Differential environmental attributability in global re-emergence of zoonotic and anthroponotic diseases from 1990 to 2010

by

Joseph August Lewnard

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of B.A. in Geography

Department of Geography
McGill University
Montreal, Quebec, Canada

April 2013

ABSTRACT

Introduction

Dynamic environments favor survival and persistence of pathogens with wide host breadth. It is unclear whether this explains current spatiotemporal variation in the global distribution of human infectious diseases. To assess whether environmental selection for ecological generalist- and specialist pathogens gives rise to present distribution of such diseases, I assess how environmental disruption varied with incidence of zoonotic and anthroponotic diseases from 1990 to 2010.

Methods

I use autoregressive zero-inflated negative binomial models to assess how change in physical environment and socio-economic conditions over time mitigated or facilitated the occurrence of excess cases of seventeen diseases within 207 national populations.

Results

Zoonotic transmission has greater environmental sensitivity than anthroponotic transmission. Physical environmental conditions appear prerequisite to the occurrence of autochthonous zoonoses within populations, while socio-economic change amplifies the transmission of these pathogens once they are endemic. Among anthroponotic diseases, clear generalities across multiple diseases and across multiple transmission mechanisms do not arise.

Discussion

Taken together, these results indicate that environmental change differentially mediated global re-emergence of zoonotic and anthroponotic diseases from 1990 to 2010, and specifically that global re-emergence over this period reflected evolutionary dynamics which hold true at the level of microbes' community ecology.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Lea Berrang-Ford, for her tireless support through this and through other endeavors. I would like to thank my reader, Dr. Nancy Ross, for her critical appraisal of this text, and for her instruction in that subject during GEOG 503.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: JUSTIFICATIONS, AIM, AND OBJECTIVES

1.1: Justifications

The emergence of novel pathogens from ecological sources has accompanied transitional periods throughout the natural history of *Homo sapiens*. Shifts in the behavioral, social, and environmental configurations of human activity allow new exposures to microbes, whose rapid evolution and reproduction favor opportunism in the presence of naive hosts and dynamic environments.[1,2] The increasing frequency with which novel infections emerge[3] and the rising incidence of once-controlled infections[4] indicate disruptions in the ecological balance among humans and their microbiota, occurring as global environments encounter unprecedented urbanization, landscape degradation, and spatial interconnectedness.

A small body of epidemiologic research has evaluated cross-cutting ecological drivers for global-scale variation in the distribution of human infectious diseases. Taylor et al. (2001)[5] provide the first quantitative analysis of population-scale risk factors for microbes' acquisition of human infectivity, indicating the ecological nature of the emerging disease paradigm by demonstrating that animal-borne agents are the progenitors for most pathogenic human microbes. Wolfe et al. (2007)[6] implicate climate, geography, and ecology as regulators of the processes by which enzootic pathogens evolve to persist among human hosts, concluding that characteristics of modern diseases owe to environmental circumstances surrounding their emergence over evolutionary time. Jones et al. (2008)[3] determine that local social, environmental, and ecological conditions differentially govern novel zoonotic and nonzoonotic disease evolution and introduction; Woolhouse & Gowtage-Sequeria (2005)[7] suggest that the significance of certain environmental drivers to the emergence or re-emergence of diseases depends upon whether diseases are zoonotic or nonzoonotic, but fail to account for these differences causally. Although less focused upon pathogen novelty than diversity, Dunn et al. (2010)[8] and Guernier et al. (2004)[9] demonstrate that biodiversity, climate, and human activity contribute to spatial variability in global pathogen richness and prevalence.

Universal factors underlying temporal and spatial variation in the incidence of emerging and re-emerging infections are notably absent from this literature. It is unclear from the studies listed

above whether incidence, representing the amplification and transmission of diseases within populations, shares causes with the evolution of novel human pathogens. While the identified studies indicate circumstances conducive to pathogens' initial spillover into human populations, all but one[7] neglect to consider the burden such infections exert following biological introduction. Whereas environmental determinants for the evolution of novel and diverse pathogen species are defined within these global observational studies, and are corroborated by experimental work,[10,11] we lack an empirical understanding of whether identified drivers of pathogen evolution aid in sustaining transmission. This spillover from a source population and subsequent amplification within the new population represent different processes, both mathematically and conceptually, and are governed by competing evolutionary dynamics.[6,10,12,13,14,15]

It is well known that temporally-variable environments select for generalist traits among organisms, allowing species to exploit diverse resources for survival when access to a narrow resource pool becomes unreliable.[16,17] Applying this principle to disease emergence, we may consider wide host breadth, or a microbe's ability to infect multiple species, to embody such ecological generalism.[10,18,19] Environmental change interrupting circulation within a host species disrupts a pathogen's selection for host-specialized traits as a once-stable circulation scheme erodes, causing adaptive generalization enabling the pathogen to exploit a wider array of hosts within the increasingly-unpredictable environment. Under such a scenario, specialized selection for infectivity among members of the new host species may be unfavorable as long as environmental conditions remain dynamic. However, innate maladaptions prior to the cross-species spillover event will generally diminish the pathogen's ability to persist and circulate among members of the novel host species. For this reason, variable environments favoring transmission of zoonotic diseases among humans and animals are likely to differ from more-constant environments, where anthroponotic transmission is sufficiently reliable to allow a pathogen population to persist and specialize for human infectivity.[16,17]

Often denoted R_0, the basic reproductive ratio provides a numeric estimation of whether a host species serves as a "source" or as a "sink" in a pathogen's transmission dynamics.[15] The ratio indicates, on average, how many individuals within a host population will acquire infection from

a single infected member of the same species. An outbreak among hosts with $R_0 < 1$ must necessarily be self-limited, in that continued transmission requires exposure to an ecological source species with better reservoir competence.[6] Defining zoonotic diseases as those for which $R_0 < 1$ among human cases, and anthroponotic diseases as those for which $R_0 > 1$, one readily sees that humans' roles as ecological sinks for zoonotic pathogens, or as a resource niche for anthroponotic pathogens, render zoonotic and anthroponotic circulation dissimilar. We presently lack understanding of the factors sustaining transmission of both zoonotic and anthroponotic diseases within populations, and do not know whether these differ from the better-defined environmental processes underlying spillover.[1,3,5,7,8]

In the simplest scenario (depicted in Figure 1), the process whereby diseases proceed from circulating among animals to circulating among humans can be depicted as biphasic, involving an initial stage wherein enzootic pathogens acquire the ability to infect a human host, and a second stage selecting for traits favoring zoonotic and ultimately anthroponotic transmission over enzootic circulation.[1,6,10-12,17-21] Specialization for transmission among human hosts can only commence when biological traits allowing human infection already exist, and thus must follow the initial cross-species spillover event. Space-time variation in the moments where spillover occurs,[1,3] and in the distribution of zoonotic and anthroponotic diseases,[6] compel us to consider whether consistent causes underlie these processes at a global scale.

Existing empirical work indicates that stochastic evolutionary changes widening a pathogen's host breadth accumulate deterministically in temporally-variable environments;[10-12,17-19] ecological, environmental, and human demographic circumstances coalesce to allow cross-species spillover events wherein pathogens diverge from their enzootic progenitors and successfully infect human hosts in discrete environmental foci.[3] Wolfe et al. (2007)[6] indicate that the tropics currently harbor more diseases of this generalist class than do temperate locations; this observation that tropical diseases continue to rely upon enzootic and zoonotic transmission cycles is supported by evidence that environmental circumstances at low latitudes favor host and pathogen species richness[8,9] and spillover.[3] Zoonoses are less likely to evolve within temperate regions, and those that do have more often specialized to persist among humans than have their tropical counterparts.[6]

The differential capacity for tropical and temperate regions to support pathogen evolution from enzootic to anthroponotic stages[6] and the constrained circulation of zoonotic pathogens at increasing latitudes[8,9] provide observational confirmation of evolutionary theory that specialization for anthroponic transmission is environmentally-mediated rather than stochastic.[10,11,15-21] Surveillance and prevention efforts stand to benefit from integrating consideration of determinants for pathogen origination via spillover with the unique and yet-uncharacterized determinants for sustained circulation and re-emergence. Although novel pathogens remain a threat to human health and security,[22] a pathogen's exploitation of our current interactions via anthroponotic transmission affords higher relative burden to anthroponotic diseases than to self-limited, maladapted diseases spilling over from animal sources.[23] Globally-applicable environmental causes for transmission and re-emergence of extant human pathogens thus merit specific empirical attention.

It is likely that unaddressed conflicts in the meaning of disease "emergence" have forestalled investigation of, and differentiation among, factors underlying the dissimilar processes of spillover and circulation. Whereas the term arose in reference to evolutionarily novel microbes,[24] it has since been appropriated by the epidemiological community in reference to any "infection that has newly appeared in a population, or has existed but is rapidly increasing in incidence or geographic range." The latter definition[12] makes no mention of a pathogen's innate biological novelty, but rather considers frequency of occurrence. Having appeared in the inaugural issue of the journal *Emerging Infectious Diseases*, this definition has since been widely upheld within qualitative literature employing the terms "emerging" and "re-emerging" so synonymously that differential causes for these processes are often not distinguishable, or are taken to be indistinguishable.[4,25-29] Inherently, a re-emerging disease is not a new entity and cannot fulfill the criteria of the earlier evolutionary novelty definition; factors selecting for its pre-adapted transmission scheme among a community of hosts likely differ from circumstances that would cause fundamental evolutionary change.[10,11] The empirical bias towards an evolutionary event definition for disease emergence[3,5,7] has prevented validation of qualitative literature conflating emergence with re-emergence in an epidemiological disease-transmission context. This circumstance is incompatible with the need for rigorous reviews[30] and other non-technical

literature[31] making empirical outcomes accessible to a broader audience of public health and environmental management officials. Determining where circulating enzootic pathogens are likely to infiltrate human populations is of little consequence to global health if locations at risk for spillover lack critical characteristics for sustaining either zoonotic or anthroponotic disease transmission. Thus aside from creating academic confusion, unaddressed and unresolved conflict in the definition for emergence poses risk of impeding the design of policy for infectious disease burden reduction.

The research gap regarding fundamental environmental drivers behind increasing incidence of infectious diseases has arisen from the complexity of conceptualizing and attributing causation to distal determinants, from concern for the appropriateness of conducting such analyses at a global scale, and from a historical absence of uniform, high-quality disease reporting datasets.[32,33] International and longitudinal disparities in case-surveillance capacity inevitably bias statistical models, and challenge the validation of mathematical models. As data repositories become increasingly comprehensive,[34] new surveillance platforms become available,[35] and monitoring by international administrative agencies becomes institutionalized,[36] exploratory research into drivers of global variation in incidence is becoming feasible.

1.2: Aim

Here my purpose is to characterize environmental determinants for the presence and amplification of authochthonous circulation of zoonotic and anthroponotic diseases. Considering extant human pathogens that have evolved from enzootic progenitors, I assess whether environmental factors differentially drive human incidence at two stages along the spectrum of evolutionary fitness: zoonotic infections (with basic reproductive ratios less than one) are classed as maladapted on the basis of self-limited transmission among humans, while anthroponotic infections (with basic reproductive ratios greater than one) are classed as fit on the basis of human reservoir competence.

1.3: Objectives

I seek to quantitatively compare the environmental determinants of national, annual incidence between selected diseases of global importance by:

1. Identifying the environmental drivers of national, annual incidence for the individual diseases;
2. Comparing trends in environmental determination of incidence between zoonotic and anthroponotic pathogens; and
3. Evaluating the validity and potential of using global datasets to quantify variation in disease incidence.

I appraise whether the determinants selected within my models resemble those that have been identified to drive evolution of novel human pathogens. I assess how the spatial distribution of environmental risk for zoonotic and anthroponotic disease transmission corresponds to the distribution of risk for novel pathogen emergence.

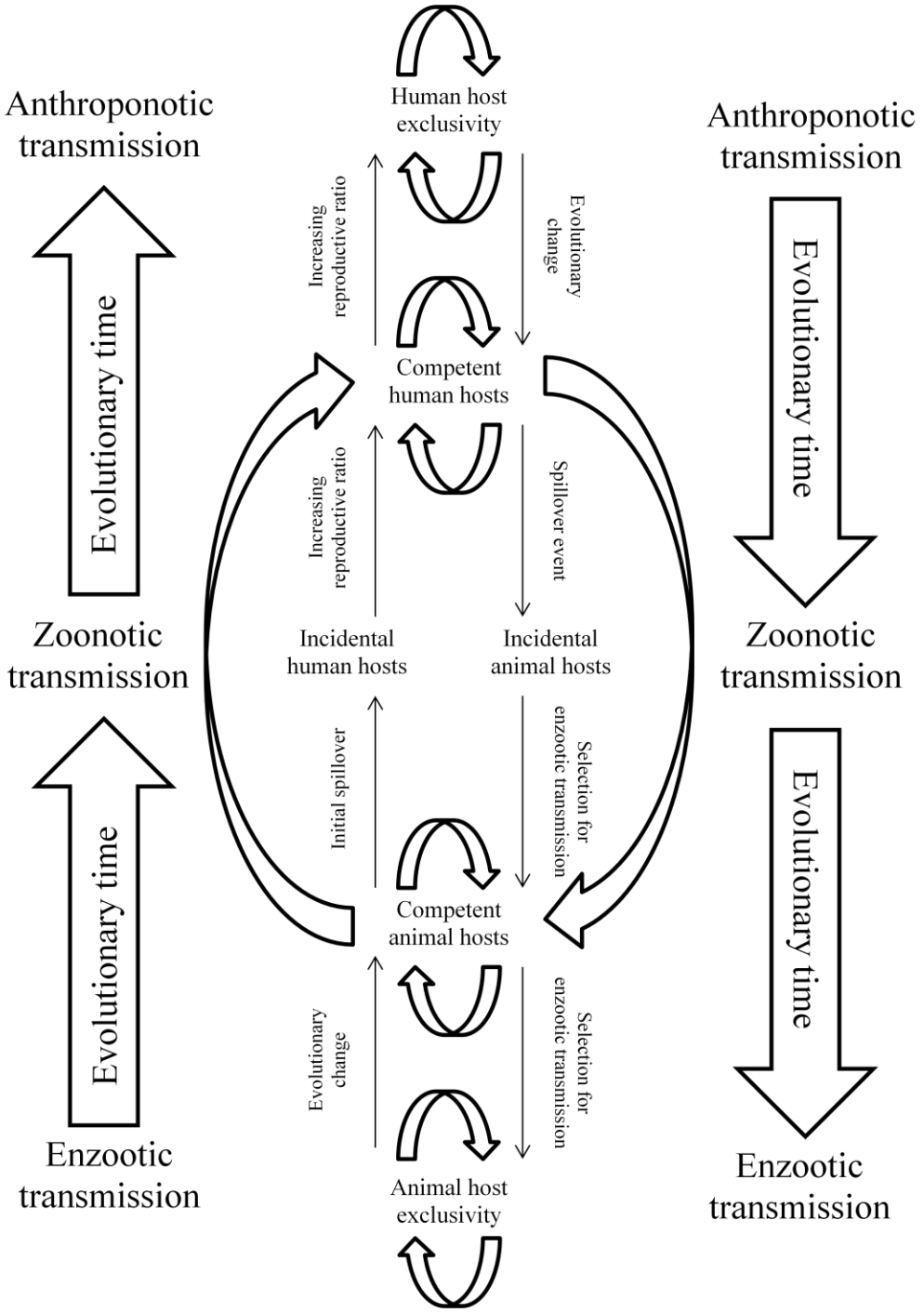

Figure 1.1: CONCEPTUAL OUTLINE OF THE BIPHASIC EMERGENCE PROCESS: Evolutionary events allowing sporadic cross-species spillover set the stage for a second evolutionary phase selecting for reproductive efficiency within the new host species. My work considers the upward trajectory (movement from enzootic to anthroponotic transmission, left-hand side in figure), but we note that temporally-variable environments and selective pressures for generalism likewise allow anthroponotic diseases to spill over into animal populations (right-hand side in figure). The biphasic scheme is advocated by Morse (1996)[12] and Wolfe (2007)[6] and validated within experimental work.[10,11,14-19]

CHAPTER 2: LITERATURE REVIEW

2.1: Quantitative comparative studies

A search of the Web of Science and Medline databases for articles with ((emerg* OR re-emerg*) AND (infecti* OR pathogen*)) identifies 5,562 peer-refereed articles. Nonetheless, only three original analyses apply uniform quantitative methods to evaluate environmental determinants of these processes at a global scale and across multiple diseases. I review these studies in detail with respect to their methodologies and their outcomes, identifying a lack of empirical attention to time-space variation in global disease incidence and its putative environmental causes. Nonetheless, many qualitative studies and reviews focus upon such variation in characterizing the environmental etiology of disease emergence and re-emergence.

2.1.1: Spatial risk models: Jones et al. (2008), Guernier et al. (2004), Dunn et al. (2010)

Jones et al. (2008)[3] provide a foundational comparative study of global determinants for the emergence of novel pathogens and re-introduction of pathogens already known to exist. The authors primarily model the number of emerging infectious disease (EID) events, defined as the first time a pathogen appears among human hosts, within spatially-delineated areas against time-invariant covariate data aggregated to the same spatial grid. The authors control for reporting bias by including a covariate indicating a country's representation among *Journal of Infectious Diseases* contributors from 1973 to 2007. The authors use a narrow selection of environmental covariates, including two human factors (population density in 2000 and change in population density from 1990-2000), as well as latitude, mean annual rainfall, and wildlife host richness. Through iterative logistic regressions upon random draws from EID events over the grid, the authors demonstrate that EID events are most likely to be of zoonotic origin, and that determinants for emergence vary categorically within the authors' taxonomy of emerging infectious disease classes. For instance, human population growth drives the emergence of drug-resistant pathogens more significantly than other pathogens; higher latitude and mean annual rainfall drive the emergence of non-wildlife zoonoses; time-invariant mammalian species richness drives the emergence of wildlife zoonoses; and no environmental covariates except human population density in 2000 appear significant for vector-borne diseases.

A number of factors in the emergence problem remain unaddressed by Jones et al. (2008). Although employing disease emergence events over an extensive time period (1940-2004), the study relies upon time-invariant covariates which diminish the utility of statistical outcomes for causal inference. For instance, mean annual rainfall cannot link time-discrete emergence events to extreme weather accounting for unusual human-animal contact that would allow spillover at a given point in time. Likewise, the timescale of the population data is mismatched to the timescale of the disease emergence events; conceptually, population change from 1990-2000 cannot be taken as a cause for emergence events in 1945, as growth trajectories within the developed and developing worlds shifted substantially during the twentieth century. While innovative, the time-invariant control for reporting bias shares this problem; it extends only through half of the period of analysis, and fails to capture year-by-year variation in reporting effort that may occur amidst development, or amidst collapses within a country's public health infrastructure (for instance during periods of war or repression). Since such institutional collapses may be important to allowing diseases to emerge, the robustness of the findings to reporting bias remains incomplete. Although the EID events encompass introduction of disease to a population, they fail to represent more gradual increases in circulation of diseases over time within populations. Increases in transmission extending beyond what would be expected given R_0 and rates of population change indeed represent a compelling emergence phenomenon lacking an effective, temporally-discrete analog. Because the study does not model changes in the frequency with which diseases occur, it remains uncertain whether the identified factors predicting the presence of a novel or introduced disease likewise drive transmission within populations. Lacking a frequency measure, the findings are not directly translatable to audiences seeking to reduce disease burden within populations; without causal utility and translatability to burden reduction, the study leaves space for others considering time-variant covariates as well as numeric measures for disease frequency.

Guernier et al. (2004)[9] evaluate environmental determinants of pathogen species richness across the globe; pathogen richness is a measure of the diversity of pathogens occurring within a spatial unit, and is hypothesized to be important to disease emergence in that environmental parameters promoting pathogen richness must broadly promote pathogen persistence or evolution. Dunn et al. (2010)[8] additionally evaluate pathogen richness, and incorporate a separate model for

pathogen prevalence. Both studies identify spatial correlations in species richness of pathogens and of other organisms across the planet, with the highest diversity occurring in the proximity of the equator. Guernier et al. (2004) implicate climate in driving the latitudinal gradient of pathogen richness, and suggest that warm conditions within wet-dry seasonal cycles are requisite to biotic interactions favoring the transmission of pathogens among multiple host species. Dunn et al. (2010) indicate that per-capita health expenditure is associated with low pathogen prevalence after controlling for the positive effect of pathogen richness. The authors suggest that increased spending in equatorial regions, where pathogen richness and prevalence are both high, may help contain disease burden. Both studies employ a backward stepwise procedure to select environmental covariates.

Guernier et al. (2004) and Dunn et al. (2010) identify significant partial effects from climatic and biodiversity variables on pathogen richness and pathogen prevalence; parameters of biological productivity thus appear important for determining circulation of human pathogens. Although the studies incorporate a wider array of covariates than do Jones et al. (2008), all share inferential limitations. None incorporate temporal variation, and thus none account for how environmental change over time causes the observed outcomes. Although pathogen richness indicates potential risk for emergence of novel microbes, pathogen richness is itself only tangentially related to indicators for disease emergence by the Morse definition.[12] While prevalence relates more directly to frequency within populations, Dunn et al. (2010) use a weak prevalence measure ranking countries by endemicity for a disease on an ordinal scale with only three units: endemic, sporadic, and non-endemic. The scale fails to account for the wide numerical variation in endemicity observed across countries and over time, and in the absence of temporal variation the authors are unable to account for emergence in the sense of increasing frequency over time (for instance, from sporadic to endemic amidst a critical environmental change). Although Guernier et al. (2004) and Dunn et al. (2010) characterize the global map as it stands today, the effects of temporal variability in the studies' environmental exposures remain uncertain.

2.1.2: Systematic review: Woolhouse & Gowtage-Sequeria (2005)

Woolhouse & Gowtage-Sequeria (2005)[7] update a prior literature survey[37] to identify 1,407 recognized human pathogen species and classify 177 of them as emerging or re-emerging.

Unlike the studies described above focusing upon presence or absence of pathogens within populations, Woolhouse & Gowtage-Sequeria (2005) distinguish factors believed to drive increased pathogen incidence within populations following initial evolutionary introduction. Through their literature review, the authors identify primary drivers believed to increase the incidence of individual infections and categorize the drivers as belonging to one or more of ten categories. They tabulate the number of infections whose emergence or re-emergence is thus attributed to each category of drivers, and rank the drivers as follows according to their relative importance to more or fewer pathogen species:

1. Changes in land use or agricultural practices
2. Changes in human demographics and society
3. Poor population health (e.g. HIV, malnutrition)
4. Hospitals and medical procedures
5. Pathogen evolution (e.g. antimicrobial drug resistance, increased virulence)
6. Contamination of food sources or water supplies
7. International travel
8. Failure of public health programs
9. International trade
10. Climate change

The authors fail to identify significant differences in drivers of emergence across assigned pathogen categories they assess; these include groupings by bacterial, helminth, protozoan, viral, or fungal nature; groupings by zoonotic or non-zoonotic transmission; and (among zoonoses) groupings by transmissibility, defining a natural break at R_0 greater than or less than one for human-to-human transmission.

Although the literature review is an informative summary tool, the over-representation of zoonotic diseases and subjectivity of outcomes to prevailing thematic interests within the literature compromise robustness of the categorical ranking of drivers. Changes in land use and agricultural practices likely appear high within the list because a large proportion of emerging pathogens are acquired through zoonotic transmission, and because unique pathogens merit unique publications;[8] because zoonoses generally have lower incidence than diseases transmitted among people,[6] it is misleading from a burden perspective to rank changing land use and

agricultural practice so high within a list of this nature, which appears ostensibly to set priorities for environmental intervention. Additionally, each included study does not evaluate the isolated effect sizes of all ten classes of drivers; high- and low-rankings for categories within this list may be a better indicator of published research topics than of the true etiology of disease emergence. Changes in land use or in agricultural practices may figure more prominently into research on emerging zoonoses than do other determinants, for instance stochastic pathogen evolution and covariance with socioeconomic inequality; without assessing such representation biases and without limiting their review to studies evaluating the partial effects of each determinant on disease incidence, the authors come short of demonstrating that the ranked scheme of environmental drivers accounts for anything more than publication bias. While Gowtage & Sequeria (2005) offer the only systematic evaluation of incidence determinants across multiple emerging and re-emerging infectious diseases, a comparative modeling approach quantifying the effects of multiple drivers across multiple diseases remains necessary.

2.2: Other global comparative studies

Numerous studies review evidence for environmental determinants of emergence qualitatively and without *a priori* methodological frameworks. While such work plays an important role summarizing outcomes from the large body of disease-specific literatures, there is a risk for qualitative reviews to generalize across more diverse diseases and at broader space-time scales than the studies they cite. The need for validation is re-enforced by the tendency for qualitative surveys to equate determinants for pathogen evolution, introduction, and amplification among human hosts.[4,25-29] Because global amplification processes remain unaddressed within the quantitative literature[3,7-9] and are likely to differ from better-defined processes driving pathogen evolution, a survey of global environmental determinants for the incidence of extant infectious diseases is overdue. Here I evaluate high-impact qualitative reviews regarding environmental determinants of emerging diseases, identifying findings requiring validation by a comprehensive survey associating global time-space variation between environmental conditions and the incidence of multiple diseases.

