivirus*	*toga*	1942	Washington, D. C., USA	38.91	-77.04	city	-	Y	Transmissible in humans	N
Hepatitis delta virus [207]	*deltavirus*	*Unassigned*	1977	Turin, Piedmont region, Italy	45.07	7.69	city	-	N	Transmissible in humans	N

DRC, Democratic Republic of the Congo; CDP, census-designated place; [#] locations inferred from maps given in the discovery papers; * location of the research field; lat and lon are the latitude and longitude of the centroid of the discovery location; [†] transmitted only via iatrogenic or maternal routes

Appendix B Supplementary material for chapter 3: Predictors of human-infective RNA virus discovery in the United States, China, and Africa

B.1 Supplementary Figures

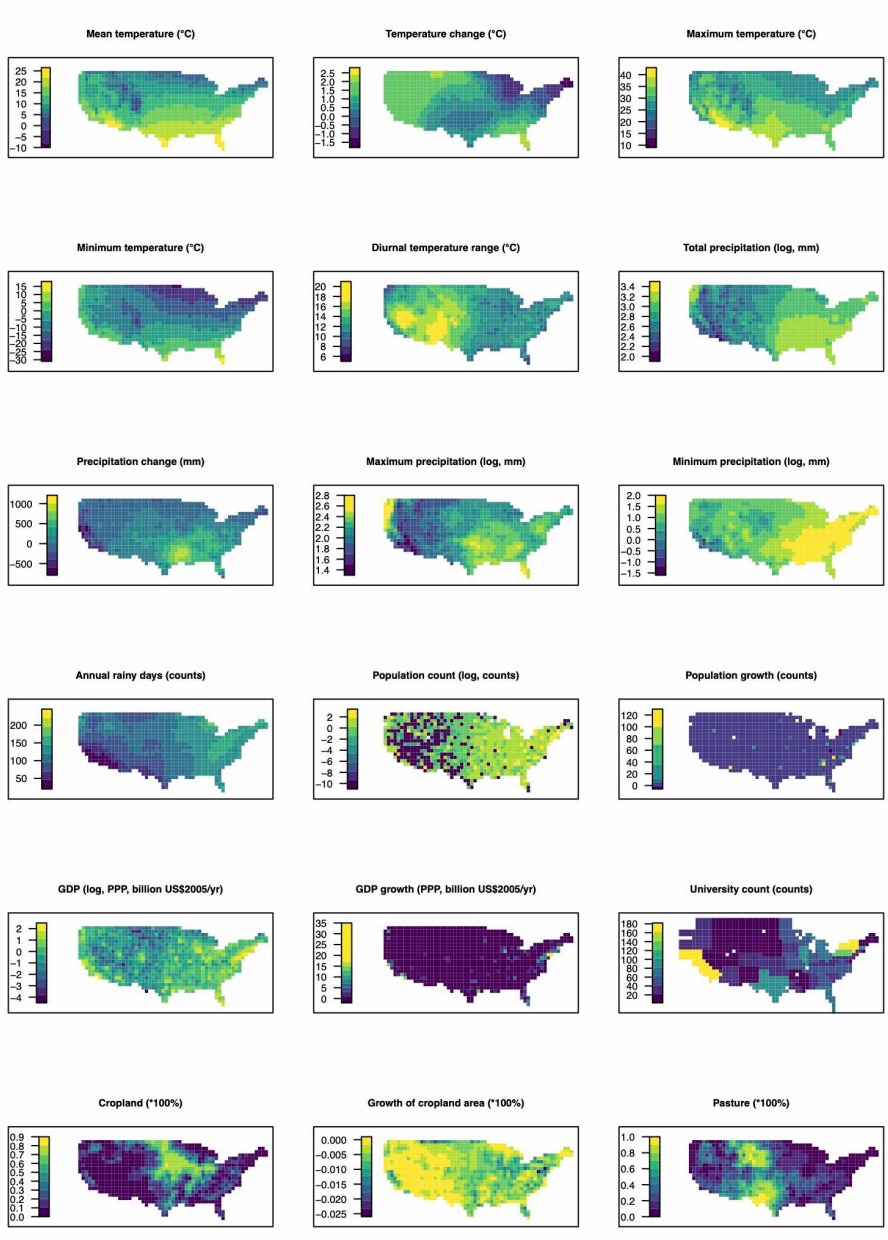


Figure B.1 Distribution maps of 32 predictors in 2015 in the United States

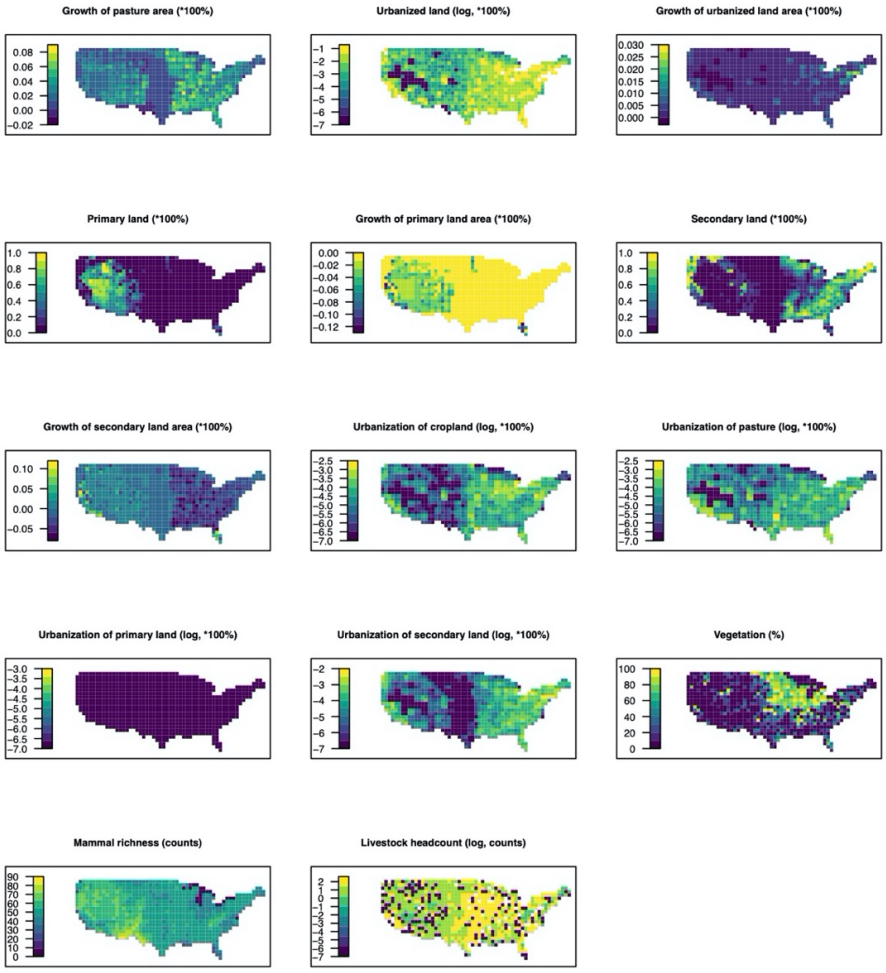


Figure B.1 Continued

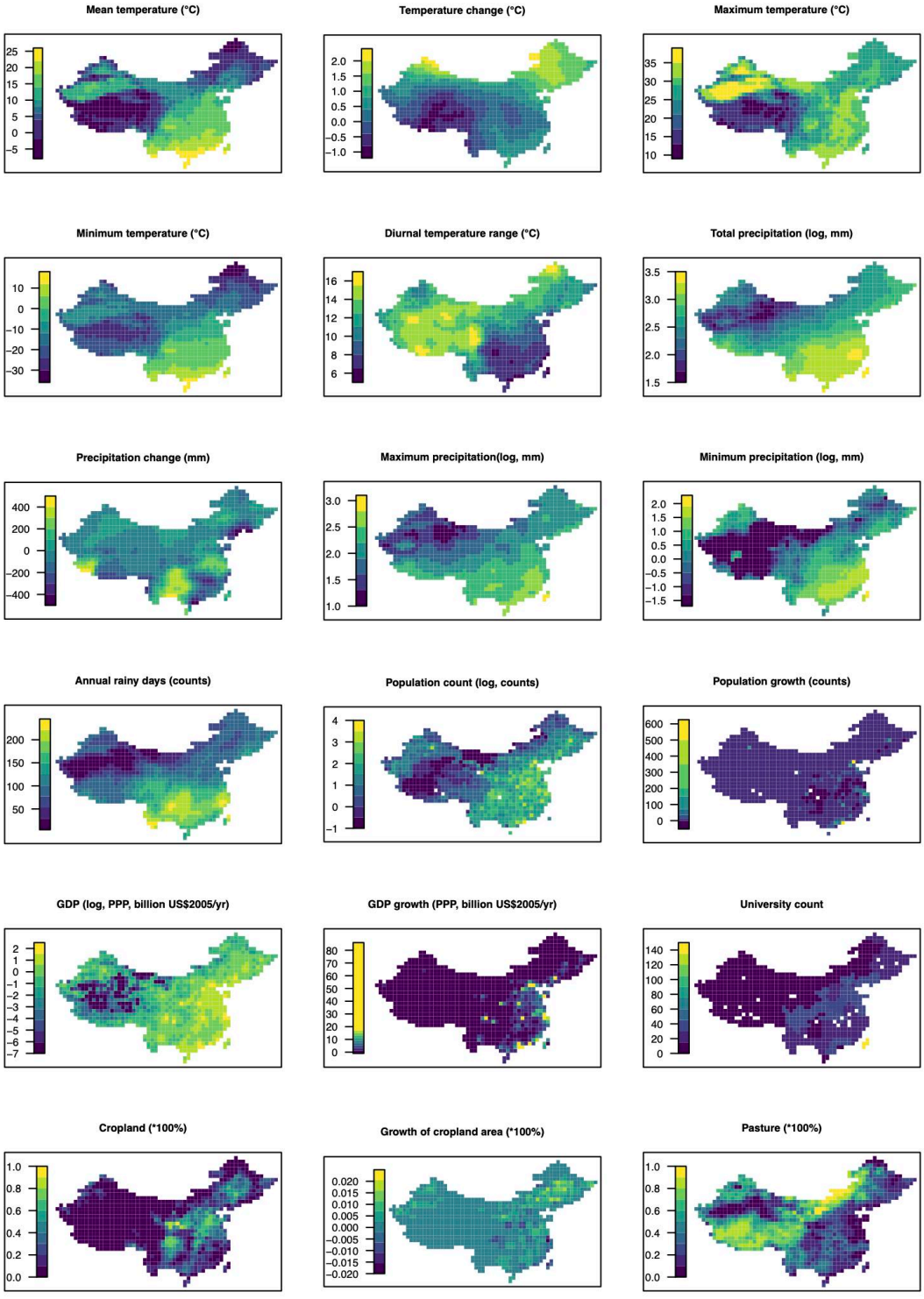


Figure B.2 Distribution maps of 32 predictors in 2015 in China

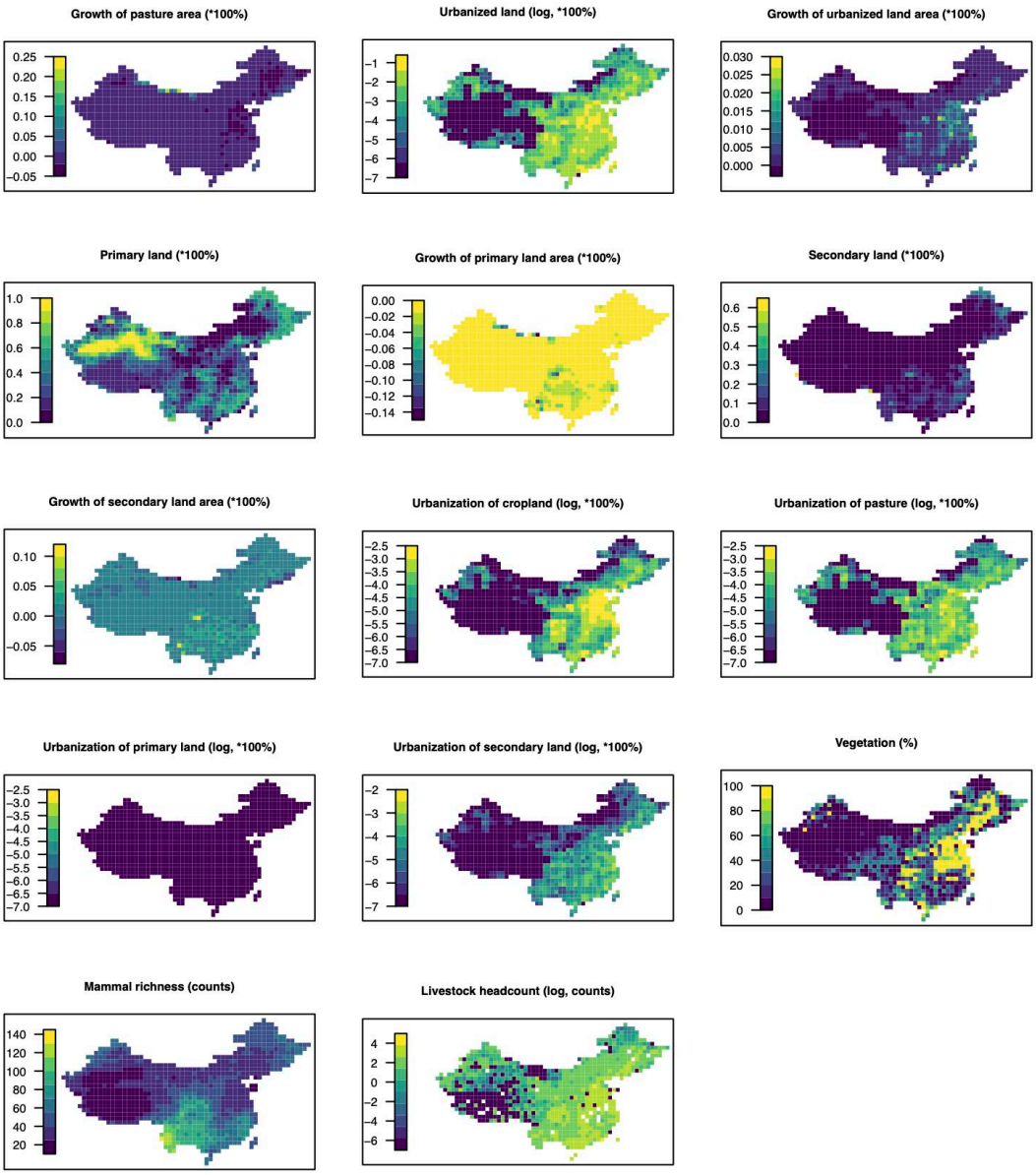


Figure B.2 Continued

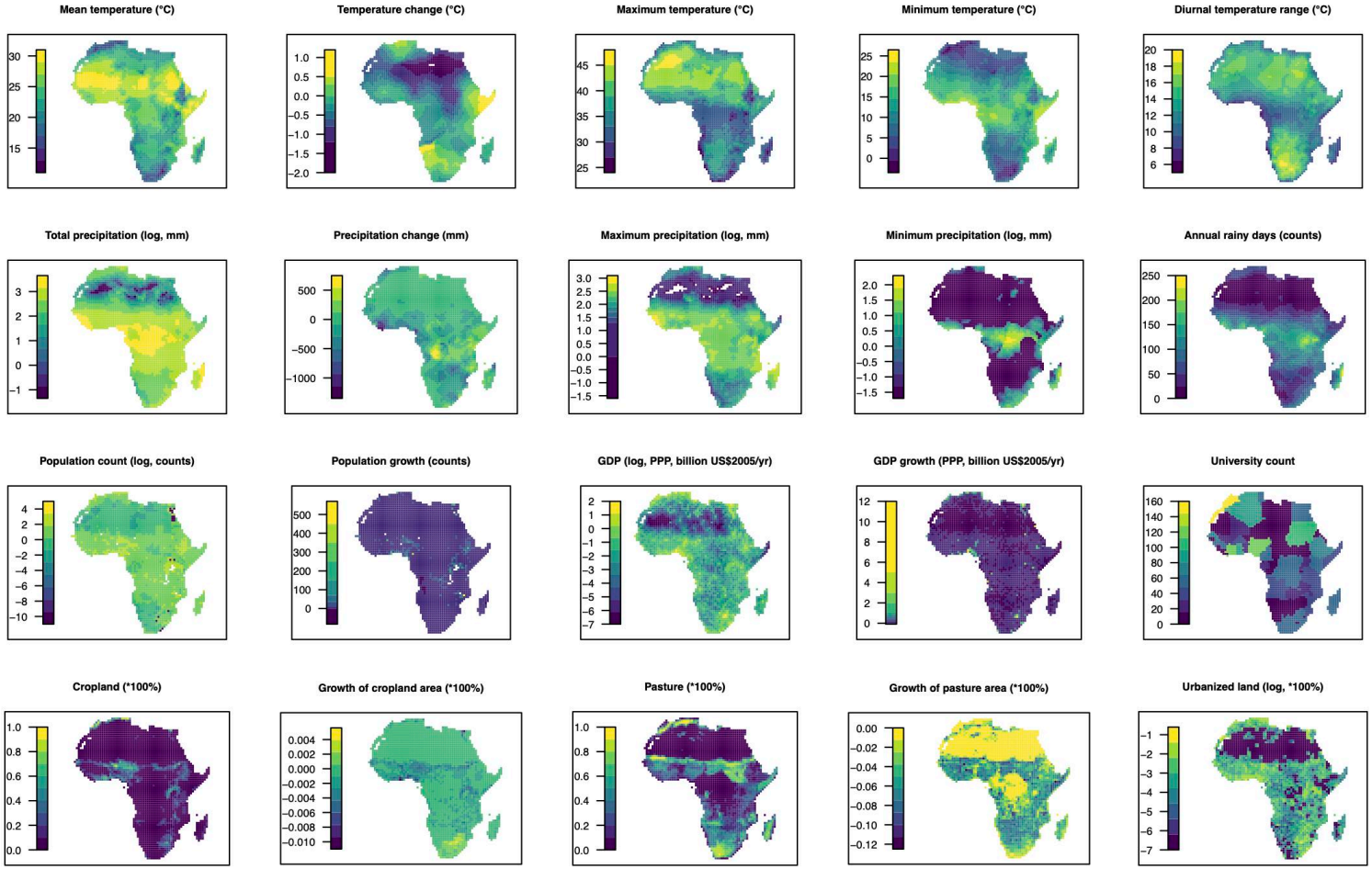


Figure B.3 Distribution maps of 32 predictors in 2015 in Africa

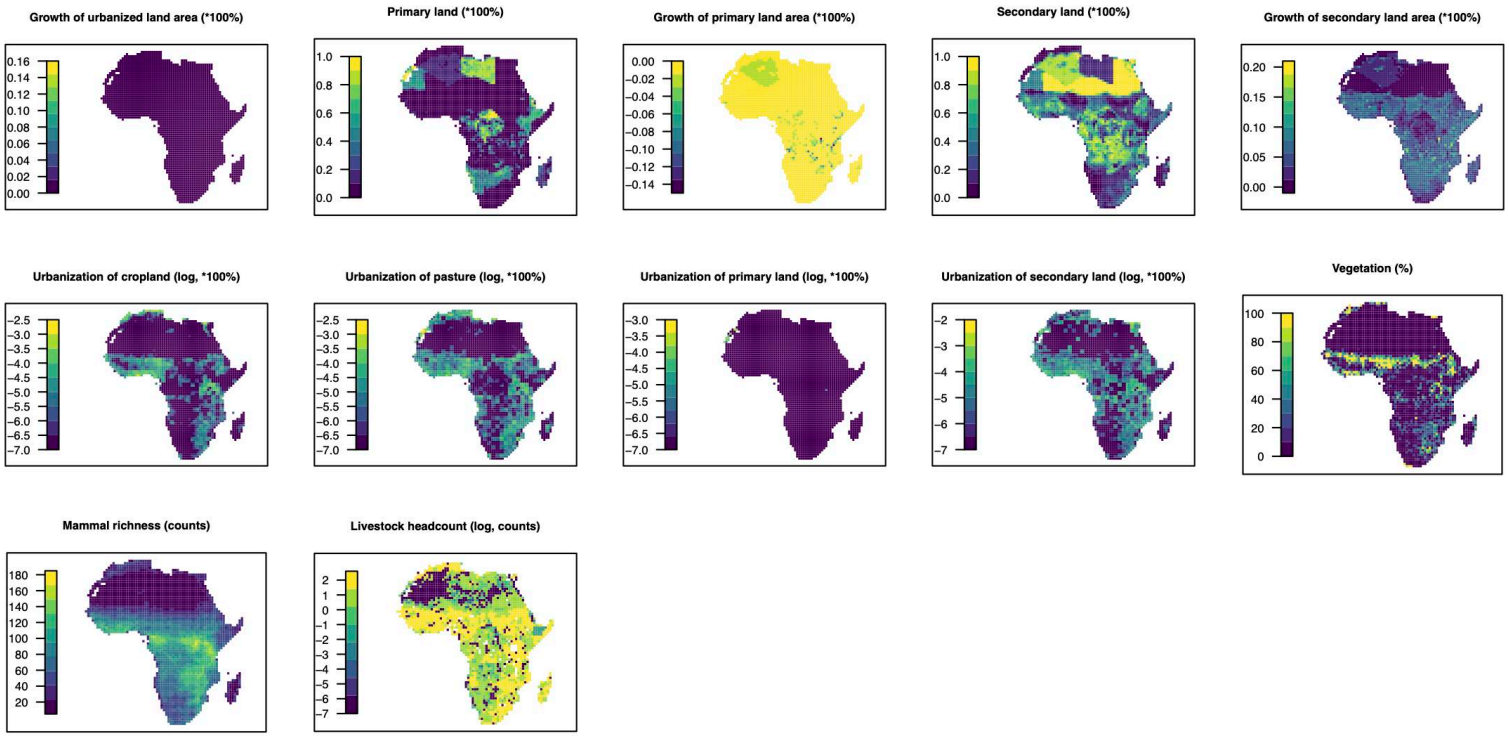


Figure B.3 Continued

B.2 Supplementary Tables

Table B.1 Summary of the human-infective RNA virus discovery database in the United States, Africa and China

Species	Original discovery year	United States					China					Africa				
		Reported?	Discovery year	location	Lat	Lon	Reported?	Discovery year	location	Lat	Lon	Reported?	Discovery year	location	Lat	Lon
Argentinian mammarenavirus	1958	No					No					No				
Brazilian mammarenavirus	1994	Yes [208]	1995	New Haven, Connecticut	41.31	-72.93	No					No				
Cali mammarenavirus	1971	Yes [3]	1974	Houston, Texas	29.76	-95.37	No					No				
Chapare mammarenavirus	2008	No					No					No				
Guanarito mammarenavirus	1991	No					No					No				
Lassa mammarenavirus	1970	Yes [6]	1970	New Haven, Connecticut	41.31	-72.93	No					Yes [6]	1970	Lassa, Borno State, Nigeria	10.69	13.27
Lujo mammarenavirus	2009	No					No					Yes [7]	2009	Lusaka, Zambia	-15.39	28.32
Lymphocytic choriomeningitis mammarenavirus	1934	Yes [8]	1934	St. Louis county, Missouri	38.61	-90.41	No					No				
Machupo mammarenavirus	1964	No					No					No				
Mobala mammarenavirus	1985	No					No					Yes [10]	1985	Bouboui and Gomoka village, Boali town, Central African Republic	4.89	18.14
Whitewater Arroyo mammarenavirus	2000	Yes [11]	2000	Alameda County, California	37.60	-121.72	No					No				
Mamastrovirus 1	1975	Yes [209]	1981	Martin County, California	40.22	-123.10	Yes [210]	1981	Guangzhou, Guangdong	23.13	113.26	Yes [211]	1981	Lebowa, South Africa	-23.5	29.5
Mamastrovirus 6	2008	Yes [212]	2009	St. Louis, Missouri	38.63	-90.20	Yes [213]	2010	Hong Kong	22.40	114.11	Yes [214]	2009	Maiduguri, Borno State, Nigeria	11.83	13.15

Mamastrovirus 8	2009	Yes [14]	2009	St. Louis, Missouri	38.63	-90.20	Yes [215]	2013	Nanjing, Jiangsu and Lanzhou, Gansu	31.95	118.78	Yes [214]	2009	Maiduguri, Borno State, Nigeria	11.83	13.15
Mamastrovirus 9	2009	Yes [15]	2009	Accomack and Northampton Counties, Virginia	37.71	-75.81	Yes [216]	2019	Jinan, Shandong	36.68	117.11	Yes [214]	2009	Maiduguri, Borno State, Nigeria	11.83	13.15
Mammalian 1 orthobornavirus	1985	Yes [16]	1985	Philadelphia, Pennsylvania	39.95	-75.17	Yes [217]	1999	Taiwan	23.70	120.96	Yes [218]	1992	Rural area of East Africa	-1.28	34.53
Mammalian 2 orthobornavirus	2015	No					No					No				
Norwalk virus	1972	Yes [18]	1972	Norwalk, Ohio	41.24	-82.62	Yes [219]	1995	Henan	33.88	113.48	Yes [220]	1993	Pretoria, Gauteng province, South Africa	-25.75	28.23
Sapporo virus	1980	Yes [221]	1988	Houston, Texas	29.76	-95.37	Yes [221]	1988	Shanghai	31.23	121.47	Yes [222]	1997	Pretoria, Gauteng province, South Africa	-25.75	28.23
Vesicular exanthema of swine virus	1998	Yes [20]	1998	Corvallis, Oregon	44.56	-123.26	No					No				
Alphacoronavirus 1	2007	No					No					No				
Human coronavirus 229E	1966	Yes [22]	1966	Chicago, Illinois	41.88	-87.63	Yes [223]	1975	Kunming, Yunnan	25.07	102.68	Yes [224]	1998	Kumasi, Ghana	6.70	-1.62
Human coronavirus NL63	2004	Yes [225]	2005	New Haven, Connecticut	41.31	-72.93	Yes [226]	2005	Hong Kong	22.40	114.11	Yes [227]	2008	Cape Town, Western Cape Province, South Africa	-33.90	18.57
Betacoronavirus 1	1967	Yes [24]	1967	Bethesda, Maryland	38.98	-77.09	Yes [226]	2005	Hong Kong	22.40	114.11	Yes [228]	2011	Pretoria, Gauteng province, South Africa	-25.75	28.23
Human coronavirus HKU1	2005	Yes [229]	2006	New Haven, Connecticut	41.31	-72.92	Yes [25]	2005	Hong Kong	22.40	114.11	Yes [228]	2011	Pretoria, Gauteng province, South Africa	-25.75	28.23
Middle East respiratory syndrome-related coronavirus	2012	Yes* [230]	2014	Lake county, Indiana	41.45	-87.37	Yes* [231]	2015	Huizhou, Guangdong	23.09	114.40	Yes* [232]	2014	Monastir, Tunisia	35.79	10.82
Severe acute respiratory syndrome-related coronavirus	2003	Yes* [233]	2003	Atlanta, Georgia	33.75	-84.39	Yes [27]	2003	Hong Kong	22.40	114.11	Yes [234]	2004	Pretoria, Gauteng province, South Africa	-25.75	28.23

Human torovirus (been abolished)	1984	No					No					No				
Bundibugyo ebolavirus	2008	No					No					Yes [227]	2008	Bundibugyo and Kikyo town, Bundibugyo District, Western Uganda	0.71	30.06
Reston ebolavirus	1991	Yes [30]	1991	Reston, Fairfax County, Virginia	38.96	-77.35	No					No				
Sudan ebolavirus	1977	No					No					Yes [31]	1977	Maridi, South Sudan	4.91	29.45
Tai Forest ebolavirus	1995	No					No					Yes [32]	1995	Abidjan, Cote-d'Ivoire	5.36	-4.01
Zaire ebolavirus	1977	No					No					Yes [33]	1977	Yambuku village, Democratic Republic of the Congo	2.83	22.22
Marburg marburgvirus	1968	Yes* [235]	2009	Denver county, Colorado	39.55	-105.78	No					Yes [236]	1975	Johannesburg, South Africa	-26.20	27.90
Aroa virus	1971	No					No					No				
Bagaza virus	2009	No					No					No				
Banzi virus	1959	No					No					Yes [37]	1959	Maponde's Kraal(Usutu river), South Africa	-26.52	31.67
Cacipacore virus	2011	No					No					No				
Dengue virus	1907	Yes [237]	1914	Savannah, Georgia	32.02	-81.12	Yes [238]	1967	Southwest Taiwan	23.06	120.59	Yes [239]	1927	Durban, KwaZulu-Natal Province, South Africa	-29.86	31.02
Edge Hill virus	1985	No					No					No				
Gadgets Gully virus	1991	No					No					No				
Ilheus virus	1947	No					No					No				
Japanese encephalitis virus	1933	Yes* [240]	1953	Waltham, Massachusetts	42.38	-71.24	Yes [241]	1941	Beijing	40.01	116.41	Yes [242]	2017	Cunene, Angola	-16.28	15.28
Kokobera virus	1964	No					No					No				
Kyasanur forest disease virus	1957	No					Yes [243]	2009	Hengduanshan Mountain, Yunnan	27.50	99.00	Yes [244]	2014	Djibouti, Republic of Djibouti	11.57	43.15
Langat virus	1956	No					No					No				
Louping ill virus	1934	Yes [46]	1934	New York	40.71	-74.01	No					No				

Murray Valley encephalitis virus	1952	No					No					No				
Ntaya virus	1952	No					No					Yes [48]	1952	Bwamba county, Uganda	0.75	30.02
Omsk hemorrhagic fever virus	1948	No					No					No				
Powassan virus	1959	Yes [245]	1973	Middlesex County, New Jersey	40.54	-74.37	No					No				
Rio Bravo virus	1962	Yes [51]	1962	Dallas city, Texas	32.78	-96.80	No					No				
Saint Louis encephalitis virus	1933	Yes [52]	1933	St. Louis City, Missouri	38.63	-90.20	No					No				
Tembusu virus	1975	No					Yes [246]	2013	Shandong	36.40	118.77	No				
Tick-borne encephalitis virus	1938	Yes* [247]	1979	Cleveland, Ohio	41.51	-81.69	Yes [248]	1956	Bali village, Wuchang, Heilongjiang	44.91	127.16	No				
Uganda S virus	1952	No					No					Yes [55]	1952	Bwamba county, Uganda	0.75	30.02
Usutu virus	2009	No					No					No				
Wesselsbron virus	1957	No					No					Yes [57]	1957	Lake Simbu region, Maputaland, KwaZulu-Natal, South Africa	-27.36	32.32
West Nile virus	1940	Yes [249]	2001	New York	40.71	-74.01	Yes [250]	2013	Jiashi County, Xinjiang	39.58	77.18	Yes [58]	1940	Omogo, West Nile district, Uganda	0.42	33.21
Yellow fever virus	1901	Yes [251]	1904	Laredo, Texas	27.51	-99.51	Yes* [252]	2016	Beijing	40.01	116.41	Yes [253]	1928	Larteh, Ghana	5.94	-0.07
Zika virus	1952	Yes* [254]	2011	Northern Colorado	39.55	-105.78	Yes* [255]	2016	Gan County, Ganzhou city, Jiangxi	25.86	115.02	Yes [60]	1952	Zika, Uganda	0.12	32.53
Hepacivirus C	1989	Yes [61]	1989	Emeryville, California	37.83	-122.29	Yes [256]	1990	Qidong county, Jiangsu	31.88	121.72	Yes [257]	1990	Johannesburg, South Africa	-26.20	27.90
Pegivirus C	1995	Yes [258]	1995	Chapel Hill, North Carolina; Rochester, Minnesota; Dallas, Texas	35.91	-79.06	Yes [259]	1996	Beijing	40.01	116.41	Yes [258]	1995	Cairo, Egypt	30.04	31.24
Pegivirus H	2015	Yes [63]	2015	New York city, New York	40.71	-74.01	Yes [260]	2018	Guangzhou, Guangdong	23.13	113.26	Yes [261]	2019	Ebolowa, Cameroon	2.92	11.15

Pestivirus A	1988	Yes [262]	1989	Whiteriver, Arizona	33.83	-109.97	No					Yes [64]	1988	Zambia	-13.13	27.85
Andes orthohantavirus	1996	No					No					No				
Bayou orthohantavirus	1995	Yes [66]	1995	Louisiana	30.98	-91.96	No					No				
Black creek canal orthohantavirus	1995	Yes [67]	1995	Miami-Dade County, Florida	25.76	-80.34	No					No				
Choclo orthohantavirus	2000	No					No					No				
Dobrava-Belgrade orthohantavirus	1992	No					No					No				
Hantaan orthohantavirus	1978	No					Yes [263]	1980	Zhejiang	29.14	119.79	No				
Laguna Negra orthohantavirus	1997	No					No					No				
Puumala orthohantavirus	1980	No					No					No				
Sangassou orthohantavirus	2010	No					No					Yes [73]	2010	Sangassou village, Macenta district, Forest Guinea	8.24	-9.32
Seoul orthohantavirus	1982	Yes [264]	1987	Mississippi	32.57	-89.88	Yes [74]	1982	Jiangsu	33.14	119.79	Yes [265]	1986	Jos, Nigeria	9.90	8.86
Sin Nombre orthohantavirus	1993	Yes [75]	1993	New Mexico	34.52	-105.87	No					No				
Thailand orthohantavirus	2006	No					No					No				
Thottopalayam thottimvirus	2007	No					No					No				
Tula orthohantavirus	1996	No					No					No				
Orthohepevirus A	1983	Yes* [266]	1987	Los Angeles County, California	34.05	-118.24	Yes [267]	1989	Kashi county, Kashi city, Xinjiang	39.46	75.99	Yes [268]	1985	Medea town, Algeria	36.26	2.75
Orthohepevirus C	2018	No					Yes [80]	2018	Hong Kong	22.40	114.11	No				
Crimean-Congo haemorrhagic fever orthonairovirus	1967	No					Yes [269]	1985	Bachu, southern Xinjiang	39.79	78.55	Yes [81]	1967	Kisangani, Tshopo province, Democratic Republic of the Congo	0.53	25.19
Dugbe orthonairovirus	1969	No					No					Yes [82]	1969	Ibadan, Nigeria	7.35	3.88

