PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Research

Cite this article: Wood CL, McInturff A, Young HS, Kim DH, Lafferty KD. 2017 Human infectious disease burdens decrease with urbanization but not with biodiversity. *Phil. Trans. R. Soc. B* **372**: 20160122. http://dx.doi.org/10.1098/rstb.2016.0122

Accepted: 28 January 2017

One contribution of 13 to a theme issue 'Conservation, biodiversity and infectious disease: scientific evidence and policy implications'.

Subject Areas:
health and disease and epidemiology

Keywords:
Infectious disease, disability-adjusted life year, dilution effect, global change

Author for correspondence:
Chelsea L. Wood
e-mail: chelwood@uw.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3717517.

THE ROYAL SOCIETY PUBLISHING

Human infectious disease burdens decrease with urbanization but not with biodiversity

Chelsea L. Wood[1,2], Alex McInturff[3], Hillary S. Young[4], DoHyung Kim[5] and Kevin D. Lafferty[6]

[1]Department of Ecology and Evolutionary Biology and Michigan Society of Fellows, University of Michigan, Ann Arbor, MI 48104, USA
[2]School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, USA
[3]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA
[4]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106, USA
[5]Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA
[6]US Geological Survey, Western Ecological Research Center, c/o Marine Science Institute, University of California, Santa Barbara, CA 93106, USA

CLW, 0000-0003-2738-3139; HSY, 0000-0003-0449-8582; KDL, 0000-0001-7583-4593

Infectious disease burdens vary from country to country and year to year due to ecological and economic drivers. Recently, Murray *et al*. (Murray CJ *et al*. 2012 *Lancet* **380**, 2197–2223. (doi:10.1016/S0140-6736(12)61689-4)) estimated country-level morbidity and mortality associated with a variety of factors, including infectious diseases, for the years 1990 and 2010. Unlike other databases that report disease prevalence or count outbreaks per country, Murray *et al*. report health impacts in per-person disability-adjusted life years (DALYs), allowing comparison across diseases with lethal and sublethal health effects. We investigated the spatial and temporal relationships between DALYs lost to infectious disease and potential demographic, economic, environmental and biotic drivers, for the 60 intermediate-sized countries where data were available and comparable. Most drivers had unique associations with each disease. For example, temperature was positively associated with some diseases and negatively associated with others, perhaps due to differences in disease agent thermal optima, transmission modes and host species identities. Biodiverse countries tended to have high disease burdens, consistent with the expectation that high diversity of potential hosts should support high disease transmission. Contrary to the dilution effect hypothesis, increases in biodiversity over time were not correlated with improvements in human health, and increases in forestation over time were actually associated with increased disease burden. Urbanization and wealth were associated with lower burdens for many diseases, a pattern that could arise from increased access to sanitation and healthcare in cities and increased investment in healthcare. The importance of urbanization and wealth helps to explain why most infectious diseases have become less burdensome over the past three decades, and points to possible levers for further progress in improving global public health.

This article is part of the themed issue 'Conservation, biodiversity and infectious disease: scientific evidence and policy implications'.

1. Introduction

A person born today in Japan can expect to live 84 years, whereas an average Sierra Leonean will live to 50 [1]. These country-to-country differences stem from many factors, but infectious diseases are among the most important, accounting for 16% of global deaths and causing tens of millions of years of healthy life to be lost annually, primarily in low-income countries [2]. These

spatial patterns in disease can also shift over time, with some diseases declining (e.g. smallpox eradication), and others increasing in prevalence (e.g. HIV/AIDS) or expanding in geographic range (e.g. dengue). The Sustainable Development Goal agenda focuses on improving global health and reducing health disparities [1]. Understanding the factors that drive disease patterns in space and time is a crucial first step in these efforts, as this approach may identify effective levers for improving human health.

There are some clear geographic patterns in the distribution of infectious disease burden. For instance, wealthy, temperate countries have fewer kinds of infectious disease (lower pathogen richness) than do poor, tropical countries [3,4], explaining disease hotspots in sub-Saharan Africa, South America and Southeast Asia [5]. But which of the multiple, often correlated, potential drivers actually create these patterns? Past studies on global-scale patterns of disease have explored a wide range of potential drivers. Four hypothesized (and interconnected) drivers recur in these studies: human demography (e.g. population density or size), poverty (e.g. *per capita* wealth), environment (e.g. climate, often proxied by latitude) and biotic factors (i.e. biodiversity, reservoir host or vector biodiversity, or proxies thereof). For instance, previous work shows that population size [6,7] and species richness [3,6–10] are positively associated with overall human pathogen richness or number of outbreaks in a country. But the number of disease agents or disease outbreaks does not necessarily reflect disease burden; for instance, while the relationship between biodiversity and pathogen richness is hypothesized to be positive (e.g. [6]), the link between biodiversity and disease burden is often hypothesized to be negative (e.g. [4]). Focusing specifically on disease burden, climate (i.e. temperature) has been positively linked to pathogen prevalence, while *per capita* healthcare spending, *per capita* wealth, and human population size have been negatively linked to prevalence [4,6]. Our study builds on these results with new data and new approaches. By also accounting for change *over time* in these putative drivers and in disease burden, our analysis avoids a serious pitfall of previous analyses: it is easier to observe spurious relationships among variables measured at a single time point than among those measured at multiple time points. For example, wealth might have strong negative associations with disease in 2010, but if increasing wealth does not produce a decrease in disease between 1990 and 2010, we would have reason to question whether the two variables are causally associated.

(a) Disease drivers

In this analysis, we investigate associations between disease burden and demography, economics, environment and biotic factors. All four associations have complex underlying predictions. Wealth is frequently considered to be the most important disease driver (e.g. [4,11]). Specifically, most diseases are more common among the poor within a given country and more common in poor countries than in wealthy countries [12]. Wealth has many health benefits, including clean water and sanitation, education and reduced contact with the environment (i.e. reduced exposure to disease vectors or reservoir hosts; [13–15]). By contrast, poverty and disease can reinforce one another other, with poor health eroding the human capital that otherwise might be used to generate wealth and escape disease, leading to a 'poverty trap' [11,16]. The mutually reinforcing nature of poverty and disease makes it difficult to assess the direction of causality between these factors, and the importance of each in reinforcing the other [4,11,12,16]. Wealth also correlates with enhanced investment in and access to healthcare. Consistent with this, previous global analyses have found that human disease prevalence is lowest where *per capita* healthcare spending is highest [6]. However, elevated *per capita* healthcare spending could also reflect a high-priority disease burden or be allocated in response to a recent uptick in disease. To a certain extent, both could be true, and the net effect might therefore be difficult to detect.