In his article for the inaugural issue of the *Journal of Emerging Infectious Diseases* Morse (1995)[12] implicates environment in the etiology of diseases newly appearing within populations,

and diseases whose incidence within populations had been observed to rise around the time of publication. Defining "emerging infections" as diseases fulfilling either of these two epidemiological criteria, Morse conceptualizes a biphasic model of disease emergence as a series of: 1) environmental events causing a novel pathogen to spill over into a human population; and 2) environmental events amplifying that pathogen's transmission among susceptible hosts. While the article has had a wide impact on the accepted definition of emerging diseases and has motivated work toward understanding global drivers for the appearance of diseases among new host populations,[3] its premise that determinants of pathogen evolution and amplification differ remains unverified in the absence of global comparative surveys for drivers of incidence.

McMichael (2004)[1] and Weiss & McMichael (2004)[38] conceptualize major historical transitions in human-environment interactions as causes for novel pathogens to appear among human hosts, and argue that the subsequent spread of new diseases is attributable to a combination of factors pertaining to 1) demographics, 2) land use, 3) human consumption, 4) human physiological fitness, and 5) human risk behaviors. The authors suggest that observed increases in the frequency with which new infections appear owe to modern globalization's status as the fourth major transition in human history, following the establishment of sedentary agricultural, the initiation of trans-continental trade among Eurasian civilizations, and the conquest of the New World by European empires. McMichael (2004) reasons that the five putative drivers of pathogen amplification have played different roles in shaping the dissemination of diseases during each period, but indicates that all are important within the present period and supports his argument with anecdotal examples of individual determinants relevant to individual diseases. Weiss & McMichael (2004) argue that these same drivers of incidence represent changes in human ecology, and may thus be factors in the continued evolution of novel pathogens during the current "global" transition. The suitability of the biphasic model for emergence, the relative weights of the identified drivers in the expansion of multiple diseases, and the sharing of parameters for novel pathogen evolution and pathogen amplification remain issues requiring empirical validation.

Two articles by Morens, Folkers, and Fauci (2004; 2008)[4,29] consider a selection of novel and re-emerging infections and evaluate the present and historical significance of environmental factors

in disease emergence. Both articles argue that modern pandemic disease spread is attributable to the same phenomena that drove the amplification of historical plagues. The authors indicate that the increasingly rapid pace of environmental change and expansion of humans' spatial networks put current and future generations at risk for complex disease emergence scenarios; surveillance and control are taken to be more difficult when pathogens have the opportunity to disseminate across the globe before being detected. Likewise, the linkage of high-risk demographic and ecological processes to entrenched regimes of global market capitalism make interventions into environmental factors for emergence difficult to execute. The focus within these articles upon drivers of incidence for individual emerging and re-emerging diseases calls attention to the current lack of quantitative studies assessing how ubiquitously-relevant such environmental risk factors are to a suite of diseases.

Wilcox & Colwell (2005)[25] apply a biocomplexity perspective to the emerging disease paradigm, indicating that the expanding incidence of infections is attributable to interconnected human and ecological processes, which together create environments favorable for pathogen transmission. As in the other reviews, the authors illustrate their argument through qualitative case studies of individual diseases. The biocomplexity focus attracts attention to the narrow suite of environmental drivers currently assessed within the quantitative literature; notably, anthropogenic drivers other than population density, growth rates, gross national product, and per-capita health expenditure remain absent from existing global risk models.[3,8,9] Given the connection of demographic, social, and economic trends to alterations in human-environment interactions under the biocomplexity framework,[25,39] there is a need for greater attention to specific anthropogenic drivers within empirical studies. Such an approach may identify whether ongoing demographic, social, and economic trends are globally relevant to the emergence process, or apply only to case studies in the emergence of specific diseases within particular locations.

Daszak et al. (2000)[22] conceptualize three categories of emerging infectious diseases: those involving zoonotic transmission from a wildlife host; those involving zoonotic transmission from a domestic animal host; and those transmitted among human hosts. The authors recognize that most diseases do not fit neatly into one of these definitions, and that host range is critical to

many pathogens' ability to maintain circulation within an environment. The authors propose that different environmental factors drive incidence of diseases humans share with wildlife, with domestic animals, and with other humans. Reviewing factors in the emergence of a representative sample of wildlife zoonoses, the authors cite urbanization, agricultural expansion, climate change, and other anthropogenic disturbances of ecological systems as unique determinants for incidence of wildlife zoonoses among humans. Other qualitative surveys, however, implicate the same factors as either ubiquitous determinants for infectious disease emergence[1,12,40] or as determinants for diseases that do not circulate between humans and wildlife.[4,7,25,29,40] The opportunity for such contradictions to arise within reviews drawing upon anecdotal evidence from individual diseases justifies a comparative study empirically assessing how environmental drivers impact global incidence of multiple diseases.

2.3: Integrating findings from the literature into an evolutionary framework

As discussed in the justification, I conceptualize a biphasic emergence model involving distinct environmentally-mediated processes for enzootic pathogens' initial spillover and subsequent amplification as zoonotic and anthroponotic disease within human populations. Differences between the two processes remain poorly understood, and my literature review (see Table 1) indicates that various authors have evaluated determinants for the two processes both separately and indiscriminately.

Wolfe et al. (2007)[6] implicate environmental factors in zoonotic pathogens' evolutionary trajectories, defining a 5-stage evolutionary scheme whereby zoonoses' host range converges toward human exclusivity (See Table 2). A similar process, though less neatly compartmentalized, is characterized within other sources.[10-12,15-21] Under this framework, shorter evolutionary distances connect low-stage zoonotic diseases to their progenitors than those connecting mature-stage anthroponotic diseases to their progenitors. Conceptualizing the evolutionary process selecting for human infectivity in terms of basic reproductive ratios R_0, the transition from animal-triggered outbreaks (Stages 2-4) to human host exclusivity (Stage 5) requires the mean effective R_0 observed across human cases to exceed one. If the mean $R_0<1$, then on average, cases will eventually fail to be replaced following subjects' death or recovery.

Thus exposure to a competent animal host will be required to sustain or restore human incidence for zoonotic diseases.[37,41-43]

Wolfe et al. (2007)[6] identify a latitudinal gradient in selection for human reservoir competence, presumably indicating that conditions within temperate climates and their inhabitant human societies have favored the evolution of mature-stage crowd diseases, which are directly transmissible among human hosts with $R_0 > 1$. In contrast, tropical areas are more likely to support lower-stage infections requiring persisting enzootic and zoonotic transmission cycles, and conferring chronic illness upon human hosts. Thus $0 < R_0 < 1$ persists within tropical environments. The finding is corroborated by observations that greater pathogen species richness occurs in the tropics;[8,9] since original maladaptions generally impede selection for anthroponotic transmission,[44] one would expect fewer diseases to circulate in relatively time-invariant environments favoring anthroponotic transmission.

$R_0 = 1$ is a critical break within a pathogen's evolutionary history. For simplicity, I use this value to differentiate human diseases as zoonotic or anthroponotic and to characterize processes underlying the biphasic emergence scheme involving pathogens' introduction and amplification within populations. Diseases at presently at Stages 2-4 effectively retain $0 < R_0 < 1$ according to the Wolfe (2007)[6] characterization of self-limited outbreaks or requisite animal exposure. I evaluate differences among diseases with $0 < R_0 < 1$ and with $R_0 > 1$ to evaluate whether zoonotic emergence processes involving cross-species spillover differ from emergence processes for diseases with no required animal host.

2.4: Determinants of novel pathogen emergence and extant pathogen introduction and amplification

Here I tabulate determinants across the reviewed studies for the appearance of novel pathogens and for the incidence of existing infections. From this list and from the available data sources, I generate a list of environmental factors to assess as determinants for diseases that retain zoonotic and enzootic transmission, and for diseases which have progressed to circulate exclusively among human hosts.

Table 1: Determinants identified

	Study	Determinants
Determinants for emergence events	Jones et al. (2008)	Human population density(2000) and change in density (1990-2000) Latitude Mean annual rainfall Wildlife host richness
	Guernier et al. (2004)	*For pathogen species richness (emergence event proxy)* Latitude Mean annual temperature and monthly temperature range Mean annual precipitation and maximum range of precipitation
	Dunn et al. (2010)	*For pathogen species richness within countries (emergence event proxy)* Wildlife host richness
	McMichael (2004)	Transitional events in natural human history: 1. Onset of sedentary agrarianism (c.10,000-5,000BP) 2. Trans-continental trade in Eurasia (c.3000-1500BP) 3. Rise of overseas empires (c.1500-1900AD) 4. Globalization (present)
	Daszak et al. (2000)	Habitat encroachment Ecological manipulation Technology and industry
Determinants for disease incidence	Dunn et al. (2010)	*For pathogen prevalence* Mean maximum annual temperature Net primary productivity (actual evapotranspiration) Pathogen richness
	Woolhouse & Gowtage-Sequeria (2005)	Changes in land use or agricultural practices > Changes in human demographics and society > Poor population health (e.g. HIV, malnutrition) > Hospitals and medical procedures > Pathogen evolution (e.g. antimicrobial drug resistance, increased virulence) > Contamination of food sources or water supplies > International travel > Failure of public health programs > International trade > Climate change
	Morse (1995)	Ecological changes Technology and industry Breakdown in public health measures
	Daszak et al. (2000)	Agricultural intensification Global travel Urbanization Biomedical manipulation
Shared determinants	Weiss & McMichael (2004); McMichael (2004)	Shifting human ecology, as indicated by: Travel, trade, migration War, conquest Environmental change Technology Human-animal relationships Demographic/social conditions
	Morse (1995)	Human demographics, behavior International travel and commerce Microbial adaptation and change
	Dunn et al. (2010)	Total human population Health spending per capita
	Morens et al. (2008)	International trade and commerce Human demographics and behavior Human susceptibility to infection Poverty and social inequality War and famine Breakdown of public health measures

		Technology and industry Changing ecosystems Climate and weather Intent to harm Lack of political will Microbial adaptation and change Economic development and land use
	Wilcox & Colwell (2005)	Urbanization Agricultural intensification Habitat alteration Public health infrastructure divestment Climate (constant and shifts)

Table 2: Integrating Wolfe et al. (2007) with the reproductive ratio

Stage	Characteristics	Transmission to humans	R_0
1	Agent only in animals	None	0
2	Primary infection	Only from animals	$0 < R_0 < 1$
3	Limited outbreak	From animals or (limited to a few cycles) humans	$0 < R_0 < 1$
4	Long outbreak	From animals or (repeating across many cycles) humans	$0 < R_0 < 1$
5	Exclusive human agent	Only from humans	$R_0 > 1$

CHAPTER 3: METHODS

3.1 Overview

I identify country-year aggregates for variables appearing consistently in Table 1, and construct a dataset including national-annual aggregates for these factors as well as incidence of seventeen infectious diseases, selected for their extensive geographic range and high reporting effort. I multiply impute missing data via iterative chained equations and use a congenial model to estimate the effects of covariates on excess annual disease incidence within countries. I select covariates via backward stepwise regression, and identify trends arising among zoonotic and anthroponotic diseases across the seventeen individual-disease models.

3.2 Materials

An extensive description of the dependent and independent variables follows the thesis in the Appendix. I include details regarding my choice of particular measurements and /or proxies, and individual accounts of the mechanisms by which each variable has been documented to facilitate transmission. I additionally summarize the nature and present distribution of each disease, and explain where on the Wolfe (2007)[6] evolutionary framework each disease currently stands to justify my designation as zoonotic or anthroponotic. Table 3 itemizes the independent variables included in the analysis, indicating their sources and the applicable measurement units. Table 4 itemizes the diseases I investigate, indicating their transmission mechanisms, status under the Wolfe (2007)[6] evolutionary framework, inferred fitness as zoonotic or anthroponotic, and (if zoonotic) other host species of importance.

Table 3: Independent variables

	Variable	Description	Unit	Source	Obtained from
Mobility and connectivity	Air passengers	Passengers carried on domestic and international air carriers registered in the country	$\frac{Passengers\ carried}{Population}$	International Civil Aviation Organization	World Bank Data Catalog (WBDC)[45]
	Fraction of paved roads	% of total roads, as measured by length, that are paved with crushed stone and hydrocarbon binder or bituminized agents, with concrete, or with cobblestones	$\frac{Paved\ road\ network}{Total\ road\ network}$	International Road Federation	WBDC[45]
	Incoming refugees	People recognized as refugees in accordance with UNHCR statute, people granted refugee-like humanitarian status, and people provided temporary protection in the destination country	$\frac{Incoming\ refugees}{Population}$	United Nations High Commissioner for Refugees (UNHCR)	WBDC[45]
Population demographics	Urban population	People living in urban areas as defined by national statistical offices	$\frac{Urban\ population}{Total\ population}$	World Bank estimates (based on UN World Urbanization Prospects)	WBDC[45]
	Rate of urbanization	Annual % urban population growth, in reference to people living in urban areas as defined by national statistical offices	$\frac{\%_t - \%_{t-1}}{\%_{t-1}}$	World Bank estimates (based on UN World Urbanization Prospects)	WBDC[45]
	Gender inequality index	Scale of 0 (indicating men and women fare equally) to 1 (indicating that women fare as poorly as possible) with respect to reproductive health, empowerment, and the labor market	0 to 1 scale	United Nations Development Program	Human Development Reports[46]
	Education index	Scale of 0 (indicating poor achievement) to 1 (indicating high achievement) compiling adult literacy rate (2/3 weighting) and combined primary, secondary, and tertiary gross enrollment ratio (1/3 weighting)	0 to 1 scale	United Nations Development Program	WBDC[45]
	Standardized GINI coefficient	Comparable Gini indices of income inequality, developed with a	0 to 100 scale	Standardized World Income Inequality	SWIID[47]

		custom missing-data algorithm		Database Version 3.1, December 2011.	
	Adult HIV prevalence	Percentage of people ages 15-49 who are infected with HIV	$\frac{HIV\ positive\ 15-49}{Total\ population\ 15-49}$	UNAIDS	WBDC[45]
	Rate of population growth	Annual % (exponential rate of growth of midyear population from year t-1 to t, expressed as a percentage)	$\left[\frac{Pop_t}{Pop_{t-1}} e^t - 1\right] * 100\%$	Derived from total population; UN Statistical Division	WBDC[45]
	GDP	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 products in constant 2000 US dollars	$\frac{GDP_{2000\ USD}}{Population}$	World Bank national accounts data	WBDC[45]
	Political terror	Mean score from two sources on a scale of 1 (secure rule of law) to 5 (where terror has expanded to affect the whole population)	PTS = [1, 1.5, 2, ... , 5]	Compiled from Amnesty International and US State Department	Political Terror Scale[48]
	Population of concern	People recognized as refugees in accordance with UNHCR statute, people granted refugee-like humanitarian status, and people provided temporary protection in the country of origin	$\frac{Total\ population\ of\ concern}{Population}$	UNHCR	WBDC[45]
Infrastructure	Improved water source	Population for which 20 liters per person per day of improved water are available within 1 kilometer of the dwelling	$\frac{Population\ with\ access}{Total\ population}$	World Health Organization and UN Children's Fund Joint Measurement Program (JMP)	WBDC[45]
	Surface water source	Population reliant upon surface water sources	$\frac{Population\ reliant}{Total\ population}$	JMP	WBDC[45]
	Improved sanitation access	Population with access to excreta disposal facilities that effectively prevent human, animal, and insect contact with excreta	$\frac{Population\ with\ access}{Total\ population}$	JMP	WBDC[45]
	Open defecation reliance	Population defecating openly	$\frac{Population\ reliant}{Total\ population}$	JMP	WBDC[45]
	Agricultural mechanization	Extent of tractor usage, as measured by density	$\frac{Total\ tractors * 100}{Hectares\ agricultural\ land}$	Food and Agriculture Organization of the	WBDC[45]

				United Nations (FAO)	
Geography and environment	Forested area	% of land area under natural or planted stands of trees at least 5 meters in situ, whether productive or not; excludes tree stands in agricultural production systems	$\frac{Forested\ area}{Total\ area}$	FAO	WBDC[45]
	Agricultural area	% of land area that is arable, under permanent crops, and under permanent pastures	$\frac{Agricultural\ area}{Total\ area}$	FAO	WBDC[45]
	Island country	Country is an island (1) or not an island (0)	Island = [0, 1]	Central Intelligence Agency (CIA) World Factbook	CIA[49]
	Landlocked country	Country is landlocked (1) or not landlocked (0)	Landlocked = [0, 1]	CIA World Factbook	CIA[49]
Weather and climate	Precipitation	Mean cm of rainfall during typical month of maximum rainfall for that country across *N* reporting stations	$\frac{1}{N}\sum_{n=i}^{N} Rain$	Derived from FAOCLIM-2 database	FAOCLIM-2[50]
	Precipitation deviation	Z-score of above, relative to values observed from 1980 to 2010	$\frac{\frac{1}{N_t}\sum_{n=i}^{N} Rain_t - \sum_{i=1980}^{T=2010}\frac{1}{N}\sum_{n=i}^{N} Rain}{SD\ Rain_{1980-2010}}$	Derived from FAOCLIM-2 database	derived
	Temperature	Mean daily temperature on 1 July (Northern hemisphere) or 1 January (Southern hemisphere) across reporting *n* stations	$\frac{1}{N}\sum_{n=i}^{N} Temperature$	Derived from NNDC-GSOD database	NNDC-GSOD[51]
	Temperature deviation	Z-score of above, relative to values observed from 1980 to 2010	$\frac{\frac{1}{n_t}\sum_{i}^{N} Temp_t - \sum_{i=1980}^{T=2010}\frac{1}{n}\sum_{i}^{N} Temp}{SD\ Temp_{1980-2010}}$	Derived from FAOCLIM-2 database	derived
Controls	Total population	Total population	Population	United Nations Population Division, World Population Prospects	WBDC[45]
	Tuberculosis case notification rate	Proportion of cases notified	$\frac{\frac{Cases\ notified}{Population/100{,}000}}{\frac{Cases\ estimated}{Population/100{,}000}}$	World Health Organization (WHO)	WHO[52]
	Vaccination coverage (measles, pertussis, tuberculosis only)	Population covered by the MCV, DTP3, or BCG vaccines	$\frac{Vaccinated\ population}{Population}$	WHO	WHO[53]

Table 4: Diseases (dependent variables)

Disease	Dominant transmission mechanism	Wolfe et al. (2007) status	Anthroponotic or Zoonotic	Host species
Cholera	Feco-oral	4	Zoonotic	Crustaceans, amoeba, zooplankton, birds, small herbivores
Dengue fever	Vector-borne (Aedes aegyptae)	4	Zoonotic	Nonhuman primates
Hantavirus syndromes	Tactile or respiratory exposure to host bodily fluids	2(3)	Zoonotic	Rodents (vary by Hantavirus species)
Lyme borreliosis	Vector-borne (Ticks)	2	Zoonotic	Mammals, birds, lizards (vary by Borrelia species)
Leishmaniasis	Vector-borne (Phlebotomine sandflies)	2(5)	Zoonotic	Jackals, foxes, other sylvatic mammals, canines
Leptospirosis	Tactile or respiratory exposure to host bodily fluids	2	Zoonotic	Nearly all mammals
Malaria	Vector-borne (Anopheles mosquitoes)	4(5)	Zoonotic	Nonhuman primates (vary by Plasmodium species)
Yersiniosis (Plague)	Vector-borne (fleas) and tactile or respiratory to host bodily fluids	3	Zoonotic	Rodents
Rabies	Tactile or respiratory exposure to host bodily fluids, most often through bites	2/3	Zoonotic	Nearly all mammals
Non-typhoidal Salmonellosis	Feco-oral	3/4	Zoonotic	Many animals, notably wild and domestic birds and mammals
Gonorrhea	Sexual	5	Anthroponotic	
Measles	Respiratory	5	Anthroponotic	
Pertussis	Respiratory	5	Anthroponotic	
Shigellosis	Feco-oral	(4)5	Anthroponotic	Nonhuman primates (zoonotic transmission is rare)
Syphilis	Sexual	5	Anthroponotic	
Tuberculosis	Respiratory	5	Anthroponotic	
Typhoid fever	Feco-oral	5	Anthroponotic	

3.3: Methods, detail

3.3.1: Data compilation

I compiled a time series array across 207 countries listing the variables indicated in Tables 3 and 4 on a country-year level from 1990 to 2010. With the exception of weather variables, all values were presented as country-year aggregates by the original sources.

To generate country-year weather values, I georeferenced weather stations by a join function between their World Meteorological Organization Station Identification Numbers and their listed geographical locations, as supplied in separate documentation.[54] I designate the typical month of greatest rainfall for each country as the month appearing most frequently within the top 10% of readings reported within that country between 1980 and 2010 using the 30,000-station FAOCLIM-2[50] database. I calculate the mean monthly rainfall across all reporting stations annually during the typical month of greatest precipitation. For temperature, I retrieve mean daily readings on 1 July or 1 January for countries with latitude midpoints in the Northern and Southern hemispheres, respectively, from the NCDC-GSOD database.[51] I calculate a mean national temperature across reporting stations on the representative "warm" day. I use the averages to compute normal scores for rainfall and for temperature within each country-year to measure a given year's departure from 31-year (1980-2010) historical climate normals derived from the same databases.

This method of aggregating weather variables at the country-year level seeks to represent the mean weather conditions experienced among the population of a country, rather than the mean conditions observed across the country's unevenly-populated land surface. The weather stations represented are located within human settlements, comprising a geographic network roughly coincident with population density within countries[55] as demonstrated in Figure 3.1. The approach is flexible in cases where the distribution of human settlement across a country changes over time so as to lead to the opening and closing of weather stations; for instance, northward population migration within a country will over time favorably bias the weather mean toward representing conditions in the country's North. I thus expect mean observed value across

reporting stations at any given time to be population-weighted according to the geographic distribution of a country's inhabitants.

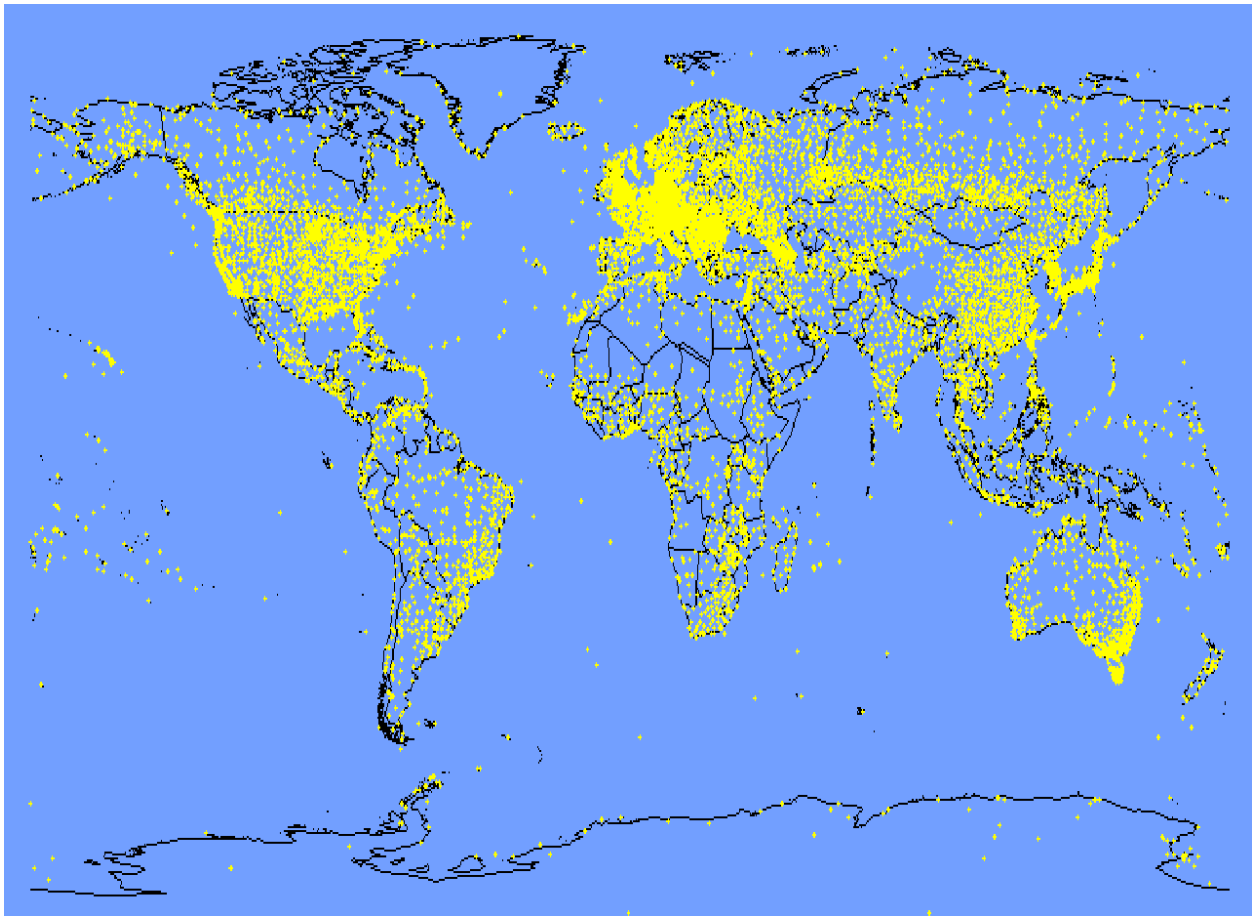

Figure 3.1: CURRENT DISTRIBUTION OF GLOBAL WEATHER STATIONS: This graphic, illustrating the stations comprising the current NCDC-GSOD network, indicates spatial coincidence with within-country settlement density. Thus although the per-capita representation of each weather station varies by country, equal weighting of observations within countries may be taken to represent mean "experienced" weather within the populace for each country.