Nairobi sheep disease orthonairovirus	1969	No					No					Yes [270]	1991	Mombasa; Malindi; and Kilifi, Coast Province, Kenya	-3.34	39.57
Thiafora orthonairovirus	1989	No					No					No				
Influenza A virus	1933	Yes [271]	1935	Philadelphi, Pennsylvania	39.95	-75.17	Yes [272]	1950	Beijing	40.01	116.41	Yes [273]	1951	Johannesburg, South Africa and Cape Town, South Africa	-26.20	27.90
Influenza B virus	1940	Yes [86]	1940	Irvington village, Greenburgh town, Westchester County, New York	41.03	-73.87	Yes [274]	1957	Beijing	40.01	116.41	Yes [275]	1970	Arusha, Arusha Region, Tanzania	-3.37	36.69
Influenza C virus	1950	Yes [87]	1950	Ann Arbor city, Michigan	42.28	-83.74	Yes [276]	1957	Beijing	40.01	116.41	Yes [277]	1968	Johannesburg, South Africa	-26.20	27.90
Dhori thogotovirus	1985	No					No					No				
Thogoto thogotovirus	1969	No					No					Yes [82]	1969	Ibadan, Nigeria	7.35	3.88
Avian orthoavulavirus 1	1943	Yes [89]	1943	Washington, D. C.	38.91	-77.04	No					No				
Hendra henipavirus	1995	No					No					No				
Nipah henipavirus	1999	No					No					No				
Canine morbillivirus	1955	Yes [92]	1955	Buffalo, New York	42.89	-78.88	No					No				
Measles morbillivirus	1911	Yes [93]	1911	Washington, D. C.	38.91	-77.04	Yes [278]	1958	Beijing	40.01	116.41	Yes [279]	1963	Dakar, Senegal	14.72	-17.47
Human respirovirus 1	1958	Yes [94]	1958	Washington, D. C.	38.91	-77.04	Yes [280]	1964	Zhejiang	29.14	119.79	Yes [281]	1963	Cape Town, Western Cape Province, South Africa	-33.90	18.57
Human respirovirus 3	1958	Yes [94]	1958	Washington, D. C.	38.91	-77.04	Yes [282]	1987	Guangzhou, Guangdong	23.13	113.26	Yes [281]	1963	Cape Town, Western Cape Province, South Africa	-33.90	18.57
Achimota pararubulavirus 2	2013	No					No					Yes [95]	2013	Volta, Ghana	6.05	0.37
Human orthorubulavirus 2	1956	Yes [96]	1956	Cincinnati, Ohio	39.10	-84.51	Yes [283]	1975	Nanjing, Jiangsu	31.95	118.78	Yes [284]	1967	Accra, Ghana	5.60	-0.19

Human orthorubulavirus 4	1960	Yes [97]	1960	Bethesda, Maryland	38.98	-77.09	Yes [285]	2005	Hong Kong	22.40	114.11	Yes [286]	2010	Ndiop village, Sine Saloum region, Senegal	15.18	-16.74
Mammalian orthorubulavirus 5	1959	Yes [98]	1959	Stanford, California	37.42	-122.17	No					No				
Menangle pararubulavirus	1998	No					No					No				
Mumps orthorubulavirus	1934	Yes [100]	1934	Nashville, Tennessee	36.16	-86.78	Yes [287]	1958	Beijing	40.01	116.41	Yes [288]	1955	Johannesburg, South Africa	-26.20	27.90
Simian orthorubulavirus	1968	No					No					No				
Sosuga pararubulavirus	2014	No					No					Yes [102]	2014	-	3.76	32.82
Tioman pararubulavirus	2007	No					No					No				
Bunyamwera orthobunyavirus	1946	Yes [289]	1964	Southern Florida	26.92	-81.21	No					Yes [104]	1946	Bwamba County, Uganda	0.75	30.02
Bwamba orthobunyavirus	1941	No					No					Yes [105]	1941	Bwamba county, Western Province of Uganda	0.75	30.02
California encephalitis orthobunyavirus	1952	Yes [106]	1952	Kern county, California	35.49	-118.86	Yes [290]	1984	Longhua, Shanghai	31.22	121.43	Yes [291]	1961	Uganda	1.37	32.29
Caraparu orthobunyavirus	1961	No					No					No				
Catu orthobunyavirus	1961	No					No					No				
Guama orthobunyavirus	1961	No					No					No				
Guaroa orthobunyavirus	1959	No					No					No				
Kairi orthobunyavirus	1967	No					No					No				
Madrid orthobunyavirus	1964	No					No					No				
Marituba orthobunyavirus	1961	No					No					No				
Nyando orthobunyavirus	1965	No					No					Yes [111]	1965	Kisumu, Kenya	-0.09	34.77
Oriboca orthobunyavirus	1961	No					No					No				
Oropouche orthobunyavirus	1961	No					No					No				

Patois orthobunyavirus	1972	No					No					No				
Shuni orthobunyavirus	1975	No					No					Yes [114]	1975	Ibadan, Nigeria	7.38	3.95
Tacaiuma orthobunyavirus	1967	No					No					No				
Wyeomyia orthobunyavirus	1965	No					No					No				
Candiru phlebovirus	1983	No					No					No				
Punta Toro phlebovirus	1970	No					No					No				
Rift Valley fever phlebovirus	1931	No					Yes* [292]	2016	Beijing	40.01	116.41	Yes [118]	1931	Rift Valley of Kenya Colony	-0.28	36.07
Sandfly fever Naples phlebovirus	1944	No					No					Yes [293]	1951	Cairo, Egypt	30.04	31.24
Heartland banyangvirus	2012	Yes [294]	2012	Andrew and Nodaway Counties, Missouri	39.82	-94.59	No					No				
Huaiyangshan banyangvirus	2011	No					Yes [120]	2011	Huaiyangshan	31.37	115.39	No				
Uukuniemi phlebovirus	1970	No					No					No				
Human picobirnavirus	1988	Yes [295]	1993	Atlanta, Georgia	33.75	-84.39	Yes [296]	2000	Lulong County, Hebei	39.94	116.94	No				
Equine rhinitis A virus	1962	No					No					No				
Foot-and-mouth disease virus	1965	No					Yes [297]	1999	Guangzhou	23.13	113.26	Yes [298]	2002	Alexandria Governorate, Egypt	30.74	29.74
Cardiovirus A	1947	Yes [299]	1961	New Orleans, Louisiana	29.95	-90.07	Yes [300]	2015	Changchun, Jilin	43.87	125.34	Yes [301]	1948	Entebbe, Uganda	0.05	32.46
Cardiovirus B	1963	Yes [302]	2007	San Diego, California	32.72	-117.16	Yes [303]	2009	Lanzhou, Gansu	36.06	103.79	Yes [304]	2009	Cameroon	5.03	12.40
Cosavirus A	2008	No					Yes [305]	2010	Shanghai	31.23	121.47	Yes [129]	2012	Maiduguri, Borno State, Nigeria	11.83	13.15
Cosavirus B	2008	No					Yes [306]	2016	Zhenjiang, Jiangsu	32.19	119.43	No				
Cosavirus D	2008	No					No					Yes [129]	2012	Maiduguri, Borno State, Nigeria	11.83	13.15
Cosavirus E	2008	No					No					Yes [129]	2012	Maiduguri, Borno State, Nigeria	11.83	13.15

Cosavirus F	2012	No					No					No				
Enterovirus A	1949	Yes [130]	1949	New York	43.30	-74.22	Yes [307]	1985	Tianjin	39.34	117.36	Yes [288]	1955	Johannesburg, South Africa	-26.20	27.90
Enterovirus B	1949	Yes [130]	1949	Wilmington	39.74	-75.54	Yes [308]	1960	Fuzhou, Fujian	26.07	119.30	Yes [309]	1953	Middelburg, Transvaal, South Africa	-25.77	29.46
Enterovirus C	1909	Yes [131]	1909	New York city, New York	40.71	-74.01	Yes	1941	Bejing	39.90	116.41	Yes [310]	1933	Monrovia, Liberia	6.29	-10.76
Enterovirus D	1967	Yes [132]	1967	Berkeley, California	37.87	-122.27	Yes [311]	1979	Shanghai	31.23	121.47	Yes [312]	1973	Morocco	31.79	-7.09
Enterovirus E	1961	Yes [133]	1961	Denver, Colorado	39.74	-104.99	No					No				
Enterovirus H	1965	No					No					No				
Rhinovirus A	1953	Yes [313]	1956	Baltimore, Maryland	39.29	-76.61	Yes [314]	1975	Guangzhou, Guangdong	23.13	113.26	Yes [315]	1963	Cape Town, Western Cape Province, South Africa	-33.90	18.57
Rhinovirus B	1960	Yes [316]	1961	Chicago, Illinois	41.88	-87.63	Yes [317]	2008	Beijing	40.01	116.41	Yes [318]	2008	Pretoria, Gauteng province, South Africa	-25.75	28.23
Rhinovirus C	2006	Yes [319]	2006	New York city, New York	40.71	-74.01	Yes [137]	2007	Hong Kong	22.40	114.11	Yes [318]	2008	Pretoria, Gauteng province, South Africa	-25.75	28.23
Erbovirus A	2005	No					No					No				
Hepatovirus A	1973	Yes [139]	1973	Bethesda, Maryland	38.98	-77.09	Yes [320]	1978	Shanghai	31.23	121.47	Yes [321]	1977	Dakar, Senegal	14.72	-17.47
Aichivirus A	1991	Yes [322]	2013	Cincinnati, Ohio	39.10	-84.51	Yes [323]	2009	Shanghai	31.23	121.47	Yes [324]	2008	Monastir, Tunisia	35.77	10.82
Parechovirus A	1958	Yes [141]	1958	Cincinnati, Ohio	39.10	-84.51	Yes [325]	2009	Shanghai	31.23	121.47	Yes [129]	2012	Ouagadougou, Burkina Faso	12.24	-1.56
Parechovirus B	2003	No					No					No				
Salivirus A	2009	Yes [143]	2009	Northern California	38.84	-120.90	Yes [326]	2010	Shanghai	31.23	121.47	Yes [327]	2009	Maiduguri, Borno State, Nigeria	11.83	13.15
Avian metapneumovirus	2011	Yes [144]	2011	Memphis, Tennessee	35.15	-90.05	No					No				
Human metapneumovirus	2001	Yes [328]	2003	Rochester, New York	43.16	-77.61	Yes [329]	2003	Hong Kong	22.40	114.11	Yes [330]	2003	Johannesburg, South Africa	-26.20	27.90
Human orthopneumovirus	1957	Yes [146]	1957	Baltimore, Maryland	39.29	-76.61	Yes [331]	1975	Kunming, Yunnan	25.07	102.68	Yes [332]	1965	Cape Town, Western Cape Province, South Africa	-33.90	18.57

Colorado tick fever virus	1946	Yes [147]	1946	Denver, Colorado	39.74	-104.99	Yes [333]	1996	Nanjing, Jiangsu	31.95	118.78	No				
Eyach virus	1980	No					No					No				
Corriparta virus	1967	No					No					No				
Great Island virus	1963	No					No					No				
Lebombo virus	1975	No					No					Yes [114]	1975	Ibadan, Nigeria	7.38	3.95
Orungo virus	1976	No					No					Yes [334]	1976	Ibadan, Nigeria	7.38	3.95
Mammalian orthoreovirus	1954	Yes [151]	1954	Cincinnati, Ohio	39.10	-84.51	Yes [335]	1995	Xuzhou, Jiangsu	34.26	117.19	Yes[336]	1963	Johannesburg, South Africa	-26.20	27.90
Nelson Bay orthoreovirus	2007	No					Yes* [337]	2009	Hong Kong	22.40	114.11	No				
Rotavirus A	1973	Yes [338]	1976	Washington, D. C.	38.90	-77.04	Yes [339]	1979	Beijing	40.01	116.41	Yes [334]	1976	Johannesburg, South Africa	-26.20	27.90
Rotavirus B	1984	Yes [340]	1985	Baltimore, Maryland	39.29	-76.61	Yes [154]	1984	Jinzhou, Liaoning	41.10	121.13	Yes [341]	1987	Kenya	-0.02	37.91
Rotavirus C	1986	Yes [342]	1995	Providence, Rhode Island	41.82	-71.41	Yes [343]	1999	Beijing	40.01	116.41	Yes [344]	1999	Pretoria, Gauteng province, South Africa	-25.75	28.23
Rotavirus H	1987	No					Yes [345]	1987	Huaihua, Hunan Province	27.55	109.96	No				
Banna virus	1990	No					Yes [157]	1990	Xishuangbanna, Yunnan Province	21.90	100.80	No				
Primate T-lymphotropic virus 1	1980	Yes [158]	1980	Bethesda, Maryland	38.98	-77.09	Yes [154]	1984	Shenyang, Liaoing	41.80	123.38	Yes [346]	1984	Ibadan, Nigeria	7.38	3.95
Primate T-lymphotropic virus 2	1982	Yes [159]	1982	Seattle, Washington	47.61	-122.33	Yes [347]	2013	Henan and Hubei	32.21	112.96	Yes [348]	1991	Franceville, Gabon	-1.63	13.60
Primate T-lymphotropic virus 3	2005	No					No					Yes [160]	2005	Océan department, South Province, Cameroon	2.50	10.50
Human immunodeficiency virus 1	1983	Yes [349]	1984	Washington, D. C.	38.90	-77.04	Yes [350]	1986	Hong Kong	22.40	114.11	Yes [351]	1984	Kisangani, Tshopo province, Democratic Republic of the Congo	0.53	25.19
Human immunodeficiency virus 2	1986	Yes* [352]	1988	New Jersey	40.06	-74.41	Yes* [353]	2000	Fuzhou, Fujian	26.07	119.30	Yes [162]	1986	Dakar, Senegal	14.72	-17.47

Simian immunodeficiency virus	1992	Yes [163]	1992	Atlanta, Georgia	33.75	-84.39	No					Yes [160]	2005	Cameroon	7.37	12.35
Central chimpanzee simian foamy virus	2012	No					No					Yes [164]	2012	Near Dja Nature Reserves, Southern Cameroon	4.50	13.50
Eastern chimpanzee simian foamy virus	1971	No					No					Yes [165]	1971	Kenya	-0.02	37.91
Grivet simian foamy virus	1997	No					No					No				
Guenon simian foamy virus	2012	No					No					Yes [164]	2012	Near Iolodrof, Southern Cameroon	3.23	10.73
Taiwanese macaque simian foamy virus	2002	No					Yes [354]	2012	Yunnan	25.18	101.86	No				
Australian bat lyssavirus	1998	No					No					No				
Duvenhage lyssavirus	1971	No					No					Yes [169]	1971	Pretoria, Gauteng province, South Africa	-25.75	28.23
European bat Yeslyssavirus	1989	No					No					No				
European bat 2 lyssavirus	1986	No					No					No				
Irkut lyssavirus	2013	No					Yes [172]	2013	Tonghua county, Jilin	41.68	125.76	No				
Mokola lyssavirus	1972	No					No					Yes [173]	1972	Ibadan, Nigeria	7.38	3.95
Rabies lyssavirus	1903	Yes [355]	1910	Southern California	34.57	-116.76	Yes [356]	1981	Beijing	40.01	116.41	Yes [357]	1933	Carolina, Mpumalanga, South Africa	-26.07	30.12
Bas-Congo tibrovirus	2012	No					No					Yes [175]	2012	Mangala village, Boma Bungu Health Zone, Democratic Republic of Congo (DRC)	-4.04	21.76
Ekpoma Yestibrovirus	2015	No					No					Yes [176]	2015	Irrua, Edo State, Nigeria	6.74	6.22
Ekpoma 2 tibrovirus	2015	No					No					Yes [176]	2015	Irrua, Edo State, Nigeria	6.74	6.22
Alagoas vesiculovirus	1967	No					No					No				
Chandipura vesiculovirus	1967	No					No					No				

Cocal vesiculovirus	1964	No					No					No				
Indiana vesiculovirus	1958	Yes [180]	1958	Beltsville, Prince George's County, Maryland	39.05	-76.90	No					No				
Isfahan vesiculovirus	1977	No					No					No				
Maraba vesiculovirus	1984	No					No					No				
New Jersey vesiculovirus	1950	Yes [183]	1950	Madison, Wisconsin	43.07	-89.40	No					No				
Piry vesiculovirus	1974	No					No					No				
Barmah Forest virus	1986	No					No					No				
Chikungunya virus	1956	Yes* [358]	2006	Minnesota	46.44	-93.36	Yes [238]	1967	Southwest Taiwan	23.06	120.59	Yes [186]	1956	Newala district, Tanzania	-10.64	39.24
Eastern equine encephalitis virus	1938	Yes [187]	1938	Southwestern Massachusetts	42.19	-73.09	No					No				
Everglades virus	1970	Yes [188]	1970	Homestead, Florida	25.47	-80.48	No					No				
Getah virus	1966	No					Yes [359]	1992	Baoting County, Hainan	18.98	109.83	No				
Highlands J virus	2000	Yes [190]	2000	Florida	27.66	-81.52	No					No				
Madariaga virus	1972	No					No					No				
Mayaro virus	1957	Yes* [360]	1999	Ohio	40.42	-82.91	No					No				
Mosso das Pedras virus	2013	No					No					No				
Mucambo virus	1965	No					No					No				
Ndumu virus	1961	No					No					Yes [195]	1961	Ndumu, Maputaland, KwaZulu-Natal, South Africa	-26.93	32.26
Onyong-nyong virus	1961	No					No					Yes [196]	1961	Entebbe, Uganda	0.05	32.46
Pixuna virus	1991	No					No					No				
Rio Negro virus	1993	No					No					No				
Ross River virus	1972	No					Yes [361]	1999	Hainan	19.16	109.94	No				

Semliki Forest virus	1979	No					No					Yes [362]	1990	Bangui, Central Africa	4.36	18.58
Sindbis virus	1955	No					No					Yes [201]	1955	Cairo, Egypt	30.04	31.24
Tonate virus	1976	No					No					No				
Una virus	1963	No					No					No				
Venezuelan equine encephalitis virus	1943	Yes [204]	1943	New York	40.71	-74.01	No					No				
Western equine encephalitis virus	1938	Yes [187]	1938	Fresno, California	36.75	-119.77	No					No				
Whataroa virus	1964	No					No					No				
Rubella virus	1942	Yes [206]	1942	Washington, D. C.	38.91	-77.04	Yes [363]	1979	Hangzhou, Zhejiang	29.87	119.33	Yes [364]	1963	Cape Town, Western Cape Province, South Africa	-33.90	18.57
Hepatitis delta virus	1977	Yes [365]	1979	New Jersey	40.06	-74.41	Yes [366]	1980	Taipei, Taiwan	24.96	121.51	Yes [367]	1984	Harare, Zimbabwe	-17.83	31.03

Appendix C Supplementary material for chapter 4: Predictors of geographical extent and disappearance of human-infective RNA viruses

C.1 Supplementary Tables

Table C.1 Geographical extent and the state of disappearance for each human-infective RNA virus species

Species	Geographical distribution	Geographical extent	Last record from ProMED-mail	Last report from scientific database	Last record (combine two sources)	Disappeared
Argentinian mammarenavirus	Argentina, South America [368]	restricted	2016	-	2016	No
Brazilian mammarenavirus	Brazil, South America (one natural infection and one lab infection) and USA, North America (one lab infection) [208]	restricted	-	1995 [208]	1995	Yes
Cali mammarenavirus	USA, North America (lab infection) [3]	restricted†	-	1974 [3]	1974	Yes
Chapare mammarenavirus	Bolivia, South America [4]	restricted†	2008	2008 [4]	2008	Yes
Guanarito mammarenavirus	Central Venezuela, South America [369]	restricted	2017	-	2017	No
Lassa mammarenavirus	West African countries, especially Sierra Leone, the Republic of Guinea, Nigeria, and Liberia; And imported cases in North America (USA), Europe (UK, Sweden, Germany, and Netherlands), and Africa (South Africa) [370]	wide	2019	-	2019	No
Lujo mammarenavirus	Zambia and South Africa, Africa [7]	restricted	2009	2009 [7]	2009	Yes
Lymphocytic choriomeningitis mammarenavirus	Europe (especially island of Vir in Croatia), North America, South America, and Australia [371]	wide	2016	2019 [372]	2019	No
Machupo mammarenavirus	Beni Department, Bolivia, South America [373]	restricted	2013	-	2013	No
Mobala mammarenavirus	Central African Republic, Africa [10]	restricted†	-	1985 [10]	1985	Yes

Whitewater Arroyo mammarenavirus	Western USA, North America [374]	restricted	-	2011 [374]	2011	No?
Mamastrovirus 1	Worldwide (all continents except Antarctica) [375]	wide	2013	2019 [376]	2019	No
Mamastrovirus 6	Worldwide (all continents except Antarctica) [375]	wide	2013	2019 [216]	2019	No
Mamastrovirus 8	Asia (India, Pakistan, , Japan, China); North America (USA); and Africa (Nigeria, Egypt) [375]	wide	2013	2019 [216]	2019	No
Mamastrovirus 9	North America (USA); Europe (UK, Netherlands); Africa (Nigeria); Asia (Pakistan, Nepal, China) [375]	wide	2013	2019 [216]	2019	No
Mammalian 1 orthobornavirus	Worldwide (all continents except Antarctica) [377]	wide	2018	-	2018	No
Mammalian 2 orthobornavirus	Germany, Europe [378]	restricted	-	2019 [378]	2019	No
Norwalk virus	Worldwide (all continents except Antarctica) [379]	wide	2019	-	2019	No
Sapporo virus	Worldwide (all continents except Antarctica) [380]	wide	2013	2019 [381]	2019	No
Vesicular exanthema of swine virus	Northwestern USA, North America [382]	restricted†	-	2006 [382]	2006	Yes
Alphacoronavirus 1	Japan, Asia [21]	restricted†	-	2007 [21]	2007	Yes
Human coronavirus 229E	Worldwide (all continents except Antarctica) [383]	wide	-	2019 [384]	2019	No
Human coronavirus NL63	Worldwide (all continents except Antarctica) [383]	wide	2005	2019 [384]	2019	No
Betacoronavirus 1	Worldwide (all continents except Antarctica) [383]	wide	2017	2019 [384]	2019	No
Human coronavirus HKU1	Worldwide (all continents except Antarctica) [383]	wide	-	2019 [384]	2019	No
Middle East respiratory syndrome-related coronavirus	Over 21 countries in Asia, North America, Europe, and Africa [385]	wide	2019	-	2019	No
Severe acute respiratory syndrome-related coronavirus	Worldwide (all continents except Antarctica) [386]	wide	2019	-	2019	Yes
Bundibugyo ebolavirus	Uganda and Congo, Africa [387]	restricted	2012	2012 [388]	2012	No
Reston ebolavirus	Virginia, USA, North America (imported cases) and Manila, Philippines, Asia (natural infection) [30]	restricted	2009	2009 [389]	2009	Yes
Sudan ebolavirus	Sudan and Uganda, Africa [390]	restricted	-	2012 [391]	2012	No
Tai Forest ebolavirus	Cote d'Ivoire, Africa [32]	restricted†	-	1995 [32]	1995	Yes

Zaire ebolavirus	Africa (Democratic Republic of the Congo, Gabon, and the Republic of the Congo and West Africa etc.) [390]	restricted	-	2018 [392]	2018	No
Marburg marburgvirus	Africa (DRC, Angola, Uganda); And imported cases in other African countries (South Africa, Kenya), Europe (Germany, Yugoslavia, Russia, Netherlands), and USA, North America [393]	wide	2018	-	2018	No
Aroa virus	Panamá and Colombia, South America [394]	restricted†	-	1973 [394]	1973	Yes
Bagaza virus	India, Asia [36]	restricted†	-	2009 [36]	2009	Yes
Banzi virus	Africa [395]	restricted	-	1965 [396]	1965	Yes
Cacipacore virus	Brazil, South America [38]	restricted	-	2011 [38]	2011	No?
Dengue virus	Worldwide (all continents except Antarctica) [386]	wide	2019	-	2019	No
Edge Hill virus	New South Wales, Australia [397]	restricted†	-	1993 [398]	1993	Yes
Gadgets Gully virus	Australia [41]	restricted†	-	1991 [41]	1991	Yes
Ilheus virus	Central and South America (Brazil, Bolivia, Trinidad, West Indies, Panamá, Colombia, Oaxaca, Mexico, French Guyana, Ecuador) [399]	restricted	-	2012 [399]	2012	No
Japanese encephalitis virus	Asia, Australia, and Angola, Africa; And imported cases in North Ameirca (USA), and Europe (UK, Belgium) [400]	wide	2019	-	2019	No
Kokobera virus	Australia and Papua New Guinea [401]	restricted†	-	1993 [402]	1993	Yes
Kyasanur forest disease virus	Asia (India, China, Saudi Arabia), Africa (Djibouti); And imported cases in Europe (Italy) [403]	wide	-	2019 [404]	2019	No
Langat virus	Malaysia and India, Asia [405]	restricted†	-	2002 [406]	2002	Yes
Louping ill virus	British Isles, Europe; And lab infections in North America (USA), Europe (Germany, UK), and Asia (Philippines) [407]	wide	-	1991 [408]	1991	Yes
Murray Valley encephalitis virus	Australia and Papua New Guinea; And imported cases in Europe (Germany), and North America (Canada) [409]	wide	2018	-	2018	No
Ntaya virus	Africa (Uganda, Cameroon, DRC, Kenya, Nigeria and Zambia etc.), Asia (India, Singapore and Borneo); And imported cases in Europe (UK) [410]	wide	-	1978 [411]	1978	Yes
Omsk hemorrhagic fever virus	Western Siberia, Russia, Europe [412]	restricted	-	2007? [412]	2007	Yes
Powassan virus	North America (the Great Lakes and Northeast regions of USA and eastern Canada) and in the Russian Far East, Asia [413]	restricted	2018	-	2018	No
Rio Bravo virus	USA, North America and serological evidence of human infection in Trinidad, South America [414]	restricted†	-	1978 [415]	1978	Yes

Saint Louis encephalitis virus	North America (USA, Canada, Mexico) and South America (Argentina) [416]	restricted	2018	-	2018	No
Tembusu virus	Malaysia and China, Asia [246]	restricted	-	2013 [246]	2013	No
Tick-borne encephalitis virus	Europe and Northern Asia; And imported cases in North America (USA and Canada) [417]	wide	2019	-	2019	No
Uganda S virus	Uganda and Nigeria, Africa [55]	restricted	-	1999 [418]	1999	Yes
Usutu virus	Africa (Senegal and Burkina Faso) and Europe (Italy, Germany, France, Croatia, Austria) [419]	restricted	2019	2019 [420]	2019	No
Wesselsbron virus	Africa [421]	restricted	-	2017 [422]	2017	No
West Nile virus	Worldwide (all continents except Antarctica) [423]	wide	2019	-	2019	No
Yellow fever virus	South America, Africa; North America, and Europe; And imported case in Asia [424]	wide	2019	-	2019	No
Zika virus	Worldwide (all continents except Antarctica) [425]	wide	2019	-	2019	No
Hepacivirus C	Worldwide (all continents except Antarctica) [426]	wide	2013	2019 [427]	2019	No
Pegivirus C	Worldwide (all continents except Antarctica) [428]	wide	2015	2019 [429]	2019	No
Pegivirus H	North America (USA), Asia (China, Iran, Vietnam), Europe (UK, Germany), Africa (Cameroon), and Australia [261]	wide	2018	2019 [430]	2019	No
Pestivirus A	Europe, Africa, North America, South America, and Oceania (New Zealand) [431]	wide	-	2006 [432]	2006	Yes
Andes orthohantavirus	South America (Chile and Argentina, Brazil) [433]	restricted	2019	2019 [434]	2019	No
Bayou orthohantavirus	Texas and Louisiana, USA, North America [435]	restricted	-	2007 [77]	2007	Yes
Black creek canal orthohantavirus	Florida, USA, North America [435]	restricted†	-	1995 [67]	1995	Yes
Choclo orthohantavirus	Panama, North America [436]	restricted	2019	-	2019	No
Dobrava-Belgrade orthohantavirus	Eastern Europe (Balkans) [437]	restricted	2017	2018 [438]	2018	No
Hantaan orthohantavirus	Eastern Asia (China, South Korea, Far East part of Russia) [439]	restricted	-	2019 [440]	2019	No
Laguna Negra orthohantavirus	South America (Western Paraguay, Bolivia, Argentina, Brazil) [441]	restricted	-	2019 [434]	2019	No
Puumala orthohantavirus	Europe (Austria, Finland, Germany, Belgium, France, Hungary, Norway and Sweden etc.) and Asia (Japan and Korea) [441 439]	restricted	2018?	2019 [442]	2019	No
Sangassou orthohantavirus	Africa (Guinea and the South Africa) [443]	restricted	-	2016 [444]	2016	No

Seoul orthohantavirus	Asia, Europe (UK, France, Belgium, and Sweden), and North America (USA) [441]	wide	-	2019 [445]	2019	No
Sin Nombre orthohantavirus	North America (USA, Canada, Mexico) [441]	restricted	2019	2019 [446]	2019	No
Thailand orthohantavirus	Asia (Thailand, Sri Lanka) [447]	restricted	-	2019 [447]	2019	No
Thottopalayam thottimvirus	Thailand [77]	restricted†	-	2007 [77]	2007	Yes
Tula orthohantavirus	Europe (Czech Republic, Switzerland, Germany, France) [441]	restricted	-	2015 [448]	2015	No
Orthohepevirus A	Worldwide (all continents except Antarctica) [449]	wide	2019	2019 [450]	2019	No
Orthohepevirus C	Asia (Hong Kong, China) and North America (Canada) [451]	restricted	2019	2019 [451]	2019	No
Crimean-Congo haemorrhagic fever orthonairovirus	Asia, Eastern Europe, and Africa [452]	wide	2019	-	2019	No
Dugbe orthonairovirus	Africa [453]	restricted	-	1996 [453]	1996	Yes
Nairobi sheep disease orthonairovirus	West Africa and Asia (India) [454]	restricted	1998	-	1998	No
Thiafora orthonairovirus	Europe (France, Germany, Netherlands and Czech Republic) [455]	restricted	-	2000 [456]	2000	Yes
Influenza A virus	Worldwide (all continents except Antarctica) [457]	wide	-	2019 [458]	2019	No
Influenza B virus	Worldwide (all continents except Antarctica) [459]	wide	-	2019 [458]	2019	No
Influenza C virus	Worldwide (all continents except Antarctica) [460]	wide	-	2019 [461]	2019	No
Dhori thogotovirus	Europe (Portugal, Russia) and Asia (India) [462]	restricted	-	1987 [463]	1987	Yes
Thogoto thogotovirus	Europe (Portugal) and Africa [462]	restricted	-	1985 [88]	1985	Yes
Avian orthoavulavirus 1	Australia, North America (USA, Canada), Asia (Palestine, Israel, Pakistan), and Europe (Germany, France, and Italy) [464]	wide	-	2013 [465]	2013	No
Hendra henipavirus	Australia [466]	restricted	2009	-	2009	Yes
Nipah henipavirus	Asia (Malaysia, Singapore, Bangladesh and India) [466]	restricted	2011	2019 [467]	2019	No
Canine morbillivirus	North America (USA), Europe (UK) [468]	restricted	-	2015 [469]	2015	No
Measles morbillivirus	Worldwide (all continents except Antarctica) [470]	wide	2019	-	2019	No
Human respirovirus 1	Worldwide (all continents except Antarctica) [471]	wide	-	2019 [472]	2019	No
Human respirovirus 3	Worldwide (all continents except Antarctica) [471]	wide	2009	2019 [472]	2019	No