Human population density should also be linked to disease prevalence. Basic epidemiological models suggest that increasing host density (here, humans) increases disease transmission [17]. However, a previous global analysis found no relationship between human density and disease prevalence at the global scale [6]. There are a few reasons why disease might not increase with population density. In particular, transmission rates, which should increase with density, are not the same as host infection risk, which might increase or decrease in response to host density. This is the case for encounter-dilution effects, such as might occur for vector-transmitted diseases or for diseases that spill over from wildlife [18,19]. Alternatively, contact rates (and therefore transmission) might saturate with increasing host density [20], or disturbance caused by increased human density might degrade the ecological systems upon which infectious diseases depend for transmission (e.g. [21]; reviewed in [22]). Pathogenic diseases could reduce human density by increasing mortality rates, reversing the expected positive relationship between density and disease (e.g. [23,24]). Finally, human migration patterns might modify or even supersede any density effects. Density is highest in cities, but urbanization might reduce exposure to non-human hosts (e.g. by reducing human contact with biodiverse forest habitat), while also placing people in locales where healthcare and sanitation infrastructure are close at hand [25,26]. For example, immigrants to the city of Naples, Italy, bring intestinal parasites from their countries of origin, but lose those infections during the years that they spend living in sanitary conditions with access to medical treatment; they fail to pass their infections even to other members of their own households [27]. Urbanization and population density effects could cancel each other out or have variable effects across diseases. Analyses that partition urbanization from population density, wealth and biodiversity might help reveal these effects.

Climate is another potential driver of disease burden. As for free-living (i.e. non-parasitic) species, infectious agents have ecological niches constrained by temperature and precipitation. Although many infectious diseases are more prevalent in tropical climates, for each disease there are hypothetical conditions that are too hot or too cold, too wet or too dry [28]. Cumulatively, this suggests that relationships between climate and disease (and between climate *change* and disease) could be complex. Because climate change might affect hosts, vectors and parasites in different ways, climate effects on disease are likely to be idiosyncratic and difficult to predict [29].

The 'diversity dilution effect' is the most controversial hypothesis about potential infectious disease drivers (e.g. [22,30–40]). According to the dilution effect hypothesis, biodiverse communities (often indicated by intact forest habitat) tend to contain non-competent hosts that interfere with

pathogen transmission or regulate reservoir host density; declines in biodiversity should reduce these 'diluting' species, leading to increased disease burden in low-diversity habitats [41]. Although dilution effects are commonly reported in the conservation biology literature [36], diversity's direct relevance for public health has rarely been tested in real-world contexts [42]. Among the few studies that correlate diversity and human disease risk, results are mixed [43]: some find negative associations between biodiversity measures and disease measures (e.g. [44]), some show increasing biodiversity increasing disease risk (e.g. [21]), and others indicate that whether biodiversity increases or decreases disease risk can depend on other factors, such as human behaviour (e.g. [45–47]). The evidence suggests that both positive and negative associations between biodiversity and disease burden are possible, depending on the disease agent, the biodiversity measure and the spatial scale [22,32,38], but the question remains: which of these outcomes predominates [48]?

In addition to the complexity within each putative driver of infectious disease burden, a further complication is that independent variables can interact with each other or be driven by a common factor. For instance, wealth, biodiversity, temperature and disease are all associated with latitude [4]; forests depend on precipitation and biodiversity can be associated with forest; health spending depends on wealth, while wealth and human density could reduce forestation; and so on.

(b) New approaches

Previous studies examining human pathogen geography have used indirect data on human disease burden [5], such as outbreak frequency (e.g. [8,49]), pathogen richness (e.g. [6,50]) or aggregated prevalence (e.g. bins of endemic, sporadic or not endemic summed across pathogens with different health impacts and ecologies [6]). However, human disease burden is best measured as disability-adjusted life-years (DALYs), or the sum of years of life lost and years lived with disability [13,14,51]. This summary measure permits an 'apples-to-apples' comparison of disease burden across disease types, allowing deadly diseases to be compared in a coherent way to diseases that are disabling but rarely fatal. Since their introduction in the mid-1990s [52], DALYs have become a standard way to measure health [2]. In the past few years, the release of new country-level data spanning two decades [51] has made it possible to investigate broad-scale spatial and temporal patterns in disease burden. Here, we use the most recent data to explore predictions about how human density, poverty, climate, and biodiversity drive temporal and spatial variation in infectious disease burden in humans.

Countries are not ideal units of replication. First, they vary in area and population, which means that statistics and indices that vary with sample size (e.g. number of outbreaks, species richness, gross national income (GNI)) are invalid for comparison. For this reason, we used only *per capita* measures or densities. A subtler limitation is the 'ecological fallacy', which occurs when patterns at one scale of organization (e.g. the country level) are presumed to reflect patterns at another scale of organization (e.g. the individual [53,54]). Ecological fallacies can also occur when patterns at one spatial scale (e.g. China at 9 400 000 km^2) are presumed to be comparable to patterns at another spatial scale (e.g. El Salvador at 21 000 km^2 [53,54]). Taking China and El Salvador as examples of large and small countries, respectively, processes that occur at the local scale for China could be the processes that occur at the country scale for El Salvador, and processes occurring at the country scale for China might have no analogue in El Salvador [55,56]. Furthermore, it is more likely that people living in El Salvador will experience its average climate, wealth and biodiversity, whereas the average conditions in China combine regions that vary greatly from one another in climate, wealth, biodiversity and other factors. We therefore chose to exclude large countries from our analysis, to increase the probability of detecting trends at the country level, and reduce the probability that those trends would be swamped by the interference of different processes in larger countries. Another problem with comparing countries is that accuracy in health statistics and reporting varies from country to country. Although Murray *et al.* [51] do their best to account for accuracy, residual differences might add error to DALY estimates, reducing statistical power. Furthermore, because countries can occur near other countries, they are not always independent replicates. To address the potential for such spatial autocorrelation, we assessed independence among countries using Moran's I. A final problem with using countries as replicates concerns the unknown, country-specific factors that could affect infectious diseases, including historical, political, or cultural effects. One way to account for this is to use many descriptive variables, but using too many variables leads to overly complicated statistical models. Unlike any other study that has used countries as replicates in spatial associations between human disease burden and potential drivers, we were also able to use countries as replicates in temporal comparisons. Temporal comparisons help control for most geographical, historical, political or cultural effects by comparing countries (i.e. in 1990) against themselves (i.e. in 2010). On the other hand, our two time-point comparison (1990 and 2010) does not possess the resolution needed to investigate dynamics and is subject to spurious correlations driven by unmeasured temporal trends. Because temporal and spatial comparisons each have their limitations, we give more weight to drivers associated with similar spatial and temporal variation in disease.