To avoid analyzing disease cases that preceded their environmental triggers during a given year, I apply a one-year lag to all covariates such that I analyze incidence *y* within country *c* in year *t* against determinants observed within the same country no later than year $t - 1$. Because the variables under analysis are assumed to operate at different temporal scales in determining disease incidence, my database additionally includes three-year lagged covariates ($t - 3$) and 5-

year moving averages maximizing the signal-to-noise ratio during a five-year period from $t - 5$ to $t - 1$.

3.3.2: Multiple imputation

The data exhibit a volume of missingness ranging between <1% and 29% for given variables. Taken across the full array of variables, this volume reduces the number of complete cases so extensively as to make complete case analysis unfeasible. I additionally take complete case analysis to be unfavorable given the high probability that missingness in the dependent variables could be correlated with their unobserved values. I thus apply the conventional missing-at-random (MAR) assumption,[56,57] whereby I assume that relationships among observed values hold true among unobserved values. Under MAR,

$$\Pr(\mathbf{r}|\mathbf{y}_{\mathrm{o}}, \mathbf{y}_{\mathrm{m}}) = \Pr(\mathbf{r}|\mathbf{y}_{o})$$

where $\mathbf{r}$ denotes the vector indicating missingness or nonmissingness, and is independent of the unobserved values $\mathbf{y}_{\mathrm{m}}$ given the observed values $\mathbf{y}_{\mathrm{o}}$. MAR dictates imputation upholding observed relationships within the dataset produces observable data, whereas the missing-completely-at-random (MCAR) assumption underlying complete-case analysis would require uniform nonresponse where missingness is taken to be independent of observed as well as unobserved values:

$$\Pr(\mathbf{r}|\mathbf{y}_{\mathrm{o}}, \mathbf{y}_{\mathrm{m}}) = \Pr(\mathbf{r})$$

An complete-case estimating equation $U(\mathbf{y}, \theta)$, where θ denotes a given population parameter, would remain consistent when data are missing only if the missing-completely-at-random assumption were fulfilled such that

$$\mathbf{E}U(\mathbf{Y}, \theta) = \mathbf{E}U(\mathbf{Y}_{o}, \theta)$$

Reporting effort varies with wealth and with public health capacity such that it is improbable that nonresponse in independent and dependent variables would be uniform. This circumstance justifies my use of multiple imputation.

I follow established guidelines[58,59] to determine that $m = 20$ imputations are sufficient to provide efficient and unbiased estimates of regression parameters relative to infinite imputations. I use iterative chained equations to impute the values of missing cells. The procedure designates a separate imputation model for each variable to be imputed, such that categorical and ordinal variables, for instance, may be imputed by appropriate logistic functions, continuous variables by

linear functions, etc. The imputation begins by filling in all missing cells of each variable with values obtained by random-sampling-with-replacement from that variable's observed values. Observed values of each variable to be imputed are then regressed against the filled-in values of all other variables contained within the imputation model. Missing values within each dependent variable are then replaced, in turn, by random draws with replacement from the variable's *k*-dimensional posterior predictive distribution given a stated regression relationship with *k* other covariates. A single cycle entails one regression for each variable to be imputed. For each of the 20 imputed datasets, I expect reasonable convergence in imputed values following 10 cycles.

I execute a congenial imputation approach specifying count variables as continuous variables with truncation at zero to avoid imputing negative incidence. Various sources[59-68] support the adoption of congenial, non-nested models in spite of the resultant violation of the requirement for an analysis model to be nested within an imputation model, indicating that inference upon the congenially-imputed data in fact converges toward an identical procedure with increasing *m*.[61] I additionally ensure robustness against error imposed by the unmet multivariate normality assumption by using bootstrap samples to estimate regression parameters within the imputation models.[69] The bootstrap approach resamples the observed data randomly with replacement, generating 1,000 datasets matching the number of observations within the dataset so as to approach the unknowable population distribution from which the sample is drawn.

Specifying congenial models is frequent in multiple imputation because the multivariate noramality (MVN) assumption rarely holds for real-world data, and parameter estimates tend to be robust to any introduced error.[59-64] Demirtas, Freels, and Yucel (2008)[65] demonstrate the validity of specifying a congenial linear model when the MVN assumption is clearly violated by variables subject to density flatness, heavy tails, non-zero peakedness, skewness, and multimodality even when case missingness *r* is high (75%) provided sample size is not small ($n>40$). In my case, sample size exceeds their threshold by orders of magnitude and case missingness is well below 75%. Graham and Schafer (1999)[66] indicate excellent performance of non-normal variables imputed under the MVN assumption without transformation in analyses even when sample size is small. Graham and Hofer (2000)[67] confirm this finding in an application of real-world data. Lee and Carlin (2010)[68] go so far as to justify linear imputation of

binary and ordinal variables under the violated MVN assumption. While appropriate *m* and the suitability of congenial imputation models are contested topics within the statistical literature, I view any biases imposed by the imputation method as insignificant relative to bias imposed by my use of international disease surveillance data. For the exploratory nature of this analysis, the imputation methods are sufficient.

To optimize prediction by the imputation models, I add (but withhold from the "causal" analysis models) an ordinal measure for the endemicity of each disease in each country (values 1-6) as provided by the Global Disease Epidemiology Online Network (GIDEON), which additionally serves as the source for national annual incidence.[70] I additionally include two geographic variables to stabilize imputed weather values; these are each country's WHO regional designation and the absolute value of each country's latitude midpoint.[49]

Autochthonous incidence of the diseases I investigate does not occur in all countries of the world. The global reporting of all diseases other than tuberculosis, gonorrhea, and syphilis follows a semicontinuous distribution, in that a high number of countries are non-endemic to given diseases and frequently report zero annual cases. To maintain original distributions and avoid variance distortion by the excess zeros, I pursue a validated two-step imputation approach[60] preserving zero-inflation:

1. I create a binary variable *a* corresponding to each disease with
$$a = \begin{cases} 1 \text{ if } y > 0 \text{ } or \text{ } y = \text{missing} \\ 0 \text{ if } y = 0 \end{cases}$$
then converted to missing all values of $y = 0$ from the y time series;
2. I generated a new incidence variable $i = ya$ post-imputation to restore original zero values for the analysis.

3.3.3: Analysis model specification

I employ zero-inflated negative binomial regression[71] in modeling the number of cases of each disease within each country annually. This technique accounts for spatiotemporally-coincident processes generating incidence *i* either equal to zero by a logistic function given the vector **z**, or drawn from the negative binomial distribution conditional on the vector **x**:

$$i \sim \begin{cases} 0 \text{ with probability } \varphi \\ g(i|\mathbf{x}) \text{ with probability } 1 - \varphi \end{cases}$$

such that

$$P\{I = i|\mathbf{x}\} = P(I = i|\mathbf{x}, \mathbf{z}) = \begin{cases} \varphi(\gamma'\mathbf{z}) + \{1 - \varphi(\gamma'\mathbf{z})\}g(0|\mathbf{x}) \text{ if } i = 0 \\ \{1 - \varphi(\gamma'\mathbf{z})\}g(i|\mathbf{x}) \text{ if } i > 0 \end{cases}$$

The mean and variance of the zero-inflated negative binomial model may thus be represented as

$$E(i|\mathbf{x}, \mathbf{z}) = \mu(1 - \varphi)$$

$$V(i|\mathbf{x}, \mathbf{z}) = \mu(1 - \varphi)(1 + \mu(\varphi + \alpha))$$

The model allows for overdispersion with the term α, which we may relate to unobserved heterogeneity τ with mean 0 and variance $\alpha = \frac{1}{\theta}$. The posterior values of i given the observed covariate vector **x** differ from a Poisson distribution according to an unobserved heterogeneity term τ:

$$E(i|\mathbf{x}, \tau) = e^{\mathbf{x}'\boldsymbol{\beta}}\tau$$

Taking $\tau \sim \Gamma(0, \frac{1}{\theta}) \equiv \text{Gamma}(k, \theta)$ where *k* denotes shape and θ denotes scale, we obtain a negative binomial distribution for *I* conditional on **x**:

$$P(I = i|\mathbf{x}) = \frac{\theta^{\theta}\mu^{i}\Gamma(\theta + i)}{\Gamma(i + 1)\Gamma(\theta)(\mu + \theta)^{\theta+i}}$$

Employing $\alpha = \frac{1}{\theta}$ as the overdispersion parameter indicates that the negative binomial posterior distribution of *i* converges to a Poisson distribution as $\alpha \to 0$. For each model, I perform a significance test for the overdispersion parameter α to validate the choice of the negative binomial model.

As per standard procedure,[58] I pool regression coefficients across the imputed datasets as the arithmetic mean of coefficients observed across individual datasets such that:

$$\hat{\beta} = \frac{1}{m}\sum_{m=1}^{M}\hat{\beta}_m$$

while the single variance estimate (V_β) captures within-imputation (*W*) as well as between-imputation (*B*) variance:

$$V_\beta = W + \left(1 + \frac{1}{m}\right)B$$

where, with M = 20,

$$W = \frac{1}{m}\sum_{m=1}^{M} s_m^2 \text{ and } B = \frac{1}{m-1}\sum_{m=1}^{M}(\hat{\beta}_m - \hat{\beta})^2$$

To isolate the effects of the covariates, I control for population and for reporting quality, as estimated by the WHO's assessed tuberculosis case notification rate.[52] This measure presents numerous advantages for the analysis at hand:

- The assessment is performed by a single agency with near-global administrative authority and monitoring capacity;
- The spatial and temporal coverage extents of the WHO assessment exceed my own, thus data are not widely extrapolated;
- Tuberculosis is endemic to all countries of the globe, such that case notification can be measured (i.e. is not subject to division by zero);
- Tuberculosis surveillance is less problematic within the temporal range than surveillance for HIV, another globally-distributed notified disease for which reporting estimates are available. Nations' HIV surveillance and reporting has been downwardly-biased by political and cultural factors that do not necessarily apply to other diseases. In addition, the new arrival of HIV within countries during the early 1990s dictates that surveillance for HIV could have been consistently lower than surveillance for already-endemic diseases such as TB within certain country-years.

Tuberculosis notification ratios serve as a coarse proxy for reporting of other diseases. Controlling for individual countries' time-variant trends in general surveillance with this measure, I expect excess error to arise from country-level differences in the reporting of TB and of other diseases due to national and international TB prioritization. To account for this remaining bias, I employ a clustered error term (described in detail below) absorbing country-level effects on a disease-specific basis. In the cases of pertussis, tuberculosis, and measles, I additionally control for vaccination coverage.

The data are not independent and identically distributed but rather represent repeated observations within countries, so I generate models robust to temporally- and intra-nationally autocorrelated residuals. This method accommodates latent dependence among observations, and integrates variation over time into the modeling framework.[13] Individual cases of infectious disease represent non-independent outcomes, with risk arising from exposure to a pool of

already-infected individuals. I thus take national incidence within a year to be an autoregressive AR(1) process such that I_t is a function of I_{t-1}. For the count model, I control for prior-year incidence such that, for $I_t > 0$, a regression parameter δI_{t-1} absorbs the number of cases expected to occur based on the country's prior-year cases; this allows the models to focus upon environmental covariates accounting for excess incidence that would signal re-emergence. In the case of the logistic model, I again use an AR(1) process to control for whether or not $I_t = 0$ on the basis of whether or not $I_{t-1} = 0$ via the binary variable $I_{t-1} = Q_t$ and its driver ωQ_{t-1}. Note that Q_t differs from the original binary variable A_t in that Q_t incorporates zeros generated during imputation.

While the autoregressive term isolates excess annual incidence within countries, I expect a "country"-level effect to persist in the residuals due to countries' time-invariant prioritization of reporting TB relative to reporting other infections. In the presence of this persisting country-level disturbance after controlling for TB notification and longitudinal autocorrelation, I expect conventional regression to produce downwardly-biased error estimates. Because my purpose is variable selection, such overstated precision would be unacceptable. I control for country-level effects by estimating cluster-robust standard errors. This decomposes the model's global residual term ϵ into cluster-specific disturbances $\gamma_c$and idiosyncratic individual-level disturbances v_{ct} which both have mean zero and variance σ derived by

$$E(\epsilon_{ct}) = \sigma^2 = {\sigma_\gamma}^2 + {\sigma_v}^2$$

To estimate error for clustered β coefficients, one would multiply the original variance formula for each β by the variance inflator term:

$$(1 + (\bar{n}_g - 1)\rho)$$

where $\bar{n}_g$ is the mean number of observations within groups and ρ represents the intracluster correlation coefficient:

$$\rho = \frac{{\sigma_\gamma}^2}{{\sigma_\gamma}^2 + {\sigma_v}^2}$$

Criticism for the cluster-robust approach has identified the technique as a conceptually-confused post-hoc adjustment for initial model misspecification.[72,73] A pure solution may be to aggregate analysis to the level of clustering; in the case of this particular study, such an approach would prevent identification of variation over time within countries, in fact weakening the models'

ability to distinguish impacts of the chosen covariates from baseline country-level reporting error here absorbed in the autoregressive term.[74] Identifying that I account already for the bulk of country-level effects by treating incidence as an AR(1) process, that I account for temporal trends in within-country reporting quality through the TB notification time series, and that with 207 countries represented the models account for a high number of clusters, I expect the statistical gains from the clustering method to be sufficiently equivalent to what may obtained by adding a time-invariant covariate for each country's differential notification of any given disease and TB.[74,75] Such covariates are neither defined within the literature nor available from current datasets. A general rule-of-thumb within the literature is to use at least 20 clusters to provide reliable estimates; I exceed this by a factor greater than 10, and thus argue that my approach is in line with standards of best practice.[74]

I perform model selection by backward stepwise regression using $p = 0.2$ as a cutoff value for variable inclusion. Critiques of stepwise regression as an ad-hoc approach lacking a basis in statistical theory are common and relevant here;[76] however, the development of superior computational approaches to variable selection in missing-data scenarios has lagged behind the development of imputation techniques.[77,78] Perhaps the most common model selection tool to avoid overfitting, the Akaike Information Criterion fails consistently to identify a correct model following imputation of ignorably missing data as I encounter here.[77] Stepwise modeling remains common practice within eco-epidemiologic literature exploring global disease drivers, even in the absence of a missing data problem that would prevent more sophisticated approaches.[8,9] In my case, interest in main effects of specific variables outweighs interest in accounting for sequential, causal processes subject to collinearity and confounding. I seek only to identify sets of environmental factors with significant, independent effects upon the disease outcomes, rather than to mathematically characterize whole transmission systems with multiple interdependencies. Thus the stepwise procedure, which limits the effects of multiple hypothesis testing by stopping at $p = 0.2$, is likely sufficient.

Prior to model building, I select the optimum representation of each covariate (1-year lag, 3-year lag, or 5-year moving average) for each disease as that which obtains the highest level of statistical significance in a multivariate analysis including the covariate of interest and all

applicable controls. All statistical analysis are performed using STATA 11 (StataCorp, College Station, TX), with additional support from Patrick Royston's `ice` package.[79]

I assemble the findings from all diseases in a matrix indicating variable retention, significance, and positive or negative β. I compare findings among zoonoses against those among anthroponotic diseases to assess differences and similarities in the roles given drivers play for both types.

PART 4: RESULTS

The results table (see Table 5) itemizes determinants vertically, separating (in white) the outcomes of the count model for incidence greater than one, and (in grey beneath) the outcomes of the logistic model for presence or absence of autochthonous transmission. Because raw outcomes from the logistic component treating incidence = 0 as a "success" are non-intuitive in the context of the count model, I invert logistic outcomes such that the reader may more easily align risk factors between both model segments. Zoonotic and anthroponotic diseases are each given two columns within the tables. For each category, the leftmost column contains diseases which, for a given covariate, retain the variable as a risk factor at a maximum Type-I error rate of $p<0.2$; the column to the right indicates diseases for which the covariate is retained as a protective factor, i.e. the association was inverse to what would be expected.

Of the seventeen generated models (one for each disease tested), F-tests for overall fit identify only sixteen to be significant with respect to the parameters included. With $p=0.6718$, the plague model appears outstandingly insignificant overall; although stepwise variable elimination had progressively greater returns in terms of lowering the model's overall Type-I error rate, this likely occurred through multiple hypothesis testing such that I lack confidence in the model's ability to differentiate true partial effects among the included variables. Because tuberculosis models fail to achieve convergence amidst stepwise variable removal past $p=0.389$, I report the outcomes with $p<0.2$ from that stage of model fitting. This indicates that five of seventeen non-control parameters within the tuberculosis model displayed $p>0.2$ but were not removed; i.e. are retained as controls. Because these variables compete with those that are retained for explanatory power within the model, I view the significance estimates for obtained tuberculosis parameter outcomes as conservative. For all models, the overdispersion term α is significant, validating the fit of the data to a negative binomial distribution rather than the less-dispersed Poisson.

Among weather variables, most of the significant findings occur within the logistic models for zero- or nonzero incidence and are among zoonotic diseases. Threshold rainfall and temperature conditions are prerequisite for disease transmission, but there is variation among zoonoses with regard to requirements for warm, dry, cool, or wet conditions. While the presence or absence of

diseases with anthroponotic transmission tends to be non-significantly associated with temperatures, it is notable that all diseases which retain a temperature term in the logistic model indicate greater probability for diseases to be present when cooler temperatures are observed. This initially appears to be in line with the Wolfe (2007) conclusion that anthroponotic transmission evolves preferentially in temperate regions. However, because the models are autoregressive and control for the highly-significant effects of prior-year endemicity, these findings are robust to residual bias from latitude (i.e. the simple conclusion that tropical diseases require warm temperatures, or that temperate diseases require cool temperatures). The models indicate that year-by-year fluctuations in weather carry the greatest weight for predicting if disease occurs at all; thus, those countries where transmission of a given disease is currently sporadic, and where other conditions requisite for transmission are met, have the greatest vulnerability to becoming endemic under a climate change scenario. The count models are remarkably less sensitive to the effects of temperature and rainfall; this is likely because country-years are too coarse a unit of analysis to detect consistent effects of weather on disease incidence. Vulnerability of focal disease ecosystems to meteorological events is dependent upon local factors; the associations are likely obscured when incidence and weather data are aggregated across an entire country.

A similar circumstance, where logistic models are far more successful than count models in identifying significant effects, arises with respect to land surface. Among zoonotic diseases, forested area is positively and significantly associated with the presence of diseases whose animal reservoirs are sylvatic. In contrast, all anthroponotic diseases which retain forested area as a covariate are associated with deforestation. This likely indicates that forested area outperforms the collinear urbanization variables in the stepwise procedure; only measles exhibits significant main effects from both urbanization and deforestation. The count model implicates deforestation in excess leptospirosis incidence; indeed, the disease has emerged within urban slums of Latin America, and the effect of deforestation remains significant after controlling for another significant effect from the rate of urbanization. Agricultural area is significant only in the logistic models for zoonotic diseases, and is not retained in any models for anthroponotic diseases. All diseases which are positively associated with agricultural area are classed as Stage 2 in the Wolfe et al. (2007)[6] scheme, indicating that direct animal exposure would likely be

responsible for all cases. Malaria, which has a far lower occurrence of zoonotic transmission, is the only zoonosis found to be significantly associated with decreasing agricultural area.

Island nation status appears as a powerful predictor for the eradication of zoonotic and anthroponotic diseases, or for the inability of those diseases to arrive; this is unsurprising given the greater ease of implementing public health and ecological interventions for small, geographically-contained settings.[80] In contrast, the count models identify no significant associations with zoonotic diseases, but identify significant positive associations with each anthroponotic disease for which any significant eradication effect appears. Thus while complete eradication of anthroponotic disease is more feasible in islands nations, interpersonal transmission is highly pronounced within such confined populations. Although the models retain covariates to control for material disadvantage, it is likely that heightened epidemic potential of introduced anthroponoses within island populations owes in part to the lower relative social and economic conditions of life within these nations.[49,81] Because governments are small and poor, minimal public health infrastructures may exist to control epidemics of diseases that appear only sporadically when cases acquired outside are introduced.

Unlike in the case of weather and land surface, the count models are more successful than the logistic models in identifying significant main effects from access to improved water and sanitation. Although these infrastructural variables appear to be of particular significance to zoonotic diseases, there appears to be no added importance for either zoonotic or anthroponotic diseases transmitted via the feco-oral route; most zoonoses for which outcomes are significant are in fact vector-borne, and only one of four significant findings relating anthroponotic diseases to a lack of water and sanitation infrastructures applies to a feco-oral disease.

Dengue may initially appear problematic in the results table in that its absence from countries is significantly associated with increasing levels of access to water and sanitation; it is worthy to note that, unlike other tropical diseases, autochthonous dengue does occur sporadically within temperate and developed countries. This indicates that within countries which currently see sporadic autochthonous cases, aggressive and targeted control efforts extending beyond basic infrastructural improvement are necessary to prevent the disease from becoming endemic. The

same is not true for malaria, the transmission of which appears significantly impeded by improvements to water and sanitation infrastructures. Exceptional findings in the case of dengue likely owe to unique resilience of its vector; unlike *Anopheles* mosquitoes transmitting malaria, *Aedes* mosquitoes carrying dengue have been documented to live and breed within sanitation facilities such as septic tanks,[82] and withstand the removal of standing water because their eggs resist desiccation for months on end.[83] That these models can identify differences between the impacts of infrastructural interventions on mosquito-borne diseases, and at the same time identify significant effects for anthroponotic respiratory diseases (likely due to collinearity with unobserved building quality and ventilation) suggests changes to infrastructure have remarkable power in mitigating disease vulnerability.

Whereas inflation-adjusted GDP per capita exhibits highly significant main effects on the presence of zoonoses even after controlling for collinear factors including malnutrition and access to water and sanitation, it does not appear as a significant risk factor for any anthroponotic disease. Because the autoregressive control term is highly significant in the logistic models, I interpret the poverty-zoonoses associations as meaning that, within countries where cases occur sporadically, zoonotic disease transmission is acutely sensitive to fluctuations in a country's economic productivity, or to the failure of a national economy to keep pace with population growth. In contrast, poverty appears significantly protective against measles and pertussis in the count models. Both are vaccine-preventable childhood diseases for which vaccination coverage has been diminishing in wealthy countries. Given that vaccination coverage has nonlinear effects upon disease incidence at a population scale, it is possible that poverty and vaccination compete for explanatory power in the main effects-only models employed here.

Political terror appears as an almost ubiquitously-significant driver of zoonotic disease presence or incidence, in spite of the retention of confounding demographic and economic variables within our models. Although often neglected from epidemiological models, political terror here appears to vie with malnutrition and sanitation as a basic and critical component of populations' vulnerability to partially-controlled and fully-endemic zoonoses, as indicated by the logistic and count models, respectively. Among anthroponotic diseases, political terror is a powerful predictor for excess cases of sexually-transmitted infections and tuberculosis; because the latter

disease rides the back of the HIV epidemic as the leading cause of death among HIV-infected people, and because our model for tuberculosis includes up to a 5-year lag, this identified association may owe in part to sexual HIV transmission varying with respect to political terror.[84] Butler et al. (2007)[85] demonstrate empirically that sexual violence arises within countries under the same socio-political conditions that cause high scores on the Political Terror Scale. Because STI risk is higher among perpetrators and victims of sexual violence,[86,87] it is likely that political terror is associated with directly- (gonorrhea and syphilis) and indirectly- (tuberculosis via HIV) sexually-transmitted infections, but unassociated with other anthroponotic diseases, in part because of increased sexual violence in states experiencing terror. Notably, political terror outperforms UNHCR population of concern, i.e. the number of people displaced owing to circumstances to which political terror would give rise; no disease models retain UNHCR population of concern through the variable selection procedure.

Population growth rate and prevalence of malnutrition are remarkably significant predictors of zoonotic disease incidence and presence; in contrast, rate of population growth is retained in no anthroponotic disease models, and malnutrition prevalence is significant only to one anthroponotic disease. We employ population growth rate as a coarse measure capturing both natural increase and immigration within countries. Zoonoses, like crowd diseases such as measles and pertussis, circulate primarily among children; the measure primarily indicates natural increase in developing countries endemic to tropical zoonoses, whereas it indicates immigration in wealthy, temperate countries. Within developed countries, public health capacity to serve a growing population and to screen for disease among newly-entering migrants indicates the importation of disease by mobile populations may not be significant at the national level. Reasons for the differential impact of malnutrition upon zoonotic and anthroponotic diseases are less clear. While malnourished individuals may be at greater risk for some infections than for others, it is additionally likely that malnutrition, as the cumulative outcome of multiple and extreme sources of disadvantage, shares these distal causes with zoonoses that circulate in impoverished tropical settings. Lyme disease and hantavirus, zoonoses that generally occur in middle- and higher-income settings, do not retain malnutrition as a covariate.