Achimota pararubulavirus 2	Ghana, Africa [95]	restricted	-	2013 [95]	2013	No
Human orthorubulavirus 2	Worldwide (all continents except Antarctica) [473]	wide	-	2019 [472]	2019	No
Human orthorubulavirus 4	Worldwide (all continents except Antarctica) [473]	wide	-	2019 [472]	2019	No
Mammalian orthorubulavirus 5	North America (USA), Europe (UK) [474]	restricted	-	1989 [475]	1989	Yes
Menangle pararubulavirus	Australia [476]	restricted†	1999	1998 [99]	1998	Yes
Mumps orthorubulavirus	Worldwide (all continents except Antarctica) [477]	wide	2009	2019 [478]	2019	No
Simian orthorubulavirus	Japan [479]	restricted†	-	1990 [479]	1990	Yes
Sosuga pararubulavirus	Sudan and Uganda, Africa [102]	restricted	-	2014 [102]	2014	No
Tioman pararubulavirus	Malaysia [480]	restricted†	-	2007 [103]	2007	Yes
Bunyamwera orthobunyavirus	Africa, North America and South America [481]	wide	-	2017 [482]	2017	No
Bwamba orthobunyavirus	Africa (Mozambique, Tanzania and Uganda) [483]	restricted	-	2002 [483]	2002	Yes
California encephalitis orthobunyavirus	Worldwide (all continents except Antarctica) [484]	wide	-	2018 [485]	2018	No
Caraparu orthobunyavirus	South America (Brazil, Peru) and North America (Panama) [486]	restricted	-	2015 [487]	2015	No
Catu orthobunyavirus	South America (Brazil, Trinidad) [107]	restricted†	-	1974 [488]	1974	Yes
Guama orthobunyavirus	Brazil, South America [107]	restricted†	-	1961 [107]	1961	Yes
Guaroa orthobunyavirus	North America and South America [489]	restricted	-	2015 [490]	2015	No
Kairi orthobunyavirus	South America (Surinam, Argentina) [491]	restricted†	-	2009 [492]	2009	Yes
Madrid orthobunyavirus	Panama, North America [110]	restricted†	-	1964 [110]	1964	Yes
Marituba orthobunyavirus	Brazil, South America [107]	restricted†	-	1961 [107]	1961	Yes
Nyando orthobunyavirus	Africa (Kenya, Senegal) [493]	restricted†	-	1972 [494]	1972	Yes
Oriboca orthobunyavirus	Brazil, South America [107]	restricted†	-	1961 [107]	1961	Yes
Oropouche orthobunyavirus	South America (Brazil, Peru, and Trinidad and Tobago), and North America (Panama) [495]	restricted	2017	-	2017	No
Patois orthobunyavirus	Mexico, North America [113]	restricted†	-	1972 [113]	1972	Yes
Shuni orthobunyavirus	Africa (Nigeria and South Africa) [496]	restricted	-	2014 [497]	2014	No

Tacaiuma orthobunyavirus	South America (Amazon region, Central Brazil and Northern Argentina) [498]	restricted	-	1990 [499]	1990	Yes
Wyeomyia orthobunyavirus	South America [500]	restricted	-	1975 [501]	1975	Yes
Candiru phlebovirus	South America (Brazil and Peru) [502]	restricted	-	2004 * [502]	2004	Yes
Punta Toro phlebovirus	Panama, North America [503]	restricted	-	1981 [504]	1981	Yes
Rift Valley fever phlebovirus	Africa and Asia (the Arabian Peninsula); And imported cases in France [505]	wide	2019	-	2019	No
Sandfly fever Naples phlebovirus	Europe, Asia and Africa [506]	wide	2012	2018 [507]	2018	No
Heartland banyangvirus	USA, North America [508]	restricted		2019 [509]		No
Huaiyangshan banyangvirus	Asia (China, South Korea, North Korea, Japan) [510]	restricted	2014	2018 [511]	2018	No
Uukuniemi phlebovirus	Europe (Central Europe and Norway) [512]	restricted	-	1977 [512]	1977	Yes
Human picobirnavirus	Asia, Europe, Austrialia, North America and South America [513]	wide	-	2016 [514]	2016	No
Equine rhinitis A virus	Europe (UK, Austria) [138]	restricted†	-	2005 [138]	2005	Yes
Foot-and-mouth disease virus	Europe, Africa, South America, Asia [515]	wide	-	2007 [516]	2007	Yes
Cardiovirus A	Worldwide (all continents except Antarctica) [517]	wide	-	2015 [300]	2015	No
Cardiovirus B	Worldwide (all continents except Antarctica) [518]	wide	2008	2018 [519]	2018	No
Cosavirus A	Asia (Pakistan, Afghanistan, China, Nepal, Thailand, Japan, India), Africa (Nigeria, Tunisia, Ethiopia), South America (Bolivia, Venezuela, Brazil), and Europe (Scotland) [520]	wide	-	2019 [521]	2019	No
Cosavirus B	Asia (Pakistan, China, India), South America (Bolivia), and Africa (Tunisia) [522]	wide	-	2016 [306]	2016	No
Cosavirus D	Asia (Pakistan, Nepal, Thailand, India), Africa (Nigeria, Tunisia), and South America (Bolivia, Brazil, Venezuela) [523]	wide	-	2017 [523]	2017	No
Cosavirus E	Australia, South America (Bolivia, Venezuela, Brazil), Africa (Nigeria), and Europe (Italy) [524]	wide	-	2018 [525]	2018	No
Cosavirus F	Pakistan [129]	restricted	-	2012 [129]	2012	No
Enterovirus A	Worldwide (all continents except Antarctica) [526]	wide	-	2019 [527]	2019	No
Enterovirus B	Worldwide (all continents except Antarctica) [528]	wide	-	2019 [529]	2019	No
Enterovirus C	Worldwide (all continents except Antarctica) [530]	wide	-	2019 [531]	2019	No

Enterovirus D	Worldwide (all continents except Antarctica) [532]	wide	-	2019 [533]	2019	No
Enterovirus E	North America (USA), Europe/Asia (Turkey) [534]	restricted	-	2008 [534]	2008	Yes
Enterovirus H	Asia (Korea, Vietnam) [535]	restricted†	-	2000 [535]	2000	Yes
Rhinovirus A	Worldwide (all continents except Antarctica) [318]	wide	2016	2019 [536]	2019	No
Rhinovirus B	Worldwide (all continents except Antarctica) [318]	wide	2016	2019 [536]	2019	No
Rhinovirus C	Worldwide (all continents except Antarctica) [318]	wide	2016	2019 [536]	2019	No
Erbovirus A	Europe (Austria) [138]	restricted†	-	2005 [138]	2005	Yes
Hepatovirus A	Worldwide (all continents except Antarctica) [537]	wide	2019	2019 [538]	2019	No
Aichivirus A	Worldwide (all continents except Antarctica) [539]	wide	2011	2019 [540]	2019	No
Parechovirus A	Worldwide (all continents except Antarctica) [541]	wide	2016	2019 [541]	2019	No
Parechovirus B	Europe (Sweden, Finland) [542]	restricted	2016	2019 [543]	2019	No
Salivirus A	Worldwide (all continents except Antarctica) [544]	wide	-	2019 [521]	2019	No
Avian metapneumovirus	USA, North America [144]	restricted	-	2011 [144]	2011	No
Human metapneumovirus	Worldwide (all continents except Antarctica) [330]	wide	2017	2019 [545]	2019	No
Human orthopneumovirus	Worldwide (all continents except Antarctica) [546]	wide	2017	2019 [547]	2019	No
Colorado tick fever virus	North America (USA and Canada) and Asia (China) [548]	restricted	2018	2019 [549]	2019	No
Eyach virus	Europe (Czech Republic, France, Germany etc.) [550]	restricted†	-	1984 [551]	1984	Yes
Corriparta virus	Queensland, Australia [552]	restricted	-	1990 [553]	1990	Yes
Great Island virus	Europe (Former Czechoslovakia, Kemerovo region of Russia, Slovakia) [554]	restricted	-	1980 [148]	1980	Yes
Lebombo virus	Africa (Nigeria) [114]	restricted†	-	1975 [114]	1975	Yes
Orungo virus	Tropical Africa (Nigeria, Senegal, Gambia, Central African Republic, Gabon) [555]	restricted	-	1982 [556]	1982	Yes
Mammalian orthoreovirus	North America (USA, Canada) Asia (China), and Europe (Switzerland, France, Slovenia) [557]	wide	-	2019 [558]	2019	No
Nelson Bay orthoreovirus	Asia (Malaysia, Indonesia; And imported cases in Hong Kong China, and Miyazaki, Japan) [559]	restricted	-	2014 [559]	2014	No
Rotavirus A	Worldwide (all continents except Antarctica) [560]	wide	2018	2019 [561]	2019	No
Rotavirus B	Asia, Africa, and North America (USA) [562]	wide	2018	2017 [562]	2017	No

Rotavirus C	Worldwide (all continents except Antarctica) [155]	wide	2018	2019 [563]	2019	No
Rotavirus H	Asia (China, Bangladesh) [564]	restricted	2018	2008 [565]	2008	Yes
Banna virus	Yunnan and Xinjiang, China, Asia [566]	restricted†	2010	1992 [567]	1992	Yes
Primate T-lymphotropic virus 1	Worldwide (all continents except Antarctica) [568]	wide	2018	2019 [569]	2019	No
Primate T-lymphotropic virus 2	Worldwide (all continents except Antarctica) [570]	wide	2005	2019 [570]	2019	No
Primate T-lymphotropic virus 3	Cameroon, Africa [571]	restricted	-	2010 [572]	2010	No
Human immunodeficiency virus 1	Worldwide (all continents except Antarctica) [573]	wide	2019	2019 [574]	2019	No
Human immunodeficiency virus 2	Worldwide (all continents except Antarctica) [575]	wide	2019	2019 [576]	2019	No
Simian immunodeficiency virus	Cameroon, Africa, and USA, North America (2 lab infections) [577]	restricted†	-	2005 [577]	2005	Yes
Central chimpanzee simian foamy virus	Cameroon, Africa [164]	restricted	-	2012 [164]	2012	No
Eastern chimpanzee simian foamy virus	Africa, North America [578]	restricted	2013	2019 [579]	2019	No
Grivet simian foamy virus	Europe (Germany), North America (USA), and Africa [580]	wide†	-	1998 [580]	1998	Yes
Guenon simian foamy virus	Cameroon, Africa [164]	restricted	-	2012 [164]	2012	No
Taiwanese macaque simian foamy virus	North America (Canada), Asia (Southeast Asia, Bangladesh, China) [581]	restricted	-	2013 [582]	2013	No
Australian bat lyssavirus	Australia [583]	restricted	2019	2014 [584]	2014	No
Duvenhage lyssavirus	South Africa, Africa; And imported case in Netherlands, Europe from Kenya [585]	restricted†	-	2008 [586]	2008	Yes
European bat 1 lyssavirus	Europe (Ukraine and Russia) [587]	restricted†	-	1989 [170]	1989	Yes
European bat 2 lyssavirus	Europe (UK and Finland) [587]	restricted†	2002	2003 [588]	2003	Yes
Irkut lyssavirus	Asia (Far east Russia, China) [589]	restricted	-	2013 [172]	2013	No
Mokola lyssavirus	Nigeria, Africa [590]	restricted†	-	1972 [173]	1972	Yes
Rabies lyssavirus	Worldwide (all continents except Antarctica) [591]	wide	2019	-	2019	No
Bas-Congo tibrovirus	DRC, Africa [592]	restricted	-	2012 [175]	2012	No

Ekpoma 1 tibrovirus	Nigeria, Africa [593]	restricted	-	2015 [176]	2015	No
Ekpoma 2 tibrovirus	Nigeria, Africa [593]	restricted	-	2015 [176]	2015	No
Alagoas vesiculovirus	South America (Brazil, Colombia) [594]	restricted†	-	1987 [594]	1987	Yes
Chandipura vesiculovirus	India, Asia [595]	restricted	2016	2017 [596]	2017	No
Cocal vesiculovirus	Trinidad, South America [179]	restricted†	-	1964 [179]	1964	Yes
Indiana vesiculovirus	North America (USA, Panama) [597]	restricted†	-	1988 [597]	1988	Yes
Isfahan vesiculovirus	Asia (Iran, Turkmenistan) [598]	restricted†	-	1978 [599]	1978	Yes
Maraba vesiculovirus	Brazil, South America [182]	restricted†	-	1984 [182]	1984	Yes
New Jersey vesiculovirus	North America (USA, Panama) [181]	restricted†	-	1969 [600]	1969	Yes
Piry vesiculovirus	South America [184]	restricted†	-	1974 [184]	1974	Yes
Barmah Forest virus	Australia [601]	restricted	2019	-	2019	No
Chikungunya virus	Africa, Asia, Europe, North America and South America [602]	wide	2019	-	2019	No
Eastern equine encephalitis virus	USA, North America, and one imported case in the UK, Europe [603]	restricted	2018	2018 [603]	2018	No
Everglades virus	Southern Florida, USA, North America [604]	restricted†		1980 [604]	1980	Yes
Getah virus	Asia, Australia [605]	wide	-	1992 [359]	1992	Yes
Highlands J virus	Eastern USA, North America [190]	restricted†	-	2000 [190]	2000	Yes
Madariaga virus	South America (Peru, Venezuela) and North America (Panama, Haiti) [606]	restricted	2019	2019 [607]	2019	No
Mayaro virus	Central and South America; And imported cases in Europe and North America (USA) [608]	restricted	2019	-	2019	No
Mosso das Pedras virus	Argentina, South America [193]	restricted	-	2013 [193]	2013	No
Mucambo virus	South America (Brazil) and North America (Mexico) [194]	restricted†	-	1978 [609]	1978	Yes
Ndumu virus	South Africa, Africa [195]	restricted†	-	1961 [195]	1961	Yes
Onyong-nyong virus	Africa; And imported cases in Europe [610]	restricted	2004	2016* [611]	2016	No
Pixuna virus	South America (Brazil, Argentina, Paraguay) [612]	restricted	-	2018 [612]	2018	No
Rio Negro virus	South America (Argentina, Paraguay) [612]	restricted	-	2018 [612]	2018	No
Ross River virus	Australia, and Pacific Island countries and territories [613]	restricted	2019	-	2019	No
Semliki Forest virus	Africa and one lab infection in Europe (Germany) [614]	restricted	-	1990 [362]	1990	Yes

Sindbis virus	Eurasia, Africa and Australia [615]	wide	2010	2014 [616]	2014	No
Tonate virus	South America (French Guiana) [617]	restricted	-	2001 [617]	2001	Yes
Una virus	South America (Brazil, Argentina, Paraguay) [612]	restricted	-	2018 [612]	2018	No
Venezuelan equine encephalitis virus	North America and South America [618]	restricted	2015	2018 [619]	2018	No
Western equine encephalitis virus	North America and South America [620]	restricted	-	2018 [621]	2018	No
Whataroa virus	New Zealand, Oceania [205]	restricted†	-	1964 [205]	1964	Yes
Rubella virus	Worldwide (all continents except Antarctica) [622]	wide	2019	2019 [623]	2019	No
Hepatitis delta virus	Worldwide (all continents except Antarctica) [624]	wide	2004	2019 [625]	2019	No

DRC, Democratic Republic of the Congo; † Less than 5 reports of human infections from literature; *indirect reference

Appendix D Supplementary material for chapter 5: Predictors of COVID-19 epidemics in countries of the World Health Organisation African Region

D.1 Supplementary Methods

D.1.1 Data on COVID-19 test data quality

The TIBA Pandemic Response Unit (TIBAPRU) created a data set recording COVID-19 testing data up to 31 October for the WHO African Region. Thirty-nine (83%) out of 47 WHO Africa region countries were reporting their national COVID testing data (Figure D.1W and Table D.7). Testing data for 37 of the countries in the region has been compiled by Our World in Data (OWiD, https://ourworldindata.org/coronavirus-testing). Run by a team of economists and other social scientists at the University of Oxford, OWiD draws mostly on government websites (health ministries and public health institutes) to obtain COVID-19 testing data. Testing data for 15 countries has also been provided by the United Nations Economic Commission for Africa (UNECA, https://digitallibrary.un.org/record/3863276?ln=en). The remaining countries of the WHO African Region for which the testing data was available, publish them either on their websites or on their social media feeds. Lastly, no testing data could be found for 8 countries (79%) in the WHO Africa region (Figure D.1W and Table D.7).

The data provided from the various sources was not always consistent; in some instances, it referred to the number of tests performed while in other instances it referred to the number of individuals who have been tested. Regarding data provided by OWiD, the TIBAPRU found that often the positivity rate along with their number of tests were out of sync by a day or two, or completely different therefore the TIBAPRU went to the primary source listed and corrected it.

Data quality was placed into four categories (no data, basic, satisfactory, and good; Table 5.1). COVID-19 testing data remains basic or not available for 21 countries (44.7%). They are often published irregularly, lack clarity on whether or not they count tested samples or tested people, do not state what the share of positive and negative tests has been, whether tests with pending results count as completed, or what the breakdown of tests by region, laboratory, gender or age.

D.1.2 Data on government mitigation responses for COVID-19

The TIBA Pandemic Response Unit collected a data set for government mitigation responses to COVID-19 for WHO Africa Region (https://git.ecdf.ed.ac.uk/epigroup/covid-19/tiba_pru_measures/blob/master/TIBA_PRU_measures.xlsx) manually or using web scraping programs. All mitigation responses fall into 5 categories and 14 subcategories (Table D.8), and the normalized strictness scores were devised for each of the 14 subcategories. Based on these normalized

strictness scores, the stringency index representing policies on containment and closure were calculated using a method developed by the Oxford COVID-19 Government Response Tracker (OxCGRT),[626] i.e. by averaging the normalized strictness values of 12 subcategories of measures, excluding all governance and socio-economic measures and surveillance and testing from public health measures.

Two variables related to the stringency index were generated: area under the curve (AUC) of stringency index scores from 25 February 2020 to 31 October 2020 and stringency index score when cumulative deaths reached 0.1 per 100K population in the first wave. A threshold of 0.1 was selected because most countries reached this threshold from March to July when deaths in the WHO African region increased exponentially (Figure 2B). Several other thresholds from 0.001 to 0.2 were also explored for validation. Two countries (Burundi and Eritrea) did not reach the threshold of 0.01 till 31 October, and I filled the missing stringency indices for the two countries following the three methods below and re-run the same model three times. 1) Using the stringency index on the median date when other 42 countries reached the corresponding mortality rate threshold; 2) Calculating the median interval from the first case to the corresponding mortality rate threshold for other 42 countries, and using the stringency index on the date from the first case plus the same median interval for missing countries; 3) Using the highest stringency index for missing countries.

D.1.3 Validation of stringency index using Google mobility data

I used Google mobility data ((https://www.google.com/covid19/mobility/)) to test the quality of stringency index collected by the TIBA Pandemic Response Unit. Google mobility data, as a proxy measure for social distancing, help public health workers understand changes in mobility patterns in response to various non-pharmaceutical interventions policies during the COVID-19 pandemic, including bans on large gatherings, school and university closures, domestic mobility restrictions, and physical isolation. It provides information on percent change of movement over time compared to a baseline by a number of specific categories including retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential. The baseline is the median value for the corresponding day of the week during the 5-week period Jan 3 – Feb 6, 2020. The residential percent change of mobility was used to validate stringency index for the following reasons: 1) The residential category has a high correlation coefficient with the other five categories of mobility (Figure D.2); 2) The location accuracy and the understanding of residential places varies less across regions than other categories, so the comparison between countries will cause less bias; 3) the intention of many mitigation response measures is to encourage people to stay in their residence.

As of 15 November 2020, 25 out of 47 WHO African countries had data for mobility (Figure D.3), which are mostly concentrated in the southern and western regions. Residential mobility is available for 24 out of the 25 countries (except Guinea-Bissau).

Time series plot of stringency index against residential mobility was shown in Figure D.4). I used generalised additive mixed model (GAMM) to estimate the relationships between the stringency index and residential mobility over time. I fitted the residential mobility as a spline function of stringency index $s(Stringency\ index)$ and a spline function of day of the year $s(doy)$ which was used to control for the temporal trend. The temporal relationship between residential mobility and stringency index can be different between countries, so I also introduced a spline function of country $s(country, bs = 're')$ as random intercepts and country and day of the year $(country, doy, bs = 're')$ as random slopes. The model was expressed as follows.

$$g(Y_{ij}) = s(Stringency\ index) + s(doy) + s(country, bs = 're') + s(country, doy, bs = 're') + \varepsilon_{ij}$$

where Y_{ij} denoted the residential mobility for the *i*th individual in the *j*th country, and ε_{ij} is the random noise. $s()$ indicated penalized spline function. $bs = 're'$ indicated the basis function is a random effect structure (basis coefficients are penalized by a ridge penalty to control the degree of smoothness). I used the default parameter settings from the R package mgcv for penalized spline function.

After controlling for the temporal trend and random effect, the stringency index was non-linearly associated with the residential mobility ($p<0.0001$), with an effective degree of freedom of 8.66 (Figure D.5A). The doy was also non-linearly associated with the residential mobility ($p<0.0001$), with an effective degree of freedom of 8.86 (Figure D.5B). The R^2 of the model is 0.77 and the

explained deviance is 77.5%. Using the fitted model, I predicted the values of residential mobility, and they have a good consistency with the actual values (Figure D.6).

D.2 Supplementary Figures

Figure D.1 Distribution of predictors of COVID-19 epidemics in the WHO African Region

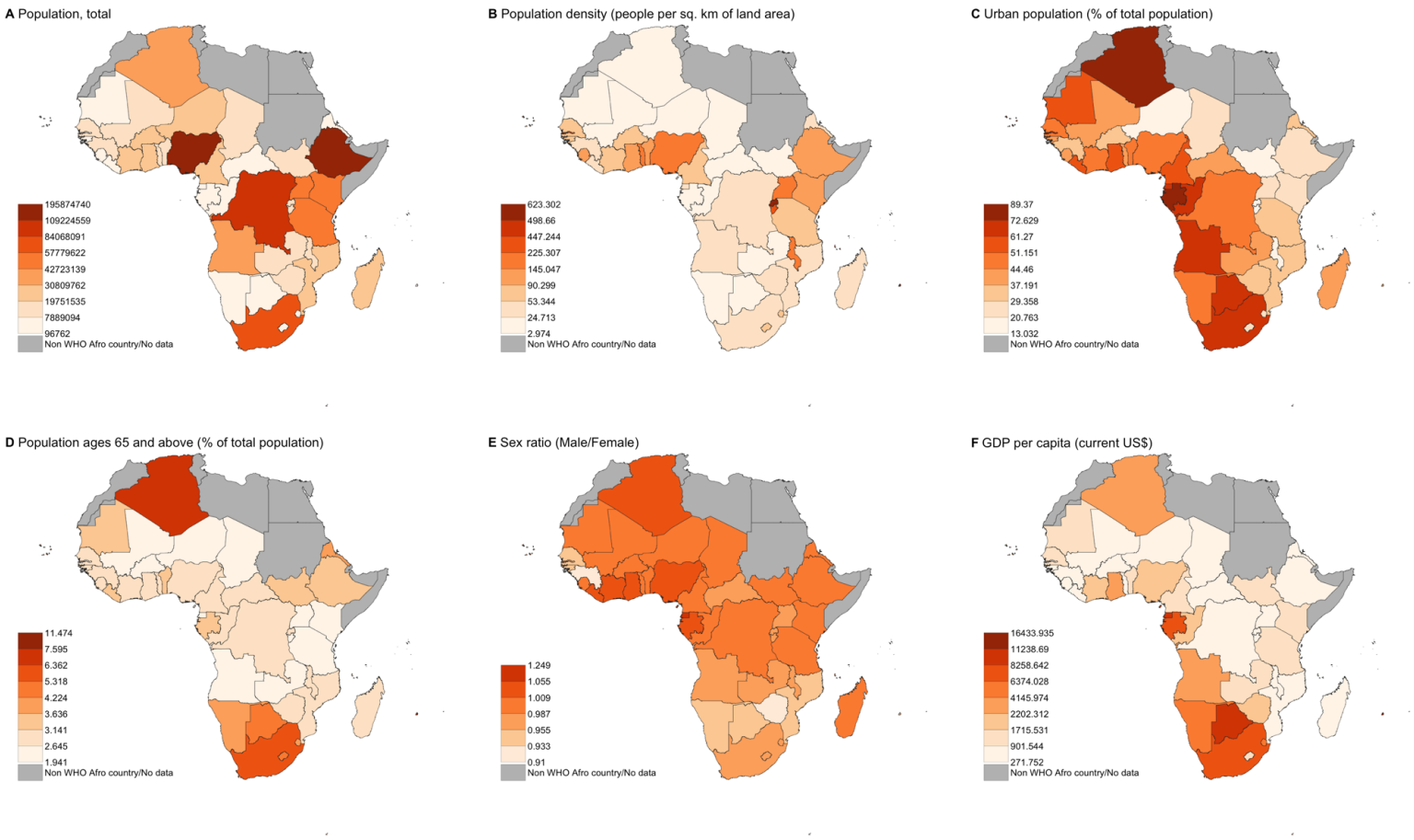
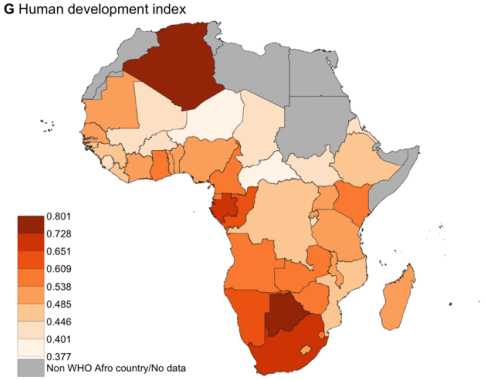
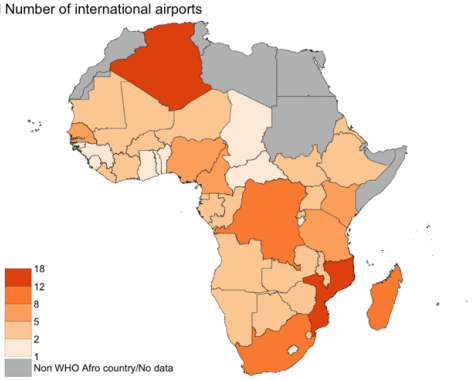


Figure D.1 (continued 1)

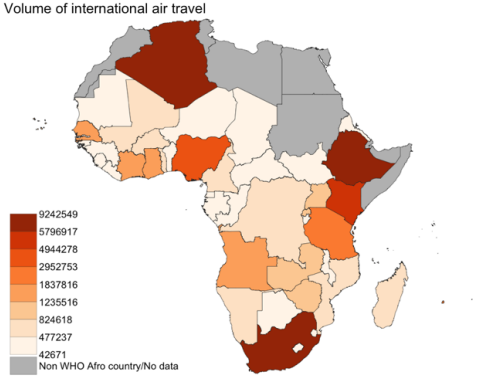
G Human development index



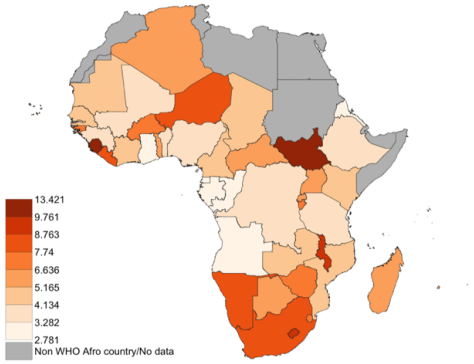
H Number of international airports



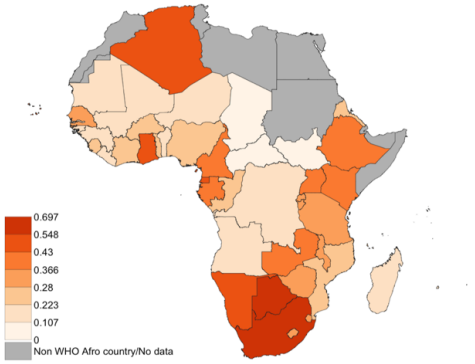
I Volume of international air travel



J Current health expenditure (% of GDP)



K Infectious disease resilience index



L DALY rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases

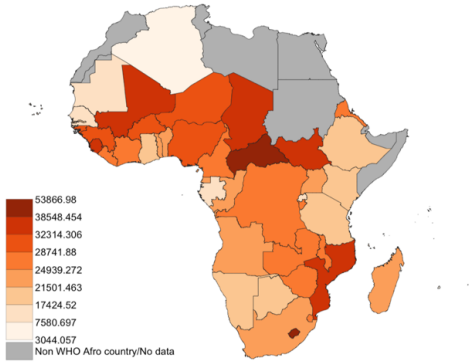
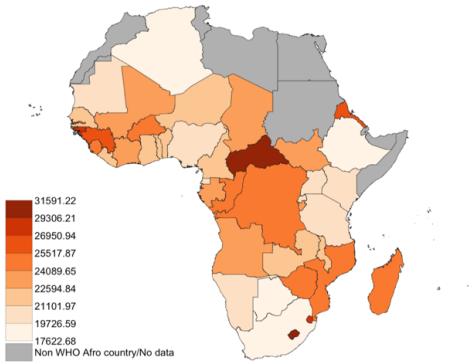
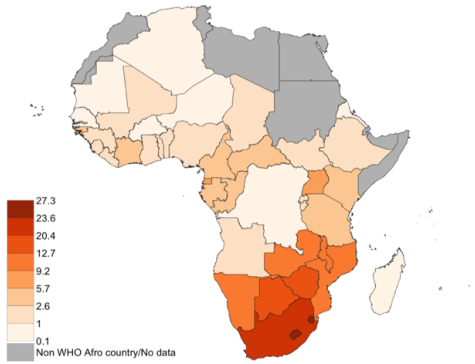


Figure D.1 (continued 2)

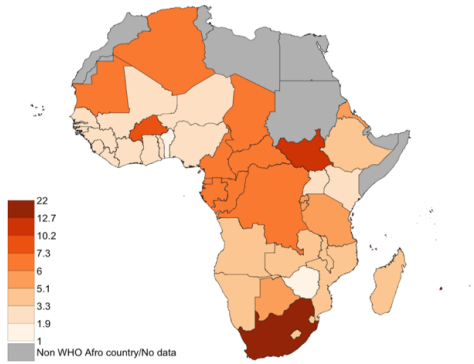
M DALY rates per 100,000 individuals from non-communicable diseases



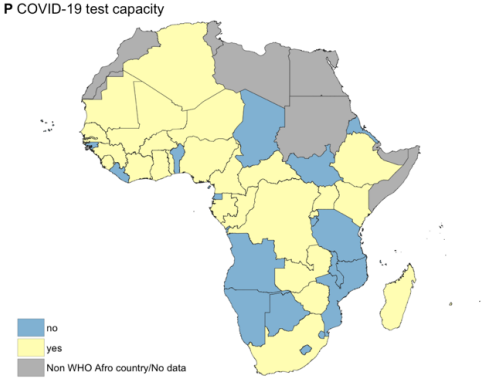
N Prevalence of HIV, total (% of population ages 15-49)



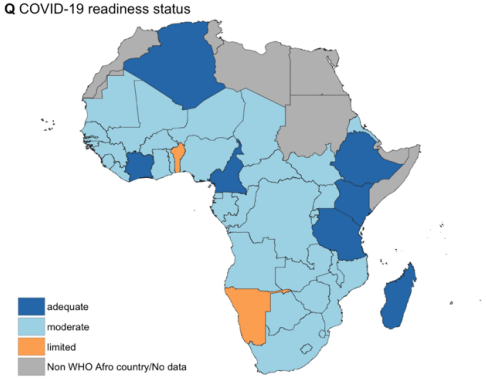
O Diabetes prevalence (% of population ages 20 to 79)



P COVID-19 test capacity



Q COVID-19 readiness status



R Number of borders

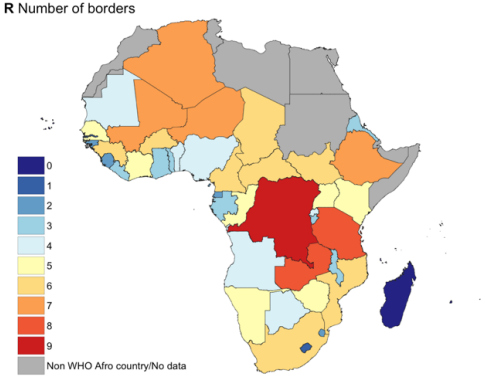


Figure D.1 (continued 3)