2. Material and methods

To examine hypotheses about what drives global infectious disease burdens, we used correlations between putative drivers and DALYs. We took two approaches: the first, a spatial analysis, used country-level estimates from 2010 as the unit of replication, and the second, a temporal analysis, used change in country-level estimates between 1990 and 2010. Because we sought to find associations between putative drivers and disease, we chose diseases and countries to maximize sample size while minimizing confounds and over-fitting. In this section, we discuss how we selected and defined the variables and calculated associations among them. The full dataset is available as the electronic supplementary material, S1.

(a) Diseases

We started with the infectious diseases tracked by the World Health Organization's Global Burden of Disease (WHO GBD) database (electronic supplementary material, S2). We excluded the WHO GBD disease classes that encompass multiple pathogens, and those that can be caused by either infectious or non-infectious processes (e.g. diarrheal diseases, respiratory infections, 'other

infectious diseases', 'other neglected tropical diseases'). We also excluded those diseases that were reported from fewer than 10 countries (African trypanosomiasis, onchocerciasis, trachoma) and one disease (yellow fever) that caused < 100 global DALYs in the included countries (see §3) in 2010 (electronic supplementary material, S2). After these exclusions, 24 diseases remained for analysis. For the spatial analysis, our response variable was DALYs per 100 000 people in 2010 for each of the included infectious diseases. For the temporal analysis, change in disease burden between 1990 and 2010 was calculated at the country level for each disease. Countries with zero values in both 1990 and 2010 were excluded from the temporal analysis (but not from the spatial analysis), because change in the burden of a disease is not possible in countries where the disease does not exist; that is, in a country with a 1990 burden of zero and a 2010 burden of zero, change over time (0) is unrelated to the putative drivers included in the statistical model, and reduces statistical power for detecting relationships between putative drivers and change in burden of other diseases where burden is more than zero. The temporal analysis controls for parallel global trends in putative drivers and disease that could lead to spurious disease–driver correlations in traditional time-series analyses. In other words, a global increase in wealth that parallels a global decline in disease would not register as a negative wealth–disease correlation in our analysis. To establish such a wealth–disease correlation in our analysis requires that countries with the greatest growth in wealth experience the greatest decline in disease.

(b) Potential drivers

For each country, we derived or found published data on the following putative disease drivers: *per capita* wealth, human population density, per cent of people living in urban environments, temperature, precipitation per unit area, forest cover per unit area and bird + mammal species richness per unit area (data sources listed in the electronic supplementary material, S2). To match disease data, driver data were from 1990 and 2010 or the closest years possible (see the electronic supplementary material, S2). We started with the 192 sovereign nations tracked by the World Bank. From these, we imposed an upper country-area cut-off (greater than 1.5 million km^2) to limit the potential for ecological fallacy and a lower country-area cut-off (less than 625 km^2) to prevent tiny nations from having a disproportionate effect on results. We imposed an area-forested cut-off because change in forestation over time has little meaning in nations that begin with little or no forest cover, and a 10% cut-off (according to satellite-derived forest cover metrics, see below) represented a natural breakpoint in the data. We also excluded countries without available population, GNI or climate data (see the electronic supplementary material, S2 for how the 192 countries were treated and electronic supplementary material, S3 for world map indicating the included countries). After these exclusions, 60 countries (more than 30% of all 192 countries) remained (electronic supplementary material, S2 and S3). For the temporal analysis, we calculated change in drivers between 1990 and 2010 as log ratios ($\log_{10}$[driver in 2010/driver in 1990]).

Biodiversity change was the most difficult driver to measure. We wanted to avoid country-level richness estimates because these increase with country area, so we created a novel integrated species richness estimate for birds and mammals by averaging bird and mammal species richness across all 10 × 10 km pixels within each country's terrestrial grid [57]. We found no well-resolved global biodiversity estimates across multiple time points, nor did we expect that reported range changes or extinctions between 1990 and 2010 would have been a meaningful way to measure biodiversity change. Therefore, to estimate biodiversity change, we first found the most recent and most highly resolved biodiversity dataset available ([57]; electronic supplementary material, S2) and averaged bird and mammal species richness across all 10 × 10 km pixels within each country's terrestrial grid. This yielded country-level biodiversity estimates for 2012. To generate country-level biodiversity estimates for 1990, we used a least-squares regression to predict the slope and intercept that best fit the relationship between country-level biodiversity in 2012 and satellite-derived forest cover in 2012. We then used the coefficients from this regression model to predict the biodiversity in each country in 1990 and 2012, given each country's satellite-derived forest cover estimate for 1990 and 2012. We divided each country's 1990 biodiversity prediction by its 2012 biodiversity prediction to obtain an estimated inverse change in biodiversity for each country. We then multiplied this predicted inverse change by the observed 2012 biodiversity to hindcast biodiversity in 1990 for each country. This was done for mammals and birds separately. We acknowledge that our approach only measures biodiversity change indirectly, that some biodiversity loss might be associated with change in forest attributes other than cover [58], and that our forest-based metric will fail to capture change in non-forest terrestrial biodiversity, freshwater biodiversity, marine biodiversity and diversity of organisms other than birds and mammals. However, within this limited scope (i.e. forest mammals and birds), there is evidence for a strong positive link between forestation and biodiversity (e.g. [59–61]). Our approach also represents a significant improvement over previous efforts, all of which have used biodiversity estimates at only one point in time (e.g. [3,6,8]).

Data on forest cover in 1990 and 2012 were derived from published Landsat-based estimates for forested cells at 30-m resolution [62,63]. This fine resolution was inconsistent with our biodiversity data, so forest cover data were aggregated to 5 km grid cells. In each cell, we denoted per cent forest cover while accounting for cloud and snow cover. These larger cells were used to estimate the forest cover for each country. Forest cover data at this spatial resolution were not available for 2010 (the year in which we have data for human disease burden), but were available for 2012 [62]. Satellite-derived forest cover estimates for 1990 [63] and 2012 [62] were not produced with consistent methods: the 1990 estimates define as 'forested' any 30-m pixel with greater than or equal to 30% canopy cover, whereas the 2012 estimates use a 25% cut-off. However, because both methods were used to derive estimates for the year 2000 [62,63], we were able to reconcile the 1990 and 2012 values by regressing estimates for 2000 produced with the 30% cut-off against estimates for 2000 produced with the 25% cut-off ($F_{1,58} = 1823$, $p < 2 \times 10^{-16}$, $R^2 = 0.9692$) and applying this correction factor ($y = 1.04x - 0.32$) to the 2012 estimates. Although satellite-derived data can produce forest cover estimates at unprecedented spatial and temporal resolutions, this method also has an important constraint: satellite data might incorrectly classify tree plantations (e.g. oil palm plantations) as forest [64]. While these plantations can have a high canopy cover, the biodiversity they contain is probably less than that in natural forest [65–68]. At the present moment, there are no global-scale forest cover data that parse forest from tree plantations (only regional datasets, e.g. [65]). To test whether our results were robust to this constraint on the forest cover data, we also considered a second, independent measure for forest: natural forest cover (i.e. excluding plantation) reported by governments to the UN Food and Agriculture Organization (FAO). These data are, in some ways, inferior to satellite-derived forest cover data, because estimates are self-reported and therefore might be inaccurate and because 'forest' definitions differ among countries. However, FAO data do parse plantation forest area from natural forest area in country-level forest cover estimates. Although we suspected that FAO forest estimates would be inferior, our use of structural equation modelling (SEM, see §2c) allowed us to use both FAO and satellite-derived forestation to inform a latent variable for 'forestation'.