Gender inequality and education indeces appear significant primarily in count models for anthroponotic diseases; here they are retained, but rarely with more than marginal significance. Although STIs were highly sensitive to political terror, none of these diseases retain gender inequality through the stepwise selection procedure. In contrast, both anthroponotic feco-oral diseases retained gender inequality at the highest level of statistical significance. It is unclear whether this circumstance is an artefact of the data or confirms qualitative accounts[88,89] that young children whose mothers face institutionalized disadvantage are vulnerable to disease acquisition. Diarrheal infections exact their highest burden among infants and young children;[90] likewise, the greater severity of cases among children, and higher rates of mortality, may dictate that pediatric cases are more likely to be reported, thus comprising the bulk of our data. The argument for childhood vulnerability in gender-inequitable societies is weakened by the absence of significant relationships with childhood-related zoonoses. Education appears to significantly reduce risk for vaccine-preventable anthroponotic childhood diseases. This may be rationalized by of an education index simultaneously accounting for the prevalence of education among a country's adults and children; global vaccination efforts generally target schoolchildren,[91] and countries without in-school vaccination schedules often rely upon parental education as a means of encouraging uptake.[92] Thus children's and parents' educations account significantly for vaccine uptake in various developing countries,[93,94] and may exhibit a main effect on incidence after controlling for the main effect of vaccination rates.

Variables related to mobility perform poorly relative to expectations; the majority of significant findings among zoonoses and nonzoonoses indicated greater risk in nations with less-mobile populations. This is likely due to confounding by development status, as people within resource-poor, underdeveloped nations lack the basic infrastructures and economic assets necessary to enable frequent, international travel. Likewise, nations of asylum for refugees tend to be highly-developed and to have stringent screening and vaccination requirements for immigrants. The effects of mobility are more likely to be identified within analyses at more local levels, where individual outbreaks and contact networks associated with imported cases may be identifiable.

Table 5: Results

			Zoonotic diseases $n = 10$		**Anthroponotic diseases** $n = 7$	
			+	-	+	-
Mobility and connectivity	Air passengers	$i > 0$	Rabies†	Lyme* Salmonella**		Measles† Gonorrhea* Syphilis* Tuberculosis†
		$i \neq 0$		Malaria Rabies†		
	Paved roads fraction	$i > 0$				Typhoid* Shigellosis
		$i \neq 0$	Plague Salmonella† Leptospirosis	Dengue**	Measles** Pertussis	
	Incoming refugees	$i > 0$		Salmonella*		Measles Gonorrhea* Syphilis
		$i \neq 0$		Malaria	Measles	
Population and demographics	Urban population	$i > 0$			Pertussis**	Typhoid
		$i \neq 0$				Measles
	Rate of urbanization	$i > 0$	Leptospirosis*	Lyme†		
		$i \neq 0$	Malaria Rabies**	Lyme**	Measles*	
	Gender inequality index	$i > 0$	Hantavirus Leptospirosis†		Typhoid** Shigellosis**	
		$i \neq 0$	Rabies†			Typhoid
	Education index (inverse)	$i > 0$	Malaria†		Typhoid** Shigellosis Measles**	
		$i \neq 0$	Leishmaniasis Malaria	Dengue† Lyme* Salmonella		
	GINI coefficient	$i > 0$	Dengue	Rabies	Shigellosis Syphilis	
		$i \neq 0$	Dengue** Malaria**	Hantavirus Salmonella†		Measles* Shigellosis†
	Adult HIV prevalence	$i > 0$	Leishmaniasis† Malaria* Rabies		Pertussis Tuberculosis**	Syphilis
		$i \neq 0$		Dengue** Leptospirosis Lyme†	Measles	Pertussis** Typhoid†
	Rate of population growth	$i > 0$	Leishmaniasis* Lyme Plague			
		$i \neq 0$	Cholera* Leishmaniasis† Plague*	Dengue		
	Malnutrition prevalence	$i > 0$	Leishmaniasis† Rabies* Salmonella*		Pertussis* Gonorrhea† Tuberculosis	
		$i \neq 0$	Cholera Dengue** Malaria† Plague** Rabies		Measles	

'**' denotes $p < 0.01$; '*' denotes $p < 0.05$; '†' denotes $p < 0.1$

Geopolitical circumstances	Poverty (GDP/capita at const. USD inverse)	$i > 0$	Cholera* Dengue			Measles* Pertussis**
		$i \neq 0$	Leishmaniasis* Leptospirosis** Malaria** Plague**	Dengue	Measles	
	Political terror scale	$i > 0$	Dengue** Leishmaniasis Leptospirosis** Salmonella		Measles Gonorrhea** Syphilis** Tuberculosis**	
		$i \neq 0$	Cholera Dengue† Malaria* Rabies**			Shigellosis†
	UNHCR population of concern	$i > 0$				
		$i \neq 0$				
Infrastructure	Improved water (inverse)	$i > 0$	Cholera* Lyme† Malaria* Rabies		Pertussis†	
		$i \neq 0$	Malaria*	Dengue** Lyme†	Measles* Typhoid	
	Improved sanitation (inverse)	$i > 0$	Hantavirus Leptospirosis** Malaria*		Pertussis† Shigellosis**	
		$i \neq 0$	Cholera	Dengue*		
	Agricultural mechanization (inverse)	$i > 0$	Salmonella	Lyme*	Shigellosis	
		$i \neq 0$	Dengue	Plague**	Typhoid†	
Weather and climate	Rainfall (month max rain)	$i > 0$				
		$i \neq 0$	Cholera† Malaria*	Lyme*	Typhoid†	Pertussis†
	z-score	$i > 0$			Tuberculosis	
		$i \neq 0$		Leptospirosis		
	Temperature (month max. temp.)	$i > 0$	Cholera		Typhoid*	Pertussis Tuberculosis†
		$i \neq 0$	Dengue** Malaria**	Cholera* Lyme** Leptospirosis		Measles* Pertussis† Typhoid Shigellosis
	z-score	$i > 0$	Malaria*		Shigellosis Syphilis	Measles
		$i \neq 0$		Cholera		
Environmental change and geography	Forested area (%)	$i > 0$		Leptospirosis*		Shigellosis*
		$i \neq 0$	Dengue** Leptospirosis Lyme*			Measles* Typhoid†
	Agricultural area (%)	$i > 0$		Leptospirosis Rabies		
		$i \neq 0$	Leptospirosis Lyme* Rabies*	Malaria*		
	Island country	$i > 0$	Leishmaniasis Leptospirosis		Measles* Pertussis* Typhoid†	Tuberculosis*
		$i \neq 0$	Dengue**	Cholera Lyme* Malaria** Plague† Rabies**		Measles** Pertussis** Typhoid†
	Landlocked country	$i > 0$		Malaria*		Pertussis
		$i \neq 0$	Hantavirus† Shigellosis			

'**' denotes $p < 0.01$; '*' denotes $p < 0.05$; '†' denotes $p < 0.1$

CHAPTER 5: DISCUSSION

5.1: Principal findings

Physical environment, as measured by annual weather, island nation status, and changes to land vegetative surface, is of greatest importance in determining the presence or absence of zoonotic diseases. Such conditions thus appear prerequisite for autochthonous zoonotic pathogen circulation. When physical environmental conditions necessary for sustaining zoonotic disease are achieved within a country, social and human development factors tend to play a more important role governing the number of infections that occur. Social and human development factors are relevant to diseases transmitted by animals and among humans; with the exception of deforestation, physical environmental change tends to be unassociated with anthroponotic diseases. Notably, zoonotic diseases are more sensitive to the effects of both categories of environmental predictors than are anthroponotic diseases. Taken together, the results indicate that disease ecologies lose their environmental sensitivity as specialization for anthroponotic transmission increases basic reproductive ratio. Thus the finding from evolutionary biology that time-variant environments select for ecological generalists[11,16,17] appears to hold true in the global distribution of zoonotic and anthroponotic disease incidence.

Autoregressive modeling isolates excess "emergent" cases of infections for evaluation against a suite of environmental determinants. This method extends prior work addressing emergence and re-emergence as a discrete "event";[3] rather than limiting outcomes to disease introduction, zero-inflated negative binomial regression allows me to examine separate processes accounting for disease introduction and subsequent amplification within populations. Modelling diseases individually in light of their unique underlying ecological systems, I obtain relatively consistent results within the zoonotic and anthroponotic disease categories indicating that temporal variability within environments becomes less important to human disease incidence as pathogens acquire specificity for human transmission.

5.2: Strengths and weaknesses of the present study

The study is limited by a narrow selection of globally-distributed pathogens. Having selected pathogens which appeared via GIDEON to maximize global spatial coverage and reporting

quality, the sensitivity of outcomes to the specific diseases chosen is difficult to assess. Holistically, my conclusions represent a qualitative assessment of the trends arising from all seventeen individual-disease models; thus replacing several diseases with alternatives is unlikely to fundamentally alter generalities arising across the full array of models. Moreover, the fact that the results are in line with broader evolutionary theory[11,16,17] suggests that selected diseases most likely comprise a sufficiently representative sample.

The use of international disease surveillance data is inevitably problematic in that observation- and process error are unlikely to be independent and identically-distributed. Non-uniform probability of case notification applies across diseases, across countries, within countries, and over time.[33] I seek to minimize analytical interference by employing tuberculosis notification as a time-variant control for surveillance effort within and across countries, and by allowing for clustered residuals to absorb time-invariant discrepancies in country-specific notification for tuberculosis and for other diseases. Ultimately, the overriding bias imposed by differential reporting capacity supports the null hypothesis of no association with the covariates, and thus makes my inference only more conservative. Because disease surveillance and prevention both fall within the jurisdiction of a country's public health infrastructure, I expect case prevention capacity to be correlated with case notification such that countries lacking the public health resources to prevent outbreaks likewise lack the technical capacity to monitor incidence. This assumption applies not only across nations, such that one would expect poorer surveillance in less-developed countries, but also within nations over time; an environmental or humanitarian crisis causing prevention efforts to collapse would by this logic compromise disease surveillance as well during the same year. Reasoning that the country-years with highest risk for disease also have highest risk for under-reporting incidence, I expect the parameter estimates to be immeasurably biased toward the null. Thus although the variables and controls are coarse, the overriding source of bias favors Type-II error and conservative inference regarding distal environmental causes.

In the absence of uniform, disease-specific case notification ratios, I premise the conclusions in part upon the assumption that time-variant tuberculosis notification rates capture secular trends toward superior or inferior disease surveillance effort within countries over the period of

analysis. Given the autoregressive model specification, error attributable to within-country reporting heterogeneity for a given disease is more threatening than heterogeneous reporting effort across countries, which would be largely captured at baseline levels within the autoregressive term. Controlling for prior-year incidence allows isolation of the proportion of "emergent" cases of infection within a country. This quantity will be biased under the condition that tuberculosis case notification rates vary over time in a manner that is uncorrelated with variation in the notification rate for a disease at hand. Under this circumstance, the clustered errors absorbing systematic difference in reporting for TB and for the disease at hand will assume a time-invariant value that should rather vary over time. Such a circumstance could arise if a country were to stage a dramatic TB intervention over a short period of time, without accompanying this intervention with improvements in other public health systems. Here, the notification rates for TB during the intervention period would be systematically higher than notification rates for other diseases, which may hold constant or even drop as limited public health resources are re-allocated to TB surveillance. Although it is likely that such events occur throughout the dataset,[95] the marginal effects within individual models are likely to be small given n=4316. Because I seek only to generalize holistically across the suite of seventeen individual-disease models, the overall approach is suitably robust against this hazard.

Stepwise model selection techniques as applied here are subject to amplifying Type-I error through multiple hypothesis testing such that there is a danger of overstating significance estimates.[76] Emphasizing the outcomes of an individual candidate model is a common pitfall of stepwise regression; many candidate models will have similar goodness of fit, and parameter estimates within a given model may differ from those obtained by alternative models.[96] Admitting the present study is subject to these problems, I justify the stepwise approach in that the analysis neither: 1) focuses in great depth upon any individual model generated by the stepwise approach, nor 2) considers parameter estimates in any detail beyond whether they are positive or negative. Moreover, the application of reverse stepwise modeling within two[8,9] of three[3,8,9] prior global comparative studies in the eco-epidemiology of emerging infections suggests the approach and its sources of bias are common within this area of research, and well understood by its audience. Broader inference across outcomes from seventeen individual-disease models is in line with best practice for interpreting stepwise regression outcomes,[76,96]

although I must ultimately consider only one candidate model for each individual disease. Assuming Type-I error amplification for a given covariate to be stochastic across the seventeen individual-disease models, and acknowledging the pervasive conservative bias toward Type-II error within the analysis, I accept the methodological compromises that arise through a stepwise procedure in view that this particular source of error is mitigated by overriding bias toward the null.

The causal reliability of the approach and of the results rests upon the premise that aggregate country-year measures are appropriate surrogates for environmental exposures among infected individuals and populations. This premise is confounded by the fact that exposures and infections are not uniformly distributed within national populations, but are rather dependent upon factors such as individuals' geographic location and socioeconomic status.[97] This circumstance likely explains the retention of significant meteorological associations only within logistic models; variable weather is inherently localized in discrete space and time. In this way, national annual incidence of an already-endemic disease may be robust to focalized spatial outbreaks just as nation-wide meteorological averages are robust to local variability. I am limited by considering only heat on a representative day and precipitation during a representative month; meteorological determinants for human activity and ecological dynamics are diverse and context-dependent, and unlikely to be encapsulated by these summary measures alone. In the case of mobility variables, variation at the country-year level is more indicative of development status than disease risk, such that observed relationships are in fact inverse to those observed at more proximal units of analysis. Higher risk for disease within highly-mobile populations may be better identified through population-stratified approaches compartmentalizing individuals by their mobility status, or through contact-network models tracing pathogen transmission to individual imported cases. In either event, the scale must be reduced from that presently employed. Likewise, exposure to gender inequality, poverty, and educational opportunity are likely to vary categorically among social groups within populations such that summary measures aggregated to a national, annual unit are inadequately representative of individual risk for disease.

National aggregate measures may be more appropriate in the narrower case of individual risk factors which are nearly universal within populations. Infectious disease outcomes among individuals are non-independent, and are rather functions of exposure to a population-level reservoir of disease among infected individuals. Widespread malnutrition, HIV prevalence, political terror, and poor sanitation within a country are likely to achieve significance within population-level models because, when widely distributed, their magnified effects upon transmission dynamics affect people who are not individually exposed to these risk factors. Consider, for instance, poor sanitation; although one individual may have access to clean water and toilet facilities, the marginal protective effect of that resource is compromised by the individual's exposure to an outside community of individuals who, lacking sanitary facilities, have high prevalence of infection. Because common risk factors may wield population-level effects in the absence of uniform individual exposure, this nationally-aggregated approach is likely to be fruitful for covariates which are so prevalent within populations that their effect upon transmission dynamics outweighs the individual-level protective effect of non-exposure. Inferences arising across a suite of my statistical models may identify such variables, informing compartmental mathematical modeling for more precise estimations of threshold exposure requirements and corresponding effect sizes.

5.3: Strengths and weaknesses relative to other studies

This represents the first global comparative study accounting for space-time variation in environment and in incidence of diverse infectious diseases. The literature review identifies three studies that have rigorously tested the role of environment in the global distribution of pathogens; of these, only one has specifically evaluated emergent disease outcomes. Defining a case-incidence dependent variable, the present study operationalizes emergence with a measure that is concrete and compelling to public health audiences seeking to reduce disease burden. Empirically, the measure additionally represents the isolated process of pathogen transmission, rather than the dissimilar, and dissimilarly-mediated, processes of evolution and transmission evaluated together in prior work. I model global incidence of "re-emerging" infections at the expense of failing to represent novel Stage-2 pathogens[6] appearing in discreet spatial foci, as emphasized by Jones et al. (2008).[3] Because Stage-2 pathogens represented within the current literature propagate low disease burden relative to globally re-emerging pathogens, the present

study is an overdue indication of how the emerging disease paradigm matters in applied global health. Pinpointing associations between environmental change and subsequent expansion in the incidence of extant, high-burden infections, I indicate that factors in the evolution of novel pathogens contribute directly to the transmission of more-mature zoonoses, while anthroponotic disease reduction requires tailored interventions outside the general environmental sphere.

While a small sample of diseases relative to prior quantitative[3,7-9] and qualitative[4,22,26-29] studies dictates that this analysis sacrifices a degree of generalizability, my evaluation of a relatively wide selection of both human and physical environmental factors allows for numerous gains. From the broad pool of environmental factors and processes suggested within the qualitative literature to drive the emergence of infections, I establish quantitative links at a global scale for zoonotic as well as anthroponotic diseases. I indicate the pronounced effect of under-explored social determinants, notably political terror, malnutrition, population growth, poverty, and access to sanitation, on the incidence of zoonotic disease. Although widely referenced within the qualitative literature, the impact of these socio-environmental factors remains poorly assessed within a quantitative literatures, which generally concern physical and ecological factors to a greater extent. I indicate not only that these circumstances drive zoonotic and anthroponotic disease incidence, but also find, perhaps counter-intuitively, that social variables have greater influence over zoonoses than over anthroponotic diseases.

While this study is innovative its use of incidence data and in its consideration of variation over time, the findings must be interpreted in view of: 1) the coarseness with which underlying data are measured; 2) the small, although representative, sample of diseases modeled; 3) the high degree of confounding among covariates; and 4) the implications of country-year aggregation as a measure for exposure, as previously discussed. While empirical validation of ideas put forth in the qualitative literature is useful, I caution that findings from the present study are subject to statistical bias, namely from stepwise selection, inconsistent reporting, and national aggregation of observations. As an exploratory comparative study, the findings indicate differential roles of social and physical environment in the causation of zoonotic and anthroponotic disease re-emergence within countries.

5.4: Implications

A generally stronger association between time-variant environments and zoonotic disease incidence than between such environments and anthroponotic disease incidence suggests that evolutionary dynamics produce similar effects at the global scale and within the microenvironment of a pathogen community. The pronounced role of environmental change in introducing or amplifying zoonotic disease transmission within countries indicates that host breadth is an asset to pathogen survival and persistence at the level of national populations, as well as among progeny of a single parent cell, when external environments are dynamic. Selection for generalist pathogens within time-variant environments indicates that settings where enzootic diseases are driven to expand their host breadth and spill into human hosts are likewise at risk for sustaining zoonotic transmission of extant diseases. Mitigating human-animal exposures and stabilizing ecological and social conditions within such settings may accomplish the dual objectives of preventing opportunistic spillover events by novel pathogens and curbing transmission of extant, higher-burden zoonotic pathogens. Initial maladaptions to human infectivity however remain an impediment to persistence of novel microbes, and the ability for time-variant environments to sustain spillover as well as zoonotic circulation does not indicate that novel pathogens will successfully attain human infectivity within these settings.[10-12]

Anthroponotic disease transmission generally appears less volatile amidst changes to humans' socio-economic systems and physical environments; unlike in the case of zoonotic diseases, singular variables do not appear outstandingly significant across wide majorities of anthroponotic diseases included within the study. Whereas it is clear that resolving socio-economic circumstances such as rapid population growth, urbanization, malnutrition, poverty, political terror, land surface change, and infrastructural shortcomings could aid in reducing transmission of numerous zoonoses and in reducing nations' risk amidst uncontrollable processes such as climate change, a general finding of this nature does not arise categorically across anthroponotic diseases. While diminished environmental sensitivity among anthroponotic pathogens is a logical evolutionary circumstance given their specialization with regard to host breadth, future work covering a broader selection of anthroponotic infections should address whether this finding owes to the sample at hand or truly reflects a translation of microbial evolutionary principles to broader geographic scales.

REFERENCES

1. McMichael AJ. Environmental and social influences on emerging infectious diseases: past, present, and future. Phil Trans R Soc Lond B. 2004;359:1049-58.
2. McNeill W. Plagues and People. 1976; Middlesex: Penguin.
3. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D Gittleman JL, et al. Global trends in emerging infectious diseases. Nature. 2008;451:990-4.
4. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. Nature. 2004;430:242-9.
5. Taylor LH, Latham SM, Woolhouse MEJ. Risk factors for human disease emergence. Phil Trans R Soc Lond B. 2001;356:983-9.
6. Wolfe ND, Panosian Dunavan C, Diamond J. Origins of major human infectious diseases. Nature. 2007;447:279-83.
7. Woolhouse MEJ, Gowtage-Sequeria S. Host range and emerging and re-emerging pathogens. Emerg Infect Dis. 2005;11(12):1842-7.
8. Dunn RR, Davies TJ, Harris NC, Gavin MC. Global drivers of human pathogen richness and prevalence. Proc R Soc B. 2010;277:2587-95.
9. Guernier V, Hochberg ME, Guegan JF. Ecology drives the worldwide distribution of human diseases. PLOS Biol. 2004;2(6):e141.
10. Turner PE, Morales NM, Alto BW, Remold SK. Role of evolved host breadth in the initial emergence of an RNA virus. Evolution. 2010;64(11):3273-86.
11. Kassen R. The experimental evolution of specialists, generalists, and the maintenance of diversity. J Evol Biol. 2002;15:173-90.
12. Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis. 1996;1(1):7-15.
13. LaDeau SL, Glass GE, Thompson Hobbs N, Latimer A, Ostfeld RS. Data-model fusion to better understand emerging pathogens and improve infectious disease forecasting. Ecol Applications. 2011;21(5):1443-60.
14. Morse SS. Regulating viral traffic. Issues Sci Technol. 1990;7:81-4.
15. McCormack RK, Allen LJS. Disease emergence in deterministic and stochastic models for host and pathogen. Applied math and computation. 2005;168(2):1281-1305.
16. Levins R. Theory of fitness in a heterogeneous environment. 1. The fitness set and the adaptive function. Am Nat. 1962;96:361-73.
17. Levins R. Evolution in changing environments. Princenton, NJ: Princeton Univ Press; 1968.
18. Dennehy JJ, Friedenberg NA, McBride RC, Holt RD, Turner PE. Experimental evidence that source genetic variation drives pathogen emergence. Proc R Soc B. 2010;277:3113-21.
19. Ogbunugafor CB, Basu S, Morales NM, Turner PE. Combining mathematics and empirical data to predict emergence of RNA viruses that differ in reservoir use. Phil Trans R Soc B. 2010;365:1919-30.
20. Conway DJ, Roper C. Micro-evolution and emergence of pathogens. Int J Parasitol. 2000;30(12):1423-30.

21. Emerging parasite zoonoses: the role of host-parasite relationship. Int J Parasitol. 2000;30(12):1361-7.
22. Daszak P, Cunningham AA, Hyatt AD. Emerging infectious diseases of wildlife—threats to biodiversity and human health. Science. 2000;287:443-9.
23. Mathers C, Ma Fat D, Boerma JT. The global burden of disease: 2004 update. World Health Organ. 2008.
24. Morse SS, Schluedenberg. Emerging viruses: the evolution of viruses and viral diseases. J Infect Dis. 1990;162:1-7.
25. Wilcox BA, Colwell RR. Emerging and reemerging infectious diseases: Biocomplexity as an interdisciplinary paradigm. EcoHealth. 2005;2(4):244-57.
26. Fauci AS. Emerging and reemerging infectious diseases: The perpetual challenge. Academ Med. 2005;80(12):1079-85.
27. Murphy FA. New, emerging, and reemerging infectious diseases. Advances Vir Research. 1994. 43;1-52.
28. Snowden FM. Emerging and reemerging diseases: a historical perspective. Immunol Rev. 2008;225(1):9-26.
29. Morens DM, Folkers GK, Fauci AS. Emerging diseases: A perpetual challenge. Lancet Infect Dis. 2008;8:710-9.
30. Mulrow CD. Rationale for systematic reviews. BMJ. 1994;309(6954):597-9.
31. Woolf SH. The meaning of translational research and why it matters. JAMA. 2008;299(2):211-3.
32. Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, Bhatt S, et al. Global mapping of infectious disease. Phil Trans R Soc Lond B Biol Sci. 2013;368(1614):20120250.
33. World Health Organization. Research priorities for zoonoses and marginalized infections. World Health Organ Tech Rep Ser. 2012;971:ix-xi,1-119.
34. Yu VL. Global infectious diseases and epidemiology network: A world wide web-based program for diagnosis and informatics in infectious diseases. Clin Infect Dis. 2005;401(1):123-6.
35. Keller M, Blench M, Tolentino H, Freifeld CC, Mandl KD, Mawudeku A, et al. Use of unstructured event-based reports for global infectious disease surveillance. Emerg Infect Dis. 2009;15(5):689-95.
36. Katz R, Allen H. Domestic understanding of the revised international health regulations. Public Health Rep. 2009;124(6):806-12.
37. Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. 2005;20(5):238-44.
38. Weiss RA, McMichael AJ. Social and environmental risk factors in the emergence of infectious diseases. Nature Med. 2004;10:S70-6.
39. Diez Roux AV. Integrating social and biologic factors in health research: a systems view. Ann Epidemiol. 2007;17(7):569-74.