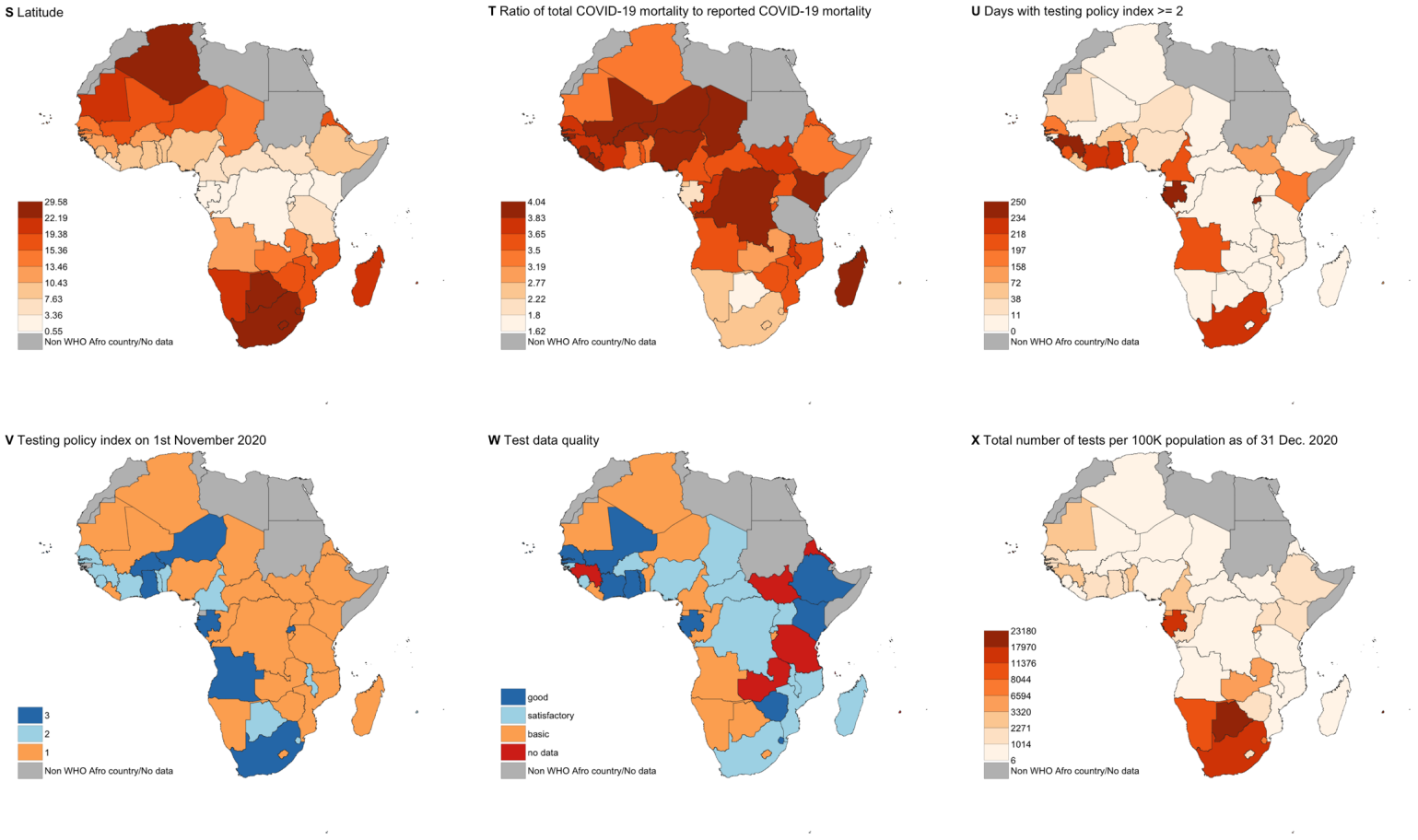
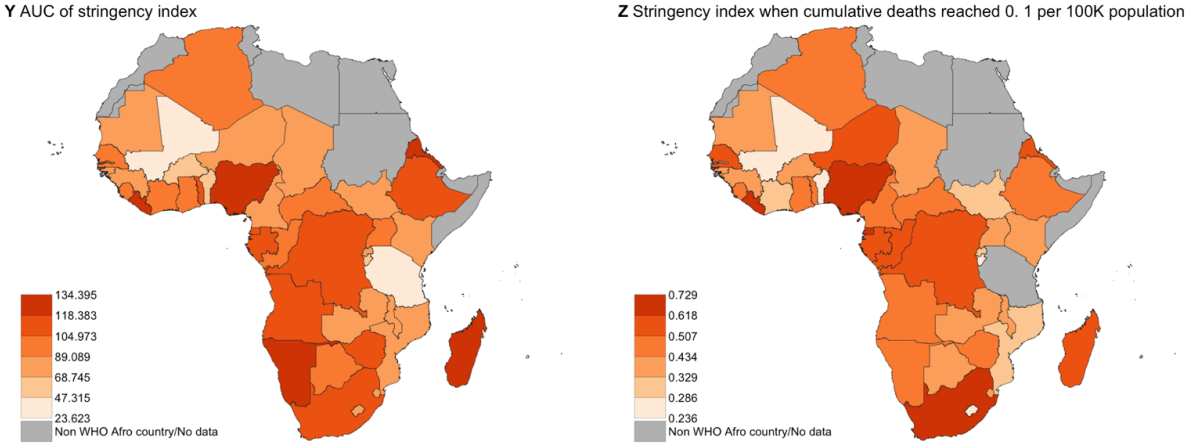


Figure D.1 (continued 4)



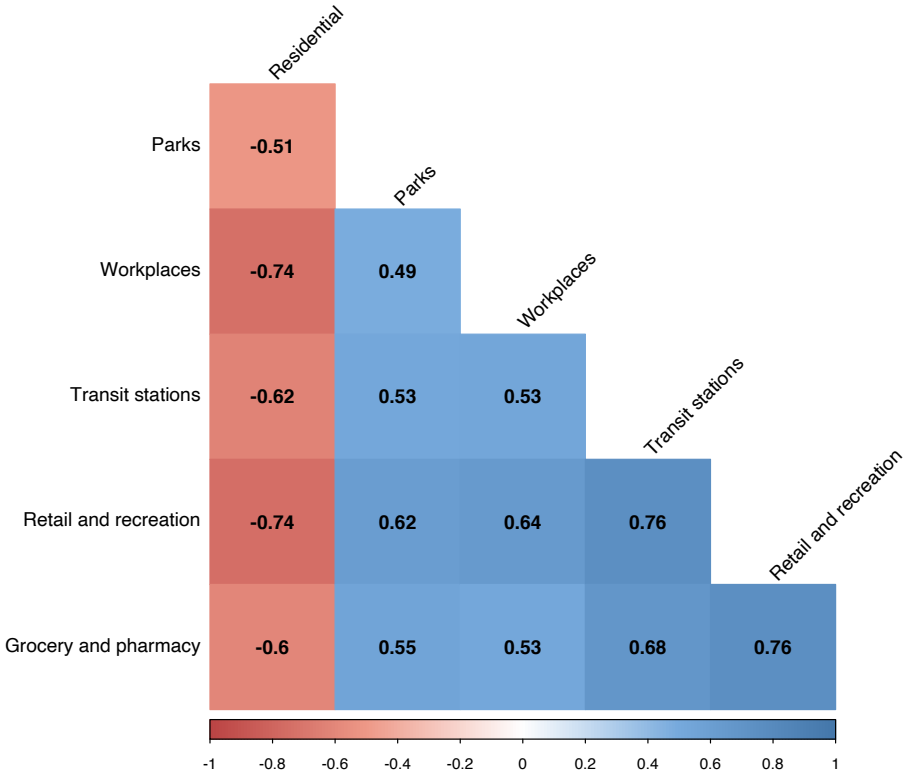


Figure D.2 Correlation matrix for the six categories of Google mobility

Positive correlations are displayed in blue and negative correlations in red colour. Colour intensity is proportional to the correlation coefficients.

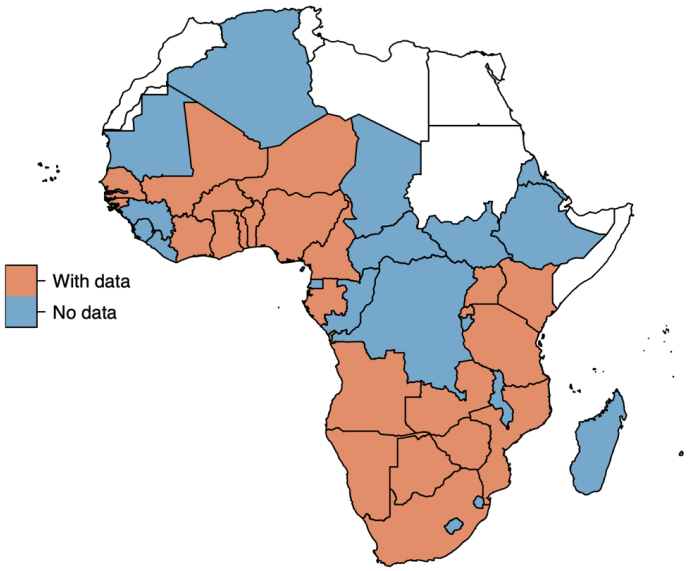


Figure D.3 Data availability for Google mobility in the WHO African Region

Orange indicates countries with Google mobility data, and blue indicates countries without Google mobility data.

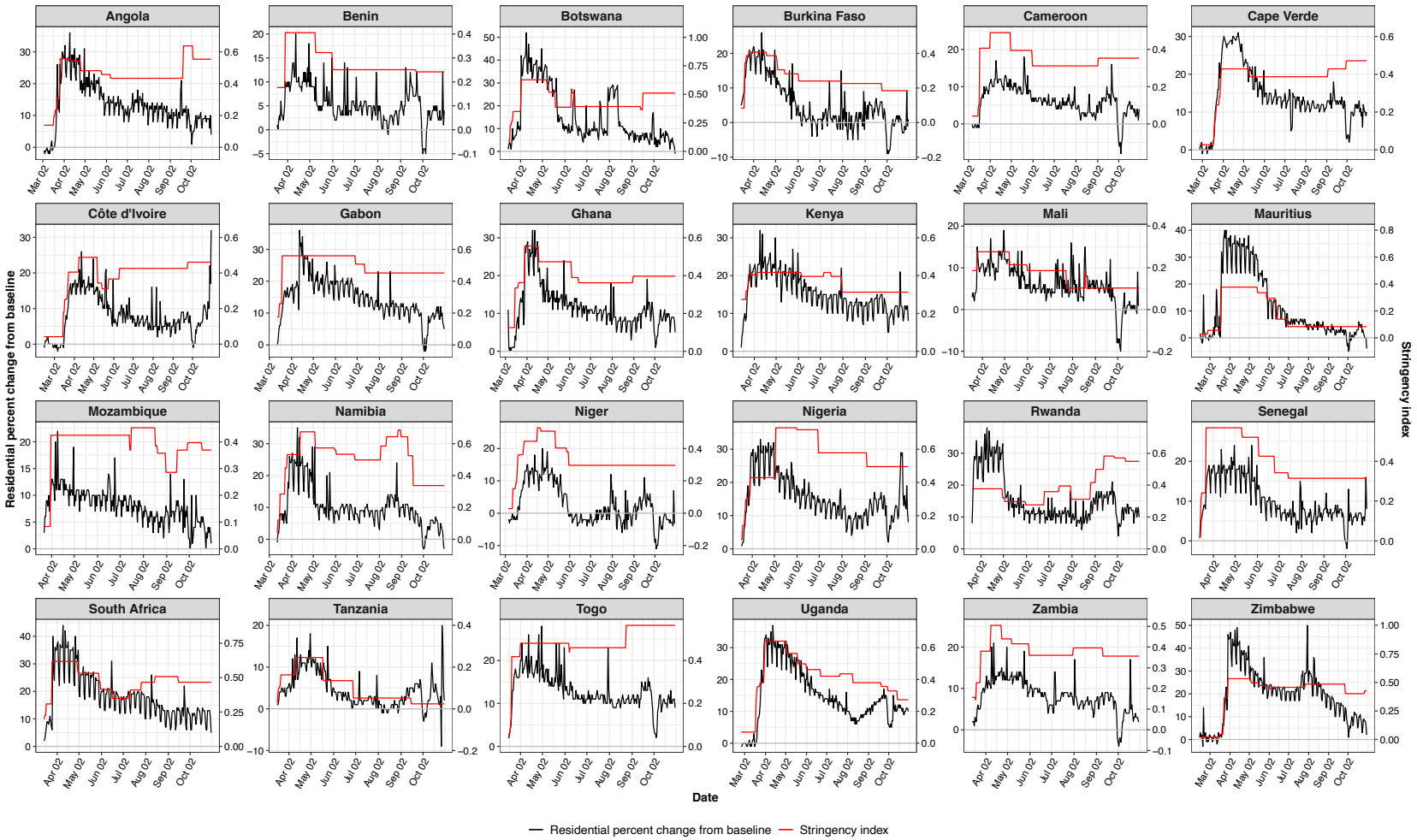
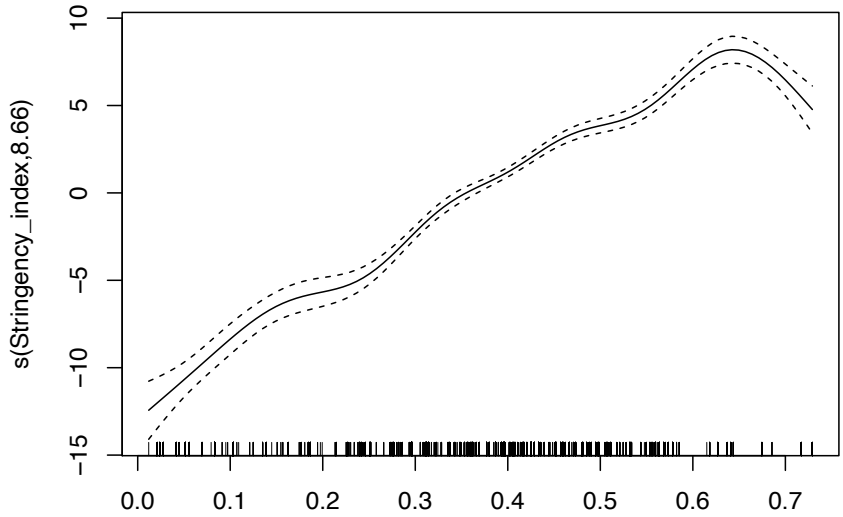


Figure D.4 Residential percent change from baseline and stringency index over time in 24 countries of the WHO African Region

A Stringency index



B Day of the year

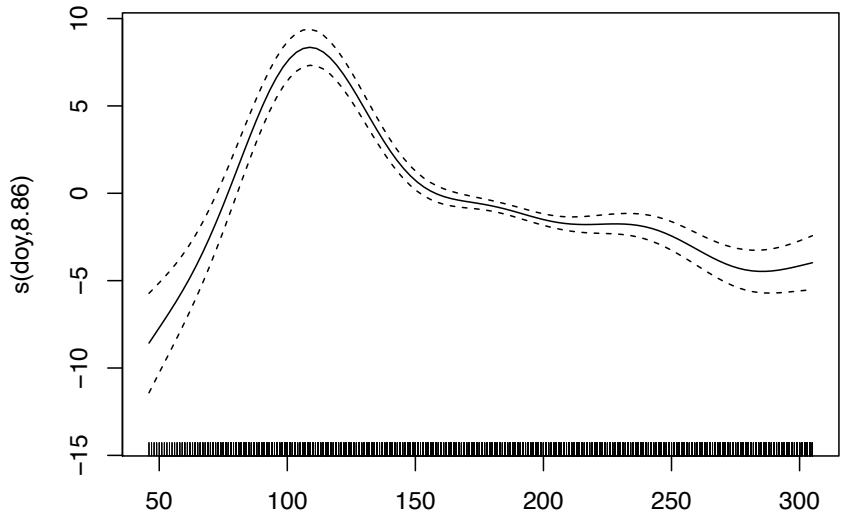


Figure D.5 Partial effect plots of stringency index (A) and day of the year (doy) (B) on residential mobility

The region between dashed lines represents the 95% confidence levels. The rugs at the bottom showed the values of stringency index and doy, respectively.

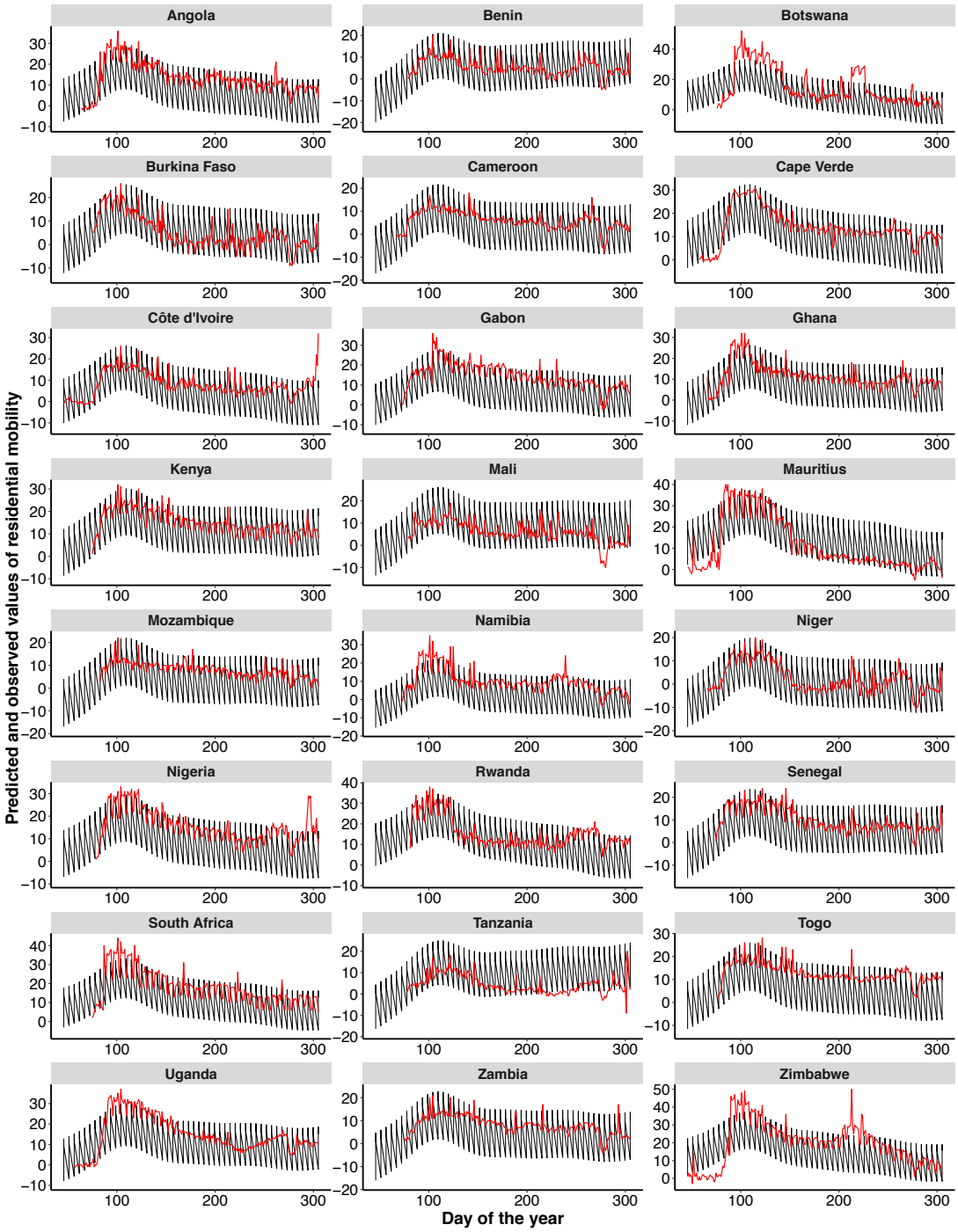


Figure D.6 Comparison of the predicted (in black) and observed (in red) values of residential mobility over time

D.3 Supplementary Tables

Table D.1 Hazard ratios and 95% confidence intervals for predictors of timing of the first COVID-19 case in univariable and multivariable cox regression model

Category	Variable	Univariable model	Multivariable model
Demographics	Population, total	2.002 (1.408-2.847)	
	Population density (people per sq. km of land area)	0.826 (0.608-1.122)	
	Urban population (% of total population)	1.265 (0.963-1.661)	1.404 (1.011-1.949)
	Population ages 65 and above (% of total population)	1.028 (0.754-1.402)	
	Sex ratio (Male/Female)	1.219 (0.900-1.652)	
Socioeconomic	GDP per capita (current US$)	1.198 (0.906-1.583)	
	Human development index	1.212 (0.904-1.626)	
Travel	Number of international airports	1.381 (1.007-1.892)	1.475 (1.017-2.139)
	Volume of international air travel	1.806 (1.358-2.403)	1.519 (1.095-2.106)
Healthcare	Current health expenditure (% of GDP)	0.603 (0.428-0.849)	0.754 (0.539-1.054)
	Infectious disease resilience index	1.253 (0.908-1.729)	
Readiness	COVID-19 test capacity, yes	3.643 (1.844-7.200)	3.861 (1.829-8.151)
	COVID-19 readiness status, adequate	2.628 (1.138-6.067)	
Geography	Number of borders, above 0	1.974 (0.817-4.773)	2.866 (1.122-7.316)
	Latitude	0.873 (0.646-1.179)	

Table D.2 Risk ratios and 95% confidence intervals for predictors of per capita COVID-19 mortality in the first wave in univariable Poisson generalized linear mixed model

Category	Variable	Univariable model
Demographics	Population density (people per sq. km of land area)	0.751 (0.546-1.032)
	Urban population (% of total population)	1.626 (1.221-2.167)
	Population ages 65 and above (% of total population)	1.192 (0.859-1.654)
	Sex ratio (Male/Female)	1.171 (0.866-1.582)
Socio-economic	GDP per capita (current US$)	1.520 (1.054-2.192)
	Human development index	1.552 (1.144-2.105)
Travel	Number of international airports	1.135 (0.835-1.543)
	Volume of international air travel	1.395 (1.050-1.855)
Healthcare	Current health expenditure (% of GDP)	1.026 (0.752-1.399)
	Infectious disease resilience index	1.653 (1.247-2.192)
Co-morbidities	DALY rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases	0.760 (0.556-1.039)
	DALY rates per 100,000 individuals from non-communicable diseases	0.908 (0.668-1.233)
	Prevalence of HIV, total (% of population ages 15-49)	1.397 (1.044-1.870)
	Diabetes prevalence (% of population ages 20 to 79)	1.052 (0.766-1.444)
Readiness	COVID-19 test capacity, yes	0.812 (0.424-1.557)
	COVID-19 readiness status, adequate	1.239 (0.508-3.019)
Geography	Number of borders, above 0	0.840 (0.288-2.456)
	Latitude	1.467 (1.092-1.970)

Table D.3 Risk ratios and 95% confidence intervals for predictors of per capita COVID-19 mortality in the first wave in multivariable Poisson generalized linear mixed model

Variable	Multivariable model (AICc=624.94)	Add day with testing policy index >= 2 (AICc=627.45)	Add testing data quality (AICc=627.46)	Add total number of tests per 100K population as of 31 Dec. 202(AICc=626.50)	Add AUC of stringency index (AICc=627.65)	Add stringency index when cumulative deaths reached 0.1 per 100K population (AICc=626.57)
Urban population (% of total population)	1.607 (1.252-2.063)	1.593 (1.238-2.049)	1.628 (1.261-2.102)	1.722 (1.306-2.271)	1.594 (1.218-2.086)	1.520 (1.163-1.986)
Volume of international air travel	1.314 (1.040-1.660)	1.308 (1.035-1.653)	1.300 (1.024-1.649)	1.324 (1.050-1.668)	1310 (1.034-1.660)	1.268 (0.997-1.612)
Prevalence of HIV, total (% of population ages 15-49)	1.402 (1.104-1.780)	1.411 (1.110-1.793)	1.406 (1.107-1.784)	1.469 (1.144-1.886)	1.402 (1.104-1.780)	1.430 (1.127-1.816)
Days with testing policy index >= 2, above median		1.128 (0.689-1.847)				
Test data quality, satisfactory to good			1.128 (0.676-1.879)			
Total number of tests per 100K population as of 31 Dec. 2020				0.846 (0.626-1.143)		
AUC of stringency index					1.027 (0.743-1.420)	
Stringency index when cumulative deaths reached 0.1 per 100K population						1.116 (0.877-1.549)

Table D.4 Odds ratios and 95% confidence intervals for predictors of outcome with respect to AUC of stringency index and COVID-19 mortality rate in univariable multinomial logistic regression model

Covid-19 readiness status was excluded from the model because there was no country with adequate Covid-19 readiness level in the reference low/low level, and putting it in the model will generate super wide 95%CIs.

Category	Variable	Univariable model		
		High/Low	Low/High	High/High
Demographics	Population density (people per sq. km of land area)	0.912 (0.437-1.903)	0.578 (0.200-1.667)	0.136 (0.014-1.327)
	Urban population (% of total population)	4.532 (1.208-16.993)	4.392 (1.176-16.403)	9.898 (2.348-41.731)
	Population ages 65 and above (% of total population)	0.345 (0.053-2.228)	0.778 (0.291-2.078)	1.004 (0.488-2.065)
	Sex ratio (Male/Female)	1.619 (0.451-5.811)	0.941 (0.245-3.615)	2.280 (0.638-8.152)
Socioeconomic	GDP per capita (current US$)	1.265 (0.307-5.212)	1.113 (0.254-4.882)	2.653 (0.768-9.167)
	Human Development Index	2.349 (0.713-7.738)	2.746 (0.832-9.068)	4.233 (1.263-14.184)
Travel	Number of international airports	1.702 (0.605-4.783)	1.226 (0.396-3.792)	1.685 (0.606-4.684)
	Volume of international air travel	1.783 (0.294-10.815)	1.831 (0.305-10.990)	3.342 (0.613-18.228)
Healthcare	Current health expenditure (% of GDP)	0.944 (0.424-2.103)	0.716 (0.297-1.728)	0.716 (0.304-1.687)
	Infectious disease resilience index	2.202 (0.676-7.172)	3.666 (1.083-12.414)	5.866 (1.646-20.903)
Co-morbidities	DALY rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases	0.710 (0.283-1.779)	0.564 (0.218-1.463)	0.604 (0.240-1.518)
	DALY rates per 100,000 individuals from non-communicable diseases	0.894 (0.376-2.127)	1.207 (0.534-2.727)	0.853 (0.362-2.005)
	Prevalence of HIV, total (% of population ages 15-49)	1.198 (0.353-4.066)	2.127 (0.731-6.193)	1.777 (0.604-5.227)
	Diabetes prevalence (% of population ages 20 to 79)	0.682 (0.268-1.739)	0.447 (0.126-1.590)	0.870 (0.418-1.809)
Readiness	COVID-19 test capacity, yes	1.945 (0.322-11.756)	1.250 (0.221-7.084)	2.222 (0.375-13.179)
Geography	Latitude	0.700 (0.279-1.758)	1.309 (0.552-3.105)	0.869 (0.366-2.065)

Table D.5 Odds ratios and 95% confidence intervals for predictors of outcome with respect to AUC of stringency index and COVID-19 mortality rate in multivariable multinomial logistic regression model

Variable	Multivariable model (AICc= 111.61)			Add days with testing policy index >= 2 (AICc= 116.55)			Add test data quality (AICc= 121.24)			Add total number of tests per 100K population as of 31 Dec. 2020 (AICc= 113.66)		
	High/Low	Low/High	High/High	High/Low	Low/High	High/High	High/Low	Low/High	High/High	High/Low	Low/High	High/High
Urban population (% of total population)	10.596 (1.441-77.928)	9.463 (1.276-70.204)	18.102 (2.260-144.997)	18.715 (2.187-160.166)	12.726 (1.587-102.050)	32.274 (3.483-299.082)	13.462 (1.627-111.404)	10.701 (1.307-87.625)	21.277 (2.401-188.595)	10.075 (1.461-69.487)	9.193 (1.295-65.249)	25.801 (3.090-215.412)
Infectious disease resilience index	4.508 (0.780-26.051)	8.015 (1.311-48.995)	10.136 (1.566-65.592)	6.969 (0.999-48.624)	11.907 (1.618-87.641)	15.274 (1.974-118.220)	4.151 (0.738-23.361)	7.447 (1.240-44.733)	9.388 (1.478-59.627)	4.165 (0.704-24.633)	7.941 (1.145-55.091)	16.323 (2.047-130.158)
Days with testing policy index >= 2, above median				0.076 (0.006-0.965)	0.313 (0.029-3.403)	0.064 (0.004-0.971)						
Test data quality, satisfactory to good							3.732 (0.348-40.057)	2.088 (0.200-21.836)	2.494 (0.205-30.406)			
Total number of tests per 100K population as of 31 Dec. 2020										0.387 (0.032-4.674)	0.348 (0.026-4.761)	0.142 (0.009-2.211)

Table D.6 Risk ratios and 95% confidence intervals for predictors of per capita COVID-19 mortality in the second wave in univariable Poisson generalized linear mixed model

Category	Variable	RR (95%CI)
Demographics	Population density (people per sq. km of land area)	0.779 (0.443-1.368)
	Urban population (% of total population)	1.109 (0.651-1.887)
	Population ages 65 and above (% of total population)	0.858 (0.473-1.554)
	Sex ratio (Male/Female)	0.739 (0.450-1.213)
Socioeconomic	GDP per capita (current US$]	1.367 (0.719-2.597)
	Human Development Index	1.945 (1.168-3.239)
Healthcare	Current health expenditure (% of GDP)	1.091 (0.654-1.823)
	Infectious disease resilience index	3.026 (1.999-4.582)
Co-morbidities	DALY rates per 100,000 individuals from communicable, neonatal, maternal & nutritional diseases	0.584 (0.349-0.978)
	DALY rates per 100,000 individuals from non-communicable diseases	0.632 (0.386-1.034)
	Prevalence of HIV, total (% of population ages 15-49)	2.767 (1.864-4.109)
	Diabetes prevalence (% of population ages 20 to 79)	0.898 (0.516-1.561)
Geography	Number of neighbours	1.247 (0.204-7.609)
	Latitude	2.370 (1.520-3.695)
First wave mortality	Per 100K population mortality in the first wave	2.221 (1.435-3.438)
First wave intervention	AUC of stringency index in the first wave	1.098 (0.577-2.088)
First wave testing	Testing policy on November 1st, above 2	0.998 (0.353-2.822)
	Test data quality in the first wave, satisfactory to good	1.038 (0.360-2.990)
	Total number of tests per 100K population as of 31 Dec. 2020	1.322 (0.760-2.300)

Table D.7 COVID-19 testing data in detail in the WHO African Region

Country	Source Quality	Type of Data	Data release frequency	Source
Algeria	Basic	N/A	Infrequently	http://www.aps.dz/sante-science-technologie/111985-coronavirus-320-nouveaux-cas-191-guerisons-et-10-deces
Angola	Basic	Samples analysed	Infrequently	OWiD, supplemented by https://www.minsa.gov.ao/TodasNoticias.aspx and https://governo.gov.ao/ao/noticias/taxa-de-positividade-e-de-6-6-por-cento/
Benin	Basic	Tests	Infrequently	https://www.gouv.bj/coronavirus/#mesures
Botswana	Basic	Tests	Infrequently	https://www.facebook.com/BotswanaGovernment/posts/3191898507559452
Burkina Faso	Satisfactory	Tests	Daily	OWiD, supplemented by https://www.sante.gov.bf/accueil
Burundi	Basic	Tests	Infrequently	http://minisante.bi/?p=735
Cabo Verde	Good		Daily	OWiD, supplemented by https://covid19.cv/conferencia-de-imprensa-sobre-a-covid-19-de-29-de-julho-de-2020/
Cameroon	Basic	Tests	Infrequently	https://twitter.com/drmanaouda
Central African Republic	Satisfactory	People tested	Daily	https://twitter.com/MSPCentrafrique?lang=en
Chad	Satisfactory	Samples analysed	Daily	https://www.facebook.com/ministeresantetchad/ and https://sante-tchad.org/communique-de-presse/
Comoros	No Data			
Congo	Basic	Tests	Infrequently	http://sante.gouv.cg/wp-content/uploads/2020/09/SITREP-N-100-COVID-19-CONGO-21-09-2020-1.pdf
Côte d'Ivoire	Good	Samples analysed	Daily	OWiD, supplemented by https://www.facebook.com/Mshpci
Democratic Republic of the Congo	Satisfactory	Samples analysed	Daily	OWiD, supplemented by https://riposte-epidemie-rdc.info/bulletins.php
Equatorial Guinea	Basic	Tests	Infrequently	https://www.guineaecuatorialpress.com/imgdb/2020/pyn9_INFORMEPERIODICO31072020Spanishfinal.pdf https://www.worldometers.info/coronavirus/country/equatorial-guinea/
Eritrea	No Data			
Eswatini	Good	Tests	Daily	http://www.gov.sz/index.php/covid-19-corona-virus/covid-19-press-statements-2020 and https://twitter.com/EswatiniGovern1
Ethiopia	Good	Tests	Daily	https://twitter.com/EPHIEthiopia

Gabon	Good	Tests	Daily	OWiD, supplemented by https://twitter.com/Covid19GOUVGA
Gambia	Good	Tests	Daily	http://www.moh.gov.gm/covid-19-report/
Ghana	Good	Samples analysed	Daily	https://www.ghanahealthservice.org/covid19/archive.php
Guinea	No Data			
Guinea-Bissau	Satisfactory	Tests	Daily	OWiD, supplemented by https://covid19gb.com/boletins/
Kenya	Good	Samples analysed	Daily	OWID, supplemented by https://twitter.com/MOH_Kenya
Lesotho	Basic	Samples analysed	Infrequently	OWID, supplemented by https://www.gov.ls/wp-content/uploads/2020/06/Update-on-Covid-19-status.pdf and https://www.gov.ls/official-statements-2/
Liberia	Basic	Samples analysed	Infrequently	http://moh.gov.lr/documents/reports/covid-19/2020/covid-19-sitrep-vol-17-2/
Madagascar	Satisfactory	Tests	Daily	http://www.sante.gov.mg/ministere-sante-publique/?s=COVID-19%3A+Situation
Malawi	Satisfactory	Tests	Daily	https://www.facebook.com/malawimoh
Mali	Good	Samples analysed	Daily	http://www.sante.gov.ml/index.php/actualites/communiques?start=84
Mauritania	Basic	Tests	Infrequently	http://www.sante.gov.mr/?lang=fr
Mauritius	No Data			
Mozambique	Satisfactory	Tests	Daily	http://www.misau.gov.mz/index.php/covid-19-boletins-diarios?limitstart=0
Namibia	Basic	Samples analysed	Infrequently	http://www.mhss.gov.na/home
Niger	Basic	People tested	Infrequently	https://reliefweb.int/report/niger/niger-coronavirus-covid-19-situation-report-08-8-21-june-2020
Nigeria	Satisfactory	Samples analysed	Daily	OWID, supplemented by https://ncdc.gov.ng/diseases/sitreps/?cat=14&name=An%20update%20of%20COVID-19%20outbreak%20in%20Nigeria
Rwanda	Satisfactory	Tests	Daily	https://www.rbc.gov.rw/index.php?id=717
São Tomé and Príncipe	Satisfactory	Tests	Daily	https://covid.ms.gov.st/st/

Senegal	Good	Tests	Daily	http://www.sante.gouv.sn/actualites
Seychelles	No Data			
Sierra Leone	Satisfactory	Tests	Daily	http://dhse.gov.sl/?s=SITREP
South Africa	Satisfactory	Tests	Daily	https://www.nicd.ac.za/diseases-a-z-index/covid-19/surveillance-reports/
South Sudan	No Data			
Togo	Good	People tested	Daily	OWiD, supplemented by: https://twitter.com/Covid19TG
Uganda	Satisfactory	Samples analysed	Daily	OWiD, supplemented by: https://twitter.com/MinofHealthUG
United Republic of Tanzania	No Data			
Zambia	No Data			
Zimbabwe	Good	Tests	Daily	http://www.mohcc.gov.zw/

OWiD, Our world in Data https://ourworldindata.org/coronavirus-testing

Table D.8 Sub-categories of measures in response to COVID-19 and their strictness scales in the WHO African Region

Category	Subcategory	Strictness Level	Score
Governance and socio-economic measures	Emergency administrative structures activated or established	None	0
		In place	1
Lockdown	Full lockdown	No lockdown	0
		Citizens required to stay at home nationwide	1
		Partial lockdown only in targeted areas of the country	2
	Partial lockdown	No partial lockdown	0
		Relaxed/Non-strict nationwide lockdown	1
		Full lockdown only in targeted areas of the country	2
Movement restrictions	Border closure	Borders open as normal	0
		Borders shut to some targeted countries, but not all affected countries	1
		Borders shut to all affected countries or a large list of countries	2
		Borders closed completely (except for repatriation, cargo and aid)	3
	Curfews	No specific curfew present	0
		Targeted curfew for some areas after 8pm	1
		Targeted curfew for some areas before 8pm	2
		Nationwide curfew after 8pm	3
		Nationwide curfew before 8pm	4
	Domestic travel restrictions	Able to travel freely throughout the whole country	0
		Very long-distance travel (e.g. domestic flights) suspended	1
		Long distance travel (e.g. between provinces) suspended	2
		Medium distance travel (outside your city) suspended	3
		Short distance travel (the next town/village) suspended	4
Public health measures	Isolation and quarantine policies	No policy	0
		Travellers from affected countries	1
		All travellers	2
		Any possible exposure	3
	Public health recommendations	None	0
		Recommendations in place	1
		No requirement	0

	Requirement to wear protective gear in public	Requirement for specific places only	1
		Requirement for most public places	2
	Surveillance and testing	No surveillance and monitoring	0
		Targeted/infrequent population surveillance	1
		Mass population surveillance	2
		Mass population testing	3
	Border health checks	No border health checks	0
		Basic test (e.g. temperature) for those from affected countries	1
		Basic (e.g. temperature) for all arrivals	2
		Advanced test for those from specific countries/Health document required from specific countries	3
		Advanced test for all arrivals/Health document required from all arrivals	4
Social distancing	Closure of businesses and public services	No closures	0
		Some businesses/services limited in hours	1
		Entertainment based businesses/services closed	2
		Entertainment based businesses/services closed, others limited in hours	3
		Entertainment and Hospitality closed	4
		Entertainment and Hospitality closed, others limited in hours	5
		All non-essential businesses closed	6
	Limit public gatherings	No limit on public gatherings	0
		Gatherings over 1000 suspended	1
		Gatherings over 100 suspended	2
		Non-religious gatherings over 50 suspended	3
		Gatherings over 50 suspended	4
		Non-religious gatherings over 10 suspended	5
		Gatherings over 10 suspended	6
		Restriction on gatherings less than 10	7
	Schools closure	All schools open	0
		Some levels of school closed	1
		All levels of school closed with exceptions for examinations	2
		All levels of school closed	3

Appendix E Publications

E.1 Publications related to this thesis

1. **Zhang F**, Karamagi H, Nsenga N, Nanyunja M, Karinja M, Amanfo S, Chase-Topping M, Calder-Gerver G, McGibbon M, Huber A, Wagner-Gamble T, Guo CG, Haynes S, Morrison A, Ferguson M, Awandare GA, Mutapi F, Yoti Z, Cabore J, Moeti MR, Woolhouse MEJ. Predictors of COVID-19 epidemics in countries of the World Health Organisation African Region. ***Nature Medicine*** 2021 Sep 3. Epub ahead of print.