Finally, we sought to control for geographic patterns set by human evolution and distribution limits for disease vectors

and reservoir hosts. Many human infectious diseases originated in Africa, and some, like Chagas disease, occur only in the Americas [69]. For this reason, we included a binary variable (Western versus Eastern Hemisphere) to assess the extent to which diseases were distributed in the New World (the Americas) versus the Old World. This variable was included in the spatial analysis only, because the temporal analysis controls for such country-level factors by comparing each country to itself.

(c) Analyses

Many of the putative drivers of disease burden correlated with one another. To address this non-independence, we used SEM (a form of path analysis) to consider relationships among drivers and between drivers and disease burden, using the software package SmartPLS (v. 3.2.4). Partial least-squares SEM (PLS-SEM) is most suitable when data do not meet normality assumptions, when there are many potential factors and when sample size is small [70]. In contrast to most statistical approaches, which assume that drivers are independent from each other, SEM allows users to designate potential relationships among drivers and to construct latent variables that result from several measures. For instance, we assumed that rain could affect forests, which could affect biodiversity, which could affect disease, and we also expected that forests and rain could each affect disease directly. We used species richness information on birds and mammals to form a composite latent variable to represent biodiversity. As described in §2b, we used two independent forest estimates (expressed as forest area divided by country area) to inform a composite latent variable called forestation. Finally, we used information on *per capita* GNI to represent wealth. (Because *per capita* spending on healthcare was nearly perfectly correlated with GNI, and so did not add extra explanatory power, we did not include it.) A spatial and temporal PLS-SEM model was constructed for each disease. Standard errors were calculated with bootstrapping. The bootstrapped subsample was used to estimate the parameters of each PLS path model 5000 times; standard errors, *t*-statistics, and *p*-values were then calculated from the 5000 estimates for each parameter [70]. This re-sampling strategy makes PLS-SEM robust to violations of normality. We then used standardized regression coefficients from the PLS-SEM output as inputs to a meta-analysis (described below), which summarized patterns across all the included diseases. We recognize that PLS-SEM is a new technique; although we believe it to be the superior choice for the reasons stated above (especially because it can explicitly account for relationships among independent variables), we also ran a more traditional panel analysis. This analysis produced qualitatively similar results and is presented in the electronic supplementary material, S4.

For each disease, we built an interaction network among the putative drivers. One network was developed for spatial correlations and a similar network was developed for temporal correlations. To reduce over-fitting and increase power, we simplified networks in two steps. First, if there was no logical reason to expect that a factor could drive a particular disease, we eliminated it *a priori*. For instance, we assumed that temperature, precipitation, forest cover and biodiversity did not affect HIV transmission (see table 1 for the *a priori* path exclusions). Second, we further pruned each model by removing the least significant paths in sequence until only marginally significant paths ($p < 0.10$) remained. After running the pruned model, we checked the adjusted R^2 to assess whether the remaining factors explained significant variation in disease burden. We used Ward's method for hierarchical cluster analysis to produce trees that indicate similarity among disease agent interaction networks. These trees were intended to summarize and illustrate the results of the PLS-SEM by indicating diseases with similar responses to putative drivers.

We then used meta-analysis on the standardized regression coefficients from PLS-SEM models to detect general patterns among diseases. To avoid biasing general patterns by excluding those coefficients that were non-significant and therefore had little influence on disease, we used coefficients from the *a priori* simplified models from which illogical paths had been excluded, but not the pruned models from which non-significant paths were excluded. We derived standard errors for each standardized regression coefficient using the standard deviation associated with the SEM model divided by the square root of the number of countries for which that model was run (table 1). We calculated a cumulative effect size for each driver across all diseases, using a random-effects model weighted by the inverse of the variance for each effect size. All meta-analyses were performed in the *metafor* package in R.

To compare the spatial and temporal results and, in particular, to assess their consistency, we plotted the mean spatial coefficient and mean temporal coefficient for each driver. Consistency in spatial and temporal results suggests robust findings. Specifically, we had more confidence in results for putative drivers for which the spatial and temporal coefficients were similar. However, we were also interested in factors for which the temporal and spatial patterns differed, because this could indicate areas where results of previous analyses—all of which have used spatial comparisons exclusively—are misleading. For example, using spatial analysis alone, we might observe that warm countries have more disease than cold countries and conclude that increasing temperature could increase disease. Observing change over time allows us to test whether that spatial pattern accurately reflects underlying processes, and could easily reveal the opposite—that increasing temperature decreases disease.

3. Results

In the spatial PLS-SEMs, there were several interactions among the putative drivers (figure 1*a*, top panels). Consistent with well-recognized patterns, high-latitude countries tended to have greater wealth, lower temperature and lower biodiversity than low-latitude countries. Wealthier countries tended to be more urbanized than poorer countries. Biodiversity was higher in forested countries and forest cover was positively associated with precipitation and negatively associated with human population density.

In the spatial PLS-SEMs, all diseases had at least one significant association in the pruned model, and more than 60% of diseases had three or four significant associations (figure 1*a*, lower right panel). Hemisphere (Western), wealth, per cent urbanization and density were significantly (and generally negatively) associated with 7–15 diseases, temperature was positively associated with seven diseases, and biodiversity was positively associated with four diseases. By contrast, associations between disease burden and precipitation or forestation were mixed, including some strongly negative and some strongly positive outcomes. Three infectious diseases (rabies, hepatitis A and varicella) showed significant spatial autocorrelation, suggesting that we over-estimated degrees of freedom when assessing *p*-values for these diseases (electronic supplementary material, S5). Furthermore, as indicated by the adjusted R^2, models failed to explain significant variation in disease burden for four diseases: hepatitis C, diphtheria, dengue and food-borne trematodiases. We note that although this approach is not well suited for assessing statistical significance for these particular infectious diseases, it does reliably reveal patterns across diseases.