40. Morse SS, Mazet JAK, Woolhouse M, Parrish CR, Carroll D, Karesh WB, et al. Prediction and prevention of the next pandemic zoonosis. Lancet. 2012;380(9857):1956-65.
41. Karesh WB, Dobson A, Lloyd-Smith JO, Lubroth J, Dixon MA, Bennett M, et al. Ecology of zoonoses: natural and unnatural histories. Lancet. 2012;380:1936-45.
42. Dobson AP, Carper ER. Infectious diseases and human population history. Biosci. 1996;46:115-26.
43. Molyneux DH. "Neglected" diseases but unrecognized successes—challenges and opporutunities for infectious disease control. Lancet. 2004;364:380-3.
44. Dennehy JJ, Friedenberg NA, Holt RD, Turner PE. Viral ecology and the maintenance of novel host use. Am Nat. 2006;165:429-39.
45. World DataBank [database on the Internet]. Washington, DC: The World Bank. 2012 [cited 2013 Jan 9]. Available from http://databank.worldbank.org/data/home.aspx
46. Gender Inequality Index [database on the Internet]. New York (NY): Human Development Reports Office, United Nations Development Program. 2011 [cited 2013 Jan 9]. Available from http://hdr.undp.org/en/statistics/gii/
47. Solt F. Standardizing the World Income Inequality Database. Soc Sci Quarter. 2009;90(2):231-42. SWIID Version 3.1, 2011.
48. Wood RM, Gibney M. Political Terror Scale [database on the Internet]. 2012 [cited 2013 Jan 9]. Available from http://www.politicalterrorscale.org/download.php
49. Central Intelligence Agency. The World Factbook. Washington, DC. 2012 [cited 2013 Jan 9]. Available from: https://www.cia.gov/library/publications/the-world-factbook/
50. Food and Agriculture Organization of the Untied Nations. FAOCLIM 2: world-wide agroclimatic data. Rome, 2013. [cited 2013 Jan 9]. Available from: http://www.fao.org/nr/climpag/pub/EN1102_en.asp
51. National Climatic Data Center. Global Surface Summary of Day. 2013 [cited 2013 Jan 9]. Available from: https://www.ncdc.noaa.gov/cgi-bin/res40.pl#Z
52. Global Tuberculosis Database [database on the Internet]. Geneva: World Health Organization. 2012 [cited 2013 Jan 9]. Available from http://www.who.int/tb/country/data/download/en/index.html
53. Global Routine Vaccination Coverage [database on the Internet]. Geneva: World Health Organization. 2010 [cited 2013 Jan 9]. Available from http://www.who.int/immunization_monitoring/data/en/
54. Vasquez T. Master location identifier database [database on the Internet]. Garland (TX): Weather Graphics. 2010 [updated 2012 Aug 1; cited 2013 Jan 9]. Available from http://www.weathergraphics.com/identifiers/
55. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005 Dec;25(15):1965-78.
56. Lu G, Copas JB. Missing at random, likelihood ignorability, and model completeness. Ann Statistics. 2004;32(2):754-65.
57. Rubin DB. Inference and missing data. Biometrika. 1976;63:581-92.

58. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: J Wiley & Sons; 1987.
59. Schafer JL. Analysis of Incomplete Multivariate Data. London: Chapman & Hall; 1997.
60. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb;30(4):377-99.
61. Meng XL. Multiple-imputation inferences with uncongenial sources of input. Statist. Sci. 1994;9(4):538-58.
62. Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychol. Meth. 2002;7:147-77.
63. Horton NJ, Lipsitz SR. Multiple imputation in practice: Comparison of software packages for regression models with missing variables. American Statistician. 2001;55(3):244-54/
64. Johnson DR, Young R. Toward best practices in analyzing datasets with missing data: Comparisons and recommendations. J Marriage Fam. 2011 Oct;73(5):926-45.
65. Demirtas H, Freels SA, Yucel RM. The plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: A simulation assessment. J Stat Comp Sim;78(1):69-84.
66. Graham JW, Schafer JL. On the performance of multiple imputation for multivariate data with small sample size. In: Hoyle R, editor. Statistical Stragegies for Small Sample Research. Thousand Oaks (CA): Sage; 1999.
67. Graham JW, Hofer SM. Multiple imputation in multivariate research. In: Little TD, Schnabel KU, Baumert J, editors. Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples. Mahway (NJ): Lawrence Erlbaum; 2000.
68. Lee KJ, Carlin JB. Multiple imputation for missing data: Fully conditional specification versus multivariate normal imputation. Am J Epidemiol. 2010 Mar 1;171(5):624-32.
69. Steele RJ, Wang N, Raftery AE. Inference from multiple imputation for missing data using mixtures of normals. Statistical Methodology. 2010;7(3):351-65.
70. Global Infectious Disease Epidemiology Online Network [database on the Internet]. Los Angeles (CA): GIDEON Informatics, Inc. 1994 [updated weekly; cited 2012 May 1 to 2012 May 31]. Available from http://web.gideononline.com/web/epidemiology/index.php
71. Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. Psychol Bull. 1995;118(3):392-404.
72. Freedman D. On the so-called "Huber Sandwich Estimator" and "Robust Standard Errors". The American Statistician. 2006;60(4):299-302.
73. King G, Roberts M. How robust standard errors expose methodological problems they do not fix. Annual Meeting of the Society for Political Methodology, Duke University. 2012.
74. Arceneaux K, Nickerson DW. Modeling certainty with clustered data: a comparison of methods. Political Analysis. 2009;17:177-90.
75. Green DP, Vavreck. Analysis of cluster-randomized experiments: a comparison of alternative estimation approaches. Political Analysis. 2008;16:138-52.

76. Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP. Why do we still use stepwise modeling in ecology and behavior? J Anim Ecol. 2006;75:1182-9.
77. Chaurasia A, Harel O. Using AIC in multiple linear regression framework with multiply imputed data. Health Serv Outcomes Res Methodol. 2012;12(2):219-33.
78. Nakagawa S, Freckleton RP. Model averaging, missing data and multiple imputation: A case study for behavioral ecology. Behav Ecol Sociobiol. 2011;65(1):103-16.
79. Royston P. Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. Stata J. 2009;9(3):466-77.
80. Hopkins DR. Disease eradication. N Engl J Med. 2013;368:54-63.
81. United Nations. Programme of action for the sustainable development of small island developing states. Barbados; United Nations Global Conference on the Sustainable Development of Small Island Developing States; 1994.
82. Barrera R, Amador M, Diaz A, Smith J, Munoz-Jordan JL, Rosario Y. Unusual productivity of *Aedes aegypti* in septic tanks and its implications for dengue control. Med Vet Entomol. 2008;22(1):62-9.
83. Rezende GL, Martins AJ, Gentile C, Farnesi LC, Pelajo-Machado M, Peixoto AA, et al. Embryonic desiccation resistance in *Aedes aegypti*: presumptive role of the chitinized Serosal Cuticle. BMC Develop Biol. 2008;8(1):82.
84. Global tuberculosis control: WHO report 2010. Geneva: World Health Organ. 2010.
85. Butler CK, Gluch T, Mitchell NJ. Security forces and sexual violence: a cross-national analysis of a principal-agent argument.
86. Maman S, Campbell J, Sweat MD, Gielen AC. The intersections of HIV and violence: directions for future research and interventions. Soc Sci Med. 2000;50(4):459-78.
87. Dunkle KL, Jewkes RK, Nduna M, Levin J, Jama N, Khuzwayo N, et al. Perpetration of partner violence and HIV risk behavior among young men in the rural Eastern Cape, South Africa. AIDS. 2006;20(16):2107-14.
88. Robert B, Jackson M, Crist A. Gender inequality, vulnerability, and disaster: Issues in theory and research. In: The gendered terrain of disaster: Through women's eyes. 1998: 27-44.
89. Okojie CEE. Gender inequalities of health in the third world. Soc Sci Med. 1994;34(9):1237-47.
90. Cabral SA, de Moura ATS, Berkelhamer JE. Overview of the global health issues facing children. Pediatrics. 2012;129(1):1-3.
91. DeRoeck D, Jodar L, Clemens J. Putting typhoid vaccination on the global health agenda. N Engl J Med. 2007;357:1069-71.
92. Forsyth KD, Campins-Marti M, Caro J, Cherry JD, Greenberg D, Guiso N. New pertussis vaccination beyond infancy: recommendations by the global pertussis initiative. Clin Infect Dis. 2004;39(12):1802-9.
93. Koenig MA, Dishai D, Ali Khan M. Health interventions and health equity: the example of measles vaccination in Bangladesh. 2001;27(2):283-302.

94. Ndikuyeze A, Cook A, Cutts FT, Bennett S. Priorities in global measles control: report of an outbreak in N'Djamena, Chad. Epidemiol Infect. 1995;115(2):309-14.
95. Baltussen R. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost effectiveness and resource allocation. 2006;4:14.
96. Mundry R, Nunn CL. Stepwise model fitting and statistical inference: turning noise into signal pollution. Am Nat. 2009;173(1).
97. Leon DA, Walt G. Poverty, inequality, and health: an international perspective. Oxford, UK: Oxford University Press; 2001.

APPENDIX 1

1: Environmental determinants

1.1: Overview

My survey of the qualitative and quantitative literatures broadly identifies the following as drivers of global disease emergence:

A. Mobility and connectivity: The ease with which goods and individuals move within their own country and between countries.
B. Demographics and Geopolitics: Descriptions of the conditions of life within a population, including how the population is distributed, its growth trajectories, and how the people relate to one another and to their state;
C. Infrastructure: Measures of the physical quality of services affecting day-to-day life;
D. Geography and environment: Characteristics of the land a population inhabits, with special attention to the population's modes of using the land or interacting with ecological surroundings;
E. Weather and climate: Annual meteorological conditions and fluctuations in meteorological conditions, as would be expected to affect human and ecological activity

Within these categories, I define the measurable factors itemized below.

1.2: Mobility and connectivity

<u>Air travel</u>

Affordable air travel facilitates movement between distant regions of the world.[1,2] Whereas an arduous transcontinental or transoceanic voyage once separated microbes from naive host populations, microbes may now travel from their home range to a wholly new and susceptible environment within a span of hours. Limiting the outcome of interest to autochthonous incidence of a disease, I evaluate whether the rapid interconnection of peoples and places by air travel impacts pathogen introduction and establishment within novel settings, incorporating cases acquired domestically following, potentially, importation due to travel.

Road building

Constructing roadways between urban centers and across peripheral frontiers impacts pathogen transmission among peoples and places. Particularly in developing countries, road building indicates the resettlement of populations into novel environments; the development of new modes of production, travel, and trade within countries; and anthropogenic environmental change, often accompanied by other construction or resource development projects.[3,4] Sparse international and longitudinal data constrain our ability to analyze the effects of road network expansion and of increased vehicular traffic on disease incidence. Minimizing data missingness, I use the proportion of paved roads as a hypothetical indicator for increasing roadway development and surface transit within countries, and hypothesize increasing paved roadways to be associated with higher disease incidence.

Incoming refugees

UNHCR population of concern

Populations have suspected migrant as harbingers of disease throughout recorded history.[5] Risk is particularly acute among those whose travel is forced, as they are likely subject to crowding, poverty, poor sanitation, insufficient health care, and other conditions of marginalization prior to leaving their homes. Such conditions may be perpetuated within countries of arrival, where migrants may encounter infectious agents to which they have heightened susceptibility in the absence of prior exposure.[6] In addition to personal risk for infection, migrants are at risk of transmitting infections to domestic populations in their countries of arrival. This risk of cross-border microbial trafficking is greatest in developing and unstable countries lacking the political and clinical infrastructures to control the entrance of sick individuals during mass population movements.[7,8]

1.3: Population demographics

Urban population

Rate of urbanization

Changes to human behavior and living conditions have accompanied rapid urbanization in recent decades.[9] The urban growth rate exceeds the pace of infrastructural development, of institutional

service expansion, and of economic development in many low-source settings;[10] populations migrating into urban slums across the global South encounter high density, poverty, poor sanitation, and have limited resources to control disease transmission.[11] I investigate how changes in the share of the population living in urban areas, and the rate at which this share changes, drive the circulation of infections within populations at large.

Gender inequality index

The effects of gender inequality on infectious disease emergence are perhaps most clearly portrayed in the case of HIV and other STIs, where women's diminished agency within society and within relationships constrains control over their and their partner's risk for STIs.[12-14] Broader effects persist at the population level and with regard to diseases that are not sexually transmitted. Complex causal pathways connect social and material deprivation to physiological stress and to risk-of-infection for an individual woman.[15] Mothers without individual material, economic, and social independence project this lack of agency, and its adverse effects, upon their children and other dependents;[16] in gender-inequitable societies, women's poor access to healthcare, transportation, and environmental or occupational safety constrains the ease of fulfilling social obligations such as childbearing and childcare.[17] The rectification of gender inequality and maternal disadvantage may reduce infectious disease burden, as children account for the majority of disease incidence and disability-adjusted life years lost to infection.[18]

Education index

Population-wide education status is a well-documented determinant of health, including incident chronic and infectious disease.[19] Primary schooling may entail direct health education, while secondary and post-secondary schooling may train individuals for health professions so as to strengthen clinical infrastructures. More broadly, educational attainment serves as a population demographic for intra- and international comparisons.[20] Qualitative literatures cite higher prevalence of emerging and re-emerging infections within populations of low educational attainment.[9-11,21,22]

The index I employ captures expected years of schooling for children as well as mean years of schooling for adults to indicate cross-generational educational status, thus including the lagged

effects of a country's failure to educate its aging population as well as the real-time effects of new or eroding educational opportunities for young generations. While the relationship between education and disease incidence may be complicated by within-school vaccination campaigns and by transmission of infections among children at school,[23-25] covariance with other development indicators suggests low education index scores will be associated with the incidence of zoonotic and anthroponotic diseases.

Poverty

Standardized GINI coefficient

Poorer nations lack surveillance and control resources for identifying and curtailing infectious disease emergence and re-emergence, and suffer a disproportionate proportion of global infectious disease burden.[4,9-11,21,22,26,27] The coincidence of material poverty with infection occurs at international and intranational geographic scales.[28] I consider as determinants for infection both national income averages, expressed as annual GDP per capita, and within-country socioeconomic inequality. Inequality has been demonstrated to exhibit a separate effect upon disease transmission[29] beyond what is attributable to poverty and associated lifestyle risks.[30,31] The association may even involve feedback at broad timescales; pathogen prevalence has been theorized to select for individualist socioeconomic systems over many generations.[32]

Adult HIV prevalence

Malnutrition prevalence

Malnutrition and HIV-induced immune suppression are critical drivers of susceptibility for infection at the individual level.[33] I seek to understand the degree to which these individually-relevant factors, when widely prevalent, amplify disease transmission within whole populations. Taking vaccinations as an example of a marginal individual-level factor with pronounced significance at population levels,[34] I expect malnutrition and HIV prevalence to be positively associated with the incidence of zoonotic and anthroponotic diseases.

Rate of population growth

Population growth encompasses several mechanisms potentially driving disease incidence within populations. Under dominant paradigms, high rates of infection are expected within societies

with high birth rates.[27] While global infectious disease emergence challenges this paradigm,[35] high rates of infection should nonetheless persist within disproportionately young, fast-growing populations of the developing world because children experience the greatest burden of infectious disease.[18] Population growth may likewise be attributable to in-migration. Mobile populations are at risk of trafficking microbes into their destination country, or of falling victim to diseases endemic to their destination country upon arrival as described previously.

UNHCR population of concern

The discussion of disease risk within refugee-asylum nations applies again to nations from which refugees are sourced; using the UNHCR population of concern as a metric for persons displaced by political terror but remaining within their home country, I investigate whether the compromised physical, social, and economic well-being of such populations drives infectious disease transmission within countries-at-large by putative mechanisms previously discussed in the case of forced migration.

Political Terror Scale

Degeneration of good governance may propagate forced population movements in extreme circumstances, but may first disrupt social services, public health infrastructures, and other public institutions at lower levels of ill political intent.[36-38] I incorporate Gibney and Dalton's Political Terror Scale,[39] an ordinal measure of the intensity and scope of human rights violations perpetrated by national governments. Berrang-Ford et al. (2011)[40] demonstrate the utility of this index at capturing excess disease burden owing to political violence and instability in cases where two-sided intra-national or international conflict does not occur.

1.4: Infrastructure

Improved water

Improved sanitation

Improvements in water and sanitation infrastructures have been integral to the global North's near-eradication of waterborne and feco-oral pandemics, and remain a priority for global health and development agencies.[9,11,21,22,41-43] Physical infrastructure is likely to impact diseases of

other transmission routes as well. Quality of water and waste infrastructures have been implicated in the etiology of vector-borne diseases, of zoonoses, of foodborne diseases, and of non-feco-oral water-washed diseases such as skin infections and eye infections.[44] Infrastructural improvements associated with urbanization and development aided malaria eradication from Canada and the US.[45] Quantitative forecasts for malaria re-emergence remain sensitive to variables indicating countries' ability to control disease through improved physical works.[46] The potential for collinearity between water infrastructure variables, generalized poverty, and domicile construction quality could allow this analysis to additionally implicate water- and sanitation covariates in the incidence of diseases that are not environmentally-borne. Indeed, transmission of respiratory disease is in part governed by poor indoor air quality attributable to the design and maintenance of built environments; here, water and sanitation infrastructures are likely to proxy overall construction quality and to appear as determinants for diseases whose transmission mechanisms are not otherwise water-related.[47,48]

Agricultural mechanization

The transmission of zoonotic pathogens between humans and domesticated animals has long been a factor in the emergence of novel infections,[4] and remains important to foodborne, waterborne, and vector-borne outbreaks as well as influenza strain evolution.[49] Behavioral risk factors associated with the acquisition of these and other zoonoses place primary-sector employees at higher risk of infection in settings where agricultural systems remain underdeveloped, as evidenced by a lack of technological capital.[50-52]

1.5: Geography and environment

Forested area

Agricultural area

Direct and synergistic effects of land-use change on disease ecologies threaten to exacerbate the emergence of zoonotic and anthroponotic infections. Manipulations of vector- and host habitats, and corresponding changes in human-animal contact opportunities, dictate that perturbations of landscape ecology may allow increases in zoonotic disease risk.[4,9,11,21,22,53,54] Mass rural-to-urban migrations bringing about deforestation and the abandonment of agricultural lands align land

surface change with tangential urbanization, putting populations at risk for socio-economic disruptions tied to the emergence of diseases. Among these disruptions are increased malnutrition, and the clustering of poverty within underdeveloped megacities.[54] Specific trajectories of ecological disruption requisite to the emergence of infectious diseases vary by disease and by setting; for instance, forested land areas are expected to be positively associated with Lyme disease risk in temperate North America,[55] but tropical deforestation is associated with increasing incidence of malaria, leishmaniases, and various arboviruses.[56]

Island country

Landlocked country

Island ecosystems, including disease ecosystems, are small and self-contained relative to their counterparts on continental mainlands. Discontinuity in vector and host habitats has been proposed to protect islands from emerging vector-borne diseases prevalent in nearby continental areas.[57] Environmental disease interventions including host species management have additionally been more effective in island environments than other environments;[58-61] ecological and geographic containment facilitate control of host species in the case of zoonoses. The implications of island status for other diseases are less easy to infer, but island status and associated remoteness from economic partners suppresses economic development and may thus relate to the social gradient of health and disease.[62] The eco-environmental effects of landlocked status on disease risk are less clear. However, landlocked nations consistently display lower economic productivity than their counterparts,[62,63] and depend upon neighbors' transit systems, infrastructure, administrative practices, and political stability for global integration.[64] Given this disadvantage and the general gradient of population health with respect to resource access, one may expect landlocked status to predict higher incidence of zoonotic and anthroponotic diseases.

1.6: Weather and climate

Rainfall during month of maximum rainfall

Z-score

Variable precipitation accounts for heterogeneity among global biomes at like latitudes and elevations. Seasonal precipitation cycles[65] drive the observed latitudinal distribution of pathogen

richness, favoring the proliferation of disease species in tropical regions with starkly-contrasting wetness and dryness periods throughout the year. In addition to altering population ecologies of host species, high-precipitation events overwhelm water and sanitation infrastructures to compromise microbial containment in urban environments.[69] Although global studies have not evaluated the year-by-year effects of extreme precipitation events on disease incidence, various articles tie such events to waterborne,[66] rodent-borne,[67] and vector-borne[68] diseases outbreaks at smaller geographic scales. Extreme precipitation events may additionally cause adverse economic consequences, harming crop growth, ease of travel, and secondary-sector production.[69,70]

Temperature during month of maximum temperature
Z-score

The impacts of secular warming trends upon infectious disease incidence are an area of active research. Increased risk to vector-borne and zoonotic diseases are readily understood in terms of the impacts of seasonal temperatures upon host/vector geography, reproductive rates, and parasite incubation periods.[71] Further, novel zoonotic pathogens are apt to evolve within warm tropical latitudes amidst environmental disruption, as may be expected under a global warming scenario.[65,72-74] Additional risks apply in the case of waterborne diseases susceptible to increasing sea surface temperatures.[75] Human risk is not limited to infection by pathogens whose transmission is ecologically- or environmentally-mediated; there is potential for malnutrition stresses to increase wholescale pathogen susceptibility as agriculture becomes unreliable in rapidly-warming regions. Forced economic migration, for instance from sub-economic farmland to urban areas, may additionally expose populations to the risk factors associated with refugeeism and urbanization.[76,77]

2: Diseases

2.1: Overview

I proceed here to list diseases selected for their global or near-global geographic distribution and notification. The list encompasses zoonotic- and anthroponotic diseases transmitted via representatively-broad means of conveyance (i.e. respiratory, fecal-oral, vector-borne, sexual,

and water-borne). The categorization adheres to the scheme established by Wolfe et al. (2007).[72] Where those authors do not establish a numerical score for a given disease, I infer one according to their methods and provide supporting evidence.

2.2: Zoonotic diseases

Cholera

Wolfe et al. (2007)[72] designate cholera a zoonotic disease because small aquatic animals including crustaceans, amoeba, and zooplankton harbor the bacterium *Vibrio cholerae*; in the terrestrial sphere, birds and herbivores may harbor the disease as well. Most transmission occurs among humans via the fecal-oral route. Although the disease is distributed across the globe and is estimated to cause millions of cases annually, it is not equally endemic to all countries; international variation in incidence is explained with fair reliability by the proportion of a country's population with access to improved sanitation.[78] I assess how aberrant weather and changes to biotic, built, and social environments compound this existing risk.

Dengue fever

Dengue viruses originated within nonhuman primates in Southeast Asia or West Africa and have since disseminated to East Africa and Latin America. Most human transmission is isolated from enzootic transmission, and owes to feeding by the peridomestic mosquito *Aedes aegypti*.[78] Feeding by other mosquitoes occupying the forest canopy and gallery forest allows occasional human spillover from persisting sylvatic dengue cycles which involve nonhuman primate reservoirs.[79,80] Experimental assays indicate similar transmissibility of enzootic and human-endemic dengue strains, and enzootic strains have been documented to cause fever and hemorrhagic syndrome in humans just as the anthroponotic disease.[81-85] The clinically indistinct pathophysiology of sylvatic and human-endemic dengue viruses, taken together with the low capacity for molecular biosurveillance in endemic world regions, could allow a major species-crossover event to occur undetected.[85-,87]

Hantavirus

The limited global burden of hantaviruses relative to other infections prevent Wolfe et al. (2007)[72] from considering the viruses in depth. Hantaviruses appear overwhelmingly to be Stage-2 pathogens; the majority of infections causing hemorrhagic fever with renal syndrome, as well as cardiopulmonary syndrome, can be traced to exposure to a rodent source. Limited interpersonal transmission of Andes virus has been documented, but mechanisms for this transmission remain poorly characterized.[88-90] New- and Old-World hantaviruses circulate primarily within disadvantaged rural and urban locales, respectively, and incidence has been tied to meteorological events favoring the reproduction of rodent hosts.[91-93] Identifying global hantaviruses as low-stage zoonotic pathogens with documented ability to evolve toward higher stages, identifying how social and ecological factors coalesce to drive their incidence and endemicity within countries is important to preventing increases in global burden.

Lyme borreliosis

Borreliosis is one of few vector-borne, zoonotic diseases with a high focal burden in developed countries and at northern latitudes. Disease occurs throughout Asia, Europe, the Americas, and Northern Africa in individuals infected by tick-borne spirochetes of the genus *Borrelia*.[94] The causative bacteria have complex life cycles involving dozens of mammalian and avian reservoir species.[95] Vertical transmission is thought not to occur within host and vector species; additionally, humans and other large-mammal hosts are reservoir-incompetent due to the mature life stage of tick vectors feeding upon them; younger ticks acquire spirochetes from smaller hosts, and in turn infect larger hosts after molting.[96] The disease has thus remained a Stage-2 zoonosis,[72] with transmission and range expansion governed by ecological factors affecting vector and reservoir species. Anthropogenic environmental change has been implicated in the recent emergence of *B. burgdorferi* within the northeastern United States.[53,97] Because the burden of borreliosis remains highest in developed countries, the likeness of its drivers to those of other zoonoses and vector-borne diseases endemic to the global South could indicate risk for environmentally-mediated "tropical" diseases in developed, temperate settings.