2. **Zhang F**, Chase-Topping M, Guo CG, van Bunnik B, Brierley L, Woolhouse MEJ. Global discovery of human-infective RNA viruses: a modelling analysis. ***PLOS Pathogens*** 2020; 16(11): e1009079.

3. **Zhang F**, Chase-Topping M, Guo CG, Woolhouse MEJ. Predictors of human-infective RNA virus discovery in the United States, China and Africa, an ecological study (bioRxiv Preprint) doi: 10.1101/2021.09.13.460031.

E.2 Other publications

1. van Bunnik BAD, Morgan ALK, Bessell PR, Calder-Gerver G, **Zhang F**, Haynes S, Ashworth J, Zhao S, Cave RNR, Perry MR, Lepper HC, Lu L, Kellam P, Sheikh A, Medley GF, Woolhouse MEJ. Segmentation and shielding of the most vulnerable members of the population as elements of an exit strategy from COVID-19 lockdown. ***Philosophical transactions of the Royal***

Society of London. Series B, Biological sciences. 2021; 376(1829):20200275.

2. Guo CG, Tian L, **Zhang F**, et al. Associations of seasonal variations and meteorological parameters with incidences of upper and lower gastrointestinal bleeding. ***Journal of Gastroenterology and Hepatology*** 2021 Jul 21. Epub ahead of print.

3. Guo CG, **Zhang F**, Wu JT, et al. Divergent trends of hospitalizations for upper and lower gastrointestinal bleeding based on population prescriptions of aspirin, proton pump inhibitors and Helicobacter pylori eradication therapy Trends of upper and lower gastrointestinal bleeding. ***United European Gastroenterology Journal*** 2021 May 6. Epub ahead of print.

4. Guo CG, Cheung KS, **Zhang F**, et al. Delay in retreatment of Helicobacter pylori infection increases risk of upper gastrointestinal bleeding. ***Clinical Gastroenterology and Hepatology*** 2021;19(2):314-322.

5. Guo CG, Cheng KS, **Zhang F**, et al. Incidences, temporal trends and risks of hospitalization for gastrointestinal bleeding in new or chronic low-dose aspirin users after treatment for Helicobacter pylori: a territory-wide cohort study. ***Gut*** 2020; 69(3): 445-452.

6. Guo CG, Cheng KS, **Zhang F**, et al. Risks of hospitalization for upper gastrointestinal bleeding in selective serotonin reuptake inhibitors users after Helicobacter pylori eradication therapy: a propensity score matching analysis. ***Alimentary Pharmacology & Therapeutics*** 2019; 50(9):1001-1008.

7. Liu Z, **Zhang F**, Zhang Y, et al. Association between floods and infectious diarrhea and their effect modifiers in Hunan province, China: A two-stage model. ***The Science of the total environment*** 2018; 626: 630-637.

8. Lu L, Brierley L, Robertson G, **Zhang F**, Lycett S, Smith D, Chase-Topping M, Simmonds P, Woolhouse MEJ. Evolutionary origins of epidemic potential among human RNA viruses (bioRxiv preprint) doi: 10.1101/771394.

References

1. Parodi AS, Greenway DJ, Rugiero HR, et al. [Concerning the epidemic outbreak in Junin]. *El Dia medico* 1958; **30**(62): 2300-1.
2. Lisieux T, Coimbra M, Nassar ES, et al. New arenavirus isolated in Brazil. *Lancet* 1994; **343**(8894): 391-2.
3. Buchmeier M, Adam E, Rawls WE. Serological evidence of infection by Pichinde virus among laboratory workers. *Infection and immunity* 1974; **9**(5): 821-3.
4. Delgado S, Erickson BR, Agudo R, et al. Chapare virus, a newly discovered arenavirus isolated from a fatal hemorrhagic fever case in Bolivia. *PLoS Pathog* 2008; **4**(4): e1000047.
5. Salas R, de Manzione N, Tesh RB, et al. Venezuelan haemorrhagic fever. *Lancet* 1991; **338**(8774): 1033-6.
6. Buckley SM, Casals J. Lassa fever, a new virus disease of man from West Africa. 3. Isolation and characterization of the virus. *Am J Trop Med Hyg* 1970; **19**(4): 680-91.
7. Briese T, Paweska JT, McMullan LK, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog* 2009; **5**(5): e1000455.
8. Armstrong C, Lillie RD. Experimental Lymphocytic Choriomeningitis of Monkeys and Mice Produced by a Virus Encountered in Studies of the 1933 St. Louis Encephalitis Epidemic. *Public Health Reports (1896-1970)* 1934; **49**(35): 1019-27.
9. Mackenzie RB, Beye HK, Valverde L, Garron H. Epidemic Hemorrhagic Fever in Bolivia. I. A Preliminary Report of the Epidemiologic and Clinical Findings in a New Epidemic Area in South America. *Am J Trop Med Hyg* 1964; **13**: 620-5.
10. Georges AJ, Gonzalez JP, Abdul-Wahid S, Saluzzo JF, Meunier DM, McCormick JB. Antibodies to Lassa and Lassa-like viruses in man and mammals in the Central African Republic. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1985; **79**(1): 78-9.
11. Enserink M. Emerging diseases. New arenavirus blamed for recent deaths in California. *Science* 2000; **289**(5481): 842-3.
12. Appleton H, Higgins PG. Letter: Viruses and gastroenteritis in infants. *Lancet* 1975; **1**(7919): 1297.
13. Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, Wang D. Metagenomic analysis of human diarrhea: viral detection and discovery. *PLoS Pathog* 2008; **4**(2): e1000011.
14. Finkbeiner SR, Holtz LR, Jiang Y, et al. Human stool contains a previously unrecognized diversity of novel astroviruses. *Virology journal* 2009; **6**: 161.
15. Finkbeiner SR, Li Y, Ruone S, et al. Identification of a novel astrovirus (astrovirus VA1) associated with an outbreak of acute gastroenteritis. *Journal of virology* 2009; **83**(20): 10836-9.
16. Rott R, Herzog S, Fleischer B, et al. Detection of serum antibodies to Borna disease virus in patients with psychiatric disorders. *Science* 1985; **228**(4700): 755-6.
17. Hoffmann B, Tappe D, Hoper D, et al. A Variegated Squirrel Bornavirus Associated with Fatal Human Encephalitis. *The New England journal of medicine* 2015; **373**(2): 154-62.
18. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *Journal of virology* 1972; **10**(5): 1075-81.
19. Chiba S, Sakuma Y, Kogasaka R, et al. Fecal shedding of virus in relation to the days of illness in infantile gastroenteritis due to calicivirus. *J Infect Dis* 1980; **142**(2): 247-9.
20. Smith AW, Berry ES, Skilling DE, et al. In vitro isolation and characterization of a calicivirus causing a vesicular disease of the hands and feet. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 1998; **26**(2): 434-9.
21. Terao Y, Takagi H, Phan TG, Okitsu S, Ushijima H. Identification of antibody against porcine coronavirus in human milk. *Clinical laboratory* 2007; **53**(3-4): 129-30.

22. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine* 1966; **121**(1): 190-3.
23. Fouchier RA, Hartwig NG, Bestebroer TM, et al. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci U S A* 2004; **101**(16): 6212-6.
24. McIntosh K, Dees JH, Becker WB, Kapikian AZ, Chanock RM. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proc Natl Acad Sci U S A* 1967; **57**(4): 933-40.
25. Woo PC, Lau SK, Chu CM, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of virology* 2005; **79**(2): 884-95.
26. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *The New England journal of medicine* 2012; **367**(19): 1814-20.
27. Peiris JS, Lai ST, Poon LL, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 2003; **361**(9366): 1319-25.
28. Beards GM, Hall C, Green J, Flewett TH, Lamouliatte F, Du Pasquier P. An enveloped virus in stools of children and adults with gastroenteritis that resembles the Breda virus of calves. *Lancet* 1984; **1**(8385): 1050-2.
29. Towner JS, Sealy TK, Khristova ML, et al. Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog* 2008; **4**(11): e1000212.
30. Miranda ME, White ME, Dayrit MM, Hayes CG, Ksiazek TG, Burans JP. Seroepidemiological study of filovirus related to Ebola in the Philippines. *Lancet* 1991; **337**(8738): 425-6.
31. Bowen ET, Lloyd G, Harris WJ, Platt GS, Baskerville A, Vella EE. Viral haemorrhagic fever in southern Sudan and northern Zaire. Preliminary studies on the aetiological agent. *Lancet* 1977; **1**(8011): 571-3.
32. Le Guenno B, Formenty P, Wyers M, Gounon P, Walker F, Boesch C. Isolation and partial characterisation of a new strain of Ebola virus. *Lancet* 1995; **345**(8960): 1271-4.
33. Johnson KM, Lange JV, Webb PA, Murphy FA. Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. *Lancet* 1977; **1**(8011): 569-71.
34. Martini GA, Knauff HG, Schmidt HA, Mayer G, Baltzer G. A hitherto unknown infectious disease contracted from monkeys. "Marburg-virus" disease. *German medical monthly* 1968; **13**(10): 457-70.
35. Srihongse S, Johnson CM. The first isolation of Bussuquara virus from man. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1971; **65**(4): 541-2.
36. Bondre VP, Sapkal GN, Yergolkar PN, et al. Genetic characterization of Bagaza virus (BAGV) isolated in India and evidence of anti-BAGV antibodies in sera collected from encephalitis patients. *The Journal of general virology* 2009; **90**(Pt 11): 2644-9.
37. Smithburn KC, Paterson HE, Heymann CS, Winter PA. An agent related to Uganda S virus from man and mosquitoes in South Africa. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 1959; **33**: 959-62.
38. Batista WC, Tavares Gda S, Vieira DS, Honda ER, Pereira SS, Tada MS. Notification of the first isolation of Cacipacore virus in a human in the State of Rondonia, Brazil. *Revista da Sociedade Brasileira de Medicina Tropical* 2011; **44**(4): 528-30.
39. Ashburn PM, Craig CF. Experimental Inverstigations Regarding the Etiology of Dengue Fever. *The Journal of Infectious Diseases* 1907; **4**(3): 440-75.
40. Hawkes RA, Boughton CR, Naim HM, Wild J, Chapman B. Arbovirus infections of humans in New South Wales. Seroepidemiology of the flavivirus group of togaviruses. *Med J Aust* 1985; **143**(12-13): 555-61.
41. Humphery-Smith I, Cybinski DH, Byrnes KA, St George TD. Seroepidemiology of arboviruses among seabirds and island residents of the Great Barrier Reef and Coral Sea. *Epidemiology and infection* 1991; **107**(2): 435-40.
42. Laemmert HW, Jr., Hughes TP. The virus of Ilheus encephalitis; isolation, serological specificity and transmission. *Journal of immunology* 1947; **55**(1): 61-7.
43. Hayashi M. Übertragung des Virus von Encephalitis epidemica japonica auf Affen*. *Psychiatry and Clinical Neurosciences* 1933; **1**(1): 419-65.

44. Work TH, Trapido H. Summary of preliminary report of investigations of the Virus Research Centre on an epidemic disease affecting forest villagers and wild monkeys of Shimoga District, Mysore. *Indian journal of medical sciences* 1957; **11**(5): 341-2.
45. Smith CE. A virus resembling Russian spring-summer encephalitis virus from an ixodid tick in Malaya. *Nature* 1956; **178**(4533): 581-2.
46. Rivers TM, Schwentker FF. Louping Ill in Man. *The Journal of experimental medicine* 1934; **59**(5): 669-85.
47. French EL. Murray Valley encephalitis isolation and characterization of the aetiological agent. *Med J Aust* 1952; **1**(4): 100-3.
48. Smithburn KC. Neutralizing antibodies against certain recently isolated viruses in the sera of human beings residing in East Africa. *Journal of immunology* 1952; **69**(2): 223-34.
49. Chumakov MP. Results of a study made of Omsk hemorrhagic fever (OL) by an expedition of the Institute of Neurology. *Vestn Acad Med Nauk SSSR* 1948; **2**(19).
50. Mc LD, Donohue WL. Powassan virus: isolation of virus from a fatal case of encephalitis. *Canadian Medical Association journal* 1959; **80**(9): 708-11.
51. Sulkin SE, Burns KF, Shelton DF, Wallis C. Bat salivary gland virus: infections of man and monkey. *Texas reports on biology and medicine* 1962; **20**: 113-27.
52. Webster LT, Fite GL. A Virus Encountered in the Study of Material from Cases of Encephalitis N the St. Louis and Kansas City Epidemics of 1933. *Science* 1933; **78**(2029): 463-5.
53. Bowen ET, Simpson DI, Platt GS, et al. Arbovirus infections in Sarawak, October 1968-February 1970: human serological studies in a land Dyak village. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1975; **69**(2): 182-6.
54. Solowjow WD. Zur Aetiologie der Zecken- Encephalitis. *Acta Med Urss* 1938; **1**(4): 484-92.
55. Dick GW, Haddow AJ. Uganda S virus; a hitherto unrecorded virus isolated from mosquitoes in Uganda. I. Isolation and pathogenicity. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1952; **46**(6): 600-18.
56. Pecorari M, Longo G, Gennari W, et al. First human case of Usutu virus neuroinvasive infection, Italy, August-September 2009. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2009; **14**(50).
57. Smithburn KC, Kokernot RH, Weinbren MP, De Meillon B. Studies on arthropod-borne viruses of Tongaland. IX. Isolation of Wesselsbron virus from a naturally infected human being and from Aedes (Banksinella) circumluteolus Theo. *The South African journal of medical sciences* 1957; **22**(2-3): 113-20.
58. Smithburn KC, Hughes TP, Burke AW, Paul JH. A Neurotropic Virus Isolated from the Blood of a Native of Uganda. *The American Journal of Tropical Medicine and Hygiene* 1940 **s1-20**(4): 471 - 92.
59. Reed W, Carroll JS, Agramonte A. The etiology of yellow fever: An additional note. *Journal of the American Medical Association* 1901; **XXXVI**(7): 431-40.
60. Dick GW. Zika virus. II. Pathogenicity and physical properties. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1952; **46**(5): 521-34.
61. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989; **244**(4902): 359-62.
62. Yoshiba M, Okamoto H, Mishiro S. Detection of the GBV-C hepatitis virus genome in serum from patients with fulminant hepatitis of unknown aetiology. *Lancet* 1995; **346**(8983): 1131-2.
63. Kapoor A, Kumar A, Simmonds P, et al. Virome Analysis of Transfusion Recipients Reveals a Novel Human Virus That Shares Genomic Features with Hepaciviruses and Pegiviruses. *mBio* 2015; **6**(5): e01466-15.
64. Giangaspero M, Wellemans G, Vanopdenbosch E, Belloli A, Verhulst A. Bovine viral diarrhoea. *Lancet* 1988; **2**(8602): 110.
65. Lopez N, Padula P, Rossi C, Lazaro ME, Franze-Fernandez MT. Genetic identification of a new hantavirus causing severe pulmonary syndrome in Argentina. *Virology* 1996; **220**(1): 223-6.

66. Morzunov SP, Feldmann H, Spiropoulou CF, et al. A newly recognized virus associated with a fatal case of hantavirus pulmonary syndrome in Louisiana. *Journal of virology* 1995; **69**(3): 1980-3.
67. Ravkov EV, Rollin PE, Ksiazek TG, Peters CJ, Nichol ST. Genetic and serologic analysis of Black Creek Canal virus and its association with human disease and Sigmodon hispidus infection. *Virology* 1995; **210**(2): 482-9.
68. Vincent MJ, Quiroz E, Gracia F, et al. Hantavirus pulmonary syndrome in Panama: identification of novel hantaviruses and their likely reservoirs. *Virology* 2000; **277**(1): 14-9.
69. Gligic A, Dimkovic N, Xiao SY, et al. Belgrade virus: a new hantavirus causing severe hemorrhagic fever with renal syndrome in Yugoslavia. *J Infect Dis* 1992; **166**(1): 113-20.
70. Lee HW, Lee PW, Johnson KM. Isolation of the etiologic agent of Korean Hemorrhagic fever. *J Infect Dis* 1978; **137**(3): 298-308.
71. Johnson AM, Bowen MD, Ksiazek TG, et al. Laguna Negra virus associated with HPS in western Paraguay and Bolivia. *Virology* 1997; **238**(1): 115-27.
72. Brummer-Korvenkontio M, Vaheri A, Hovi T, et al. Nephropathia epidemica: detection of antigen in bank voles and serologic diagnosis of human infection. *J Infect Dis* 1980; **141**(2): 131-4.
73. Klempa B, Koivogui L, Sylla O, et al. Serological evidence of human hantavirus infections in Guinea, West Africa. *J Infect Dis* 2010; **201**(7): 1031-4.
74. Song G, Qiu XZ, Ni DS, Zhao JN, Kong BX. Etiological studies of epidemic hemorrhagic fever. I. Virus isolation in Apodemus agrarius from non-endemic area and its antigenic characterization [Article in Chinese]. *Zhongguo yi xue ke xue yuan xue bao Acta Academiae Medicinae Sinicae* 1982; **4**(2): 73-7.
75. Nichol ST, Spiropoulou CF, Morzunov S, et al. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 1993; **262**(5135): 914-7.
76. Pattamadilok S, Lee BH, Kumperasart S, et al. Geographical distribution of hantaviruses in Thailand and potential human health significance of Thailand virus. *Am J Trop Med Hyg* 2006; **75**(5): 994-1002.
77. Okumura M, Yoshimatsu K, Kumperasart S, et al. Development of serological assays for Thottapalayam virus, an insectivore-borne Hantavirus. *Clinical and vaccine immunology : CVI* 2007; **14**(2): 173-81.
78. Vapalahti O, Lundkvist A, Kukkonen SK, et al. Isolation and characterization of Tula virus, a distinct serotype in the genus Hantavirus, family Bunyaviridae. *The Journal of general virology* 1996; **77 (Pt 12)**: 3063-7.
79. Balayan MS, Andjaparidze AG, Savinskaya SS, et al. Evidence for a virus in non-A, non-B hepatitis transmitted via the fecal-oral route. *Intervirology* 1983; **20**(1): 23-31.
80. Sridhar S, Yip CCY, Wu S, et al. Rat Hepatitis E Virus as Cause of Persistent Hepatitis after Liver Transplant. *Emerg Infect Dis* 2018; **24**(12): 2241-50.
81. Simpson DI, Knight EM, Courtois G, Williams MC, Weinbren MP, Kibukamusoke JW. Congo virus: a hitherto undescribed virus occurring in Africa. I. Human isolations--clinical notes. *East African medical journal* 1967; **44**(2): 86-92.
82. Causey OR, Kemp GE, Madbouly MH, Lee VH. Arbovirus surveillance in Nigeria, 1964-1967. *Bulletin de la Societe de pathologie exotique et de ses filiales* 1969; **62**(2): 249-53.
83. Dandawate CN, Work TH, Webb JK, Shah KV. Isolation of Ganjam virus from a human case of febrile illness: a report of a laboratory infection and serological survey of human sera from three different states of India. *The Indian journal of medical research* 1969; **57**(6): 975-82.
84. Chastel C, Main AJ, Richard P, Le Lay G, Legrand-Quillien MC, Beaucournu JC. Erve virus, a probable member of Bunyaviridae family isolated from shrews (Crocidura russula) in France. *Acta virologica* 1989; **33**(3): 270-80.
85. Smith W, Andrewes CH, Laidlaw PP. A virus obtained from influenza patients. *Lancet* 1933; **222**(5732): 66-8.
86. Thomas Francis J. Differentiation of Influenza A and Influenza B by the Complement-Fixation Reaction. *Proceedings of the Society for Experimental Biology and Medicine* 1940; **45**(3): 861-3.
87. Francis T, Jr., Quilligan JJ, Jr., Minuse E. Identification of another epidemic respiratory disease. *Science* 1950; **112**(2913): 495-7.

88. Filipe AR, Calisher CH, Lazuick J. Antibodies to Congo-Crimean haemorrhagic fever, Dhori, Thogoto and Bhanja viruses in southern Portugal. *Acta virologica* 1985; **29**(4): 324-8.
89. Burnet FM. Human infection with the virus of Newcastle disease of fowls. *The Medical Journal of Australia* 1943; **30**(6): 313-4.
90. Selvey L, Sheridan J. Outbreak of Severe Respiratory Disease in Humans and Horses Due to a Previously Unrecognized Paramyxovirus. *Journal of travel medicine* 1995; **2**(4): 275.
91. Lee KE, Umapathi T, Tan CB, et al. The neurological manifestations of Nipah virus encephalitis, a novel paramyxovirus. *Annals of neurology* 1999; **46**(3): 428-32.
92. Karzon DT. Studies on a neutralizing antibody against canine distemper virus found in man. *Pediatrics* 1955; **16**(6): 809-18.
93. Goldberger J, Anderson JF. The nature of the virus of measles. *Journal of the American Medical Association* 1911; **LVII**(12): 971-2.
94. Chanock RM, Parrott RH, Cook K, et al. Newly recognized myxoviruses from children with respiratory disease. *The New England journal of medicine* 1958; **258**(5): 207-13.
95. Baker KS, Todd S, Marsh GA, et al. Novel, potentially zoonotic paramyxoviruses from the African straw-colored fruit bat Eidolon helvum. *Journal of virology* 2013; **87**(3): 1348-58.
96. Chanock RM. Association of a new type of cytopathogenic myxovirus with infantile croup. *The Journal of experimental medicine* 1956; **104**(4): 555-76.
97. Johnson KM, Chanock RM, Cook MK, Huebner RJ. Studies of a new human hemadsorption virus. I. Isolation, properties and characterization. *American journal of hygiene* 1960; **71**: 81-92.
98. Schultz EW, Habel K. SA virus; a new member of the myxovirus group. *Journal of immunology* 1959; **82**(3): 274-8.
99. Philbey AW, Kirkland PD, Ross AD, et al. An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats. *Emerg Infect Dis* 1998; **4**(2): 269-71.
100. Johnson CD, Goodpasture EW. An Investigation of the Etiology of Mumps. *The Journal of experimental medicine* 1934; **59**(1): 1-19.
101. Itoh H, Morimoto Y, Doi Y, Sanpe T. Studies on simian viruses some properties of SV 41 grown in vero cell cultures and search for serum neutralizing antibodies in humans and various animals. *Uirusu* 1968; **18**(6): 495-503.
102. Albarino CG, Foltzer M, Towner JS, et al. Novel paramyxovirus associated with severe acute febrile disease, South Sudan and Uganda, 2012. *Emerg Infect Dis* 2014; **20**(2): 211-6.
103. Yaiw KC, Crameri G, Wang L, et al. Serological evidence of possible human infection with Tioman virus, a newly described paramyxovirus of bat origin. *J Infect Dis* 2007; **196**(6): 884-6.
104. Smithburn KC, Haddow AJ, Mahaffy AF. A neurotropic virus isolated from Aedes mosquitoes caught in the Semliki forest. *Am J Trop Med Hyg* 1946; **26**: 189-208.
105. Smithburn KC, Mahaffy AF, Paul JH. Bwamba Fever and Its Causative Virus. *The American Journal of Tropical Medicine and Hygiene* 1941; **s1-21**(1): 75 - 90.
106. Hammon WM, Reeves WC. California encephalitis virus, a newly described agent. *California medicine* 1952; **77**(5): 303-9.
107. Causey OR, Causey CE, Maroja OM, Macedo DG. The isolation of arthropod-borne viruses, including members of two hitherto undescribed serological groups, in the Amazon region of Brazil. *Am J Trop Med Hyg* 1961; **10**: 227-49.
108. Groot H, Oya A, Bernal C, Barreto-Reyes P. Guaroa virus, a new agent isolated in Colombia, South America. *Am J Trop Med Hyg* 1959; **8**: 604-9.
109. van Tongeren HA. Occurrence of arboviruses belonging to the C-, Bunyamwera and Guama groups, and of Oropouche, Junin, Tacaiuma and Kwatta viruses in man in the province of Brokopondo, Surinam: a serological survey. *Tropical and geographical medicine* 1967; **19**(4): 309-25.
110. Derodaniche E, Paesdeandrade A, Galindo P. ISOLATION OF TWO ANTIGENICALLY DISTINCT ARTHROPOD-BORNE VIRUSES OF GROUP C IN PANAMA. *Am J Trop Med Hyg* 1964; **13**: 839-43.
111. Williams MC, Woodall JP, Corbet PS. Nyando Virus: A Hitherto Undescribed Virus Isolated from Anopheles Funestus Giles Collected in Kenya. *Archiv fur die gesamte Virusforschung* 1965; **15**: 422-7.

112. Anderson CR, Spence L, Downs WG, Aitken TH. Oropouche virus: a new human disease agent from Trinidad, West Indies. *Am J Trop Med Hyg* 1961; **10**: 574-8.
113. Scherer WF, Anderson K, Dickerman RW, Ordonez JV. Studies of Patois group arboviruses in Mexico, Guatemala, Honduras, and British Honduras. *Am J Trop Med Hyg* 1972; **21**(2): 194-200.
114. Moore DL, Causey OR, Carey DE, et al. Arthropod-borne viral infections of man in Nigeria, 1964-1970. *Annals of tropical medicine and parasitology* 1975; **69**(1): 49-64.
115. Sirhongse S, Johnson CM. Wyeomyia Subgroup of Arbovirus: Isolation from Man. *Science* 1965; **149**(3686): 863-4.
116. Travassos da Rosa AP, Tesh RB, Pinheiro FP, Travassos da Rosa JF, Peterson NE. Characterization of eight new phlebotomus fever serogroup arboviruses (Bunyaviridae: Phlebovirus) from the Amazon region of Brazil. *Am J Trop Med Hyg* 1983; **32**(5): 1164-71.
117. Sather GE. Punta Toro (PT) strain. D 4021A. *Am J Trop Med Hyg* 1970; **19**(6): Suppl:1103-4.
118. Daubney R, Hudson JR, Garnham PC. Enzootic hepatitis or rift valley fever. An undescribed virus disease of sheep cattle and man from east africa. *The Journal of Pathology and Bacteriology* 1931; **34**(4): 545-79.
119. Sabin AB, Philip CB, Paul JR. PHLEBOTOMUS (PAPPATACI OR SANDFLY) FEVER: A DISEASE OF MILITARY IMPORTANCE SUMMARY OF EXISTING KNOWLEDGE AND PRELIMINARY REPORT OF ORIGINAL INVESTIGATIONS. *Journal of the American Medical Association* 1944; **125**(9): 603-6.
120. Zhang YZ, Zhou DJ, Xiong Y, et al. Hemorrhagic fever caused by a novel tick-borne Bunyavirus in Huaiyangshan, China. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* 2011; **32**(3): 209-20.
121. Sekeyová M, Gresíková M, Stúpalová S. Serological study on distribution of Uukuniemi virus in man. *Folia Parasitologica* 1970; **17**(4): 341-3.
122. Pereira HG, Fialho AM, Flewett TH, Teixeira JM, Andrade ZP. Novel viruses in human faeces. *Lancet* 1988; **2**(8602): 103-4.
123. Plummer G. An equine respiratory virus with enterovirus properties. *Nature* 1962; **195**: 519-20.
124. Pilz W, Garbe HG. [Further cases of foot-and-mouth disease virus infections in man]. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene 1 Abt Medizinisch-hygienische Bakteriologie, Virusforschung und Parasitologie Originale* 1965; **198**(1): 154-7.
125. Smadel JE, Warren J. The virus of encephalomyocarditis and its apparent causation of disease in man. *The Journal of clinical investigation* 1947; **26**(6): 1197.
126. Casals J. Immunological Characterization of Vilyuisk Human Encephalomyelitis Virus. *Nature* 1963; **200**: 339-41.
127. Kapoor A, Victoria J, Simmonds P, et al. A highly prevalent and genetically diversified Picornaviridae genus in South Asian children. *Proc Natl Acad Sci U S A* 2008; **105**(51): 20482-7.
128. Holtz LR, Finkbeiner SR, Kirkwood CD, Wang D. Identification of a novel picornavirus related to cosaviruses in a child with acute diarrhea. *Virology journal* 2008; **5**: 159.
129. Kapusinszky B, Phan TG, Kapoor A, Delwart E. Genetic diversity of the genus Cosavirus in the family Picornaviridae: a new species, recombination, and 26 new genotypes. *PLoS One* 2012; **7**(5): e36685.
130. Sickles GM, Dalldorf G. Serologic differences among strains of the Coxsackie group of viruses. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine* 1949; **72**(1): 30.
131. Flexner S, Lewis PA. The transmission of acute poliomyelitis to monkeys. *Journal of the American Medical Association* 1909; **LIII**(20): 1639-.
132. Schieble JH, Fox VL, Lennette EH. A probable new human picornavirus associated with respiratory diseases. *American journal of epidemiology* 1967; **85**(2): 297-310.
133. Moscovici C, Laplaca M, Maisel J, Kempe H. Studies of bovine enteroviruses. *American journal of veterinary research* 1961; **22**: 852-63.
134. Conrad ME. Infectious hepatitis: An unresolved military problem. *Military medicine* 1965; **130**: 594-600.