Most diseases had unique path models in the spatial analysis (figure 1*a*), but there was some clustering among disease models (figure 2*a*). Malaria, schistosomiasis and

Table 1. Diseases selected for analysis from the WHO GBD database. Excluded from consideration due to low country frequency were onchocerciasis, African trypansomiasis and trachoma. Excluded from consideration due to low DALYs was yellow fever. Also indicated are the number of countries reporting each disease, transmission mode and category of each disease, any factors that were eliminated from the model based on first principles, and the sum of global DALYs in 2010. Diseases are grouped by transmission category and then placed in order of descending global DALYs in the included countries in 2010. See the electronic supplementary material, table S1 for additional details.

disease	no. countries	transmission mode	transmission category	*a priori* model simplification	DALYs 2010
malaria	60	mosquito	vector/zoonotic		8 448 885
schistosomiasis	14	snail	vector/zoonotic		524 883
lymphatic filariasis	22	mosquito	vector/zoonotic		373 381
dengue	60	mosquito	vector/zoonotic		312 308
rabies	60	animal bite	vector/zoonotic		217 773
leishmaniasis	34	sand fly	vector/zoonotic		216 541
food-borne trematodiases	24	food	vector/zoonotic		169 017
Chagas disease	20	bug	vector/zoonotic		119 545
cysticercosis	44	ingestion, autoinfection	vector/zoonotic		95 188
echinococcosis	30	ingestion	vector/zoonotic		17 222
HIV/AIDS	58	sex, IV drugs	direct	no temp, rain, biodiversity, forests	15 491 418
tuberculosis	60	contact	direct	no biodiversity, forests	8 152 330
typhoid	60	faecal contamination	direct		1 853 655
hookworm	42	contact with contaminated soil	direct	no biodiversity	983 123
hepatitis B	60	sex, IV drugs	direct	no temp, rain, biodiversity, forests	594 729
measles	60	contact	direct	no temp, rain, biodiversity, forests	581 949
ascariasis	43	ingestion	direct	no biodiversity	492 865
pertussis	60	contact	direct	no temp, rain, biodiversity, forests	480 554
whipworm	39	ingestion	direct	no biodiversity	427 150
hepatitis A	60	contact, faecal contamination	direct	no biodiversity	401 122
varicella	60	contact	direct	no temp, rain, biodiversity, forests	95 151
hepatitis C	60	IV drugs	direct	no temp, rain, biodiversity, forests	69 080
diphtheria	60	contact	direct	no biodiversity, forests	15 247
leprosy	60	contact	direct	no biodiversity	519

lymphatic filariasis all had negative associations with population density, were more burdensome in the Old World, and were more burdensome in warmer countries. Diphtheria shared these associations but also was less burdensome in urbanized countries, and cysticercosis shared these associations but was additionally less burdensome in forested countries and more burdensome in countries with high levels of biodiversity. These five diseases formed the first of three clusters. In the second cluster were most of the directly transmitted diseases, including tuberculosis, leprosy, HIV/AIDS, pertussis, measles, hookworm, hepatitis A, hepatitis B, varicella and hepatitis C, along with two zoonoses: dengue and rabies. These diseases had strong negative associations with urbanization. The final cluster contained two geo-helminth diseases (ascariasis, whipworm), one directly transmitted disease spread by faecal contamination (typhoid), and the remaining zoonoses (leishmaniasis, echinococcosis, food-borne trematodiases, Chagas disease). The geo-helminth diseases were positively associated with precipitation and the others were positively associated with biodiversity.

Disease burden and its putative drivers changed over time. As reported at the global scale by Murray *et al.* [51], DALYs per

(a) spatial analysis

full model

	latitude	wealth	pop	rain	forest	biodiv
latitude		0.78		−0.93		−0.50
wealth			0.56			
pop					−0.43	
rain					0.53	
forest						0.33

pruned model

	latitude	wealth	pop	rain	forest	biodiv
latitude		0.78		−0.93		−0.50
wealth			0.56			
pop					−0.43	
rain					0.53	
forest						0.33

	full model: west	wealth	urban	pop	temp	rain	forest	biodiv	pruned model: west	wealth	urban	pop	temp	rain	forest	biodiv	Adj R^2	p-value
zoonotic and vector–borne diseases	malaria	−0.47	0.05	−0.10	−0.24	0.67	−0.38	−0.03	0.16	−0.47			−0.27	0.75	−0.37			0.40	0.0001
	schistosomiasis	−0.41	0.00	−0.03	−0.29	0.54	−0.23	−0.21	0.16	−0.42			−0.27	0.43				0.20	0.0200
	lymphatic filariasis	−0.54	0.02	−0.06	−0.21	0.58	−0.19	−0.13	0.20	−0.55			−0.23	0.54				0.32	0.0001
	dengue	−0.09	0.06	−0.17	0.05	0.09	−0.21	0.27	0.13			−0.18				0.23		0.06	0.4400
	rabies	−0.29	−0.06	−0.35	0.06	0.06	−0.22	0.18	0.19	−0.30		−0.36					0.25	0.45	0.0001
	leishmaniasis	−0.16	−0.15	−0.14	0.23	0.03	−0.12	−0.34	0.61			−0.27				−0.45	0.55	0.30	0.0010
	food–borne trematodiasis	0.17	−0.23	0.15	0.16	−0.09	−0.26	0.16	0.26								0.27	0.06	0.2100
	Chagas disease	0.42	0.01	−0.03	−0.11	−0.14	0.20	−0.11	0.28	0.42							0.26	0.32	0.0001
	cysticercosis	−0.21	−0.12	−0.15	−0.39	0.26	0.03	−0.26	0.30	−0.26			−0.38	0.45		−0.28	0.34	0.33	0.0001
	echinococcosis	−0.12	−0.63	0.01	−0.09	−0.30	−0.38	−0.30	0.10		−0.46				−0.54	−0.23		0.38	0.0001
non–zoonoses	HIV/AIDS	−0.11	−0.19	−0.19	−0.30						−0.14	−0.22	−0.30					0.13	0.0200
	tuberculosis	−0.40	−0.15	−0.34	−0.19	0.39	−0.23			−0.39		−0.38	−0.20	0.49	−0.24			0.46	0.0010
	typhoid	−0.45	−0.25	−0.28	0.21	0.18	−0.13	−0.12	0.33	−0.44	−0.31	−0.31	0.23				0.26	0.54	0.0010
	hookworm	−0.14	−0.10	−0.28	−0.04	0.10	0.33	0.00				−0.38			0.36			0.31	0.0020
	hepatitis B	−0.32	−0.27	−0.38	0.04					−0.32	−0.27	−0.38						0.36	0.0010
	measles	−0.38	−0.38	−0.03	−0.22					−0.39	−0.40		−0.22					0.20	0.0010
	ascariasis	−0.19	−0.23	0.09	0.17	0.23	0.18	0.13			−0.27				0.31			0.21	0.0020
	pertussis	−0.29	−0.23	−0.28	−0.25					−0.29	−0.23	−0.28	−0.25					0.23	0.0010
	whipworm	−0.22	−0.16	0.16	−0.05	0.06	0.50	−0.07							0.45			0.19	0.0450
	hepatitis A	−0.27	−0.05	−0.47	−0.11	−0.11	0.20	0.07		−0.20		−0.49						0.26	0.0010
	varicella	−0.25	−0.15	−0.33	0.02					−0.20		−0.42						0.19	0.0400
	hepatitis C	0.14	−0.08	−0.20	0.03							−0.25						0.04	0.6300
	diptheria	−0.33	0.11	−0.18	−0.26	0.47	−0.14			−0.35		−0.16	−0.26	0.31				0.18	0.1300
	leprosy	−0.46	−0.15	−0.23	−0.09	0.45	−0.32	0.20		−0.51				0.72	−0.38	0.24		0.39	0.0010