Leishmaniases

I include cases of visceral, cutaneous, and mucocutaneous leishmaniasis. Phlebotomine sandflies carry various species of *Leishmania* throughout Latin America, Africa, southern Europe, and central and eastern Asia. Autochthonous transmission has been reported within US states, Canadian provinces, and Australia, suggesting that the disease is at risk for becoming globally endemic.[98,99] Small- and large-mammal reservoirs for cutaneous and mucocutaneous leishmania are primarily sylvatic, and occupy peridomestic areas to a lesser extent. In contrast, canine reservoirs contribute to the majority of zoonotic visceral leishmaniasis transmission; jackals, foxes, and other sylvatic reservoirs are less important to human incidence of the visceral disease.[100] Although outbreaks have occurred among humans without zoonotic transmission, the majority of leishmania causing disease in humans result from zoonotic exposure.[101,102] Rapid geographic expansion of autochthonous leishmaniasis has been attributed to climatic, environmental, and demographic change, as well as to syndemic effects from HIV co-infection.[100,102]

Leptospirosis

Leptospirosis is a globally-distributed zoonosis caused by infection with one of several pathogenic species of spirochetes from the genus *Leptospira*. Nearly all mammals are competent hosts for *Leptospira*; hosts of particular importance to human transmission include rodents, cattle, swine, and canines. Virtually no anthroponotic transmission occurs.[103] Leptospires are shed in animal urine and may persist within soil and aquatic environments without a host for extended periods of time.[104] The predominance of animal exposure as a risk factor has long dictated an occupational classification for leptospirosis; high rates of infection occur among farmers and others employed in the primary sector within developed- and developing settings.[105] Urban outbreaks within developing countries have acquired increasing attention, and demonstrate a shifting geographic and social distribution of human disease risk.[105,106] Although disease persists throughout the year in humid, low-latitude settings, risk is especially acute within both tropical and temperate locations during wet seasons and extreme weather events.[107] Global climatic, environmental, and demographic change may thus threaten to expand current incidence.

Malaria (all strains)

Because most reported malaria is not subtyped, I consider infection by all species of the genus *Plasmodium*, of which the species *falciparum* and *vivax* are most prevalent globally. Autochthonous malaria incidence is distributed throughout Africa, Latin America, and Asia, and has been eradicated from formerly-endemic temperate regions including Australia and North America.[108] The evolutionary history of *P. falciparum* is contested; although divergence from simian infections is known to have occurred recently, mitochondrial genomes fail to indicate whether hominid reservoirs acquired competence for the avian progenitor thousands or millions of years ago.[109] The time at which *P. vivax* diverged from simian-infecting species likewise remains unclear, although there is reasonable consensus placing the event within the past 100,000 years.[110-112] *P. knowlesi*, endemic to southeast Asia, is a simian malaria species with frequent zoonotic transmission to humans; the difficulty of characterizing *P. knowlesi* infections in the absence of molecular diagnostic assays suggests that the relative importance of this zoonotic infection remains grossly underestimated.[113,114]

Wolfe et al. (2007)[72] rightly characterize falciparum and vivax malarias as well-evolved pathogens with near-exclusive anthroponotic transmission. The discovery of *P. knowlesi* as a zoonotic species causing a high share of Asian malaria[113-115] following publication of the Wolfe et al. (2007)[72] article has been a major scientific event. Human-infecting *P. malariae* within South America likewise circulates enzootically among monkey species.[116] Whereas Wolfe et al. (2007)[72] consider only the highest-burden malarias within their analysis, we seek to represent the full spectrum of cases currently resulting in malaria diagnoses; these include properly-, improperly-, and non-subtyped infections. It thus appears prudent to designate malarias within these analyses as late-stage zoonoses, noting that Wolfe and other authors[117] characterize evolutionary emergence as occurring along a spectrum from zoonotic to anthroponotic stages. Within this scheme, there is more space for overlapping human- and enzootic transmission, as seen among malarias, within the "zoonotic" class (i.e. Stages 2-4) than within the anthroponotic class (Stage 5 only).

Although malaria incidence is now greatest in tropical locales, the disease's former endemicity within the US and Canada, as well as its current endemicity in eastern Russia, indicate its

geographic range may not be constrained by climatic factors. Sanitation, development, and vector control play important roles in malaria risk projections within models accounting for forecasted climate change.[118,119] My analysis thus accounts for the conflation of human and environmental factors driving global variation in malaria incidence.

Plague

Plague (yersiniosis) caused by the bacterium *Yersinia pestis* is found globally in rodents and in flea vectors feeding upon them. Global incidence of plague is low relative to that of other pathogens included within our present study; the focal distribution of outbreaks across space and time justifies our consideration of environmental drivers causing incidence to cluster while overall prevalence remains low.[120,121] Plague is among few neglected tropical diseases simultaneously endemic to low- and middle-income countries and the US. The relative weights of human- and eco-environmental determinants behind plague re-emergence remain unaddressed within the disease's sparse epidemiologic literature.[122]

Rabies

Rabies is a near-globally endemic viral encephalitis absent only from a small number of island and Scandinavian states, where it was either never introduced or has been successfully eradicated. Nucleotide sequencing has identified multiple and heterogeneous causative viruses, loosely grouped into the genus *Lyssavirus*.[123] Humans acquire rabies through exposure to infected saliva when bitten by a rabid animal; globally and especially in developing countries, dog bites account for the majority of human infections.[124] Although infection within endemic developed countries is managed in part by controlling enzootic transmission through the vaccination of wildlife and domestic animals, tens of thousands of potentially-exposed individuals receive prophylaxis annually within the US alone. Such veterinary and medical interventions are not possible in poorer nations; thus the disease's 2 million DALYs and economic losses from livestock death are felt most acutely in resource-poor settings.[125] Rabies viruses replicate with poor fidelity, and can thus diminish progeny fitness within the span of a few generations; this evolutionary circumstance gives rise to stochastic variation in disease prevalence within microenvironments, and may obscure broader environmental determinants of disease transmission across complex physical and ecological landscapes.[126]

Non-typhoid Salmonellosis

Nontyphoidal salmonella are primarily-foodborne bacterial species which additionally may be transmitted directly among humans and animals, or within contaminated water and in iatrogenic situations via the fecal-oral route.[127] Although many animals are reservoir-competent, the majority of humans infections owe to consumption of, or other contact with, infected eggs; avian hosts are subject to vertical, trans-ovarian transmission.[128,129] Emerging resistance to broad-spectrum cephalosporins within industrial animal husbandry operations has complicated efforts to control disease in developed and developing settings alike.[130] The increasing presence of salmonella in runoff from agricultural land, as well as in urban and peridomestic settings, indicates heightened human risk during wet seasons and extreme weather events.[131-134] I herein aim to identify how environmental and demographic changes coalesce with such meteorological events to amplify human risk in individual years.

2.3: Anthroponotic diseases

Gonorrhea

Syphilis

I choose to evaluate gonorrhea and syphilis rather than HIV/AIDS because poor international surveillance for the latter disease during its emergent phases in the early 1990s makes analysis from 1990-2010 difficult to validate.[135] Existence of the diseases presented here was not politically contentious during the period of analysis; *Neisseria gonorrhea* and *Treponema pallidum* are sexually-transmitted bacteria causing gonorrhea and syphilis, respectively, and achieved global endemicity centuries ago. In spite of declining STI incidence at the dawn of the HIV/AIDS epidemic, national and international control efforts for gonorrhea and syphilis have been broadly ineffective in recent decades.[136-138] Molecular- and case-surveillance programs indicate rapid dissemination of novel strains among high-risk groups in developed and developing nations alike.[139,140] The emergence and global diaspora of antimicrobial-resistant strains challenges future abilities to control the disease even in resource-rich settings.[141] Increasing prevalence and treatment failures have ascribed a high economic cost to managing the two diseases; treatment is particularly important for young mothers, as vertical transmission can

occur for both diseases. Particular interest in reducing gonorrhea incidence owes in part to the disease's role in facilitating acquisition and transmission of HIV.[142]

Economic development underlying stages within the human demographic transition in part governs the ecology of these diseases; changing sexual behaviors, explosive population growth, and shifting population distributions within the urbanizing, developing world have been identified as drivers of disease incidence.[143,144] Stigma and illegality associated with high-risk behaviors have additionally complicated surveillance within low-resource settings.[146,147] I situate the diseases within a broad context of social, demographic, and environmental change to pinpoint how factors affecting the arrangement of human life and activity may underlie epidemic STI transmission.

Measles

Pertussis

Global resurgence in the incidence of measles and pertussis has occurred in spite of vaccine availability, and in spite of high vaccine coverage in wealthy countries.[147,148] Both diseases are among the most-transmissible human infections, having evolved from ancestral pathogens affecting domestic animals.[72] Low or decreasing vaccine coverage rates have received extensive attention as determinants for infection; molecular techniques have additionally identified vaccine-driven strain evolution and genotype replacement as drivers of outbreaks within high vaccine-coverage settings.[149-151] Although historic literatures indicate critical community sizes and densities for the effective transmission of measles and pertussis, the impact of such environmental factors remains poorly quantified in present empirical literature.[152,153] Controlling for vaccination to isolate excess disease generally attributable to population dynamics and demographics, I seek to implicate destabilizing social and environmental conditions in the re-emergence and continued incidence of these diseases.

Shigellosis

Shigellosis is caused by enteric bacteria and transmitted via the feco-oral route through water, food, or direct contact. Although various Shigella bacterial species circulate enzootically among primates and have potential to spill over into human populations, Wolfe et al. (2007)[72] classify S.

sonnei, the cause of most infections in developed countries, as a Stage-5 pathogen. S. flexneri causes a higher proportion of cases in developing countries, and certain serotypes retain the potential to infect nonhuman primates; however, documented cases of S. flexneri in such animals have been attributed to human-animal transmission, and nonhuman primates lack reservoir competence for shigellae circulating among humans.[154-156] Shigellosis is one of few Stage-5 enteric pathogens, and additionally one of few Stage-5 pathogens with the potential for environmentally-mediated transmission.[72] Because risk factors and transmission mechanisms for shigellosis are similar to those for zoonotic enteric infections, including shigellosis within the analysis helps to ensure that varying transmission mechanisms for zoonotic and anthroponotic diseases are not alone in accounting for differences among identified environmental drivers.[157]

Tuberculosis

Active tubercle bacillus (TB) is a life-threatening condition affecting the lungs and other organs among select individuals infected by *Mycobacterium tuberculosis*; the infection is thought to be prevalent among one third of the global population.[158] Resurgent incidence of symptomatic illness during the late twentieth century is attributed to widespread immune suppression amidst the HIV epidemic.[159] National and international responses have been mounted to control TB such that the timespan under analysis includes periods of both rising and falling global incidence.[160] Pathogen evolution selecting for vaccine- and antimicrobial-resistant strains has made clinical intervention only partially effective; however, other biological, social and environmental determinants such as indoor air quality, host malnutrition, poor living conditions, and macroeconomic trends are significant predictors of disease within populations, indicating that disease eradication requires management outside the clinical sphere.[161] Controlling for the effects of vaccination, I seek to understand how ongoing social and environmental processes contribute to variation in the incidence of TB.

Typhoid fever

Typhoid fever is caused by a bacterium transmitted via the fecal-oral route; the pathophysiological response is systemic, but includes severe diarrhea that facilitates transmission in settings where sanitation is poor. Although case mortality is estimated to be only 1%, typhoid contributes greatly to global disease burden because of its high prevalence within developing

countries.[162] I limit case definition to those cases specified as *Salmonella enterica typhi*, which infects only humans; the closely-related bacterium *S. paratyphi* causes paratyphoid, but infects other mammals and birds and may thus be acquired via zoonotic transmission.[163] While typhoid was once endemic to the globe, incidence and burden decreased during the twentieth century in developed countries due to improved sanitation practices and infrastructures, as well as the pasteurization of milk and chlorination of municipal drinking waters.[164] By evaluating year-by-year data across many nations, I seek to identify how typhoid incidence remains sensitive to these and other environmental factors.

3: References

1. Mangili A, Gendreau MA. Transmission of infectious diseases during commercial air travel. Lancet. 2005 Mar 12;365(9463):989-96.
2. Tatem AJ, Rogers DG, Hay SI. Global transport networks and infectious disease spread. Adv Parasitol. 2006;62:293-343.
3. Eisenberg JNS, Cevallos W, Ponce K, Levy K, Bates SJ, Scott JC, et al. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. PNAS. 2006 Dec 19;103(51):19460-5.
4. McMichael AJ. Environmental and social influences on emerging infectious diseases: past, present and future. Phil Trans R Soc Lond B. 2004;359:1049-58.
5. Barnett ED, Walker PF. Role of immigrants and migrants in emerging infectious diseases. Med Clin N Am. 2008;92:1447-58.
6. Stauffer WM, Weinberg M. Emerging clinical issues in refugees. Curr Opin Infect Dis. 2009;22(5):436-42.
7. Barnett ED. Infectious disease screening for refugees resettled in the United States. Clin Infect Dis. 2004;39(6):833-41.
8. Prothero RM. Forced movements of population and health hazards in tropical Africa. Int J Epid. 1994;23(4):657-64.
9. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. Nature. 2004;430:242-9.
10. Binder S, Levitt AM, Sacks JJ, Hughes JM. Emerging infectious diseases: Public health issues for the 21st Century. Science. 1999;284(5418):1311-3.
11. Weiss RA, McMichael AJ. Social and environmental risk factors in the emergence of infectious diseases. Nature Med. 2004;10:S70-6.
12. Farmer P, Lindenbaum S, Good MJ. Women, poverty, and AIDS: an introduction. Cult Med Psychiatry. 1993;17(4):387-97.
13. Dunkle K, Jewkes RK, Brown HC, Gray GE, McIntyre JA, Harlow SD. Gender-based violence, relationship power, and risk of HIV infection in women attending antenatal clinics in South Africa. Lancet. 2004;363(9419):1415-21.
14. Zhihong S, Larsen U. Gender inequality increases women's risk of HIV infection in Moshi, Tanzania. J Biosoc Sci. 2008;40(4):505.
15. Osmani S, Sen A. The hidden penalties of gender inequality: Fetal origins of ill-health. Econ Human Biol. 2003;1(1):105-21.
16. Robert B, Jackson M, Crist A. Gender inequality, vulnerability, and disaster: Issues in theory and research. In: The gendered terrain of disaster: Through women's eyes. 1998: 27-44.
17. Okojie CEE. Gender inequalities of health in the third world. Soc Sci Med. 1994;34(9):1237-47.
18. Mathers C, Fat DM, Boerma JT. The global burden of disease: 2004 update. World Health Organization. 2008.
19. Kickbusch I. Health literacy: addressing the health and education divide. Health Promot Int. 2001;15(3):289-97.
20. Wilson C. On the scale of global demographic convergence 1950-2000. Pop Develop Rev. 2001;27(1):155-71.
21. Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis. 1995;1(1):7-15.

22. Morens DM, Folkers GK, Fauci AS. Emerging infections: a perpetual challenge. Lancet Infect Dis. 2008;8:710-9.
23. Hoge CW, Fisher L, Donnell Jr. HD, Dodson DR, Tomlinson Jr. GV, Breiman RF, et al. Risk factors for transmission of Mycobacterium tuberculosis in a primary school outbreak: Lack of racial difference in susceptibility to infection. Am j Epidemiol. 1994;139(5):520-30.
24. Harley D, Harrower B, Lyon M, Dick A. A primary school outbreak of pharyngoconjunctival fever caused by adenovirus type 3. Tuberculosis notifications in Australia. 2001;1998(25).
25. Marks PJ, Vipond IB, Regan FM, Wedgwood K, Fey RE, Caul EO. A school outbreak of Norwalk-like virus: Evidence for airborne transmission. Epidemiol Infect. 2003;131(1):727-736.
26. Heymann DL, Rodier GR, WHO Operational Support Team to the Global Outbreak Alert and Response Network. Lancet Infect Dis. 2001;1(5):345-53.
27. Omran AR. The epidemiologic transition: A theory of the epidemiology of population change. Milbank Mem Fund Quart. 1971;49(4):509-38.
28. Chen LC, Berlinguer G. Health equity in a globalizing world. In: Evans T et al. (eds). Challenging inequities in health: from ethics to action. Oxford: Oxford Univ Press. 2001:34-44.
29. Farmer P. Social inequalities and emerging infectious diseases. Emerg Infect Dis. 1996;2(4):259-69.
30. Kennedy BP, Kawachi I, Prothrow-Stith D. Income distribution and mortality: Cross sectional ecological study of the Robin Hood index in the United States. BMJ. 1996;312(7037):1004-7.
31. Drain PK, Smith JS, Hughes JP, Halperin DT, Holmes K. Correlates of national HIV seroprevalence: An ecologic analysis of 122 developing countries. JAIDS. 2004;35(4):407-20.
32. Fincher CL, Thornhill R, Murray DR, Schaller M. Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. Proc R Soc B. 2008. 275(1640):1279-85.
33. Schaible UE, Stefan HE. Malnutrition and infection: complex mechanisms and global impacts. PLoS Medicine. 2007;4(5):e115.
34. Anderson RM, May RM. Vaccination and herd immunity to infectious diseases. Nature. 1985;318:323-9.
35. Harper K, Armelagos G. The changing disease-scape in the third epidemiological transition. Int J Environ Res Pub Health. 2010:7(2):675-97.
36. Gisselquist D. Impact of long-term civil disorders and wars on the trajectory of HIV epidemics in sub-Saharan Africa. J Soc Aspects HIV/AIDS. 2004;1:114-27.
37. Hoddie M, Smith J. Forms of civil war violence and their consequences for future public health. Int Stud Quarterly. 2009;53:175-202.
38. Berrang-Ford L. Civil conflict and sleeping sickness in Africa in general and Uganda in particular. Conflict and Health. 2007;1:6.
39. Gibney M, Dalton M. The political terror scale. Policy Studies and Developing Nations. 1996;4:73-84.
40. Berrang-Ford L, Lundine J, Breau S. Conflict and human African trypanosomiasis. Soc Sci Med. 2011;72(3):398-407.

41. Dowdle WR. The principles of disease elimination and eradication. Bull World Health Organ. 1998;76(S2):22-5.
42. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. Lancet Infect Dis. 2005;5(1):42-52.
43. Eisenberg JNS, Desai MA, Levy K, Bates SJ, Liang S, Naumoff K, et al. Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. Environ Health Perspect. 2007;115(8):1216-23.
44. Mara D, Feachem R. Water and excreta-related diseases: Unitary environmental classification. J Environ Eng. 1999;125(4):334-9.
45. Berrang-Ford L, MacLean JD, Gyorkos TW, Ford JD, Ogden NH. Climate change and malaria in Canada: a systems approach. Interdisc Perspect Infect Dis. 2009:1-13. doi:10.1155/2009/385487
46. Martens WMJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ. Potential impact of global climate change on malaria risk. Environ Health Perspect. 1995;103(5):458-64.
47. Burge HA. Indoor air and infectious disease. Occup Med. 1989;4(4):713-21.
48. Hoge CW, Reichler MR, Dominguez EA, Bremer JC, Mastro TD, Hendricks KA, et al. An epidemic of pneumococcal disease in an overcrowded, inadequately ventilated jail. 1994. N Engl J Med;331:643-8.
49. Christou L. The global burden of bacterial and viral zoonotic infections. Clin Microbiol Infect. 2011;17:326-30.
50. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, Lovett MA, et al. Leptospirosis: a zoonotic disease of global importance. Lancet Infect Dis. 2003;3(12):757-71.
51. Zeitz PS, Butler JC, Cheek JE, Samuel MC, Childs JE, Shands LA, et al. A case-control study of hantavirus pulmonary syndrome during an outbreak in the southwestern United States. J Infect Dis. 1995;171(4):864-70.
52. Armien B, Pascale JM, Bayard V, Munoz C, Mosca I, Guerrero G, et al. High seroprevalence of hantavirus infection on the Azuero peninsula of Panama. Am J Trop Med Hyg. 2004;70(6):682-7.
53. Patz JA, Daszak P, Tabor GM, Alonso Aguirre A, Pearl M, Epstein J, et al. Unhealthy landscapes: Policy recommendations on land use change and infectious disease emergence. Environ Health Perspect. 112(10):1092-8.
54. Patz JA, Olson SH, Uejio CK, Gibbs HK. Disease emergence from global climate and land use change. Med Clin N Am. 2008;92(6):1473-91.
55. Barbour AG, Fish D. The biological and social phenomenon of Lyme disease. Science. 1993;260:1610.
56. Walsh JF, Molyneux DH, Birley MH. Deforestation: effects on vector-borne disease. Parasitology. 1993;106(S1):S55-75.
57. Githeko AK, Lindsay SW, Confalonieri UE, Patz JA. Climate change and vector-borne diseases: a regional analysis. Bull World Health Organ. 2000;78(9):1136-47.
58. Rand PW, Lubelczyk C, Holman MS, Lacombe EH, Smith RP. Abundance of Ixodes scapularis (Acari: Ixodidae) after the complete removal of deer from an isolated offshore island, endemic for Lyme disease. J Med Entomol. 2004;41(4):779-84.
59. Crump JA, Murdoch DR, Baker MG. Emerging infectious diseases in an island ecosystem: the New Zealand perspective. Emerg Infect Dis. 2001;7(5):767-72.

60. Kaneko A, Taleo G, Kalkoa M, Yamar S, Kobayakawa T, Bjorkman A. Malaria eradication on islands. Lancet. 2000;356(9241):1560-4.
61. Wilson N, Lush D, Bell P. Lessons for zoonotic disease and vector eradication in New Zealand from past successes. J Epidemiol Community Health. 2011;65:A169-70.
62. Carr SC, McWha I, MacLachlan M, Furnham A. International-local remuneration differences across six countries: do they undermine poverty reduction work? Int J Psychol. 2010;45(5):321-40.
63. Chowdhury AK. Geography against development: a case for landlocked developing countries. 2006. United Nations: New York.
64. Faye ML, McArthur JW, Sachs JD, Snow T. The challenges facing landlocked developing countries. J Human Develop. 2004;5(1):31-68.
65. Guernier V, Hochberg ME, Guegan JF. Ecology drives the worldwide distribution of human diseases. PLOS Biol. 2004;2(6):e141.
66. Curriero FC, Patz JA, Rose JB, Lele S. The association between extreme precipitation and waterborne disease outbreaks. AJPH. 2001;91(8):1194-9.
67. Engelthaler DM, Mosley DG, Cheek JE, Levy CE, Komatsu KK, Ettestad P, et al. Climatic and environmental patterns associated with Hantavirus pulmonary syndrome, Four Corners Region, United States. Emerg Infect Dis. 1999;5(1):87-94.
68. Hales S, Weinstein P, Souares Y, Woodward A. El Nino and the dynamics of vectorborne disease transmission. Environ Health Perspect. 1999;107(2):99-102.
69. Ted Velhuis JA, Clemens FH, Sterk G, Berends BR. Microbial risks associated with exposure to pathogens in contaminated urban flood water. Water Res. 2010;44(9):2910-8.
70. Kunkel KE, Pielke RA, Changnon SA. Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: a review. Bull Am Meteorol Soc. 1999;80(6):1077-98.
71. Tabachnick WJ. Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. J Exp Biol. 2010;213:946-54.
72. Wolfe ND, Panosian Dunavan C, Diamond J. Origins of major human infectious diseases. Nature. 2007;447:279-83.
73. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature. 2008;451:990-3.
74. Dunn RR, Davies TJ, Harris NC, Gavin MC. Global drivers of human pathogen richness and prevalence. Proc R Soc B. 2010;277(1694):2587-95.
75. Semenza JC, Herbst S, Rechenburg A, Suk JE, Hoser C, Schreiber C, et al. Climate change impact assessment of food- and waterborne diseases. Crit Rev Environ Sci Tech. 2012;42(8):857-90.
76. Patz JA, Epstein PR, Burke TA, Balbus JM. Global climate change and emerging infectious diseases. JAMA. 1996;275(3):217-23.
77. McMichael C, Barnett J, McMichael AJ. An ill wind? Climate change, migration, and health. Environ Health Perspect. 2012;120(5):646-54.
78. Holmes EC, Twiddy SS. The origin, emergence, and evolutionary genetics of dengue virus.
79. Vasilakis N, Cardosa J, Hanley KA, Holmes EC, Weaver SC. Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. Nat Rev Microbiol. 2011;9(7):532-41.