135. Andrewes CH, Chaproniere DM, Gompels AE, Pereira HG, Roden AT. Propagation of common-cold virus in tissue cultures. *Lancet* 1953; **265**(6785): 546-7.
136. Tyrrell DA, Parsons R. Some virus isolations from common colds. III. Cytopathic effects in tissue cultures. *Lancet* 1960; **1**(7118): 239-42.
137. Lau SK, Yip CC, Tsoi HW, et al. Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic cluster, probably representing a previously undetected HRV species, HRV-C, associated with acute respiratory illness in children. *Journal of clinical microbiology* 2007; **45**(11): 3655-64.
138. Kriegshauser G, Deutz A, Kuechler E, Skern T, Lussy H, Nowotny N. Prevalence of neutralizing antibodies to Equine rhinitis A and B virus in horses and man. *Veterinary microbiology* 2005; **106**(3-4): 293-6.
139. Feinstone SM, Kapikian AZ, Purceli RH. Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science* 1973; **182**(4116): 1026-8.
140. Yamashita T, Kobayashi S, Sakae K, et al. Isolation of cytopathic small round viruses with BS-C-1 cells from patients with gastroenteritis. *J Infect Dis* 1991; **164**(5): 954-7.
141. Ramos-Alvarez M, Sabin AB. Enteropathogenic viruses and bacteria; role in summer diarrheal diseases of infancy and early childhood. *Journal of the American Medical Association* 1958; **167**(2): 147-56.
142. Niklasson B, Heller KE, Schonecker B, et al. Development of type 1 diabetes in wild bank voles associated with islet autoantibodies and the novel ljungan virus. *International journal of experimental diabesity research* 2003; **4**(1): 35-44.
143. Greninger AL, Runckel C, Chiu CY, et al. The complete genome of klassevirus - a novel picornavirus in pediatric stool. *Virology journal* 2009; **6**: 82.
144. Kayali G, Ortiz EJ, Chorazy ML, et al. Serologic evidence of avian metapneumovirus infection among adults occupationally exposed to Turkeys. *Vector Borne Zoonotic Dis* 2011; **11**(11): 1453-8.
145. van den Hoogen BG, de Jong JC, Groen J, et al. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat Med* 2001; **7**(6): 719-24.
146. Chanock R, Roizman B, Myers R. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *American journal of hygiene* 1957; **66**(3): 281-90.
147. Florio L, Stewart MO, Mugrage ER. The Etiology of Colorado Tick Fever. *The Journal of experimental medicine* 1946; **83**(1): 1-10.
148. Malkova D, Holubova J, Kolman JM, et al. Antibodies against some arboviruses in persons with various neuropathies. *Acta virologica* 1980; **24**(4): 298.
149. DOHERTY RL. STUDIES OF ABORIGINES AT AURUKUN AND WEIPA MISSIONS, NORTH QUEENSLAND 2. OTHER LABORATORY STUDIES. *Journal of paediatrics and child health* 1967; **3**(4): 213–8.
150. Chumakov MP, Sarmanova ES, Bychkova MV, et al. [Identification of the Virus of Kemerovo Tick-Borne Fever. Evidence of the Antigenic Independence of This Virus]. *Voprosy virusologii* 1963; **29**: 440-4.
151. Ramos-Alvarez M, Sabin AB. Characteristics of poliomyelitis and other enteric viruses recovered in tissue culture from healthy American children. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine* 1954; **87**(3): 655-61.
152. Chua KB, Crameri G, Hyatt A, et al. A previously unknown reovirus of bat origin is associated with an acute respiratory disease in humans. *Proc Natl Acad Sci U S A* 2007; **104**(27): 11424-9.
153. Bishop RF, Davidson GP, Holmes IH, Ruck BJ. Virus particles in epithelial cells of duodenal mucosa from children with acute non-bacterial gastroenteritis. *Lancet* 1973; **2**(7841): 1281-3.
154. Hung T, Chen GM, Wang CG, et al. Waterborne outbreak of rotavirus diarrhoea in adults in China caused by a novel rotavirus. *Lancet* 1984; **1**(8387): 1139-42.
155. Bridger JC, Pedley S, McCrae MA. Group C rotaviruses in humans. *Journal of clinical microbiology* 1986; **23**(4): 760-3.

156. Yang H, Chen S, Ji S. A novel rotavirus causing large scale of adult diarrhea in Shi Jiazhuang [Article in Chinese]. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* 1998; **19**(6): 336-8.
157. Xu PT, Wang YM, Zuo JM, Lin JW, Xu PM. New orbiviruses isolated from patients with unknown fever and encephalitis in Yunnan Province [Article in Chinese]. *Chinese Journal of Virology* 1990; **06**(01): 27-33+100.
158. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 1980; **77**(12): 7415-9.
159. Kalyanaraman VS, Sarngadharan MG, Robert-Guroff M, Miyoshi I, Golde D, Gallo RC. A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* 1982; **218**(4572): 571-3.
160. Calattini S, Chevalier SA, Duprez R, et al. Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. *Retrovirology* 2005; **2**: 30.
161. Barre-Sinoussi F, Chermann JC, Rey F, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983; **220**(4599): 868-71.
162. Kanki PJ, Barin F, M'Boup S, et al. New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-IIIAGM). *Science* 1986; **232**(4747): 238-43.
163. Khabbaz RF, Rowe T, Murphey-Corb M, et al. Simian immunodeficiency virus needlestick accident in a laboratory worker. *Lancet* 1992; **340**(8814): 271-3.
164. Rua R, Betsem E, Calattini S, Saib A, Gessain A. Genetic characterization of simian foamy viruses infecting humans. *Journal of virology* 2012; **86**(24): 13350-9.
165. Achong BG, Mansell PW, Epstein MA. A new human virus in cultures from a nasopharyngeal carcinoma. *The Journal of pathology* 1971; **103**(2): P18.
166. Schweizer M, Falcone V, Gange J, Turek R, Neumann-Haefelin D. Simian foamy virus isolated from an accidentally infected human individual. *Journal of virology* 1997; **71**(6): 4821-4.
167. Brooks JI, Rud EW, Pilon RG, Smith JM, Switzer WM, Sandstrom PA. Cross-species retroviral transmission from macaques to human beings. *Lancet* 2002; **360**(9330): 387-8.
168. Samaratunga H, Searle JW, Hudson N. Non-rabies Lyssavirus human encephalitis from fruit bats: Australian bat Lyssavirus (pteropid Lyssavirus) infection. *Neuropathology and applied neurobiology* 1998; **24**(4): 331-5.
169. Meredith CD, Prossouw AP, Koch H. An unusual case of human rabies thought to be of chiropteran origin. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 1971; **45**(28): 767-9.
170. Selimov MA, Tatarov AG, Botvinkin AD, Klueva EV, Kulikova LG, Khismatullina NA. Rabies-related Yuli virus; identification with a panel of monoclonal antibodies. *Acta virologica* 1989; **33**(6): 542-6.
171. Lumio J, Hillbom M, Roine R, et al. Human rabies of bat origin in Europe. *Lancet* 1986; **1**(8477): 378.
172. Liu Y, Zhang S, Zhao J, Zhang F, Hu R. Isolation of Irkut virus from a Murina leucogaster bat in China. *PLoS neglected tropical diseases* 2013; **7**(3): e2097.
173. Familusi JB, Osunkoya BO, Moore DL, Kemp GE, Fabiyi A. A fatal human infection with Mokola virus. *Am J Trop Med Hyg* 1972; **21**(6): 959-63.
174. Negri A. Beitrag zum Studium der Aetiologie der Tollwuth. *Zeitschrift für Hygiene und Infektionskrankheiten* 1903; **43**(1): 507-28.
175. Grard G, Fair JN, Lee D, et al. A novel rhabdovirus associated with acute hemorrhagic fever in central Africa. *PLoS Pathog* 2012; **8**(9): e1002924.
176. Stremlau MH, Andersen KG, Folarin OA, et al. Discovery of novel rhabdoviruses in the blood of healthy individuals from West Africa. *PLoS neglected tropical diseases* 2015; **9**(3): e0003631.
177. Federer KE, Burrows R, Brooksby JB. Vesicular stomatitis virus--the relationship between some strains of the Indiana serotype. *Research in veterinary science* 1967; **8**(1): 103-17.
178. Bhatt PN, Rodrigues FM. Chandipura: a new Arbovirus isolated in India from patients with febrile illness. *The Indian journal of medical research* 1967; **55**(12): 1295-305.

179. Jonkers AH, Shope RE, Aitken TH, Spence L. COCAL VIRUS, A NEW AGENT IN TRINIDAD RELATED TO VESICULAR STOMATITIS VIRUS, TYPE INDIANA. *American journal of veterinary research* 1964; **25**: 236-42.
180. Patterson WC, Mott LO, Jenney EW. A study of vesicular stomatitis in man. *Journal of the American Veterinary Medical Association* 1958; **133**(1): 57-62.
181. Tesh R, Saidi S, Javadian E, Loh P, Nadim A. Isfahan virus, a new vesiculovirus infecting humans, gerbils, and sandflies in Iran. *Am J Trop Med Hyg* 1977; **26**(2): 299-306.
182. Travassos da Rosa AP, Tesh RB, Travassos da Rosa JF, Herve JP, Main AJ, Jr. Carajas and Maraba viruses, two new vesiculoviruses isolated from phlebotomine sand flies in Brazil. *Am J Trop Med Hyg* 1984; **33**(5): 999-1006.
183. Hanson RP, Rasmussen AF, Jr., Brandly CA, Brown JW. Human infection with the virus of vesicular stomatitis. *The Journal of laboratory and clinical medicine* 1950; **36**(5): 754-8.
184. Pinheiro FP, Bensabath G, Andrade AH, et al. Infectious diseases along Brazil's trans-amazon highway: surveillance and research. *Bulletin of the Pan American Health Organization* 1974; **8**(2): 111-22.
185. Vale TG, Carter IW, McPhie KA, James GS, Cloonan MJ. Human arbovirus infections along the south coast of New South Wales. *The Australian journal of experimental biology and medical science* 1986; **64 (Pt 3)**: 307-9.
186. Ross RW. The Newala epidemic. III. The virus: isolation, pathogenic properties and relationship to the epidemic. *The Journal of hygiene* 1956; **54**(2): 177-91.
187. Howitt B. Recovery of the Virus of Equine Encephalomyelitis from the Brain of a Child. *Science* 1938; **88**(2289): 455-6.
188. Ehrenkranz NJ, Sinclair MC, Buff E, Lyman DO. The natural occurrence of Venezuelan equine encephalitis in the United States. *The New England journal of medicine* 1970; **282**(6): 298-302.
189. Doherty RL. Arthropod-borne viruses in Australia and their relation to infection and disease. *Progress in medical virology Fortschritte der medizinischen Virusforschung Progres en virologie medicale* 1974; **17**(0): 136-92.
190. Meehan PJ, Wells DL, Paul W, et al. Epidemiological features of and public health response to a St. Louis encephalitis epidemic in Florida, 1990-1. *Epidemiology and infection* 2000; **125**(1): 181-8.
191. Luciani K, Abadia I, Martinez-Torres AO, et al. Madariaga virus infection associated with a case of acute disseminated encephalomyelitis. *Am J Trop Med Hyg* 2015; **92**(6): 1130-2.
192. Anderson CR, Downs WG, Wattley GH, Ahin NW, Reese AA. Mayaro virus: a new human disease agent. II. Isolation from blood of patients in Trinidad, B.W.I. *Am J Trop Med Hyg* 1957; **6**(6): 1012-6.
193. Pisano MB, Oria G, Beskow G, et al. Venezuelan equine encephalitis viruses (VEEV) in Argentina: serological evidence of human infection. *PLoS neglected tropical diseases* 2013; **7**(12): e2551.
194. Demucha Macias J, S'Anchez Spindola I. Two Human Cases of Laboratory Infection with Mucambo Virus. *Am J Trop Med Hyg* 1965; **14**: 475-8.
195. Kokernot RH, McIntosh BM, Worth CB. Ndumu virus, a hitherto unknown agent, isolated from culicine mosouitoes collected in northern Natal. Union of South Africa. *Am J Trop Med Hyg* 1961; **10**: 383-6.
196. Williams MC, Woodall JP. O'nyong-nyong fever: an epidemic virus disease in East Africa. II. Isolation and some properties of the virus. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1961; **55**: 135-41.
197. Vasconcelos PF, Da Rosa JF, Da Rosa AP, Degallier N, Pinheiro Fde P, Sa Filho GC. [Epidemiology of encephalitis caused by arbovirus in the Brazilian Amazonia]. *Revista do Instituto de Medicina Tropical de Sao Paulo* 1991; **33**(6): 465-76.
198. Contigiani MS, de Basualdo M, Camara A, et al. [Presence of antibodies against Venezuelan equine encephalitis virus subtype VI in patients with acute febrile illness]. *Revista Argentina de microbiologia* 1993; **25**(4): 212-20.
199. Doherty RL, Carley JG, Best JC. Isolation of Ross River virus from man. *Med J Aust* 1972; **1**(21): 1083-4.

200. Willems WR, Kaluza G, Boschek CB, et al. Semliki forest virus: cause of a fatal case of human encephalitis. *Science* 1979; **203**(4385): 1127-9.
201. Taylor RM, Hurlbut HS, Work TH, Kingston JR, Frothingham TE. Sindbis virus: a newly recognized arthropodtransmitted virus. *Am J Trop Med Hyg* 1955; **4**(5): 844-62.
202. Digoutte JP, Girault G. [The protective properties in mice of tonate virus and two strains of cabassou virus against neurovirulent everglades Venezuelan encephalitis virus (author's transl)]. *Annales de microbiologie* 1976; **127B**(3): 429-37.
203. Causey OR, Casals J, Shope RE, Udomsakdi S. Aura and Una, Two New Group a Arthropod-Borne Viruses. *Am J Trop Med Hyg* 1963; **12**: 777-81.
204. Casals J, Curnen EC, Thomas L. Venezuelan Equine Encephalomyelitis in Man. *The Journal of experimental medicine* 1943; **77**(6): 521-30.
205. Ross RW, Miles JA, Austin FJ, Maguire T. INVESTIGATIONS INTO THE ECOLOGY OF A GROUP A ARBOVIRUS IN WESTLAND, NEW ZEALAND. *The Australian journal of experimental biology and medical science* 1964; **42**: 689-702.
206. Habel K. Transmission of Rubella to Macacus mulatta Monkeys. *Public Health Reports (1896-1970)* 1942; **57**(31): 1126-39.
207. Rizzetto M, Canese MG, Arico S, et al. Immunofluorescence detection of new antigen-antibody system (delta/anti-delta) associated to hepatitis B virus in liver and in serum of HBsAg carriers. *Gut* 1977; **18**(12): 997-1003.
208. Barry M, Russi M, Armstrong L, et al. Brief report: treatment of a laboratory-acquired Sabia virus infection. *The New England journal of medicine* 1995; **333**(5): 294-6.
209. Oshiro LS, Haley CE, Roberto RR, et al. A 27-nm virus isolated during an outbreak of acute infectious nonbacterial gastroenteritis in a convalescent hospital: a possible new serotype. *Journal of Infectious Diseases* 1981; **143**(6): 791-5.
210. Xu AY, Pang QF, Qiu FX. Detection of astrovirus in faeces of infants with gastroenteritis in autumn [Article in Chinese]. *Journal of Medical Research* 1981; (06): 32.
211. Dowling JM, Wynne H. Role of Enteric Adenoviruses and Rotaviruses in Infantile Gastroenteritis. *Lancet* 1981; **2**(8241): 305-6.
212. Finkbeiner SR, Le BM, Holtz LR, Storch GA, Wang D. Detection of Newly Described Astrovirus MLB1 in Stool Samples from Children. *Emerging Infectious Diseases* 2009; **15**(3): 441-4.
213. Chu DK, Chin AW, Smith GJ, et al. Detection of novel astroviruses in urban brown rats and previously known astroviruses in humans. *The Journal of general virology* 2010; **91**(Pt 10): 2457-62.
214. Kapoor A, Li L, Victoria J, et al. Multiple novel astrovirus species in human stool. *Journal of General Virology* 2009; **90**: 2965-72.
215. Wang Y, Li Y, Jin Y, Li DD, Li X, Duan ZJ. Recently identified novel human astroviruses in children with diarrhea, China. *Emerg Infect Dis* 2013; **19**(8): 1333-5.
216. Tao Z, Wang H, Zhang W, Xu A. Novel astrovirus types circulating in Shandong Province (Eastern China) during 2016: A clinical and environmental surveillance. *J Clin Virol* 2019; **116**: 69-73.
217. Chen CH, Chiu YL, Wei FC, et al. High seroprevalence of Borna virus infection in schizophrenic patients, family members and mental health workers in Taiwan. *Molecular psychiatry* 1999; **4**(1): 33-8.
218. Bode L, Riegel S, Lange W, Ludwig H. Human infections with Borna disease virus: Seroprevalence in patients with chronic diseases and healthy individuals. *Journal of medical virology* 1992; **36**(4): 309-15.
219. Fang ZY, Wen LY, Jin SJ, Zhao ZH. Norwalk-like virus infection found in diarrhea patients in China [Article in Chinese]. *Chinese Journal of Virology* 1995; **11**(03): 215-9.
220. Taylor MB, Schildhauer CI, Parker S, et al. Two successive outbreaks of SRSV-associated gastroenteritis in South Africa. *Journal of medical virology* 1993; **41**(1): 18-23.
221. Nakata S, Estes MK, Chiba S. Detection of Human Calicivirus Antigen and Antibody by Enzyme-Linked Immunosorbent Assays. *J Clin Microbiol* 1988; **26**(10): 2001-5.
222. Wolfaardt M, Taylor MB, Booysen HF, Engelbrecht L, Grabow WO, Jiang X. Incidence of human calicivirus and rotavirus infection in patients with gastroenteritis in South Africa. *Journal of medical virology* 1997; **51**(4): 290-6.

223. Virus Research Group of Kun Number 323 Unit TCPsLA. Isolation, identification and serological studies of a coronavirus strain [Article in Chinese]. *Acta Microbiologica Sinica* 1975; **15**(03): 231-5+62.
224. Hays JP, Myint SH. PCR sequencing of the spike genes of geographically and chronologically distinct human coronaviruses 229E. *J Virol Methods* 1998; **75**(2): 179-93.
225. Esper F, Weibel C, Ferguson D, Landry ML, Kahn JS. Evidence of a novel human coronavirus that is associated with respiratory tract disease in infants and young children. *J Infect Dis* 2005; **191**(4): 492-8.
226. Chan KH, Cheng VC, Woo PC, et al. Serological responses in patients with severe acute respiratory syndrome coronavirus infection and cross-reactivity with human coronaviruses 229E, OC43, and NL63. *Clinical and diagnostic laboratory immunology* 2005; **12**(11): 1317-21.
227. Smuts H, Workman L, Zar HJ. Role of human metapneumovirus, human coronavirus NL63 and human bocavirus in infants and young children with acute wheezing. *Journal of medical virology* 2008; **80**(5): 906-12.
228. Venter M, Lassauniere R, Kresfelder TL, Westerberg Y, Visser A. Contribution of Common and Recently Described Respiratory Viruses to Annual Hospitalizations in Children in South Africa. *Journal of medical virology* 2011; **83**(8): 1458-68.
229. Esper F, Weibel C, Ferguson D, Landry ML, Kahn JS. Coronavirus HKU1 infection in the United States. *Emerging Infectious Diseases* 2006; **12**(5): 775-9.
230. Bialek SR, Allen D, Alvarado-Ramy F, et al. First Confirmed Cases of Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Infection in the United States, Updated Information on the Epidemiology of MERS-CoV Infection, and Guidance for the Public, Clinicians, and Public Health Authorities - May 2014. *Mmwr-Morbid Mortal W* 2014; **63**(19): 431-6.
231. Gao J, Song P. China upgrades surveillance and control measures of Middle East respiratory syndrome (MERS). *Bioscience trends* 2015; **9**(3): 203-4.
232. Abroug F, Slim A, Ouanes-Besbes L, et al. Family Cluster of Middle East Respiratory Syndrome Coronavirus Infections, Tunisia, 2013. *Emerging Infectious Diseases* 2014; **20**(9): 1527-30.
233. Charles M, Cdc. Severe acute respiratory syndrome (SARS) and coronavirus testing - United States, 2003 (Reprinted from MMWR, vol 52, pg 297-302, 2003). *Jama-J Am Med Assoc* 2003; **289**(17): 2203-6.
234. Chiu W-T, Huang J-S, Ho Y-S. Bibliometric analysis of Severe Acute Respiratory Syndrome-related research in the beginning stage. *Scientometrics* 2004; **61**(1): 69-77.
235. CDC. Imported case of Marburg hemorrhagic fever - Colorado, 2008. *MMWR Morb Mortal Wkly Rep* 2009; **58**(49): 1377-81.
236. Gear JS, Cassel GA, Gear AJ, et al. Outbreake of Marburg virus disease in Johannesburg. *British medical journal* 1975; **4**(5995): 489-93.
237. Lavinder CH, Francis E. The Etiology of Dengue. An Attempt to Produce the Disease in the Rhesus Monkey by the Inoculation of Defibrinated Blood. *The Journal of Infectious Diseases* 1914; **15**(2): 341-6.
238. Clarke EJ, Jr., Suitor EC, Jr., Jenkin HM. A serologic survey for arbovirus antibodies in inhabitants of a southwest Taiwan village. *Tropical and geographical medicine* 1967; **19**(4): 326-32.
239. Edington AD. "Dengue " as seen in the Recent Epidemic in Durban. *J Med Assoc S Africa* 1927; **1**(17): 446-8.
240. Pina FP, Merikangas UR. Japanese B-Encephalitis in an American Soldier Returning from Korea. *New England Journal of Medicine* 1953; **249**(13): 531-2.
241. Yen CH. Isolation of a Virus from an Acute Encephalitis Case in Peiping. *Proceedings of the Society for Experimental Biology and Medicine* 1941; **46**(4): 609-11.
242. Simon-Loriere E, Sall AA. Autochthonous Japanese Encephalitis with Yellow Fever Coinfection in Africa. *New England Journal of Medicine* 2017; **376**(15): 1483-5.
243. Wang J, Zhang H, Fu S, et al. Isolation of kyasanur forest disease virus from febrile patient, yunnan, china. *Emerg Infect Dis* 2009; **15**(2): 326-8.
244. Andayi F, Charrel RN, Kieffer A, et al. A Sero-epidemiological Study of Arboviral Fevers in Djibouti, Horn of Africa. *PLoS neglected tropical diseases* 2014; **8**(12).

245. Goldfield M, Austin SM, Black HC, Taylor BF, Altman R. A non-fatal human case of Powassan virus encephalitis. *Am J Trop Med Hyg* 1973; **22**(1): 78-81.
246. Tang Y, Gao X, Diao Y, et al. Tembusu virus in human, China. *Transboundary and emerging diseases* 2013; **60**(3): 193-6.
247. Cruse RP, Rothner AD, Erenberg G, Calisher CH. Central European tick-borne encephalitis: an Ohio case with a history of foreign travel. *American journal of diseases of children (1960)* 1979; **133**(10): 1070-1.
248. Wang WS, Zhao CL. Isolation and identification of forest encephalitis virus [Article in Chinese]. *Journal of Harbin Medical University* 1956; (02): 44-9.
249. Nash D, Mostashari F, Fine A, et al. The outbreak of West Nile virus infection in the New York City area in 1999. *New England Journal of Medicine* 2001; **344**(24): 1807-14.
250. Li XL, Fu SH, Liu WB, et al. West nile virus infection in Xinjiang, China. *Vector Borne Zoonotic Dis* 2013; **13**(2): 131-3.
251. GUITERAS GM. THE YELLOW FEVER EPIDEMIC OF 1903 AT LAREDO, TEXAS. *Jama* 1904; **XLIII**(2): 115-21.
252. Chen J, Lu H. Yellow fever in China is still an imported disease. *Bioscience trends* 2016; **10**(2): 158-62.
253. Stokes A, Bauer JH, Hudson NP. THE TRANSMISSION OF YELLOW FEVER TO MACACUS RHESUS: PRELIMINARY NOTE. *Journal of the American Medical Association* 1928; **90**(4): 253-4.
254. Foy BD, Kobylinski KC, Chilson Foy JL, et al. Probable non-vector-borne transmission of Zika virus, Colorado, USA. *Emerg Infect Dis* 2011; **17**(5): 880-2.
255. Sun H, Jia FJ, Huang BC. Research progress and epidemic situation of the Zika Virus [Article in Chinese]. *Chin J Diagnostics (Electronic Edition)* 2016; **04**(01): 66-9.
256. Xu Z, Shen FM, Xu ZY, Huang QS. HCV infection and primary liver cell cancer [Article in Chinese]. *Tumor (Shanghai)* 1990; **10**(03): 115-05.
257. Kew MC, Houghton M, Choo QL, Kuo G. Hepatitis C virus antibodies in southern African blacks with hepatocellular carcinoma. *Lancet* 1990; **335**(8694): 873-4.
258. Simons JN, Leary TP, Dawson GJ, et al. Isolation of Novel Virus-Like Sequences Associated with Human Hepatitis. *Nat Med* 1995; **1**(6): 564-9.
259. Wang Y, Okamoto H, An P, Chen HS, Liu YL, Wang FS. Infection of hepatitis G virus among blood donors in China [Article in Chinese]. *Journal of Beijing Medical University* 1996; **28**(02): 97.
260. Wang H, Wan Z, Xu R, et al. A Novel Human Pegivirus, HPgV-2 (HHpgV-1), Is Tightly Associated With Hepatitis C Virus (HCV) Infection and HCV/Human Immunodeficiency Virus Type 1 Coinfection. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2018; **66**(1): 29-35.
261. Rodgers MA, Holzmayer V, Vallari A, et al. Hepatitis C virus surveillance and identification of human pegivirus 2 in a large Cameroonian cohort. *Journal of viral hepatitis* 2019; **26**(1): 30-7.
262. Yolken R, Dubovi E, Leister F, Reid R, Almeidohill J, Santosham M. Infantile Gastroenteritis Associated with Excretion of Pestivirus Antigens. *Lancet* 1989; **1**(8637): 517-20.
263. Lee PW, Gajdusek DC, Gibbs CJ, Xu ZY. Aetiological relation between Korean haemorrhagic fever with renal syndrome in People's Republic of China. *Lancet* 1980; **1**(8172): 819-20.
264. Forthal DN, Bauer SP, McCormick JB. Antibody to hemorrhagic fever with renal syndrome viruses (Hantaviruses) in the United States. *American journal of epidemiology* 1987; **126**(6): 1210-3.
265. Tomori O, Morikawa S, Matsuura Y, Kitamura T. Antibody to Japanese strain of haemorrhagic fever with renal syndrome (HFRS) virus in Nigerian sera. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1986; **80**(6): 1008-9.
266. Decock KM, Bradley DW, Sandford NL, Govindarajan S, Maynard JE, Redeker AG. Epidemic Non-a, Non-B Hepatitis in Patients from Pakistan. *Ann Intern Med* 1987; **106**(2): 227-30.
267. Huang RT, Wei J, Tian X, Li DR, Yin SR. Isolation of a small RNA virus from feces of a patient with enterically transmitted Non-A Non-B hepatitis in China [Article in Chinese]. *Journal of Academy of Military Medical Sciences* 1989; **13**(04): 273-7.

268. Belabbes EH, Bouguermouh A, Benatallah A, Illoul G. Epidemic Non-a, Non-B Viral-Hepatitis in Algeria - Strong Evidence for Its Spreading by Water. *Journal of medical virology* 1985; **16**(3): 257-63.
269. Yen YC, Kong LX, Lee L, et al. Characteristics of Crimean-Congo hemorrhagic fever virus (Xinjiang strain) in China. *Am J Trop Med Hyg* 1985; **34**(6): 1179-82.
270. Morrill JC, Johnson BK, Hyams C, et al. Serological evidence of arboviral infections among humans of coastal Kenya. *The Journal of tropical medicine and hygiene* 1991; **94**(3): 166-8.
271. Francis T, Magill TP. IMMUNOLOGICAL STUDIES WITH THE VIRUS OF INFLUENZA. *The Journal of experimental medicine* 1935; **62**(4): 505-16.
272. Chang HT, Chiang YT. Studies on an epidemic of influenza in Peking. *Chin Med J* 1950; **68**(7-8): 185-92.
273. Isaacs A, Andrewes CH. The Spread of Influenza. Evidence from, 1950-1951. *British medical journal* 1951: 921-7.
274. Wen C-Cü, Chu C-M. Survey of Influenza Antibodies in Normal Human Sera in Peking. *Chinese Medical Journal* 1957; **75**(10): 792-801.
275. Montefiore D, Drozdov SG, Kafuko GW, Fayinka OA, Soneji A. Influenza in East Africa, 1969-70. *Bulletin of the World Health Organization* 1970; **43**(2): 269-73.
276. Zhang Q. Process of isolating influenza virus in 1956 [Article in Chinese]. *Biological Products Newsletter* 1957; **2**(1): 80.
277. Joosting AC, Head B, Bynoe ML, Tyrrell DA. Production of common colds in human volunteers by influenza C virus. *British medical journal* 1968; **4**(5624): 153-4.
278. Tang FF, Wu SY, Huang YT, Wen ZQ. Research on the isolation of measles virus [Article in Chinese]. *Chinese Science Bulletin* 1958; (10): 314-5.
279. Baylet R, Schluep R, Cantrelle DS, Rey M. Age-Grouping in Measles in an Urban Environment (A Serological Study). *Bulletin de la Societe Medicale d'Afrique Noire de Langue Francaise* 1963; **8**(6): 771-8.
280. Chen ZH, Zhang EH, Zhang XZ, He NX. Isolation of parainfluenza type I virus by tissue culture and adsorption-hemagglutination test [Article in Chinese]. *Journal of Zhejiang University [Medical Sciences]* 1964; (03): 9-14+49.
281. Taylor-Robinson D, Tyrrell DA. Virus diseases on Tristan da Cunha. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1963; **57**: 19-22.
282. Yu JQ, Chang RX, He CJ, Guan QH, Xie JP. Serological study of 722 infants with viral pneumonia [Article in Chinese]. *Guangdong Medical Journal* 1987; **08**(04): 32-4.
283. Pathogen biology research group Jnmc. Virus isolation in 535 elderly patients with chronic bronchitis and other respiratory infections and antibody tests in some cases [Article in Chinese]. *Jiangsu Medical Journal* 1975; (01): 47-53.
284. Balestrieri A, Russo V, D'Arrigo C. Serum haemagglutination-inhibiting antibodies for haemadsorbing viruses types 1 and 3 and croup-associated in persons in Accra, Ghana. *Arch Ital Sci Med Trop e Parassit* 1967; **48**(11/12): 299-306.
285. Lau SK, To WK, Tse PW, et al. Human parainfluenza virus 4 outbreak and the role of diagnostic tests. *Journal of clinical microbiology* 2005; **43**(9): 4515-21.
286. Niang MN, Diop OM, Sarr FD, et al. Viral etiology of respiratory infections in children under 5 years old living in tropical rural areas of Senegal: The EVIRA project. *Journal of medical virology* 2010; **82**(5): 866-72.
287. Wang TJ, Sun WC, Fang Z, Du SM. Etiology of Mumps in Beijing [Article in Chinese]. *National Medical Journal of China* 1958; **44**(01): 18-29.
288. Bayer P, Gear J. Virus meningo-encephalitis in South Africa; a study of the cases admitted to the Johannesburg Fever Hospital. *South African journal of laboratory and clinical medicine Suid-Afrikaanse tydskrif vir laboratorium- en kliniekwerk* 1955; **1**(1): 22-35.
289. Work TH. Serological evidence of arbovirus infection in the Seminole Indians of southern Florida. *Science* 1964; **145**(3629): 270-2.
290. Gu HX, Spence L, Artsob H, Chia WK, Th'ng C, Lampotang V. Serological evidence of infection with California serogroup viruses (family Bunyaviridae) in residents of Long Hua, suburb of Shanghai, People's Republic of China. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1984; **78**(6): 780-1.