(b) temporal analysis

full model

	latitude	wealth	pop	rain	forest	biodiv
latitude						
wealth						
pop						
rain						
forest						0.84

pruned model

	latitude	wealth	pop	rain	forest	biodiv
latitude						
wealth						
pop						
rain						
forest						0.84

	full model: wealth	urban	pop	temp	rain	forest	biodiv	pruned model: wealth	urban	pop	temp	rain	forest	biodiv	Adj R^2	p-value
zoonotic and vector–borne diseases	malaria	−0.13	−0.10	0.32	0.09	−0.04	−0.08	0.10			0.27					0.06	0.3380
	schistosomiasis	−0.16	0.09	0.30	0.14	0.23	0.06	−0.01			0.29	0.17	0.23			0.04	0.6340
	lymphatic filariasis	0.12	0.01	0.59	0.40	0.10	0.55	−0.65			0.56	0.38		0.56	−0.68	0.31	0.0040
	dengue	0.21	−0.16	−0.08	−0.30	−0.04	0.22	−0.19				−0.28				0.06	0.0560
	rabies	−0.21	−0.33	0.15	0.32	0.05	0.23	0.05		−0.20	−0.30		0.29			0.26	0.0020
	leishmaniasis	−0.54	−0.27	0.07	0.07	0.11	0.36	−0.21		−0.56	−0.26					0.4	0.0260
	food–borne trematodiasis	0.13	−0.03	−0.02	−0.45	0.03	−0.23	0.40					−0.47			0.21	0.23	0.1380
	Chagas disease	−0.22	0.02	−0.14	−0.04	0.08	−0.07	0.13				−0.17					0.01	0.7210
	cysticercosis	−0.27	0.19	0.32	0.38	−0.11	0.33	−0.37		−0.24		0.34	0.37				0.17	0.0290
	echinococcosis	−0.13	−0.22	0.37	0.19	−0.16	0.23	−0.16							−0.24		0.04	0.5520
non–zoonoses	HIV/AIDS	0.29	0.02	0.11						0.30							0.08	0.0950
	tuberculosis	−0.27	−0.06	0.12	−0.03	0.17												n.s.	
	typhoid	0.05	0.05	0.28	−0.09	−0.04	0.15	−0.20				0.34					0.1	0.1900
	hookworm	0.06	−0.10	−0.08	−0.05	−0.13	−0.17											n.s.	
	hepatitis B	−0.16	−0.07	0.25						−0.17		0.24					0.05	0.4900	
	measles	0.17	−0.48	−0.31							−0.45	−0.31					0.32	0.0100	
	ascariasis	−0.12	0.00	0.49	0.22	0.26	0.32					0.50	0.25	0.29	0.34		0.3	0.0001	
	pertussis	0.25	−0.24	0.00							−0.21						0.03	0.6450	
	whipworm	−0.06	0.01	0.38	0.10	0.34	0.02					0.34		0.34			0.14	0.0740	
	hepatitis A	−0.39	−0.05	0.03	0.33	0.14	−0.01			−0.40			0.34				0.31	0.0001	
	varicella	−0.21	−0.39	−0.50						−0.21	−0.39	−0.50					0.53	0.0010	
	hepatitis C	−0.17	0.07	−0.10														n.s.	
	diphtheria	0.20	−0.41	−0.11	0.03	−0.05					−0.40						0.14	0.0700	
	leprosy	0.16	0.16	0.02	−0.31	−0.06	0.41						−0.37		0.41		0.26	0.0010	

Figure 1. Standardized regression coefficients from PLS-SEM full models (including all pathways except those deemed to be illogical from first principles; table 1) and pruned models (including only those pathways that were significant) for each of the 24 infectious diseases in the (*a*) spatial and (*b*) temporal analyses. The top box indicates the interactions among the drivers, where the rows are independent variables and the columns are dependent variables. Values are standardized regression coefficients and colours correspond to coefficient values (red, negative; blue, positive). Diseases are sorted into two groups (zoonoses + vector-borne diseases versus non-zoonoses) and by descending total global DALYs in the included countries in 2010 within those two groups. Drivers are sorted left from right by their tendency for positive (blue) or negative (red) coefficients in space. The bottom boxes show the direct effects of a driver on disease burden. In the full model panel, blank cells indicate untested associations assumed to be zero based on first principles. An underlined disease is one for which there was not a significant reduced model as determined by the adjusted R^2 (shown at right). The pruned model panel shows results after performing a model selection process in which any coefficient with an associated p-value < 0.10 was removed and the model re-run. Bold-font coefficients are significant.