80. Althouse BM, Lessler J, Sall AA, Diallo M, Hanley KA, Watts DM, et al. Synchrony of sylvatic dengue isolations: a multi-host, multi-vector SIR model of dengue virus transmission in Senegal. PLoS Negl Trop Dis. 2012;6(11):e1928.
81. Vasilakis N, Tesh RB, Weaver SC. Sylvatic dengue virus type 2 activity in humans, Nigeria, 1966. Emerg Infect Dis. 2008;14:502-4.
82. Franco L, Palacios G, Martinez JA, Vazquez A, Savji N, De Ory F, et al. First report of sylvatic DENV-2-associated dengue hemorrhagic fever in West Africa. PLoS Negl Trop Dis. 2011;5(8):e1251.
83. Saluzzo JF, Cornet M, Adam C, Eyraud M, Digoutte JP. Dengue 2 in eastern Senegal: serologic survey in simian and human populations 1974-85. Bull Soc Pathol Exot Filiales. 1986;79:313-22.
84. Zeller HG, Trarore-Lamizana M, Monlun E, Hervy JP, Mondo M, Digoutte JP. Dengue-2 virus isolation from humans during an epizootic in southeastern Senegal in November, 1990. Res Virol. 192;143(2):101-2.
85. Cardosa J, Ooi MH, Tio PH, Perera D, Holmes EC, Abdul Manap Z. Dengue virus serotype 2 from a sylvatic lineage isolated from a patient with dengue hemorrhagic fever. PLoS Negl Trop Dis. 2009;3(4):e423.
86. Wilder-Smith A, Renhorn KE, Tissera H, Abu Bakar S, Alphey L, Kittayapong P, et al. DengueTools: innovative tools and strategies for the surveillance and control of dengue. Glob Health Action. 2012;510.3402/gha.v5i0.17273.
87. Domingo C, Niedrig M, Teichmann A, Kaiser M, Rumer L. 2nd international external quality control assessment for the molecular diagnosis of dengue infections. PLoS Negl Trop Dis. 2010;4(10):e833.
88. Padula PJ, Edelstein A, Miguel SDL, Lopez NM, Rossi CM, Rabinovich RD. Hantavirus pulmonary syndrome outbreak in Argentina: Molecular evidence for person-to-person transmission of Andes virus. Virology. 1998;2(15):323-30.
89. Wells RM, Sosa Estani S, Yadon ZE, Enria D, Padula P, Pini N, et al. An unusual hantavirus outbreak in southern Argentina: person-to-person transmission? Hantavirus pulmonary syndrome study group for Patagonia. Emerg Infect Dis. 1997;3(2):171-4.
90. Martinez VP, Bellomo C, San Juan J, Pinna D, Forlenza R, Elder M, et al. Person-to-person transmission of Andes virus. Emerg Infect Dis. 2005;11(12):1848-53.
91. Jonsson CB, Figueiredo LT, Vapalahti O. A global perspective on hantavirus ecology, epidemiology, and disease. Clin Microbiol Rev. 2010;23(2):412-41.
92. Engelthaler DM, Mosley DG, Cheek JE, Levy CE, Komatsu KK, Ettestad P, et al. Climatic and environmental patters associated with hantavirus pulmonary syndrome, Four Corners region, United States. Emerg Infect Dis. 1999;5(1):87-94.
93. Hjelle B, Glass GE. Outbreak of hantavirus infection in the Four Corners region of the United States in the wake of the 1997-1998 El Nino-Southern Oscillation. J Infect Dis. 2000;181(5):1569-73.
94. Berger S. Lyme disease: global status 2012 edition. Gideon Informatics. 2010.
95. Gray JS. Review: The ecology of ticks transmitting Lyme borreliosis. Exp Applied Acarol. 1998;22(5):249-58.
96. Kurtenbach K, Hanincova K, Tsao JI, Margos G, Fish D, Ogden NH. Fundamental processes in the evolutionary ecology of Lyme borreliosis. Nature Rev Microbiol. 2006;4:660-9.
97. Schmidt KA, Ostfeld RS. Biodiversity and the dilution effect in disease ecology. Ecol. 2001;82:609-19.

98. Enserink M. Infectious diseases: Has leishmaniasis become endemic in the US?. Science. 2000;290:1881-3.
99. Rose K, Curtis J, Baldwin T, Mathis A, Kumar B, Sakthianandeswaren A, et al. Cutaneous leishmaniasis in red kangaroos: isolation and characterization of the causative organisms. Int J Parasitol. 2004;34:655-64.
100. Dejseux P. The increase in risk factors for leishmaniasis worldwide. Trans R Soc Trop Med Hyg. 2001;95:239-43.
101. Gramiccia M, Gradoni L. The current status of zoonotic leishmaniases and approaches to disease control. Int J Parasitol. 2005;35(11):1169-80.
102. Desjeux P. Leishmaniasis: current situation and new perspectives. Comparative Immunol Microbiol Infect Dis. 2004;27(5):305-18.
103. Adler B, de la Pena Moctezuma A. Leptospira and leptospirosis. Vet Microbiol. 2010;140(3):287-96.
104. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, Lovett MA, et al. Leptospirosis: a zoonotic disease of global importance. Lancet Infect Dis. 2003;3(12):757-71.
105. Waitkins SA. Leptospirosis as an occupational disease. British J Indust Med. 1986;43(11):721-5.
106. Ko AI, Galvao Reis M, Riebeira Dourado CM, Johnson WD, Riley LW, et al. Urban epidemic of severe leptospirosis in Brazil. Lancet. 1999;354(9181):820-5.
107. Lau CL, Smythe LD, Craig SB, Weinstein P. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? Trans R Soc Trop Med Hyg. 2010;104(10):631-8.
108. Mendis K, Rietveld A, Warsame M, Bosman A, Greenwood B, Wernsdorfer WH. From malaria control to eradication: the WHO perspective. Trop Med Internat Health. 2009;14(7):802-9.
109. Conway DJ. Tracing the dawn of *Plasmodium falciparum* with mitochondrial genome sequences. Trends Genet. 2003;19:671-4.
110. Escalante AA, Cornejo OE, Freeland DE, Poe AC, Durrego E, Collins WE, et al. A monkey's tale: the origin of *Plasmodium vivax* as a human malaria parasite. Proc Nat Acad Sci USA. 2005;102(6):180-5.
111. Carter R. Speculations on the origins of *Plasmodium vivax* malaria. Trends Parasitol. 2003;19(5):214-9.
112. Cornejo OE, Escalante AA. The origin and age of *Plasmodium vivax.* Trends Parasitol. 2006;22(12):558-63.
113. Cox-Singh J, Singh B. Knowlesi malaria: newly emergent and of public health importance? Trends Parasitol. 2008;24(9):406-10.
114. Kawai S, Hirai M, Haruki K, Tanabe K, Chigusa Y. Cross-reactivity in rapid diagnostic tests between human malaria and zoonotic simian malaria parasite *Plasmodium knowlesi* infections. Parasitol Internat. 2009;58(3):300-2.
115. Cox-Sing J, Davis TME, Lee KS, Shamsul SSG, Matusop A, Ratnam S. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. Clin Infect Dis. 2008;46(2):165-71.
116. Fandeur T, Volney B, Peneau C, de Thoisy B. Monkeys of the rainforest in French Guiana are natural reservoirs for *P. brasilianum/P. malariae* malaria. Parasitol. 2000;120:11-21.

117. Conway DJ, Roper C. Micro-evolution and emergence of pathogens. Int J Parasitol. 2000;30(12):1423-30.
118. Rogers DJ, Randolph SE. The global spread of malaria in a future, warmer world. Science. 2000;289:1763-6.
119. Martens WJM, Niessen LW, Rotmans J, Jetten TH, McMichael AJ. Potential impact of global climate change on malaria risk. Environ Health Persp. 1995;103(5):458-64.
120. Neerinckx SB, Peterson AT, Gulinck H, Deckers J, Leirs H. Geographic distribution and ecological niche of plague in sub-Saharan Africa. Internat J Heatlh Geographics. 2008;7(54).
121. Duplantier JM, Duchemin JB, Chanteau S, Carniel E. From the recent lessons of the Malagsy foci towards a global understanding of the factors involved in plague re-emergence. Vet Res. 2005;36:437-53.
122. Perry RD, Fetherstron JD. Yersinia pestis—etiologic agent of plague. Clin Microbiol Rev. 1997;10(1):35-66.
123. Rupprecht CE, Willoughby R, Slate D. Current and future trends in the prevention, treatment, and control of rabies. Expert Rev Anti-Infect Ther. 2006;4(6):1021-38.
124. Leung AKC, Davies HD, Ellis Hon KL. Rabies: Epidemiology, pathogenesis, and prophylaxis. Adv Ther. 2007;24(6):1340-7.
125. Knobel DL, Cleaveland S, Coleman PG, Fevre EM, Meltzer MI, Miranda MEG, et al. Re-evaluating the burden of rabies in Africa and Asia. Bull World Health Organ. 2005;83:360-8.
126. Finke S, Conzelmann KK. Replication strategies of rabies virus. Virus Res. 2005;111:120-31.
127. Hohmann EL. Nontyphoidal salmonellosis. Clin Infect Dis. 2001;32:263-9.
128. Gast RK, Beard CW. Research to understand and control *Salmonella enteritidis* in chickens and eggs. Poult Sci. 1993;72:1157-63.
129. Gast RK. Detection of *Salmonella enteritidis* in experimentally infected laying hens by culturing pools of egg contents. Poult Sci. Poult Sci. 1993;72:267-74.
130. Miriagou V, Tassios PT, Legakis NJ, Tzouvelekis LS. Expanded-spectrum cephalosporin resistance in non-typhoid *Salmonella*. Int J Antimicrob Agents. 2004;23(6):547-55.
131. Walters SP, Thebo AL, Boehm AB. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and *stx* genes in coastal waterbodies of central California. Water Res. 2011;45(4):1752-62.
132. Cardoso F, Shelton D, Sadeghi A, Shirmohammadi A, Pachepsky Y, Dulaney W. Effectiveness of vegetated filter strips in retention of *Escherichia coli* and *Salmonella* from swine manure slurry. J Environ Management. 2012;110:1-7.
133. Sidhu JPS, Hodgers L, Ahmed W, Chong MN, Toze S. Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. Water Res. 2012;46(20):6652-60.
134. Ndiaye ML, Dieng Y, Niang S, Pfeifer HR, Tonolla M, Peduzzi R. Effect of irrigation water on the incidence of *Salmonella* spp. on lettuces produced by urban agriculture and sold on the markets in Dakar, Senegal. Af J Microbiol Res. 2011;5(19):2885-90.
135. Merson MH, O'Malley J, Serwadda, Apisuk C. The history and challenge of HIV prevention. Lancet. 2008;372(9637):475-88.
136. Nicoll A, Hamers FF. Are trends in HIV, gonorrhea, and syphilis worsening in western Europe? BMJ. 2002;324(7349):1324-7.
137. Stamm LV, Mudrak B. Old foes, new challenges: syphilis, cholera, and TB. Future microbial. 2013;8:177-89.

138. Celum C. Sexually transmitted infections and HIV: epidemiology and interventions. Top HIV Med. 2010;18(4):138-42.
139. Unemo M, Palmer HM, Blackmore T, Herrera G, Fredlund H, Limnios A, et al. Global transmission of prolyliminopeptidase-negative *Neisseria gonorrhea* strains: implications for changes in diagnostic strategies. Sex Transm Infect. 2007;83:47-51.
140. Clara Zoni A, Alonso Gonzalez M, Sjogren HW. Syphilis in the most at-risk populations in Latin America and the Caribbean: a systematic review. Int J Infect Dis. 2013;17(2):e84-92.
141. Kirkcaldy RD, Bolan GA, Wasserheit JN. Cephalosporin-resistant gonorrhea in North America. JAMA. 2013;309(2):185-7.
142. Brown PJ, Armelagos GJ, Maes KC. Anthropology of infectious disease. In: Singer M, Erickson PI (Eds.). A companion to medical anthropology. 2011:253-70.
143. Waugh M. A global look at sexually transmitted infections. Acta Dermatovenerol Croat. 2007;15(4):228-31.
144. Holmes KK. Human ecology and behavior and sexually transmitted bacterial infections. Proc Natl Acad Sci USA. 1994;91:2448-55.
145. Donovan B. Sexually transmissible infections other than HIV. Lancet. 2004;363(9408):545-56.
146. Maan MA, Hussain F, Igbal J, Akhtar SJ. Sexually transmitted infections in Pakistan. Ann Saudi Med. 2011;31(3):263-9.
147. Libster R, Edwards KM. Re-emergence of pertussis: what are the solutions? Expert Rev Vaccines. 2012;11(11):1331-46.
148. Poland GA, Jacobson RM. The re-emergence of measles in developed countries: time to develop the next-generation measles vaccines? Vaccine. 2012;30(2):103-4.
149. Fuji N, Suzuki A, Saito M, Centeno R, Galang H, Lupisan S, et al. Interruption of the circulation of an indigenous measles genotype and the introduction of other genotypes after a mass vaccination campaign in the Philippines.
150. Santibanez S, Tischer A, Heider A, Siedler A, Hengel H. Rapid replacement of endemic measles virus genotypes. J Gen Virol. 2002;83(11):2699-708.
151. Litt DJ, Neal SE, Fry NK. Changes in genetic diversity of the *Bordatella pertussis* population in the United Kingdom between 1920 and 2006 reflect vaccination coverage and emergence of a single dominant clonal type. J Clin Microbiol. 2009;47(3):680-8.
152. Black FS. Measles endemicity in insular populations: Critical community size and its evolutionary implication. J Theroetical Biol. 1966;11(2):207-11.
153. Jekel JF. Communicable disease control and public policy in the 1970s: hot war, cold war, or peaceful coexistence?. Am J Pub Health. 1972;62(12):1578-85.
154. Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, et al. Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. Bull World Health Organ. 199;77(8):651-66.
155. Higgins R, Sauvageau R, Bonin P. *Shigella flexneri* Type 2 infection in captive nonhuman primates. Can Vet J. 1985;26:402-3.
156. Tetteh GL, Beuchat LR. Survival, growth, and inactivation of acid-stressed *Shigella flexneri* as affected by pH and temperature. Int J Food Microbiol. 2003;87(1):131-8.
157. Blake PA, Ramos S, MacDonald KL, Rassi V, Tardeli Gomes TA, et al. Pathogen-specific risk factors and protective factors for acute diarrheal disease in urban Brazilian infants. J Infect Dis. 1993;167(3):627-32.

158. Lienhardt C, Glaziou P, Uplekar M, Lonnroth K, Getahun H, Raviglione M. Global tuberculosis control: lessons learnt and future prospects. Nature Rev Microbiol. 2012;10:407-16.
159. Porter JD, McAdam KP. The re-emergence of tuberculosis. Annu Rev Public Health. 1994;15:303-23.
160. Dye C, Lonnroth K, Jaramillo E, Williams BG, Raviglione M. Trends in tuberculosis incidence and their determinants in 134 countries. Bull World Health Organ. 2009;87(9):683-91.
161. Murray M, Oxlade O, Lin HH. Modeling social, environmental and biological determinants of tuberculosis. Int J Tuberc Lung Dis. 2011;15S2:S64-70.
162. Crump JA, Lubly SP, Mintz ED. The global burden of typhoid fever. Bull World Health Organ. 2004;82:346-53.
163. Selander RK, Beltran P, Smith NH, Helmuth R, Rubin FA, Kopecko DJ, et al. Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. Infect Immun. 1990;58(7):2262-75.
164. Olsen SJ, Bleasdale SC, Magnano AR, Landrigan C, Holland BH, Tauxe RV, et al. Outbreaks of typhoid fever in the United States, 1960-99. Epidemiol Infect. 2003;130:13-21.

APPENDIX 2

Table A2.1: Cholera model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	8.02E-05	1.21E-05	6.63	0.000	5.63E-05	0.000104
Population	-3.86E-10	2.48E-10	-1.56	0.123	-8.77E-10	1.06E-10
Reporting	0.099726	0.123197	0.81	0.420	-0.14376	0.343208
GDP (2008 USD)	-3.3E-05	1.46E-05	-2.27	0.027	-6.2E-05	-3.89E-06
Improved water access	-0.83951	0.370682	-2.26	0.025	-1.57256	-0.10645
Temperature	0.0158	0.011272	1.4	0.162	-0.00641	0.038006
Intercept	7.446398	0.890661	8.36	0.000	5.692544	9.200252
Logistic model						
Prior-year endemic	-3.16431	0.203736	-15.53	0.000	-3.5667	-2.76191
Population	-1.49E-08	6.91E-09	-2.15	0.032	-2.84E-08	-1.31E-09
Reporting	0.164705	0.1671	0.99	0.325	-0.16378	0.493189
Island nation	0.900958	0.245158	3.68	0.000	0.419999	1.381917
Improved sanitation access	0.610912	0.456977	1.34	0.183	-0.29176	1.513587
Political Terror Scale	-0.16409	0.102869	-1.6	0.112	-0.36681	0.038622
Population growth	-0.17941	0.070515	-2.54	0.013	-0.31956	-0.03927
Malnutrition prevalence	-0.01839	0.012991	-1.42	0.160	-0.04417	0.00738
Rain	-0.00235	0.001198	-1.96	0.051	-0.00471	1.43E-05
Temperature	0.031405	0.013823	2.27	0.023	0.004251	0.058559
Temperature normal score	0.217847	0.14969	1.46	0.148	-0.07795	0.513645
Intercept	-0.17128	1.116072	-0.15	0.878	-2.36578	2.023222
ln(α)	1.505314	0.094477	15.93	0.000	1.320053	1.690575

Imputations: 20
Number of observations: 4316
Average RVI: 0.6431
DF(min): 58.59; (avg): 432.90; (max): 2510.05
F(5, 527.4): 10.94
P > |F| = 0.0000

Table A2.2: Dengue model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	2.82E-05	8.34E-06	3.38	0.001	1.19E-05	4.46E-05
Population	-1.01E-09	5.84E-10	-1.74	0.086	-2.18E-09	1.48E-10
Reporting	-0.22149	0.306564	-0.72	0.472	-0.83244	0.389459
GINI coefficient	0.019263	0.014451	1.33	0.185	-0.00934	0.04787
Political Terror Scale	0.313539	0.120863	2.59	0.010	0.07498	0.552097
GDP (2008 USD)	-3.1E-05	1.88E-05	-1.64	0.107	-6.9E-05	6.89E-06
Intercept	6.991914	0.823301	8.49	0.000	5.365712	8.618117
Logistic model						
Prior-year endemic	-4.65969	0.362388	-12.86	0.000	-5.37375	-3.94564
Population	-1.51E-08	5.89E-09	-2.56	0.010	-2.67E-08	-3.54E-09
Reporting	-0.37061	0.190245	-1.95	0.052	-0.74489	0.003669
Island nation	-2.64783	0.414936	-6.38	0.000	-3.46183	-1.83382
GINI coefficient	-0.05297	0.017643	-3	0.003	-0.08771	-0.01823
Forested area	-3.37164	0.665507	-5.07	0.000	-4.67709	-2.06619
Political Terror Scale	-0.24227	0.133808	-1.81	0.071	-0.50574	0.021209
Paved roads	0.026897	0.006096	4.41	0.000	0.01491	0.038884
Education	-1.66185	0.96118	-1.73	0.086	-3.55966	0.235961
Population growth	0.158634	0.098403	1.61	0.108	-0.03519	0.352454
GDP (2008 USD)	-2.4E-05	1.51E-05	-1.57	0.116	-5.3E-05	5.89E-06
HIV prevalence	15.8737	5.012389	3.17	0.002	5.960817	25.78657
Tractors	0.000146	0.000106	1.39	0.170	-6.4E-05	0.000357
Improved water access	-3.6954	1.162004	-3.18	0.002	-5.98413	-1.40668
Improved sanitation access	-1.88357	0.769267	-2.45	0.015	-3.40178	-0.36536
Malnutrition prevalence	-0.04881	0.017316	-2.82	0.006	-0.0831	-0.01452
Temperature	-0.05991	0.019251	-3.11	0.002	-0.09768	-0.02215
Intercept	17.05266	2.282073	7.47	0.000	12.57233	21.53299
ln(α)	1.52078	0.107364	14.16	0.000	1.310343	1.731218

Imputations: 20
Number of observations: 4316
Average RVI: 0.5012
DF(min): 48.04; (avg): 1597.72; (max):29034.84
F(5, 461.5): 6.29
P > |F| = 0.0000

Table A2.3: Gonorrhea model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	3.35E-05	1.27E-05	2.64	0.009	8.42E-06	5.85E-05
Population	-6.67E-09	5.01E-09	-1.33	0.185	-1.65E-08	3.21E-09
Reporting	0.006013	0.131579	0.05	0.964	-0.25941	0.271433
Air passengers	-0.11674	0.045012	-2.59	0.011	-0.20614	-0.02734
Political Terror Scale	0.222482	0.077864	2.86	0.005	0.06929	0.375673
Incoming refugees	-4.46215	2.22739	-2	0.049	-8.89813	-0.02618
Malnutrition prevalence	0.010593	0.005942	1.78	0.076	-0.00112	0.022305
Intercept	7.685492	0.347059	22.14	0.000	7.0003	8.370684
ln(α)	1.866424	0.125533			1.610647	2.122201

Imputations: 20
Number of observations: 4316
Average RVI: 1.3085
DF(min): 31.76; (avg): 143.70; (max):319.59
F(6, 370.5): 6.60
P > |F| = 0.0000

Table A2.4: Hantavirus model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	8.14E-05	5.81E-05	1.4	0.169	-3.6E-05	0.000199
Population	6.33E-10	2.30E-09	0.28	0.784	-3.97E-09	5.24E-09
Reporting	-0.47013	0.44122	-1.07	0.289	-1.34578	0.405531
Gender inequality	1.017982	0.685855	1.48	0.142	-0.34836	2.384328
Improved sanitation	-1.3599	0.848277	-1.6	0.116	-3.07028	0.350487
Intercept	8.035322	0.940706	8.54	0.000	6.115954	9.95469
Logistic model						
Prior-year endemic	-3.67947	0.847839	-4.34	0.000	-5.40496	-1.95399
Population	-1.88E-08	1.20E-08	-1.57	0.120	-4.25E-08	4.95E-09
Reporting	-0.20111	0.753319	-0.27	0.791	-1.73397	1.331745
Landlocked nation	-1.01697	0.545671	-1.86	0.065	-2.09676	0.062826
GINI coefficient	0.056919	0.03617	1.57	0.125	-0.01678	0.130623
Intercept	0.565337	1.312263	0.43	0.669	-2.08542	3.21609
ln(α)	0.967716	0.280805	3.45	0.001	0.407407	1.528025

Imputations: 20
Number of observations: 4316
Average RVI: 4.1425
DF(min): 30.69; (avg): 59.54; (max):126.88
F(4, 168.1): 2.94
P > |F| = 0.0220

Table A2.5: Leishmaniasis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	2.63E-05	8.63E-06	3.05	0.006	8.43E-06	4.41E-05
Population	-1.88E-10	5.34E-10	-0.35	0.727	-1.29E-09	9.14E-10
Reporting	0.030106	0.094865	0.32	0.753	-0.16134	0.221552
Island nation	0.19608	0.122523	1.6	0.116	-0.05028	0.442439
Political Terror Scale	0.079514	0.051333	1.55	0.129	-0.02405	0.183077
Population growth	0.061145	0.026731	2.29	0.026	0.007579	0.114712
HIV prevalence	2.605745	1.389445	1.88	0.070	-0.2228	5.434295
Malnutrition prevalence	0.009059	0.004702	1.93	0.061	-0.00044	0.018559
Intercept	8.301677	0.251022	33.07	0.000	7.789953	8.8134
Logistic model						
Prior-year endemic	-1.16491	0.260539	-4.47	0.000	-1.70028	-0.62954
Population	1.36E-10	1.15E-09	0.12	0.907	-2.20E-09	2.47E-09
Reporting	-0.13303	0.170013	-0.78	0.437	-0.4732	0.207145
Education index	0.578974	0.436469	1.33	0.188	-0.28868	1.446633
GDP (2008 USD)	1.84E-05	8.95E-06	2.06	0.044	4.97E-07	3.64E-05
Population growth	-0.11137	0.061105	-1.82	0.074	-0.23383	0.011095
Intercept	0.318249	0.382629	0.83	0.410	-0.45156	1.088058
ln(α)	0.625161	0.094538	6.61	0.000	0.438635	0.811688

Imputations: 20
Number of observations: 4316
Average RVI: 4.9837
DF(min): 22.93; (avg): 52.21; (max):182.71
F(7, 218.1): 5.15
P > |F| = 0.0000