291. Bardos V, Sefcovicova L. The Presence of Antibodies neutralizing Tahyna Virus in the Sera of Inhabitants of some European, Asian, African and Australian Countries. *Journal of Hygiene, Epidemiology, Microbiology and Immunology* 1961; **5**(4): 501-4.
292. Liu W, Sun FJ, Tong YG, Zhang SQ, Cao WC. Rift Valley fever virus imported into China from Angola. *Lancet Infect Dis* 2016; **16**(11): 1226.
293. Sabin AB. Experimental studies on Phlebotomus (pappataci, sandfly) fever during World War II. *Archiv fur die gesamte Virusforschung* 1951; **4**(4): 367-410.
294. McMullan LK, Folk SM, Kelly AJ, et al. A new phlebovirus associated with severe febrile illness in Missouri. *The New England journal of medicine* 2012; **367**(9): 834-41.
295. Grohmann GS, Glass RI, Pereira HG, et al. Enteric viruses and diarrhea in HIV-infected patients. Enteric Opportunistic Infections Working Group. *The New England journal of medicine* 1993; **329**(1): 14-20.
296. Rosen BI, Fang ZY, Glass RI, Monroe SS. Cloning of human picobirnavirus genomic segments and development of an RT-PCR detection assay. *Virology* 2000; **277**(2): 316-29.
297. Luo RH, Xie JQ, Chen YM, Yang SJ. Report of one case of hand-foot-mouth disease in human [Article in Chinese]. *New Medicine* 1999; **30**(03): 37-8.
298. Donia HA, Youssef BZ. Foot and mouth disease (FMD): serological investigation in some farms of Alexandria Governorate of Egypt. *The Journal of the Egyptian Public Health Association* 2002; **77**(3-4): 371-82.
299. Jonkers AH. Serosurvey of encephalomyocarditis virus neutralizing antibodies in southern Louisiana and Peruvian Indian populations. *Am J Trop Med Hyg* 1961; **10**: 593-8.
300. Feng R, Wei J, Zhang H, et al. National serosurvey of encephalomyocarditis virus in healthy people and pigs in China. *Archives of virology* 2015; **160**(12): 2957-64.
301. Dick GW, Best AM, et al. Mengo encephalomyelitis; a hitherto unknown virus affecting man. *Lancet* 1948; **2**(6521): 286-9.
302. Jones MS, Lukashov VV, Ganac RD, Schnurr DP. Discovery of a novel human picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin. *Journal of clinical microbiology* 2007; **45**(7): 2144-50.
303. Xu ZQ, Cheng WX, Qi HM, Cui SX, Jin Y, Duan ZJ. New Saffold cardiovirus in children, China. *Emerg Infect Dis* 2009; **15**(6): 993-4.
304. Zoll J, Erkens Hulshof S, Lanke K, et al. Saffold virus, a human Theiler's-like cardiovirus, is ubiquitous and causes infection early in life. *PLoS Pathog* 2009; **5**(5): e1000416.
305. Dai XQ, Hua XG, Shan TL, Delwart E, Zhao W. Human cosavirus infections in children in China. *J Clin Virol* 2010; **48**(3): 228-9.
306. Yang Y, Ju A, Xu X, Cao X, Tao Y. A novel type of cosavirus from children with nonpolio acute flaccid paralysis. *Virology journal* 2016; **13**(1): 169.
307. Xiao MH, Ye ZZ, Zhang ZL, Tian XQ, Zheng JM, Liu ZY. An epidemic of hand-foot-and-mouth disease due to Coxsackie A16 in Tianjin City [Article in Chinese]. *Tianjin Medical Journal* 1985; **06**(02): 355-7+2.
308. Wu JR, Che JL, Wu GQ, Lin SQ. Investigation on Coxsackie Virus Disease in Fujian Province [Article in Chinese]. *National Medical Journal of China* 1960; **46**(01): 40-8.
309. Patz IM, Measroch V, Gear J. Bornholm disease, pleurodynia or epidemic myalgia; an outbreak in the Transvaal associated with Coxsackie virus infection. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 1953; **27**(19): 397-402.
310. Hudson NP, Lennette EH. The Neutralization of Poliomyelitis Virus by the Serum of Liberian Negroes. *American journal of hygiene* 1933; **17**: 581-6.
311. Institute SEaSDPaT. Isolation and identification of acute hemorrhagic conjunctivitis virus in 1975 [Article in Chinese]. *Chinese Journal of Ophthalmology* 1979; **15**(2): 90.
312. Mirkovic RR, Kono R, Yin-Murphy M, Sohier R, Schmidt NJ, Melnick JL. Enterovirus type 70: the etiologic agent of pandemic acute haemorrhagic conjunctivitis. *Bulletin of the World Health Organization* 1973; **49**(4): 341-6.
313. Price WH. The isolation of a new virus associated with respiratory clinical disease in humans. *Proc Natl Acad Sci U S A* 1956; **42**(12): 892-6.
314. Health GIoMa. Investigation report on virus types in patients with cold in Guangzhou [Article in Chinese]. *Guangdong Medical Journal* 1975; (03): 2-6.
315. Taylor-Robinson D. Studies on some viruses (rhinoviruses) isolated from common colds. *Archiv fur die gesamte Virusforschung* 1963; **13**: 281-93.

316. Hamre D, Procknow JJ. Viruses isolated from natural common colds in the U.S.A. *British medical journal* 1961; **2**(5264): 1382-5.
317. Xiang Z, Gonzalez R, Xie Z, et al. Human rhinovirus group C infection in children with lower respiratory tract infection. *Emerg Infect Dis* 2008; **14**(10): 1665-7.
318. Briese T, Renwick N, Venter M, et al. Global distribution of novel rhinovirus genotype. *Emerg Infect Dis* 2008; **14**(6): 944-7.
319. Lamson D, Renwick N, Kapoor V, et al. MassTag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004-2005. *J Infect Dis* 2006; **194**(10): 1398-402.
320. College MRGoSFM, Hospital LoSSPs. Preliminary report on the examination of hepatitis A antigen particles by immunoelectron microscopy [Article in Chinese]. *Shanghai Medical Journal* 1978; (10): 13-5+65.
321. Szmuness W, Dienstag JL, Purcell RH, et al. The prevalence of antibody to hepatitis A antigen in various parts of the world: a pilot study. *American journal of epidemiology* 1977; **106**(5): 392-8.
322. Chhabra P, Payne DC, Szilagyi PG, et al. Etiology of viral gastroenteritis in children <5 years of age in the United States, 2008-2009. *J Infect Dis* 2013; **208**(5): 790-800.
323. Yang S, Zhang W, Shen Q, et al. Aichi virus strains in children with gastroenteritis, China. *Emerg Infect Dis* 2009; **15**(10): 1703-5.
324. Sdiri-Loulizi K, Gharbi-Khelifi H, de Rougemont A, et al. Acute infantile gastroenteritis associated with human enteric viruses in Tunisia. *Journal of clinical microbiology* 2008; **46**(4): 1349-55.
325. Shan TL, Guo W, Cui L, et al. The first detection of human parechovirus infections in China. *J Clin Virol* 2009; **45**(4): 371-2.
326. Shan T, Wang C, Cui L, et al. Picornavirus salivirus/klassevirus in children with diarrhea, China. *Emerg Infect Dis* 2010; **16**(8): 1303-5.
327. Li L, Victoria J, Kapoor A, et al. A novel picornavirus associated with gastroenteritis. *Journal of virology* 2009; **83**(22): 12002-6.
328. Falsey AR, Erdman D, Anderson LJ, Walsh EE. Human metapneumovirus infections in young and elderly adults. *J Infect Dis* 2003; **187**(5): 785-90.
329. Peiris JS, Tang WH, Chan KH, et al. Children with respiratory disease associated with metapneumovirus in Hong Kong. *Emerg Infect Dis* 2003; **9**(6): 628-33.
330. Madhi SA, Ludewick H, Abed Y, Klugman KP, Boivin G. Human metapneumovirus-associated lower respiratory tract infections among hospitalized human immunodeficiency virus type 1 (HIV-1)-infected and HIV-1-uninfected African infants. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2003; **37**(12): 1705-10.
331. Kun Number 323 Unit tCPsLA. Studies on the isolation and growth characteristics of respiratory syncytial virus [Article in Chinese]. *Acta Microbiologica Sinica* 1975; **15**(02): 125-32.
332. Doggett JE. Antibodies to respiratory syncytial virus in human sera from different regions of the world. *Bulletin of the World Health Organization* 1965; **32**(6): 849-53.
333. Yang JM, Yin GQ, Feng YH, Luo ZY, F. JJ, Zhang ZQ. Superinfection of colti virus and Japanese encephalitis virus [Article in Chinese]. *Journal of Nanjing Railway Medical college* 1996; **15**(03): 29.
334. Tomori O, Fabiyi A, Murphy F. Characterization of Orungo virus, an orbivirus from Uganda and Nigeria. *Archives of virology* 1976; **51**(4): 285-98.
335. Zhao JM, Qiang BQ, Zhao TX, et al. Detection of diarrhoea viruses in children with acute gastroenteritis. *Chinese Journal of Experimental and Clinical Virology* 1995; (01): 45-9.
336. Malherbe H, Roux P, Kahn E. THE ROLE OF ENTEROPATHOGENIC BACTERIA AND VIRUSES IN ACUTE DIARRHOEAL DISORDERS OF INFANCY AND CHILDHOOD IN JOHANNESBURG. II. 'NON-SPECIFIC' GASTRO-ENTERITIS. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 1963; **37**: 259-61.
337. Cheng P, Lau CS, Lai A, et al. A novel reovirus isolated from a patient with acute respiratory disease. *J Clin Virol* 2009; **45**(1): 79-80.
338. Kapikian AZ, Kim HW, Wyatt RG, et al. Recent advances in the aetiology of viral gastroenteritis. *Ciba Foundation symposium* 1976; (42): 273-309.
339. Pang QF, Qiu FX, Yu FR, Chen SZ. Rotavirus-the source of acute gastroenteritis in infants in autumn [Article in Chinese]. *Bulletin of Medical Research* 1979; (07): 26-7.

340. Eiden J, Vonderfecht S, Yolken RH. Evidence that a novel rotavirus-like agent of rats can cause gastroenteritis in man. *Lancet* 1985; **2**(8445): 8-11.
341. Nakata S, Estes MK, Graham DY, Wang SS, Gary GW, Melnick JL. Detection of antibody to group B adult diarrhea rotaviruses in humans. *Journal of clinical microbiology* 1987; **25**(5): 812-8.
342. Jiang B, Dennehy PH, Spangenberger S, Gentsch JR, Glass RI. First detection of group C rotavirus in fecal specimens of children with diarrhea in the United States. *J Infect Dis* 1995; **172**(1): 45-50.
343. Qiao H, Nilsson M, Abreu ER, et al. Viral diarrhea in children in Beijing, China. *Journal of medical virology* 1999; **57**(4): 390-6.
344. Sebata T, Steele AD. Human group C rotavirus identified in South Africa. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 1999; **89**(10): 1073-4.
345. Wang C, Hu C, Huang F, X. C, Hung T. A novel discovered rotavirus from adult acute diarrhoeal patients in China [Article in Chinese]. *Chinese Journal of Virology* 1987; **03**(04): 321-5+411.
346. Williams CK, Alabi GO, Junaid TA, et al. Human T cell leukaemia virus associated lymphoproliferative disease: report of two cases in Nigeria. *British medical journal (Clinical research ed)* 1984; **288**(6429): 1495-6.
347. Ma Y, Zheng S, Wang N, et al. Epidemiological analysis of HTLV-1 and HTLV-2 infection among different population in Central China. *PLoS One* 2013; **8**(6): e66795.
348. Delaporte E, Louwagie J, Peeters M, et al. Evidence of HTLV-II infection in central Africa. *AIDS (London, England)* 1991; **5**(6): 771-2.
349. Safai B, Sarngadharan MG, Groopman JE, et al. Seroepidemiological studies of human T-lymphotropic retrovirus type III in acquired immunodeficiency syndrome. *Lancet* 1984; **1**(8392): 1438-40.
350. Chang RS, Chan RC, French GL, et al. HTLV-III antibody testing in Hong Kong. *Jama* 1986; **256**(1): 41.
351. Brun-Vezinet F, Rouzioux C, Montagnier L, et al. Prevalence of antibodies to lymphadenopathy-associated retrovirus in African patients with AIDS. *Science* 1984; **226**(4673): 453-6.
352. (CDC) CfDC. AIDS due to HIV-2 infection--New Jersey. *MMWR Morb Mortal Wkly Rep* 1988; **37**(3): 33-5.
353. Yan YS, Zheng ZS, Chen G, Zheng J, Yan PP, Shao YM. Confirmation of the first HIV-2 case in China [Article in Chinese]. *Journal of Chinese AIDS&STD Prevention and Control* 2000; **06**(01): 16-8.
354. Huang F, Wang H, Jing S, Zeng W. Simian foamy virus prevalence in Macaca mulatta and zookeepers. *AIDS research and human retroviruses* 2012; **28**(6): 591-3.
355. Black SP, Powers LM. History of Rabies in Southern California. *Cal State J Med* 1910; **8**(11): 369-72.
356. Wu BQ. Report of four cases of rabies encephalitis [Artilcle in Chinese]. *New Medicine* 1981; **12**(07): 357-8.
357. Wilhelm N, Alexis T. Rabies in South Africa: Occurrence and Distribution of Cases During 1932. *Onderstepoort Journal of Veterinary Science and Animal Industry* 1933; **1**: 51-6.
358. CDC. Chikungunya fever diagnosed among international travelers--United States, 2005-2006. *MMWR Morb Mortal Wkly Rep* 2006; **55**(38): 1040-2.
359. Li XD, Qiu FX, Yang H, Rao YN, Calisher CH. Isolation of Getah virus from mosquitos collected on Hainan Island, China, and results of a serosurvey. *The Southeast Asian journal of tropical medicine and public health* 1992; **23**(4): 730-4.
360. Tesh RB, Watts DM, Russell KL, et al. Mayaro virus disease: an emerging mosquito-borne zoonosis in tropical South America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 1999; **28**(1): 67-73.
361. Xu CH, Peng YF, Bai ZJ, et al. Seroepidemiological survey of arbovirus in Hainan Province in 1998 [Article in Chinese]. *Chinese Journal of Epidemiology* 1999; **20**(04): 20.
362. Mathiot CC, Grimaud G, Garry P, et al. An outbreak of human Semliki Forest virus infections in Central African Republic. *Am J Trop Med Hyg* 1990; **42**(4): 386-93.
363. He NX, Xu TZ, Ma JY, et al. Isolation of rubella virus [Article in Chinese]. *Journal of Zhejiang University(Medical Sciences)* 1979; **8**(04): 169-72+217.

364. Selzer G. VIRUS ISOLATION, INCLUSION BODIES, AND CHROMOSOMES IN A RUBELLA-INFECTED HUMAN EMBRYO. *Lancet* 1963; **2**(7303): 336-7.
365. Rizzetto M, Shih JW, Gocke DJ, Purcell RH, Verme G, Gerin JL. Incidence and significance of antibodies to delta antigen in hepatitis B virus infection. *Lancet* 1979; **2**(8150): 986-90.
366. Rizzetto M, Purcell RH, Gerin JL. Epidemiology of HBV-associated delta agent: geographical distribution of anti-delta and prevalence in polytransfused HBsAg carriers. *Lancet* 1980; **1**(8180): 1215-8.
367. Crocchiolo PR, Caredda F, D'Arminio Monforte A, et al. The aetiology of acute hepatitis in Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1984; **78**(4): 514-8.
368. Enria DA, Briggiler AM, Sanchez Z. Treatment of Argentine hemorrhagic fever. *Antiviral research* 2008; **78**(1): 132-9.
369. Weaver SC, Salas RA, de Manzione N, et al. Guanarito virus (Arenaviridae) isolates from endemic and outlying localities in Venezuela: sequence comparisons among and within strains isolated from Venezuelan hemorrhagic fever patients and rodents. *Virology* 2000; **266**(1): 189-95.
370. Fichet-Calvet E, Rogers DJ. Risk maps of Lassa fever in West Africa. *PLoS neglected tropical diseases* 2009; **3**(3): e388.
371. Jamieson DJ, Kourtis AP, Bell M, Rasmussen SA. Lymphocytic choriomeningitis virus: an emerging obstetric pathogen? *American journal of obstetrics and gynecology* 2006; **194**(6): 1532-6.
372. Drager S, Marx AF, Pigny F, et al. Lymphocytic choriomeningitis virus meningitis after needlestick injury: a case report. *Antimicrobial resistance and infection control* 2019; **8**: 77.
373. Aguilar PV, Camargo W, Vargas J, et al. Reemergence of Bolivian hemorrhagic fever, 2007-2008. *Emerg Infect Dis* 2009; **15**(9): 1526-8.
374. Milazzo ML, Campbell GL, Fulhorst CF. Novel arenavirus infection in humans, United States. *Emerg Infect Dis* 2011; **17**(8): 1417-20.
375. Bosch A, Pintó RM, Guix S. Human Astroviruses. *Clinical Microbiology Reviews* 2014; **27**(4): 1048.
376. Tan Y, He WT, Chen MM, Mo JJ, Ju Y, Chen M. An outbreak of human astrovirus lineage 1b in a middle school in Guangxi, Southern China in 2017. *Chin Med J (Engl)* 2019; **132**(3): 336-8.
377. Chalmers RM, Thomas DR, Salmon RL. Borna disease virus and the evidence for human pathogenicity: a systematic review. *QJM : monthly journal of the Association of Physicians* 2005; **98**(4): 255-74.
378. Tappe D, Frank C, Homeier-Bachmann T, et al. Analysis of exotic squirrel trade and detection of human infections with variegated squirrel bornavirus 1, Germany, 2005 to 2018. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2019; **24**(8).
379. Ahmed SM, Hall AJ, Robinson AE, et al. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. *Lancet Infect Dis* 2014; **14**(8): 725-30.
380. Oka T, Wang Q, Katayama K, Saif LJ. Comprehensive review of human sapoviruses. *Clinical microbiology reviews* 2015; **28**(1): 32-53.
381. Supadej K, Khamrin P, Kumthip K, et al. Distribution of norovirus and sapovirus genotypes with emergence of NoV GII.P16/GII.2 recombinant strains in Chiang Mai, Thailand. *Journal of medical virology* 2019; **91**(2): 215-24.
382. Smith AW, Iversen PL, Skilling DE, Stein DA, Bok K, Matson DO. Vesivirus viremia and seroprevalence in humans. *Journal of medical virology* 2006; **78**(5): 693-701.
383. Su S, Wong G, Shi W, et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol* 2016; **24**(6): 490-502.
384. Heimdal I, Moe N, Krokstad S, et al. Human Coronavirus in Hospitalized Children With Respiratory Tract Infections: A 9-Year Population-Based Study From Norway. *J Infect Dis* 2019; **219**(8): 1198-206.
385. Mackay IM, Arden KE. MERS coronavirus: diagnostics, epidemiology and transmission. *Virology journal* 2015; **12**: 222.

386. Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus associated with severe acute respiratory syndrome. *The New England journal of medicine* 2003; **348**(20): 1953-66.
387. Albarino CG, Shoemaker T, Khristova ML, et al. Genomic analysis of filoviruses associated with four viral hemorrhagic fever outbreaks in Uganda and the Democratic Republic of the Congo in 2012. *Virology* 2013; **442**(2): 97-100.
388. Gatherer D. The 2014 Ebola virus disease outbreak in West Africa. *The Journal of general virology* 2014; **95**(Pt 8): 1619-24.
389. Barrette RW, Metwally SA, Rowland JM, et al. Discovery of swine as a host for the Reston ebolavirus. *Science* 2009; **325**(5937): 204-6.
390. Muyembe-Tamfum JJ, Mulangu S, Masumu J, Kayembe JM, Kemp A, Paweska JT. Ebola virus outbreaks in Africa: past and present. *Onderstepoort J Vet Res* 2012; **79**(2): 451.
391. Tseng CP, Chan YJ. Overview of Ebola virus disease in 2014. *Journal of the Chinese Medical Association : JCMA* 2015; **78**(1): 51-5.
392. Ebola Outbreak Epidemiology Team. Outbreak of Ebola virus disease in the Democratic Republic of the Congo, April-May, 2018: an epidemiological study. *Lancet* 2018; **392**(10143): 213-21.
393. Brauburger K, Hume AJ, Muhlberger E, Olejnik J. Forty-five years of Marburg virus research. *Viruses* 2012; **4**(10): 1878-927.
394. Srihongse S, Stacy HG, Gauld JR. A survey to assess potential human disease hazards along proposed sea level canal routes in Panama and Colombia. IV. Arbovirus surveillance in man. *Military medicine* 1973; **138**(7): 422-6.
395. Fulop LD, Barrett AD, Titball RW. Nucleotide sequence of the NS5 gene of Banzi virus: comparison with other flaviviruses. *The Journal of general virology* 1995; **76 (Pt 9)**: 2317-21.
396. Kokernot RH, Casaca VM, Weinbren MP, McIntosh BM. Survey for antibodies against arthropod-borne viruses in the sera of indigenous residents of Angola. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1965; **59**(5): 563-70.
397. Macdonald J, Poidinger M, Mackenzie JS, et al. Molecular phylogeny of edge hill virus supports its position in the yellow Fever virus group and identifies a new genetic variant. *Evolutionary bioinformatics online* 2010; **6**: 91-6.
398. Aaskov JG, Phillips DA, Wiemers MA. Possible clinical infection with Edge Hill virus. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1993; **87**(4): 452-3.
399. Venegas EA, Aguilar PV, Cruz C, et al. Ilheus virus infection in human, Bolivia. *Emerg Infect Dis* 2012; **18**(3): 516-8.
400. Gao X, Liu H, Li X, et al. Changing Geographic Distribution of Japanese Encephalitis Virus Genotypes, 1935-2017. *Vector Borne Zoonotic Dis* 2019; **19**(1): 35-44.
401. Poidinger M, Hall RA, Lindsay MD, Broom AK, Mackenzie JS. The molecular epidemiology of Kokobera virus. *Virus research* 2000; **68**(1): 7-13.
402. Hawkes RA, Pamplin J, Boughton CR, Naim HM. Arbovirus infections of humans in high-risk areas of south-eastern Australia: a continuing study. *Med J Aust* 1993; **159**(3): 159-62.
403. Shah SZ, Jabbar B, Ahmed N, et al. Epidemiology, Pathogenesis, and Control of a Tick-Borne Disease- Kyasanur Forest Disease: Current Status and Future Directions. *Frontiers in cellular and infection microbiology* 2018; **8**: 149.
404. Naren Babu N, Jayaram A, Hemanth Kumar H, et al. Spatial distribution of Haemaphysalis species ticks and human Kyasanur Forest Disease cases along the Western Ghats of India, 2017-2018. *Experimental & applied acarology* 2019; **77**(3): 435-47.
405. Hoogstraal H. Ticks in relation to human diseases caused by viruses. *Annual review of entomology* 1966; **11**: 261-308.
406. Padbidri VS, Wairagkar NS, Joshi GD, et al. A serological survey of arboviral diseases among the human population of the Andaman and Nicobar Islands, India. *The Southeast Asian journal of tropical medicine and public health* 2002; **33**(4): 794-800.
407. Jeffries CL, Mansfield KL, Phipps LP, et al. Louping ill virus: an endemic tick-borne disease of Great Britain. *The Journal of general virology* 2014; **95**(Pt 5): 1005-14.
408. Davidson MM, Williams H, Macleod JA. Louping ill in man: a forgotten disease. *The Journal of infection* 1991; **23**(3): 241-9.

409. Selvey LA, Dailey L, Lindsay M, et al. The changing epidemiology of Murray Valley encephalitis in Australia: the 2011 outbreak and a review of the literature. *PLoS neglected tropical diseases* 2014; **8**(1): e2656.
410. Dilcher M, Sall AA, Hufert FT, Weidmann M. Full-length genome sequence of Ntaya virus. *Virus genes* 2013; **46**(1): 162-4.
411. Woodruff AW, Bowen ET, Platt GS. Viral infections in travellers from tropical Africa. *British medical journal* 1978; **1**(6118): 956-8.
412. Ruzek D, Yakimenko VV, Karan LS, Tkachev SE. Omsk haemorrhagic fever. *Lancet* 2010; **376**(9758): 2104-13.
413. Corrin T, Greig J, Harding S, Young I, Mascarenhas M, Waddell LA. Powassan virus, a scoping review of the global evidence. *Zoonoses and public health* 2018.
414. Hendricks DA, Hardy JL, Reeves WC. Comparison of biological properties of St. Louis encephalitis and Rio bravo viruses. *Am J Trop Med Hyg* 1983; **32**(3): 602-9.
415. Price JL. Isolation of Rio Bravo and a hitherto undescribed agent, Tamana bat virus, from insectivorous bats in Trinidad, with serological evidence of infection in bats and man. *Am J Trop Med Hyg* 1978; **27**(1 Pt 1): 153-61.
416. Curren EJ, Lindsey NP, Fischer M, Hills SL. St. Louis Encephalitis Virus Disease in the United States, 2003-2017. *Am J Trop Med Hyg* 2018; **99**(4): 1074-9.
417. Dobler G, Gniel D, Petermann R, Pfeffer M. Epidemiology and distribution of tick-borne encephalitis. *Wiener medizinische Wochenschrift (1946)* 2012; **162**(11-12): 230-8.
418. Baba SS, Fagbami AH, Ojeh CK. Preliminary studies on the use of solid-phase immunosorbent techniques for the rapid detection of Wesselsbron virus (WSLV) IgM by haemagglutination-inhibition. *Comparative immunology, microbiology and infectious diseases* 1999; **22**(1): 71-9.
419. Gaibani P, Rossini G. An overview of Usutu virus. *Microbes Infect* 2017; **19**(7-8): 382-7.
420. Carletti F, Colavita F, Rovida F, et al. Expanding Usutu virus circulation in Italy: detection in the Lazio region, central Italy, 2017 to 2018. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2019; **24**(3).
421. Weyer J, Thomas J, Leman PA, Grobbelaar AA, Kemp A, Paweska JT. Human cases of Wesselsbron disease, South Africa 2010-2011. *Vector Borne Zoonotic Dis* 2013; **13**(5): 330-6.
422. Diagne MM, Faye M, Faye O, et al. Emergence of Wesselsbron virus among black rat and humans in Eastern Senegal in 2013. *One health (Amsterdam, Netherlands)* 2017; **3**: 23-8.
423. Chancey C, Grinev A, Volkova E, Rios M. The global ecology and epidemiology of West Nile virus. *BioMed research international* 2015; **2015**: 376230.
424. Shearer FM, Longbottom J, Browne AJ, et al. Existing and potential infection risk zones of yellow fever worldwide: a modelling analysis. *Lancet Glob Health* 2018; **6**(3): E270-E8.
425. Petersen LR, Jamieson DJ, Powers AM, Honein MA. Zika Virus. *The New England journal of medicine* 2016; **374**(16): 1552-63.
426. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *Journal of hepatology* 2014; **61**(1 Suppl): S45-57.
427. Brezzi M, Bertisch B, Roelens M, et al. Impact of geographic origin on access to therapy and therapy outcomes in the Swiss Hepatitis C Cohort Study. *PLoS One* 2019; **14**(6): e0218706.
428. Feng Y, Zhao W, Feng Y, et al. A novel genotype of GB virus C: its identification and predominance among injecting drug users in Yunnan, China. *PLoS One* 2011; **6**(10): e21151.
429. Shahid M, Idrees M, Butt AM, et al. Short article: Hepatitis C and G virus coinfection in Punjab, Pakistan: incidence and its correlation analysis with clinical data. *European journal of gastroenterology & hepatology* 2019; **31**(3): 389-92.
430. Shui J, Liu W, Liang Y, et al. Infection of human pegivirus 2 (HPgV-2) is associated with hepatitis C virus but not hepatitis B virus infection in people who inject drugs. *The Journal of general virology* 2019; **100**(6): 968-74.
431. Giangaspero M. Pestivirus Species Potential Adventitious Contaminants of Biological Products; 2013.

432. Giangaspero M, Cominardi PF. Parsonage-Turner syndrome associated with anti-bovine viral diarrhoea virus antibodies. *Veterinaria italiana* 2006; **42**(3): 255-9, 49-53.
433. Valdivieso F, Vial P, Ferres M, et al. Neutralizing antibodies in survivors of Sin Nombre and Andes hantavirus infection. *Emerg Infect Dis* 2006; **12**(1): 166-8.
434. Alonso DO, Iglesias A, Coelho R, et al. Epidemiological description, case-fatality rate, and trends of Hantavirus Pulmonary Syndrome: 9 years of surveillance in Argentina. *Journal of medical virology* 2019; **91**(7): 1173-81.
435. Knust B, Rollin PE. Twenty-year summary of surveillance for human hantavirus infections, United States. *Emerg Infect Dis* 2013; **19**(12): 1934-7.
436. Armien B, Pascale JM, Munoz C, et al. Incidence rate for hantavirus infections without pulmonary syndrome, Panama. *Emerg Infect Dis* 2011; **17**(10): 1936-9.
437. Klempa B, Avsic-Zupanc T, Clement J, et al. Complex evolution and epidemiology of Dobrava-Belgrade hantavirus: definition of genotypes and their characteristics. *Archives of virology* 2013; **158**(3): 521-9.
438. Dzagurova TK, Tkachenko EA, Ishmukhametov AA, Balovneva MV, Klempa B, Kruger DH. Severe hantavirus disease in children. *J Clin Virol* 2018; **101**: 66-8.
439. Kariwa H, Yoshimatsu K, Arikawa J. Hantavirus infection in East Asia. *Comparative immunology, microbiology and infectious diseases* 2007; **30**(5-6): 341-56.
440. Kim HK, Chung JH, Kim DM, Yun NR, Kim CM, Jalal S. Hemorrhagic Fever with Renal Syndrome as a Cause of Acute Diarrhea. *Am J Trop Med Hyg* 2019; **100**(5): 1236-9.
441. Milholland MT, Castro-Arellano I, Suzan G, et al. Global Diversity and Distribution of Hantaviruses and Their Hosts. *Ecohealth* 2018; **15**(1): 163-208.
442. Reynes JM, Carli D, Thomas D, Castel G. Puumala Hantavirus Genotypes in Humans, France, 2012-2016. *Emerg Infect Dis* 2019; **25**(1): 140-3.
443. Kruger DH, Figueiredo LT, Song JW, Klempa B. Hantaviruses--globally emerging pathogens. *J Clin Virol* 2015; **64**: 128-36.
444. Heinemann P, Tia M, Alabi A, et al. Human Infections by Non-Rodent-Associated Hantaviruses in Africa. *J Infect Dis* 2016; **214**(10): 1507-11.
445. Park Y. Epidemiologic study on changes in occurrence of hemorrhagic fever with renal syndrome in Republic of Korea for 17 years according to age group: 2001-2017. *BMC infectious diseases* 2019; **19**(1): 153.
446. Wieder Finesod A, Litachevski V, Yonath H, et al. [HANTAVIRUS PULMONARY SYNDROME DIAGNOSED IN A RETURNED TRAVELER FROM THE USA]. *Harefuah* 2019; **158**(5): 305-8.
447. Yoshimatsu K, Gamage CD, Sarathkumara YD, et al. Thailand orthohantavirus infection in patients with chronic kidney disease of unknown aetiology in Sri Lanka. *Archives of virology* 2019; **164**(1): 267-71.
448. Reynes JM, Carli D, Boukezia N, Debruyne M, Herti S. Tula hantavirus infection in a hospitalised patient, France, June 2015. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2015; **20**(50).
449. Khuroo MS, Khuroo MS, Khuroo NS. Hepatitis E: Discovery, global impact, control and cure. *World journal of gastroenterology* 2016; **22**(31): 7030-45.
450. Tulen AD, Vennema H, van Pelt W, Franz E, Hofhuis A. A case-control study into risk factors for acute hepatitis E in the Netherlands, 2015-2017. *The Journal of infection* 2019; **78**(5): 373-81.
451. Andonov A, Robbins M, Borlang J, et al. Rat hepatitis E virus linked to severe acute hepatitis in an immunocompetent patient. *J Infect Dis* 2019.
452. Ergonul O. Crimean-Congo haemorrhagic fever. *Lancet Infect Dis* 2006; **6**(4): 203-14.
453. Burt FJ, Spencer DC, Leman PA, Patterson B, Swanepoel R. Investigation of tick-borne viruses as pathogens of humans in South Africa and evidence of Dugbe virus infection in a patient with prolonged thrombocytopenia. *Epidemiology and infection* 1996; **116**(3): 353-61.
454. Yadav PD, Vincent MJ, Khristova M, et al. Genomic analysis reveals Nairobi sheep disease virus to be highly diverse and present in both Africa, and in India in the form of the Ganjam virus variant. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2011; **11**(5): 1111-20.

455. Chastel C. [Erve and Eyach: two viruses isolated in France, neuropathogenic for man and widely distributed in Western Europe]. *Bulletin de l'Academie nationale de medecine* 1998; **182**(4): 801-9; discussion 9-10.
456. Woessner R, Grauer MT, Langenbach J, et al. The Erve virus: possible mode of transmission and reservoir. *Infection* 2000; **28**(3): 164-6.
457. Cox NJ, Subbarao K. Global epidemiology of influenza: past and present. *Annual review of medicine* 2000; **51**: 407-21.
458. Xu X, Blanton L, Elal AIA, et al. Update: Influenza Activity in the United States During the 2018-19 Season and Composition of the 2019-20 Influenza Vaccine. *MMWR Morb Mortal Wkly Rep* 2019; **68**(24): 544-51.
459. Caini S, Huang QS, Ciblak MA, et al. Epidemiological and virological characteristics of influenza B: results of the Global Influenza B Study. *Influenza and other respiratory viruses* 2015; **9 Suppl 1**: 3-12.
460. Moriuchi H, Katsushima N, Nishimura H, Nakamura K, Numazaki Y. Community-acquired influenza C virus infection in children. *The Journal of pediatrics* 1991; **118**(2): 235-8.
461. Lee HS, Lim S, Noh JY, et al. Identification of influenza C virus in young South Korean children, from October 2013 to September 2016. *J Clin Virol* 2019; **115**: 47-52.
462. Hubalek Z, Rudolf I. Tick-borne viruses in Europe. *Parasitology research* 2012; **111**(1): 9-36.
463. Butenko AM, Leshchinskaia EV, Semashko IV, Donets MA, Mart'ianova LI. [Dhori virus--a causative agent of human disease. 5 cases of laboratory infection]. *Voprosy virusologii* 1987; **32**(6): 724-9.
464. Evans AS. Pathogenicity and immunology of Newcastle disease virus (NVD) in man. *American journal of public health and the nation's health* 1955; **45**(6): 742-5.
465. Ahad A, Rabbani M, Yaqub T, et al. Detection of antibody to Newcastle disease virus in human sera in Pakistan. *JAPS, Journal of Animal and Plant Sciences* 2013; **23**(4): 990-4.
466. Rockx B, Winegar R, Freiberg AN. Recent progress in henipavirus research: molecular biology, genetic diversity, animal models. *Antiviral research* 2012; **95**(2): 135-49.
467. Yadav PD, Shete AM, Kumar GA, et al. Nipah Virus Sequences from Humans and Bats during Nipah Outbreak, Kerala, India, 2018. *Emerg Infect Dis* 2019; **25**(5): 1003-6.
468. Rima BK, Duprex WP. Morbilliviruses and human disease. *The Journal of pathology* 2006; **208**(2): 199-214.
469. Zhang X, Wallace OL, Domi A, et al. Canine distemper virus neutralization activity is low in human serum and it is sensitive to an amino acid substitution in the hemagglutinin protein. *Virology* 2015; **482**: 218-24.
470. Moss WJ. Measles. *Lancet* 2017; **390**(10111): 2490-502.
471. Taylor-Robinson D. Respiratory virus antibodies in human sera from different regions of the world. *Bulletin of the World Health Organization* 1965; **32**(6): 833-47.
472. Sonawane AA, Shastri J, Bavdekar SB. Respiratory Pathogens in Infants Diagnosed with Acute Lower Respiratory Tract Infection in a Tertiary Care Hospital of Western India Using Multiplex Real Time PCR. *Indian journal of pediatrics* 2019; **86**(5): 433-8.
473. Branche AR, Falsey AR. Parainfluenza Virus Infection. *Seminars in respiratory and critical care medicine* 2016; **37**(4): 538-54.
474. Chen Z. Parainfluenza virus 5-vectored vaccines against human and animal infectious diseases. *Rev Med Virol* 2018; **28**(2).
475. Vandvik B, Norrby E. Paramyxovirus SV5 and multiple sclerosis. *Nature* 1989; **338**(6218): 769-71.
476. Wang LF. Discovering novel zoonotic viruses. *New South Wales public health bulletin* 2011; **22**(5-6): 113-7.
477. Hviid A, Rubin S, Muhlemann K. Mumps. *Lancet* 2008; **371**(9616): 932-44.
478. McDonald R, Ruppert PS, Souto M, et al. Notes from the Field: Measles Outbreaks from Imported Cases in Orthodox Jewish Communities - New York and New Jersey, 2018-2019. *MMWR Morb Mortal Wkly Rep* 2019; **68**(19): 444-5.
479. Nishio M, Tsurudome M, Bando H, Ito Y. Immunological relationships of simian virus 41 (SV41) to other paramyxoviruses and serological evidence of SV41 infection in human populations. *The Journal of general virology* 1990; **71 (Pt 9)**: 2093-7.
480. Tee KK, Takebe Y, Kamarulzaman A. Emerging and re-emerging viruses in Malaysia, 1997-2007. *Int J Infect Dis* 2009; **13**(3): 307-18.