100 000 persons declined between 1990 and 2010 (paired t-test: $t_{59} = -3.8652$, $p = 0.0006$; figure 3*i*). The satellite-derived forest cover metric and the FAO forest cover metric were strongly associated with one another in both 1990 ($t_{58} = 18.28$, $p < 0.0001$) and 2010 ($t_{58} = +15.28$, $p < 0.0001$). Between 1990 and 2010, forest cover remained stable at the

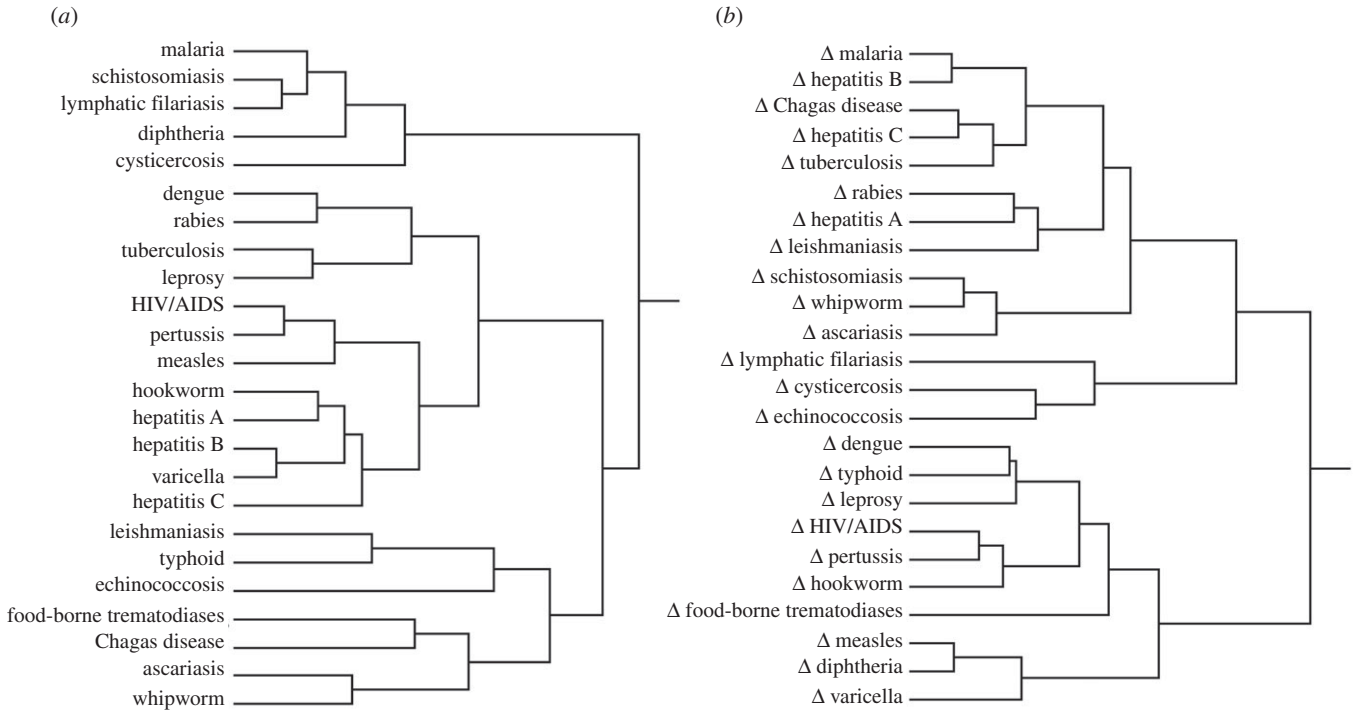

Figure 2. Clustering diseases by their associated drivers in the (*a*) spatial and (*b*) temporal analysis. This figure was produced by hierarchical Ward clustering based on the standardized regression coefficients in figure 1, where we assumed that illogical drivers had a coefficient = 0.

global level for both satellite-derived (paired *t*-test: $t_{59} = -0.8553$, $p = 0.3959$; figure 3*a*) and FAO (paired *t*-test: $t_{59} = -1.4289$, $p = 0.1583$; figure 3*b*) forest cover metrics; however, this pattern conceals substantial variability among individual countries. Countries underwent change in forest area (absolute change in forest area > 0 for satellite-derived (one sample *t*-test: $t_{59} = +5.8688$, $p < 0.0001$) and FAO (one sample *t*-test: $t_{59} = +5.0320$, $p < 0.0001$) forest data), but these increases and decreases cancel one another out at the global level. GNI (paired *t*-test: $t_{59} = 14.7734$, $p < 0.0001$; figure 3*c*), healthcare spending (paired *t*-test: $t_{59} = 15.1876$, $p < 0.0001$; figure 3*d*), population (paired *t*-test: $t_{59} = 10.5706$, $p < 0.0001$; figure 3*e*), proportion of population in cities (paired *t*-test: $t_{59} = 6.2393$, $p < 0.0001$; figure 3*f*), average precipitation (paired *t*-test: $t_{59} = 3.7101$, $p = 0.0005$; figure 3*g*) and average temperature (paired *t*-test: $t_{59} = 19.8232$, $p < 0.0001$; figure 3*h*) increased over time.

There were fewer significant temporal associations than spatial associations. This is, in part, because there was less temporal than spatial variation in disease burden and drivers. In the only interaction among temporal drivers (figure 1*b*, top panels), increasing forestation was associated with increasing biodiversity (but this was due, at least in part, to how we calculated biodiversity change). Three diseases lacked a single significant driver in the pruned model, and 14 had only one or two significant drivers (figure 1*b*, lower right panel). Change in wealth, urbanization and population had strong associations with change in disease burden. Urbanization change was negatively associated with change in burden of six diseases, wealth change was negatively associated with six diseases (and positively associated with one, HIV/AIDS), and population change was positively associated with eight diseases (and negatively associated with three). Disease burden tended to increase as temperature increased for six diseases, but it declined as disease burden increased for three other diseases. Precipitation, forestation and biodiversity change had few significant associations with disease burden. Disease burden increased as precipitation increased for two of the geo-helminth diseases (ascariasis, whipworm) and the water-borne disease schistosomiasis. Increasing biodiversity was associated with increasing disease burden of food-borne trematodiases and decreasing disease burden of lymphatic filariasis. As indicated by the adjusted R^2, the PLS-SEMs explained significant variance for lymphatic filariasis, dengue, rabies, leishmaniasis, cysticercosis, tuberculosis, hookworm, measles, ascariasis, whipworm, hepatitis A, varicella and leprosy.

In the temporal analysis, path models for disease change clustered into three broad groups (figure 2*b*). Diseases in the first cluster were united by generally weak associations with most drivers, with some responding positively to population density and precipitation (malaria, hepatitis B, Chagas disease, hepatitis C, tuberculosis, rabies, hepatitis A, leishmaniasis, schistosomiasis, whipworm, ascariasis). Within this first cluster, rabies, leishmaniasis and hepatitis A shared a negative association with increasing wealth, and schistosomiasis, whipworm and ascariasis shared a positive association with increasing precipitation. The second cluster increased as population density and temperature increased (lymphatic filariasis, cysticercosis and echinococcosis). The third cluster contained many of the directly transmitted diseases (typhoid, leprosy, HIV/AIDS, pertussis, hookworm, measles, diphtheria and varicella) along with a few zoonoses (dengue, food-borne trematodiases) and tended to decrease as urbanization increased. The two zoonoses in this group decreased as temperature increased.