Table A2.6: Leptospirosis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.000675	0.00011	6.12	0.000	0.000455	0.000896
Population	-2.65E-10	5.87E-10	-0.45	0.652	-1.43E-09	8.95E-10
Reporting	-0.22805	0.146225	-1.56	0.125	-0.52195	0.065855
Island nation	0.218307	0.162745	1.34	0.181	-0.10263	0.539241
Agricultural area	-0.51885	0.324641	-1.6	0.112	-1.15974	0.122028
Forested area	-0.56811	0.262307	-2.17	0.032	-1.08774	-0.04847
Political Terror Scale	0.189456	0.069993	2.71	0.008	0.050533	0.328378
Gender inequality	0.526051	0.286844	1.83	0.070	-0.04426	1.096367
Urban growth rate	0.07535	0.034777	2.17	0.032	0.006494	0.144205
Improved sanitation access	-0.79402	0.244882	-3.24	0.002	-1.28402	-0.30401
Intercept	5.781272	0.36155	15.99	0.000	5.066207	6.496338
Logistic model						
Prior-year endemic	-2.75117	0.234897	-11.71	0.000	-3.2267	-2.27564
Population	2.61E-10	8.80E-10	0.3	0.767	-1.47E-09	1.99E-09
Reporting	0.019627	0.172787	0.11	0.910	-0.32231	0.361563
Agricultural area	-0.78635	0.498902	-1.58	0.118	-1.77454	0.20184
Forested area	-0.69794	0.430548	-1.62	0.109	-1.55373	0.157851
Paved roads	-0.00755	0.00457	-1.65	0.106	-0.01678	0.001688
GDP (2008 USD)	2.56E-05	9.21E-06	2.78	0.007	7.20E-06	0.000044
HIV prevalence	4.262003	2.954238	1.44	0.158	-1.73711	10.26112
Rain, normal score	0.212127	0.134256	1.58	0.120	-0.05677	0.481021
Temperature	0.01465	0.009567	1.53	0.128	-0.00427	0.033574
Intercept	0.227552	0.909077	0.25	0.803	-1.57342	2.028519
ln(α)	0.551904	0.091424	6.04	0.000	0.371971	0.731838

Imputations: 20
Number of observations: 4316
Average RVI: 1.6433
DF(min): 34.73; (avg): 296.91; (max):4509.59
F(9, 557.0): 11.29
P > |F| = 0.0000

Table A2.7: Lyme borreliosis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.000181	0.000059	3.07	0.003	0.000064	0.000298
Population	9.72E-10	2.35E-09	0.41	0.679	-3.66E-09	5.60E-09
Reporting	-0.27558	0.383477	-0.72	0.474	-1.03819	0.487023
Air passengers	-0.17727	0.080297	-2.21	0.037	-0.34304	-0.01151
Urban growth rate	-0.22791	0.137379	-1.66	0.100	-0.50049	0.044666
Population growth rate	0.228559	0.166853	1.37	0.174	-0.10327	0.560387
Tractors	0.00007	3.41E-05	2.05	0.045	1.72E-06	0.000138
Improved water access	-2.73262	1.593559	-1.71	0.089	-5.88672	0.421471
Intercept	9.814725	1.563481	6.28	0.000	6.716813	12.91264
Logistic model						
Prior-year endemic	-3.56486	0.296944	-12.01	0.000	-4.15077	-2.97896
Population	-1.02E-10	1.28E-09	-0.08	0.936	-2.61E-09	2.41E-09
Reporting	-0.08465	0.337702	-0.25	0.803	-0.75856	0.589273
Island nation	1.371193	0.550965	2.49	0.013	0.287846	2.454541
Air passengers	-2.05527	0.91683	-2.24	0.026	-3.86368	-0.24686
Forested area	-2.02838	0.991095	-2.05	0.042	-3.97942	-0.07733
Education index	-2.70303	1.301183	-2.08	0.040	-5.27478	-0.13127
Urban growth rate	0.504037	0.162789	3.1	0.002	0.182664	0.825409
HIV prevalence	17.1138	9.355404	1.83	0.074	-1.71703	35.94462
Improved water access	-2.79762	1.51119	-1.85	0.065	-5.77478	0.179529
Rain	0.00389	0.001818	2.14	0.035	0.000285	0.007495
Temperature	0.065245	0.020012	3.26	0.001	0.025759	0.10473
Intercept	2.489023	2.418167	1.03	0.305	-2.29047	7.268516
ln(α)	0.501371	0.17069	2.94	0.003	0.166262	0.83648

Imputations: 20
Number of observations: 4316
Average RVI: 1.5885
DF(min): 23.88; (avg): 236.50; (max):1700.38
F(7, 243.4): 6.53
P > |F| = 0.0000

Table A2.8: Malaria model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	8.66E-07	1.49E-07	5.81	0.000	5.74E-07	1.16E-06
Population	3.04E-10	4.96E-10	0.61	0.541	-6.69E-10	1.28E-09
Reporting	-0.28105	0.172425	-1.63	0.105	-0.62164	0.059534
Landlocked nation	-0.5526	0.242907	-2.27	0.023	-1.03026	-0.07494
Education index	-0.88416	0.527951	-1.67	0.096	-1.92526	0.156942
HIV prevalence	4.926142	2.26487	2.18	0.030	0.476764	9.375521
Improved water access	-1.13048	0.533616	-2.12	0.035	-2.1799	-0.08106
Improved sanitation access	-0.88443	0.391227	-2.26	0.024	-1.65369	-0.11518
Temperature, normal score	0.303818	0.136822	2.22	0.028	0.033665	0.573971
Intercept	13.71572	0.387298	35.41	0.000	12.95564	14.4758
Logistic model						
Prior-year endemic	-4.08086	0.293243	-13.92	0.000	-4.65768	-3.50405
Population	-1.75E-08	4.00E-09	-4.38	0.000	-2.53E-08	-9.67E-09
Reporting	0.071937	0.228754	0.31	0.753	-0.37702	0.520894
GINI coefficient	-0.04355	0.016097	-2.71	0.008	-0.07537	-0.01173
Air passengers	0.231768	0.178467	1.3	0.195	-0.11971	0.583248
Agricultural area	1.521723	0.697044	2.18	0.029	0.153626	2.889821
Political Terror Scale	-0.26193	0.128725	-2.03	0.043	-0.51538	-0.00847
Incoming refugees	6.21276	4.636722	1.34	0.183	-2.97145	15.39697
Education index	1.214643	0.806082	1.51	0.134	-0.37548	2.804767
Urban growth rate	-0.10491	0.077525	-1.35	0.176	-0.25714	0.047316
GDP (2008 USD)	6.23E-05	1.59E-05	3.93	0.000	0.000031	9.35E-05
Improved water access	2.034577	0.95197	2.14	0.034	0.15817	3.910984
Malnutrition prevalence	-0.03185	0.017129	-1.86	0.066	-0.06582	0.002116
Rain	-0.00275	0.001265	-2.17	0.030	-0.00524	-0.00026
Temperature	-0.04988	0.018118	-2.75	0.006	-0.08551	-0.01424
Island nation	1.376791	0.345145	3.99	0.000	0.699249	2.054333
Intercept	5.294255	1.854362	2.86	0.005	1.646762	8.941748
ln(α)	1.471562	0.118117	12.46	0.000	1.240043	1.703082

Imputations: 20
Number of observations: 4316
Average RVI: 0.4126
DF(min): 104.11; (avg): 1184.46; (max):18796.67
F(8, 1945.1): 14.26
P > |F| = 0.0000

Table A2.9: Measles model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	7.54E-05	0.000012	6.29	0.000	5.19E-05	9.89E-05
Vaccination	-1.20209	0.380476	-3.16	0.002	-1.95167	-0.45252
Population	2.36E-10	8.36E-10	0.28	0.778	-1.40E-09	1.87E-09
Reporting	-0.0864	0.079404	-1.09	0.277	-0.24249	0.069701
Island nation	0.382713	0.171474	2.23	0.026	0.045251	0.720175
Air passengers	-0.12253	0.068326	-1.79	0.078	-0.25941	0.014353
Political Terror Scale	0.10088	0.062955	1.6	0.111	-0.02326	0.22502
Incoming refugees	-3.01502	1.918897	-1.57	0.119	-6.81651	0.786464
Education index	-0.95673	0.361707	-2.65	0.009	-1.67147	-0.242
GDP (2008 USD)	1.73E-05	7.13E-06	2.43	0.016	3.23E-06	3.14E-05
Temperature, normal score	-0.08595	0.052898	-1.62	0.107	-0.19077	0.018875
Intercept	8.808168	0.386244	22.8	0.000	8.04872	9.567616
Logistic model						
Prior-year endemic	-3.0006	0.234465	-12.8	0.000	-3.46242	-2.53878
Vaccination	1.790023	0.722754	2.48	0.014	0.367243	3.212804
Population	-2.36E-08	9.66E-09	-2.44	0.015	-4.26E-08	-4.66E-09
Reporting	-0.32178	0.254381	-1.26	0.208	-0.82444	0.18089
Island nation	0.953168	0.228008	4.18	0.000	0.502855	1.403481
GINI coefficient	0.027706	0.012068	2.3	0.023	0.003888	0.051525
Forested area	1.087719	0.444069	2.45	0.015	0.213938	1.961501
Paved roads	-0.01461	0.00404	-3.62	0.000	-0.02258	-0.00665
Incoming refugees	-4.95665	3.049158	-1.63	0.106	-10.9758	1.062469
Urban growth rate	-0.13305	0.056568	-2.35	0.020	-0.2448	-0.0213
GDP (2008 USD)	1.26E-05	9.65E-06	1.3	0.196	-6.59E-06	3.17E-05
Urban population	0.000874	0.000558	1.57	0.120	-0.00023	0.001981
HIV prevalence	-4.23277	2.878805	-1.47	0.144	-9.92863	1.463088
Improved water access	1.774376	0.831255	2.13	0.035	0.130714	3.418038
Malnutrition prevalence	-0.01897	0.011648	-1.63	0.106	-0.04203	0.004096
Temperature	0.023098	0.010901	2.12	0.035	0.001621	0.044575
Intercept	-4.58287	1.38183	-3.32	0.001	-7.30795	-1.8578
ln(α)	1.270506	0.052911	24.01	0.000	1.166779	1.374232

Imputations: 20
Number of observations: 4316
Average RVI: 0.5516
DF(min): 55.87; (avg): 1148.43; (max):18759.77
F(10, 1311.3): 12.74
P > |F| = 0.0000

Table A2.10: Pertussis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.000264	4.72E-05	5.59	0.000	0.000171	0.000356
Vaccination	-1.1165	0.44241	-2.52	0.012	-1.98522	-0.24778
Population	1.04E-09	2.84E-10	3.68	0.000	4.86E-10	1.60E-09
Reporting	-0.22834	0.158526	-1.44	0.151	-0.54004	0.083354
Island nation	0.461328	0.20735	2.22	0.027	0.053158	0.869498
Landlocked nation	-0.26907	0.187548	-1.43	0.152	-0.63698	0.09885
GDP (2008 USD)	1.85E-05	6.64E-06	2.79	0.005	5.49E-06	3.16E-05
Urban population	1.424887	0.47558	3	0.003	0.488874	2.3609
HIV prevalence	3.873636	2.623938	1.48	0.145	-1.36694	9.114212
Improved water access	-0.95441	0.50295	-1.9	0.059	-1.94733	0.038515
Improved sanitation access	-0.62	0.337083	-1.84	0.067	-1.28377	0.043771
Malnutrition prevalence	0.014904	0.006937	2.15	0.033	0.001232	0.028577
Temperature	-0.02334	0.010163	-2.3	0.024	-0.04355	-0.00313
Intercept	9.216413	0.932101	9.89	0.000	7.361675	11.07115
Logistic model						
Prior-year endemic	-2.78577	0.212012	-13.14	0.000	-3.20604	-2.36549
Vaccination	1.773801	0.689822	2.57	0.012	0.404606	3.142996
Population	-2.58E-08	1.73E-08	-1.49	0.136	-5.98E-08	8.14E-09
Reporting	-0.30857	0.202497	-1.52	0.130	-0.70884	0.091689
Paved roads	-0.0057	0.003564	-1.6	0.111	-0.01273	0.001329
HIV prevalence	8.801867	1.867072	4.71	0.000	5.120131	12.4836
Rain	0.0015	0.000825	1.82	0.071	-0.00013	0.003128
Temperature	0.020631	0.011085	1.86	0.066	-0.00139	0.042652
Island nation	1.135055	0.218708	5.19	0.000	0.704619	1.565491
Intercept	-2.58289	1.14434	-2.26	0.027	-4.858	-0.30779
ln(α)	0.992692	0.076208	13.03	0.000	0.843034	1.142349

Imputations: 20
Number of observations: 4316
Average RVI: 0.5957
DF(min): 64.86; (avg): 450.71; (max):3870.53
F(12, 2409.6): 9.98
P > |F| = 0.0000

Table A2.11: Plague (Yersiniosis) model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	5.89E-05	0.000102	0.58	0.566	-0.00015	0.000265
Population	7.90E-10	1.71E-09	0.46	0.647	-2.70E-09	4.28E-09
Reporting	-0.00641	0.212785	-0.03	0.976	-0.43452	0.4217
Population growth	0.079417	0.056895	1.4	0.169	-0.03468	0.193513
Intercept	7.486797	0.90751	8.25	0.000	5.592929	9.380665
Logistic model						
Prior-year endemic	-0.96577	0.232281	-4.16	0.000	-1.43521	-0.49634
Population	2.82E-11	2.31E-09	0.01	0.990	-4.72E-09	4.77E-09
Reporting	0.258237	0.326396	0.79	0.435	-0.40705	0.923522
Island nation	0.559681	0.314499	1.78	0.078	-0.06382	1.183186
Paved roads	-0.00698	0.004476	-1.56	0.124	-0.01593	0.001974
Population growth	-0.15009	0.066726	-2.25	0.027	-0.28272	-0.01746
GDP (2008 USD)	0.000036	1.35E-05	2.66	0.009	9.17E-06	6.27E-05
Malnutrition prevalence	-0.02887	0.008645	-3.34	0.001	-0.04604	-0.01171
Tractors	-0.00018	0.000055	-3.29	0.001	-0.00029	-7.2E-05
Intercept	2.124607	0.408246	5.2	0.000	1.307199	2.942016
ln(α)	0.703248	0.178331	3.94	0.000	0.341769	1.064727

Imputations: 20
Number of observations: 4316
Average RVI: 19.1570
DF(min): 19.86; (avg): 60.67; (max): 139.42
F(3, 95.0): 0.52
P > |F| = 0.6718

Table A2.12: Rabies model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.004132	0.000953	4.34	0.000	0.002254	0.006011
Population	5.05E-11	2.91E-10	0.17	0.862	-5.22E-10	6.23E-10
Reporting	-0.0924	0.136508	-0.68	0.501	-0.36471	0.179921
GINI coefficient	-0.01616	0.010669	-1.51	0.134	-0.03739	0.005083
Air passengers	0.292387	0.15521	1.88	0.064	-0.01705	0.601828
Agricultural area	-0.58632	0.428755	-1.37	0.174	-1.43412	0.261471
HIV prevalence	2.658023	1.828131	1.45	0.148	-0.95617	6.27222
Improved water access	-0.71877	0.544356	-1.32	0.189	-1.7938	0.35626
Malnutrition prevalence	0.020507	0.008672	2.36	0.019	0.003395	0.037619
Intercept	4.743577	0.740763	6.4	0.000	3.279439	6.207715
Logistic model						
Prior-year endemic	-3.41181	0.439377	-7.77	0.000	-4.28911	-2.5345
Population	-1.51E-08	4.16E-09	-3.62	0.000	-2.32E-08	-6.89E-09
Reporting	0.263172	0.167293	1.57	0.118	-0.06707	0.593411
Island nation	1.428887	0.350082	4.08	0.000	0.741484	2.11629
Air passengers	0.3149	0.184528	1.71	0.092	-0.05283	0.682628
Agricultural area	-1.52866	0.691535	-2.21	0.029	-2.90056	-0.15676
Political Terror Scale	-0.48942	0.150814	-3.25	0.002	-0.792	-0.18684
Gender inequality	-1.30093	0.691029	-1.88	0.065	-2.68472	0.082849
Urban growth rate	-0.21334	0.075069	-2.84	0.006	-0.36289	-0.0638
Malnutrition prevalence	-0.01964	0.012731	-1.54	0.126	-0.04489	0.005607
Intercept	4.26105	0.642965	6.63	0.000	2.980523	5.541577
ln(α)	0.882742	0.108269	8.15	0.000	0.667828	1.097656

Imputations: 20
Number of observations: 4316
Average RVI: 0.8678
DF(min): 52.39; (avg): 163.90; (max): 663.14
F(8, 824.2): 5.28
P > |F| = 0.0000

Table A2.13: Non-typhoidal Salmonellosis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	5.69E-05	1.01E-05	5.62	0.000	3.69E-05	7.68E-05
Population	-4.39E-09	2.35E-09	-1.87	0.062	-9.00E-09	2.16E-10
Reporting	-0.10032	0.112734	-0.89	0.376	-0.32445	0.123811
Air passengers	-0.1168	0.043935	-2.66	0.009	-0.20418	-0.02943
Political Terror Scale	0.076975	0.048723	1.58	0.115	-0.0187	0.17265
Incoming refugees	-4.26597	2.089732	-2.04	0.044	-8.41034	-0.1216
Tractors	-7.9E-05	5.02E-05	-1.57	0.116	-0.00018	1.96E-05
Malnutrition prevalence	0.007906	0.004018	1.97	0.050	-9.76E-06	0.015821
Intercept	8.229784	0.26644	30.89	0.000	7.704628	8.754939
Logistic model						
Prior-year endemic	-3.57996	0.220373	-16.25	0.000	-4.02489	-3.13504
Population	-1.44E-08	5.20E-09	-2.77	0.008	-2.48E-08	-3.94E-09
Reporting	-0.17449	0.230282	-0.76	0.453	-0.63865	0.289667
GINI coefficient	0.018925	0.011296	1.68	0.100	-0.00379	0.041639
Paved roads	-0.00557	0.00294	-1.89	0.061	-0.0114	0.000266
Education	-0.82848	0.585268	-1.42	0.164	-2.00865	0.351676
Intercept	1.230262	0.776034	1.59	0.122	-0.34304	2.80356
ln(α)	0.466276	0.095649	4.87	0.000	0.278218	0.654333

Imputations: 20
Number of observations: 4316
Average RVI: 1.6665
DF(min): 36.38; (avg): 327.90; (max): 2826.27
F(7, 904.1): 10.47
P > |F| = 0.0000

Table A2.14: Shigellosis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	2.42E-05	8.29E-06	2.92	0.003	7.97E-06	4.05E-05
Population	-1.88E-09	2.60E-09	-0.72	0.469	-6.97E-09	3.21E-09
Reporting	-0.02703	0.088779	-0.3	0.762	-0.20368	0.149629
GINI coefficient	0.011423	0.008046	1.42	0.158	-0.00449	0.027339
Forested area	-0.61838	0.293632	-2.11	0.036	-1.19605	-0.04072
Paved roads	-0.00445	0.002703	-1.65	0.101	-0.00978	0.000884
Gender inequality	0.904521	0.319266	2.83	0.005	0.272244	1.536799
Education index	-0.53874	0.362385	-1.49	0.138	-1.25045	0.172972
Tractors	-0.00011	7.84E-05	-1.45	0.152	-0.00027	4.29E-05
Improved sanitation	-0.80433	0.267606	-3.01	0.003	-1.33267	-0.27599
Temperature, normal score	0.121363	0.090079	1.35	0.181	-0.05767	0.3004
Intercept	9.362297	0.449242	20.84	0.000	8.476149	10.24844
Logistic model						
Prior-year endemic	-3.09316	0.290246	-10.66	0.000	-3.6865	-2.49981
Population	-2.18E-08	6.55E-09	-3.33	0.001	-3.49E-08	-8.70E-09
Reporting	-0.31892	0.275348	-1.16	0.253	-0.87403	0.236196
Landlocked nation	-0.37295	0.286406	-1.3	0.199	-0.94802	0.202112
GINI coefficient	0.023196	0.013353	1.74	0.090	-0.00375	0.050143
Political Terror Scale	0.173347	0.090794	1.91	0.061	-0.0081	0.354797
Temperature	0.023475	0.014743	1.59	0.119	-0.00631	0.053256
Intercept	-2.02306	1.194914	-1.69	0.097	-4.42889	0.382772
ln(α)	0.871501	0.097815	8.91	0.000	0.679128	1.063873

Imputations: 20
Number of observations: 4316
Average RVI: 1.4606
DF(min): 29.31; (avg): 2616.08; (max): 46344.55
F(10, 1146.6): 9.02
P > |F| = 0.0000

Table A2.15: Syphilis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.000107	3.24E-05	3.31	0.001	4.31E-05	0.000171
Population	-2.42E-09	2.81E-09	-0.86	0.388	-7.93E-09	3.08E-09
Reporting	0.073711	0.184419	0.4	0.692	-0.30036	0.447783
GINI coefficient	0.011643	0.007651	1.52	0.134	-0.00367	0.02696
Air passengers	-0.09908	0.045628	-2.17	0.033	-0.19009	-0.00808
Political Terror Scale	0.219914	0.070289	3.13	0.003	0.079232	0.360596
Incoming refugees	-3.83913	2.458663	-1.56	0.124	-8.77096	1.092712
HIV prevalence	-4.13985	2.872949	-1.44	0.160	-9.99675	1.717048
Temperature, normal score	0.13532	0.096046	1.41	0.164	-0.05673	0.327369
Intercept	6.193504	0.4639	13.35	0.000	5.270179	7.116829
ln(α)	1.899242	0.086919			1.725926	2.072558

Imputations: 20
Number of observations: 4316
Average RVI: 1.9490
DF(min): 31.33; (avg): 964.48; (max): 9939.48
F(8, 360.8): 3.51
P > |F| = 0.0006

Table A2.16: Tuberculosis model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	1.99E-05	8.15E-06	2.44	0.015	3.91E-06	3.59E-05
Vaccination	0.391482	0.343371	1.14	0.255	-0.2851	1.068066
Population	3.20E-09	3.38E-09	0.95	0.344	-3.43E-09	9.84E-09
Reporting	0.103305	0.125223	0.82	0.411	-0.14453	0.351138
Island nation	-0.4658	0.198431	-2.35	0.019	-0.8559	-0.0757
Air passengers	-0.14292	0.082803	-1.73	0.091	-0.30962	0.023792
Political Terror Scale	0.271338	0.079394	3.42	0.001	0.115475	0.427201
Paved roads	0.002921	0.002575	1.13	0.258	-0.00215	0.007987
Incoming refugees	-2.30822	1.842715	-1.25	0.212	-5.94992	1.333481
Gender inequality	0.375251	0.344147	1.09	0.277	-0.30451	1.055012
HIV prevalence	3.769241	1.217783	3.1	0.002	1.368474	6.170007
Improved water access	-0.52523	0.60853	-0.86	0.389	-1.72285	0.672399
Malnutrition prevalence	0.012909	0.008163	1.58	0.115	-0.00319	0.029005
Rain, normal score	0.092359	0.064388	1.43	0.153	-0.03481	0.219531
Temperature	-0.01538	0.008338	-1.84	0.067	-0.03182	0.001061
Intercept	8.492617	0.812522	10.45	0.000	6.891858	10.09338
ln(α)	0.980766	0.087373			0.809499	1.152034

Imputations: 20
Number of observations: 4316
Average RVI: 0.5429
DF(min): 45.65; (avg): 6219.87; (max): 63421.67
F(15, 2011.1): 9.23
P > |F| = 0.0000

Table A2.17: Typhoid model

	Coefficient	Std. Err.	t	P > |t|	[95% Conf.	Interval]
Count model						
Prior-year incidence	0.000187	2.51E-05	7.47	0.000	0.000138	0.000237
Population	7.30E-10	7.39E-10	0.99	0.328	-7.53E-10	2.21E-09
Reporting	-0.05657	0.092339	-0.61	0.542	-0.24031	0.12718
Island nation	0.282308	0.162395	1.74	0.085	-0.03897	0.603587
Paved roads	-0.00422	0.001887	-2.24	0.026	-0.00794	-0.00051
Gender inequality	0.79627	0.27206	2.93	0.005	0.252809	1.33973
Education index	-1.1832	0.339427	-3.49	0.001	-1.85976	-0.50664
Temperature	0.022825	0.009023	2.53	0.012	0.004997	0.040653
Urban population	-0.43493	0.312207	-1.39	0.166	-1.05365	0.183781
Intercept	5.734562	0.77672	7.38	0.000	4.197972	7.271151
Logistic model						
Prior-year endemic	-3.5496	0.216204	-16.42	0.000	-3.98019	-3.11901
Population	-1.35E-08	6.95E-09	-1.95	0.060	-2.77E-08	5.91E-10
Reporting	0.055578	0.201113	0.28	0.783	-0.34763	0.458787
Island nation	0.444966	0.244902	1.82	0.073	-0.04148	0.931408
Gender inequality	0.667464	0.437759	1.52	0.129	-0.19706	1.531986
Rain	-0.0023	0.001316	-1.74	0.085	-0.00492	0.000328
Temperature	0.022461	0.014154	1.59	0.115	-0.00557	0.050489
Forested area	0.848248	0.464767	1.83	0.072	-0.0771	1.7736
HIV prevalence	5.123578	2.803386	1.83	0.076	-0.55738	10.80454
Tractor	0.000125	6.71E-05	1.86	0.064	-7.14E-06	0.000257
Improved water access	1.375287	0.896046	1.53	0.133	-0.43933	3.189907
Intercept	-2.39593	1.485523	-1.61	0.111	-5.35706	0.565199
ln(α)	0.630975	0.091849	6.87	0.000	0.450909	0.811041

Imputations: 20
Number of observations: 4316
Average RVI: 1.5901
DF(min): 34.57; (avg): 312.54; (max): 4817.49
F(8, 533.8): 19.83
P > |F| = 0.0000