481. Campbell GL, Mataczynski JD, Reisdorf ES, et al. Second human case of Cache Valley virus disease. *Emerg Infect Dis* 2006; **12**(5): 854-6.
482. Wilson MR, Suan D, Duggins A, et al. A novel cause of chronic viral meningoencephalitis: Cache Valley virus. *Annals of neurology* 2017; **82**(1): 105-14.
483. Lutwama JJ, Rwaguma EB, Nawanga PL, Mukuye A. Isolations of Bwamba virus from south central Uganda and north eastern Tanzania. *African health sciences* 2002; **2**(1): 24-8.
484. Hubalek Z. Mosquito-borne viruses in Europe. *Parasitology research* 2008; **103 Suppl 1**: S29-43.
485. Byrd BD, Williams CJ, Staples JE, Burkhalter KL, Savage HM, Doyle MS. Notes from the Field: Spatially Associated Coincident and Noncoincident Cases of La Crosse Encephalitis - North Carolina, 2002-2017. *MMWR Morb Mortal Wkly Rep* 2018; **67**(39): 1104-5.
486. Nunes MR, Travassos da Rosa AP, Weaver SC, Tesh RB, Vasconcelos PF. Molecular epidemiology of group C viruses (Bunyaviridae, Orthobunyavirus) isolated in the Americas. *Journal of virology* 2005; **79**(16): 10561-70.
487. Hontz RD, Guevara C, Halsey ES, et al. Itaya virus, a Novel Orthobunyavirus Associated with Human Febrile Illness, Peru. *Emerg Infect Dis* 2015; **21**(5): 781-8.
488. Tikasingh ES, Ardoin P, Williams MC. First isolation of Catu virus from a human in Trinidad. *Tropical and geographical medicine* 1974; **26**(4): 414-6.
489. Aguilar PV, Morrison AC, Rocha C, et al. Guaroa virus infection among humans in Bolivia and Peru. *Am J Trop Med Hyg* 2010; **83**(3): 714-21.
490. Groseth A, Wollenberg KR, Mampilli V, et al. Spatiotemporal analysis of Guaroa virus diversity, evolution, and spread in South America. *Emerg Infect Dis* 2015; **21**(3): 460-3.
491. Dunlop JI, Szemiel AM, Navarro A, et al. Development of reverse genetics systems and investigation of host response antagonism and reassortment potential for Cache Valley and Kairi viruses, two emerging orthobunyaviruses of the Americas. *PLoS neglected tropical diseases* 2018; **12**(10): e0006884.
492. Tauro LB, Almeida FL, Contigiani MS. First detection of human infection by Cache Valley and Kairi viruses (Orthobunyavirus) in Argentina. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2009; **103**(2): 197-9.
493. Groseth A, Mampilli V, Weisend C, et al. Molecular characterization of human pathogenic bunyaviruses of the Nyando and Bwamba/Pongola virus groups leads to the genetic identification of Mojui dos Campos and Kaeng Khoi virus. *PLoS neglected tropical diseases* 2014; **8**(9): e3147.
494. Digoutte JP, Gagnard VJ, Bres P, Pajot FX. [Nyando virus infection in man]. *Bulletin de la Societe de pathologie exotique et de ses filiales* 1972; **65**(6): 751-8.
495. Travassos da Rosa JF, de Souza WM, Pinheiro FP, et al. Oropouche Virus: Clinical, Epidemiological, and Molecular Aspects of a Neglected Orthobunyavirus. *Am J Trop Med Hyg* 2017; **96**(5): 1019-30.
496. Golender N, Wernike K, Bumbarov V, et al. Characterization of Shuni viruses detected in Israel. *Virus genes* 2016; **52**(6): 806-13.
497. van Eeden C, Swanepoel R, Venter M. Antibodies against West Nile and Shuni Viruses in Veterinarians, South Africa. *Emerg Infect Dis* 2014; **20**(8): 1409-11.
498. de Melo ABJ, de Souza WM, Acrani GO, et al. Genomic characterization and evolution of Tacaiuma orthobunyavirus (Peribunyaviridae family) isolated in Brazil. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2018; **60**: 71-6.
499. Vasconcelos PFC, Travassos da Rosa APA, Travassos da Rosa JFS, Dégallier N. Concomitant infections by malaria and arboviruses in the Brazilian Amazon Region. *Revista Latinoamericana de Microbiología* 1990; **32**(4): 291-4.
500. Chowdhary R, Street C, Travassos da Rosa A, et al. Genetic characterization of the Wyeomyia group of orthobunyaviruses and their phylogenetic relationships. *The Journal of general virology* 2012; **93**(Pt 5): 1023-34.
501. O DESL, L DEAS, Fonseca IE, Lacerda JP. Bertioga (Guama group) and Anhembi (Bunyamwera group), two new arboviruses isolated in Sao Paulo, Brazil. *Am J Trop Med Hyg* 1975; **24**(1): 131-4.
502. Palacios G, Tesh R, Travassos da Rosa A, et al. Characterization of the Candiru antigenic complex (Bunyaviridae: Phlebovirus), a highly diverse and reassorting group of viruses affecting humans in tropical America. *Journal of virology* 2011; **85**(8): 3811-20.

503. Gowen BB, Hoopes JD, Wong MH, et al. TLR3 deletion limits mortality and disease severity due to Phlebovirus infection. *Journal of immunology* 2006; **177**(9): 6301-7.
504. Millar JW. PHLEBOTOMUS FEVER VIRUSES IN PANAMA, 1981.
505. Clark MHA, Warimwe GM, Di Nardo A, Lyons NA, Gubbins S. Systematic literature review of Rift Valley fever virus seroprevalence in livestock, wildlife and humans in Africa from 1968 to 2016. *PLoS neglected tropical diseases* 2018; **12**(7): e0006627.
506. Papa A, Konstantinou G, Pavlidou V, Antoniadis A. Sandfly fever virus outbreak in Cyprus. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2006; **12**(2): 192-4.
507. Stahn B, Sudeck H, Frickmann H, Kruger A, Burchard HG, Wiemer D. [Sandfly fever-a "neglected" disease]. *Der Hautarzt; Zeitschrift fur Dermatologie, Venerologie, und verwandte Gebiete* 2018; **69**(11): 928-37.
508. Brault AC, Savage HM, Duggal NK, Eisen RJ, Staples JE. Heartland Virus Epidemiology, Vector Association, and Disease Potential. *Viruses* 2018; **10**(9).
509. Lindsey NP, Menitove JE, Biggerstaff BJ, et al. Seroprevalence of Heartland Virus Antibodies in Blood Donors, Northwestern Missouri, USA. *Emerg Infect Dis* 2019; **25**(2): 358-60.
510. Hu J, Shi C, Li Z, et al. A cluster of cases of severe fever with thrombocytopenia syndrome bunyavirus infection in China, 1996: A retrospective serological study. *PLoS neglected tropical diseases* 2018; **12**(6): e0006603.
511. Li H, Lu QB, Xing B, et al. Epidemiological and clinical features of laboratory-diagnosed severe fever with thrombocytopenia syndrome in China, 2011-17: a prospective observational study. *Lancet Infect Dis* 2018; **18**(10): 1127-37.
512. Traavik T, Mehl R. Uukuniemi group viruses isolated in Norway. *Archives of virology* 1977; **54**(4): 317-31.
513. Malik YS, Sircar S, Saurabh S, et al. Epidemiologic Status of Picobirnavirus in India, A Less Explored Viral Disease. *The open virology journal* 2018; **12**: 99-109.
514. Sun G, Zang Q, Gu Y, Niu G, Ding C, Zhang P. Viral metagenomics analysis of picobirnavirus-positive feces from children with sporadic diarrhea in China. *Archives of virology* 2016; **161**(4): 971-5.
515. Prempeh H, Smith R, Muller B. Foot and mouth disease: the human consequences. The health consequences are slight, the economic ones huge. *BMJ (Clinical research ed)* 2001; **322**(7286): 565-6.
516. Berrios EP. [Foot and mouth disease in human beings. A human case in Chile]. *Revista chilena de infectologia : organo oficial de la Sociedad Chilena de Infectologia* 2007; **24**(2): 160-3.
517. Tesh RB. The prevalence of encephalomyocarditis virus neutralizing antibodies among various human populations. *Am J Trop Med Hyg* 1978; **27**(1 Pt 1): 144-9.
518. Tan SZ, Tan MZ, Prabakaran M. Saffold virus, an emerging human cardiovirus. *Rev Med Virol* 2017; **27**(1).
519. Itagaki T, Aoki Y, Matoba Y, et al. Detection of Saffold viruses from children with acute respiratory infections in Yamagata, Japan, between 2008 and 2015. *Journal of medical virology* 2018; **90**(1): 34-40.
520. Ayouni S, Estienney M, Hammami S, et al. Cosavirus, Salivirus and Bufavirus in Diarrheal Tunisian Infants. *PLoS One* 2016; **11**(9): e0162255.
521. Aiemjoy K, Altan E, Aragie S, et al. Viral species richness and composition in young children with loose or watery stool in Ethiopia. *BMC infectious diseases* 2019; **19**(1): 53.
522. Maan HS, Chowdhary R, Shakya AK, Dhole TN. Genetic diversity of cosaviruses in nonpolio acute flaccid paralysis cases of undefined etiology, Northern India, 2010-2011. *J Clin Virol* 2013; **58**(1): 183-7.
523. Menage L, Yodmeeklin A, Khamrin P, Kumthip K, Maneekarn N. Prevalence of human cosavirus and saffold virus with an emergence of saffold virus genotype 6 in patients hospitalized with acute gastroenteritis in Chiang Mai, Thailand, 2014-2016. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2017; **53**: 1-6.
524. Campanini G, Rovida F, Meloni F, et al. Persistent human cosavirus infection in lung transplant recipient, Italy. *Emerg Infect Dis* 2013; **19**(10): 1667-9.

525. Siqueira JD, Dominguez-Bello MG, Contreras M, et al. Complex virome in feces from Amerindian children in isolated Amazonian villages. *Nature communications* 2018; **9**(1): 4270.
526. Oberste MS, Penaranda S, Maher K, Pallansch MA. Complete genome sequences of all members of the species Human enterovirus A. *The Journal of general virology* 2004; **85**(Pt 6): 1597-607.
527. Feist H, Turowski G, Hussein K, Blocker T, Heim A. Massive Perivillous Fibrin Deposition of an Enterovirus A-Infected Placenta Associated With Stillbirth: A Case Report. *Pediatric and developmental pathology : the official journal of the Society for Pediatric Pathology and the Paediatric Pathology Society* 2019; **22**(2): 142-5.
528. Oberste MS, Maher K, Nix WA, et al. Molecular identification of 13 new enterovirus types, EV79-88, EV97, and EV100-101, members of the species Human Enterovirus B. *Virus research* 2007; **128**(1-2): 34-42.
529. Kahrs CR, Chuda K, Tapia G, et al. Enterovirus as trigger of coeliac disease: nested case-control study within prospective birth cohort. *BMJ (Clinical research ed)* 2019; **364**: l231.
530. Lukashev AN, Drexler JF, Kotova VO, et al. Novel serotypes 105 and 116 are members of distinct subgroups of human enterovirus C. *The Journal of general virology* 2012; **93**(Pt 11): 2357-62.
531. Du J, Zhu T, Zhuang L, et al. Identification and complete genome characterization of human enterovirus 117 from a child with pneumonia in China. *Archives of virology* 2019; **164**(5): 1489-92.
532. Knoester M, Helfferich J, Poelman R, Van Leer-Buter C, Brouwer OF, Niesters HGM. Twenty-nine Cases of Enterovirus-D68-associated Acute Flaccid Myelitis in Europe 2016: A Case Series and Epidemiologic Overview. *The Pediatric infectious disease journal* 2019; **38**(1): 16-21.
533. Kirolos A, Mark K, Shetty J, et al. Outcome of paediatric acute flaccid myelitis associated with enterovirus D68: a case series. *Developmental medicine and child neurology* 2019; **61**(3): 376-80.
534. Gur S, Yapkic O, Yilmaz A. Serological survey of bovine enterovirus type 1 in different mammalian species in Turkey. *Zoonoses and public health* 2008; **55**(2): 106-11.
535. Liu Z, Donahue RE, Young NS, Brown KE. Sequencing and characterization of A-2 plaque virus: A new member of the Picornaviridae family. *Virology* 2000; **272**(1): 168-76.
536. Hasegawa K, Jartti T, Bochkov YA, et al. Rhinovirus Species in Children With Severe Bronchiolitis: Multicenter Cohort Studies in the United States and Finland. *The Pediatric infectious disease journal* 2019; **38**(3): e59-e62.
537. Vaughan G, Goncalves Rossi LM, Forbi JC, et al. Hepatitis A virus: host interactions, molecular epidemiology and evolution. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2014; **21**: 227-43.
538. Foster MA, Hofmeister MG, Kupronis BA, et al. Increase in Hepatitis A Virus Infections - United States, 2013-2018. *MMWR Morb Mortal Wkly Rep* 2019; **68**(18): 413-5.
539. Khamrin P, Maneekarn N, Okitsu S, Ushijima H. Epidemiology of human and animal kobuviruses. *Virusdisease* 2014; **25**(2): 195-200.
540. Japhet MO, Famurewa O, Adesina OA, et al. Viral gastroenteritis among children of 0-5 years in Nigeria: Characterization of the first Nigerian aichivirus, recombinant noroviruses and detection of a zoonotic astrovirus. *J Clin Virol* 2019; **111**: 4-11.
541. Karelehto E, Brouwer L, Benschop K, et al. Seroepidemiology of Parechovirus A3 Neutralizing Antibodies, Australia, the Netherlands, and United States. *Emerg Infect Dis* 2019; **25**(1): 148-52.
542. Jaaskelainen AJ, Voutilainen L, Lehmusto R, et al. Serological survey in the Finnish human population implies human-to-human transmission of Ljungan virus or antigenically related viruses. *Epidemiology and infection* 2016; **144**(6): 1278-85.
543. Jaaskelainen AJ, Nurminen N, Kolehmainen P, et al. No Association Between Ljungan Virus Seropositivity and the Beta-cell Damaging Process in the Finnish Type 1 Diabetes Prediction and Prevention Study Cohort. *The Pediatric infectious disease journal* 2019; **38**(3): 314-6.
544. Reuter G, Pankovics P, Boros A. Saliviruses-the first knowledge about a newly discovered human picornavirus. *Rev Med Virol* 2017; **27**(1).
545. Pena SA, Davis SS, Lu X, et al. Severe Respiratory Illness Associated with Human Metapneumovirus in Nursing Home, New Mexico, USA. *Emerg Infect Dis* 2019; **25**(2): 383-4.

546. Shi T, McAllister DA, O'Brien KL, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet* 2017; **390**(10098): 946-58.
547. Vieira SE, Bando SY, de Paulis M, et al. Distinct transcriptional modules in the peripheral blood mononuclear cells response to human respiratory syncytial virus or to human rhinovirus in hospitalized infants with bronchiolitis. *PLoS One* 2019; **14**(3): e0213501.
548. Kadkhoda K, Semus M, Jelic T, Walkty A. Case Report: A Case of Colorado Tick Fever Acquired in Southwestern Saskatchewan. *Am J Trop Med Hyg* 2018; **98**(3): 891-3.
549. McDonald E, George D, Rekant S, et al. Notes from the Field: Investigation of Colorado Tick Fever Virus Disease Cases - Oregon, 2018. *MMWR Morb Mortal Wkly Rep* 2019; **68**(12): 289-90.
550. Charrel RN, Attoui H, Butenko AM, et al. Tick-borne virus diseases of human interest in Europe. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2004; **10**(12): 1040-55.
551. Chastel C, Main AJ, Couatarmanac'h A, et al. Isolation of Eyach virus (Reoviridae, Colorado tick fever group) from Ixodes ricinus and I. ventalloi ticks in France. *Archives of virology* 1984; **82**(3-4): 161-71.
552. Gonzalez HA, Knudson DL. Genetic relatedness of corriparta serogroup viruses. *The Journal of general virology* 1987; **68 (Pt 3)**: 661-72.
553. Boughton CR, Hawkes RA, Naim HM. Arbovirus infection in humans in NSW: seroprevalence and pathogenicity of certain Australian bunyaviruses. *Australian and New Zealand journal of medicine* 1990; **20**(1): 51-5.
554. Dilcher M, Hasib L, Lechner M, et al. Genetic characterization of Tribec virus and Kemerovo virus, two tick-transmitted human-pathogenic Orbiviruses. *Virology* 2012; **423**(1): 68-76.
555. Attoui H, Mohd Jaafar F. Zoonotic and emerging orbivirus infections. *Revue scientifique et technique (International Office of Epizootics)* 2015; **34**(2): 353-61.
556. Saluzzo JF, Ivanoff B, Languillat G, Georges AJ. [Serological survey for arbovirus antibodies in the human and simian populations of the South-East of Gabon (author's transl)]. *Bulletin de la Societe de pathologie exotique et de ses filiales* 1982; **75**(3): 262-6.
557. Lewandowska DW, Capaul R, Prader S, et al. Persistent mammalian orthoreovirus, coxsackievirus and adenovirus co-infection in a child with a primary immunodeficiency detected by metagenomic sequencing: a case report. *BMC infectious diseases* 2018; **18**(1): 33.
558. Mikuletic T, Steyer A, Kotar T, Zorec TM, Poljak M. A novel reassortant mammalian orthoreovirus with a divergent S1 genome segment identified in a traveler with diarrhea. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2019; **73**: 378-83.
559. Yamanaka A, Iwakiri A, Yoshikawa T, et al. Imported case of acute respiratory tract infection associated with a member of species nelson bay orthoreovirus. *PLoS One* 2014; **9**(3): e92777.
560. Tate JE, Burton AH, Boschi-Pinto C, Parashar UD. Global, Regional, and National Estimates of Rotavirus Mortality in Children <5 Years of Age, 2000-2013. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2016; **62 Suppl 2**: S96-s105.
561. Fujii Y, Doan YH, Suzuki Y, Nakagomi T, Nakagomi O, Katayama K. Study of Complete Genome Sequences of Rotavirus A Epidemics and Evolution in Japan in 2012-2014. *Frontiers in microbiology* 2019; **10**: 38.
562. Aung MS, Nahar S, Aida S, et al. Distribution of two distinct rotavirus B (RVB) strains in the north-central Bangladesh and evidence for reassortment event among human RVB revealed by whole genomic analysis. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2017; **47**: 77-86.
563. Joshi MS, Walimbe AM, Dilpak SP, Cherian SS, Gopalkrishna V. Whole-genome-based characterization of three human Rotavirus C strains isolated from gastroenteritis outbreaks in Western India and a provisional intra-genotypic lineage classification system. *The Journal of general virology* 2019; **100**(7): 1055-72.
564. Marthaler D, Rossow K, Culhane M, et al. Widespread rotavirus H in commercially raised pigs, United States. *Emerg Infect Dis* 2014; **20**(7): 1195-8.

565. Jiang S, Ji S, Tang Q, et al. Molecular characterization of a novel adult diarrhoea rotavirus strain J19 isolated in China and its significance for the evolution and origin of group B rotaviruses. *The Journal of general virology* 2008; **89**(Pt 10): 2622-9.
566. Lv X, Mohd Jaafar F, Sun X, et al. Isolates of Liao ning virus from wild-caught mosquitoes in the Xinjiang province of China in 2005. *PLoS One* 2012; **7**(5): e37732.
567. Li QP, Xie XC, Zhi Q, Ma L. First Isolation of 8 Strains of New Orbivirus (BANNA) from Patients with Innominate Fever in Xinjiang. *Endemic Disease Bulletin* 1992; (01): 77-81+111.
568. Bandeira LM, Uehara SNO, Puga MAM, et al. HTLV-1 intrafamilial transmission among Japanese immigrants in Brazil. *Journal of medical virology* 2018; **90**(2): 351-7.
569. Vermeulen M, Sykes W, Coleman C, et al. The prevalence of human T-lymphotropic virus type 1 & 2 (HTLV-1/2) in South African blood donors. *Vox sanguinis* 2019; **114**(5): 451-8.
570. Braco ILJ, de Sa KSG, Waqasi M, et al. High prevalence of human T-lymphotropic virus 2 (HTLV-2) infection in villages of the Xikrin tribe (Kayapo), Brazilian Amazon region. *BMC infectious diseases* 2019; **19**(1): 459.
571. Mahieux R, Gessain A. HTLV-3/STLV-3 and HTLV-4 viruses: discovery, epidemiology, serology and molecular aspects. *Viruses* 2011; **3**(7): 1074-90.
572. Zheng H, Wolfe ND, Sintasath DM, et al. Emergence of a novel and highly divergent HTLV-3 in a primate hunter in Cameroon. *Virology* 2010; **401**(2): 137-45.
573. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS (London, England)* 2006; **20**(16): W13-23.
574. Lagi F, Baldin G, Colafigli M, et al. Viro-immunological efficacy and tolerability of dolutegravir-based regimens compared to regimens based on other integrase strand inhibitors, protease inhibitors, non-nucleoside reverse transcriptase inhibitors in patients with acute HIV-1 infection: a multicenter retrospective cohort study. *International journal of antimicrobial agents* 2019.
575. Visseaux B, Damond F, Matheron S, Descamps D, Charpentier C. Hiv-2 molecular epidemiology. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2016; **46**: 233-40.
576. Requena S, Caballero E, Lozano AB, et al. Treatment outcome in dually HIV-1 and HIV-2 coinfected patients living in Spain. *AIDS (London, England)* 2019.
577. Kalish ML, Wolfe ND, Ndongmo CB, et al. Central African hunters exposed to simian immunodeficiency virus. *Emerg Infect Dis* 2005; **11**(12): 1928-30.
578. Meiering CD, Linial ML. Historical perspective of foamy virus epidemiology and infection. *Clinical microbiology reviews* 2001; **14**(1): 165-76.
579. Lambert C, Batalie D, Montange T, et al. An Immunodominant and Conserved B-Cell Epitope in the Envelope of Simian Foamy Virus Recognized by Humans Infected with Zoonotic Strains from Apes. *Journal of virology* 2019; **93**(11).
580. Heneine W, Switzer WM, Sandstrom P, et al. Identification of a human population infected with simian foamy viruses. *Nat Med* 1998; **4**(4): 403-7.
581. Pinto-Santini DM, Stenbak CR, Linial ML. Foamy virus zoonotic infections. *Retrovirology* 2017; **14**(1): 55.
582. Engel GA, Small CT, Soliven K, et al. Zoonotic simian foamy virus in Bangladesh reflects diverse patterns of transmission and co-infection. *Emerging microbes & infections* 2013; **2**(9): e58.
583. Francis JR, McCall BJ, Hutchinson P, Powell J, Vaska VL, Nourse C. Australian bat lyssavirus: implications for public health. *Med J Aust* 2014; **201**(11): 647-9.
584. Francis JR, Nourse C, Vaska VL, et al. Australian Bat Lyssavirus in a child: the first reported case. *Pediatrics* 2014; **133**(4): e1063-7.
585. van Thiel PP, de Bie RM, Eftimov F, et al. Fatal human rabies due to Duvenhage virus from a bat in Kenya: failure of treatment with coma-induction, ketamine, and antiviral drugs. *PLoS neglected tropical diseases* 2009; **3**(7): e428.
586. van Thiel PP, van den Hoek JA, Eftimov F, et al. Fatal case of human rabies (Duvenhage virus) from a bat in Kenya: The Netherlands, December 2007. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2008; **13**(2).

587. Fooks AR, Brookes SM, Johnson N, McElhinney LM, Hutson AM. European bat lyssaviruses: an emerging zoonosis. *Epidemiology and infection* 2003; **131**(3): 1029-39.
588. Nathwani D, McIntyre PG, White K, et al. Fatal human rabies caused by European bat Lyssavirus type 2a infection in Scotland. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2003; **37**(4): 598-601.
589. Chen T, Miao FM, Liu Y, et al. Possible Transmission of Irkut Virus from Dogs to Humans. *Biomedical and environmental sciences : BES* 2018; **31**(2): 146-8.
590. Paweska JT, Blumberg LH, Liebenberg C, et al. Fatal human infection with rabies-related Duvenhage virus, South Africa. *Emerg Infect Dis* 2006; **12**(12): 1965-7.
591. Hemachudha T, Ugolini G, Wacharapluesadee S, Sungkarat W, Shuangshoti S, Laothamatas J. Human rabies: neuropathogenesis, diagnosis, and management. *The Lancet Neurology* 2013; **12**(5): 498-513.
592. Chiu C, Fair J, Leroy EM. Bas-Congo virus: another deadly virus? *Future microbiology* 2013; **8**(2): 139-41.
593. Cai Y, Yu S, Jangra RK, et al. Human, Nonhuman Primate, and Bat Cells Are Broadly Susceptible to Tibrovirus Particle Cell Entry. *Frontiers in microbiology* 2019; **10**: 856.
594. Tesh RB, Boshell J, Modi GB, et al. Natural infection of humans, animals, and phlebotomine sand flies with the Alagoas serotype of vesicular stomatitis virus in Colombia. *Am J Trop Med Hyg* 1987; **36**(3): 653-61.
595. Menghani S, Chikhale R, Raval A, Wadibhasme P, Khedekar P. Chandipura Virus: an emerging tropical pathogen. *Acta tropica* 2012; **124**(1): 1-14.
596. Damle RG, Sankararaman V, Bhide VS, Mahamuni SA, Walimbe AM, Cherian SS. Genetic characterization of the glycoprotein G of Chandipura viruses in India with emphasis on an outbreak of 2015. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 2017; **55**: 112-6.
597. Quiroz E, Moreno N, Peralta PH, Tesh RB. A human case of encephalitis associated with vesicular stomatitis virus (Indiana serotype) infection. *Am J Trop Med Hyg* 1988; **39**(3): 312-4.
598. Marriott AC. Complete genome sequences of Chandipura and Isfahan vesiculoviruses. *Archives of virology* 2005; **150**(4): 671-80.
599. Gaidamovich S, Obukhova VR, Sveshnikova NA, Cherednichenko Iu N, Kostiukov MA. [Natural foci of viruses borne by Phlebotomus papatasi in the USSR according to a serologic study of the population]. *Voprosy virusologii* 1978; (5): 556-60.
600. Tesh RB, Peralta PH, Johnson KM. Ecologic studies of vesicular stomatitis virus. I. Prevalence of infection among animals and humans living in an area of endemic VSV activity. *American journal of epidemiology* 1969; **90**(3): 255-61.
601. Kurucz N, Markey P, Draper A, et al. Investigation into High Barmah Forest Virus Disease Case Numbers Reported in the Northern Territory, Australia in 2012-2013. *Vector Borne Zoonotic Dis* 2016; **16**(2): 110-6.
602. Burt FJ, Chen W, Miner JJ, et al. Chikungunya virus: an update on the biology and pathogenesis of this emerging pathogen. *Lancet Infect Dis* 2017; **17**(4): e107-e17.
603. Lindsey NP, Staples JE, Fischer M. Eastern Equine Encephalitis Virus in the United States, 2003-2016. *Am J Trop Med Hyg* 2018; **98**(5): 1472-7.
604. Calisher CH, Murphy FA, France JK, et al. Everglades virus infection in man, 1975. *Southern medical journal* 1980; **73**(11): 1548.
605. Fukunaga Y, Kumanomido T, Kamada M. Getah virus as an equine pathogen. *The Veterinary clinics of North America Equine practice* 2000; **16**(3): 605-17.
606. Carrera JP, Forrester N, Wang E, et al. Eastern equine encephalitis in Latin America. *The New England journal of medicine* 2013; **369**(8): 732-44.
607. Lednicky JA, White SK, Mavian CN, et al. Emergence of Madariaga virus as a cause of acute febrile illness in children, Haiti, 2015-2016. *PLoS neglected tropical diseases* 2019; **13**(1): e0006972.
608. Esposito DLA, Fonseca B. Will Mayaro virus be responsible for the next outbreak of an arthropod-borne virus in Brazil? *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* 2017; **21**(5): 540-4.
609. de Souza-Lopes O, de Abreu Sacchetta L. Isolation of Mucambo virus, a member of the Venezuelan equine encephalitis virus complex in the State of Sao Paulo, Brasil. *Revista do Instituto de Medicina Tropical de Sao Paulo* 1978; **20**(2): 82-6.

610. Rezza G, Chen R, Weaver SC. O'nyong-nyong fever: a neglected mosquito-borne viral disease. *Pathog Glob Health* 2017; **111**(6): 271-5.
611. Rios-Gonzalez CM. O'nyong'nyong virus: next arbovirus in Latin America? *Journal of infection and public health* 2018; **11**(1): 139.
612. Cardozo F, Konigheim B, Albrieu-Llinas G, et al. Alphaviruses: Serological Evidence of Human Infection in Paraguay (2012-2013). *Vector Borne Zoonotic Dis* 2018; **18**(5): 266-72.
613. Aubry M, Teissier A, Huart M, et al. Ross River Virus Seroprevalence, French Polynesia, 2014-2015. *Emerg Infect Dis* 2017; **23**(10): 1751-3.
614. Atkins GJ, Sheahan BJ, Liljestrom P. The molecular pathogenesis of Semliki Forest virus: a model virus made useful? *The Journal of general virology* 1999; **80 (Pt 9)**: 2287-97.
615. Adouchief S, Smura T, Sane J, Vapalahti O, Kurkela S. Sindbis virus as a human pathogen-epidemiology, clinical picture and pathogenesis. *Rev Med Virol* 2016; **26**(4): 221-41.
616. Storm N, Weyer J, Markotter W, et al. Human cases of Sindbis fever in South Africa, 2006-2010. *Epidemiology and infection* 2014; **142**(2): 234-8.
617. Talarmin A, Trochu J, Gardon J, et al. Tonate virus infection in French Guiana: clinical aspects and seroepidemiologic study. *Am J Trop Med Hyg* 2001; **64**(5-6): 274-9.
618. Weaver SC, Bellew LA, Rico-Hesse R. Phylogenetic analysis of alphaviruses in the Venezuelan equine encephalitis complex and identification of the source of epizootic viruses. *Virology* 1992; **191**(1): 282-90.
619. Carrera JP, Bagamian KH, Travassos da Rosa AP, et al. Human and Equine Infection with Alphaviruses and Flaviviruses in Panama during 2010: A Cross-Sectional Study of Household Contacts during an Encephalitis Outbreak. *Am J Trop Med Hyg* 2018; **98**(6): 1798-804.
620. Robb LL, Hartman DA, Rice L, et al. Continued Evidence of Decline in the Enzootic Activity of Western Equine Encephalitis Virus in Colorado. *Journal of medical entomology* 2019; **56**(2): 584-8.
621. Burgueno A, Frabasile S, Diaz LA, et al. Genomic Characterization and Seroprevalence Studies on Alphaviruses in Uruguay. *Am J Trop Med Hyg* 2018; **98**(6): 1811-8.
622. Lambert N, Strebel P, Orenstein W, Icenogle J, Poland GA. Rubella. *Lancet* 2015; **385**(9984): 2297-307.
623. Crooke SN, Haralambieva IH, Grill DE, Ovsyannikova IG, Kennedy RB, Poland GA. Seroprevalence and durability of rubella virus antibodies in a highly immunized population. *Vaccine* 2019; **37**(29): 3876-82.
624. Chen HY, Shen DT, Ji DZ, et al. Prevalence and burden of hepatitis D virus infection in the global population: a systematic review and meta-analysis. *Gut* 2018.
625. Groc S, Abbate JL, Le Gal F, et al. High prevalence and diversity of hepatitis B and hepatitis delta virus in Gabon. *Journal of viral hepatitis* 2019; **26**(1): 170-82.
626. Hale T, Webster S, Petherick A, Phillips T, Kira B. Oxford COVID-19 government response tracker. 2020. https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker (accessed 16th Nov 2020).