Comparing the output of the temporal and spatial PLS-SEMs for urbanization, wealth, precipitation, temperature, population and biodiversity, there was some agreement

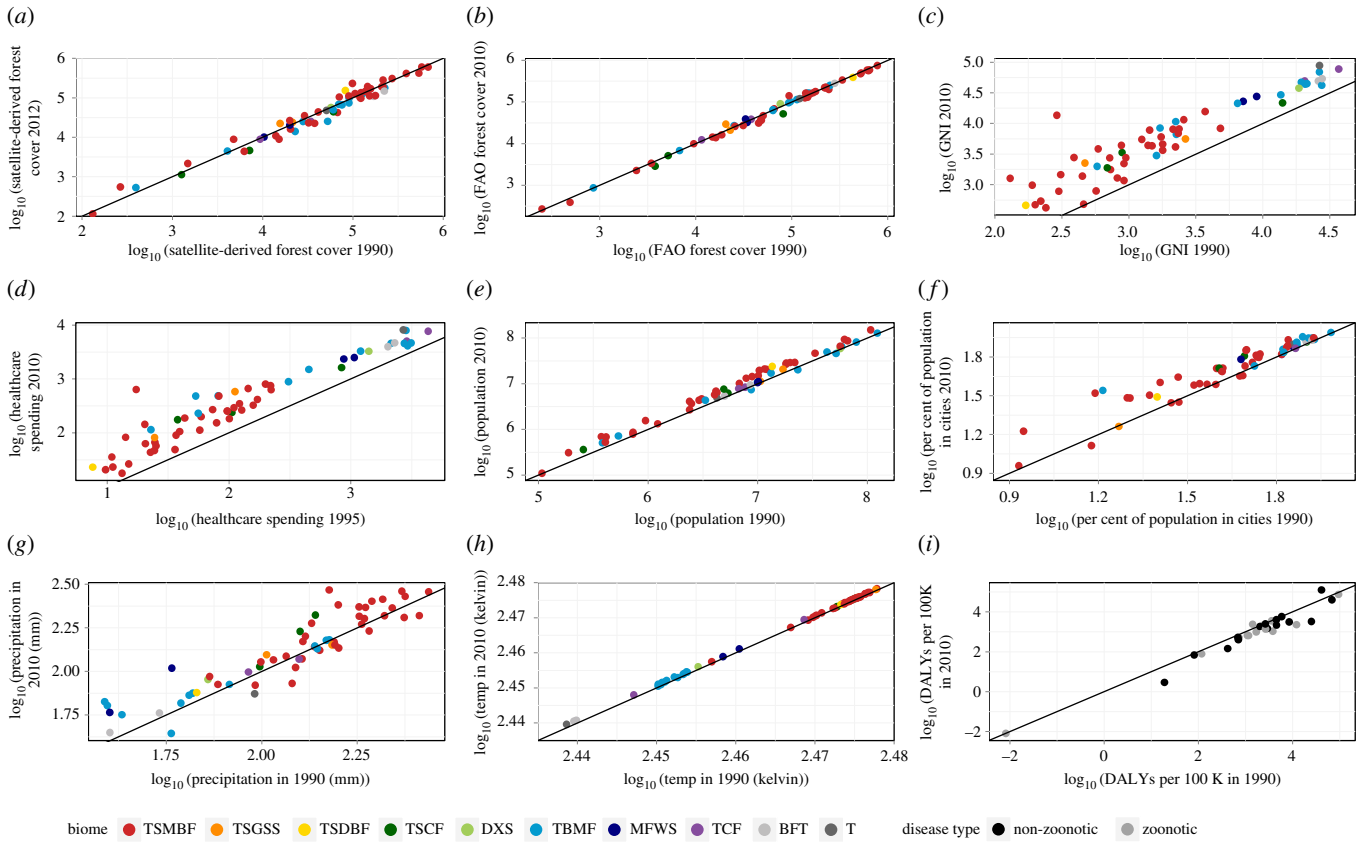

Figure 3. Trends over time. A 1 : 1 line is indicated in each plot to represent zero change; data points falling above this line represent increases over time, and those falling below the line represent decreases. Colours represent biomes in (*a*–*h*) (see legend). In *i*, black, non-zoonosis; grey, zoonosis. Biome codes: TSMBF, tropical and subtropical moist broadleaf forests; TSGSS, tropical and subtropical grasslands, savannas and shrublands; TSDBF, tropical and subtropical dry broadleaf forest; TSCF, tropical and subtropical coniferous forests; DXS, deserts and xeric shrublands; TBMF, temperate broadleaf and mixed forests; MFWS, Mediterranean forests, woodlands and scrub; TCF, temperate conifer forests; BFT, boreal forest/taiga; T, tundra.

between the net effects in space and time, with departures from agreement suggesting 'hidden' effects of the putative drivers (i.e. effects that are not detectable without data from multiple time points; figure 4). Countries with high biodiversity tended to have higher disease burdens ($p < 0.0001$; table 2), but increasing biodiversity over time had a non-significant negative effect on disease burden ($p = 0.2214$; table 2). There was no net forest cover association with disease burden in the spatial analysis ($p = 0.4729$; table 2), but countries with increasing forest cover experienced increases in disease burden ($p = 0.0081$; table 2). Urbanization had the strongest and most straightforward effect of all the putative drivers: urbanized countries had less disease in the spatial analysis ($p < 0.0001$; table 2) and increasing urbanization was associated with declining disease burden over time ($p = 0.0094$; table 2). Densely populated countries had less disease than countries with sparse populations in the spatial analysis ($p = 0.0259$; table 2), but increasing population had a non-significant positive association with change in disease burden over time ($p = 0.0552$; table 2). Wealthy countries had less disease ($p < 0.0001$), but increasing wealth did not correlate with disease change over time ($p = 0.1981$; table 2), although it was an important effect for many diseases individually (figure 1*b*). Countries with warmer climates had more disease than countries with cooler climates ($p = 0.0030$), but increasing temperature did not produce increases in disease burden ($p = 0.3256$; table 2), although—again—it was an important effect for many diseases individually (figure 1*b*). Precipitation had neither spatial ($p = 0.2050$) nor temporal ($p = 0.1535$; table 2) associations with disease burden.

4. Discussion

This analysis supports the importance of demography and economics in determining infectious disease outcomes at the country level (e.g. [14]), and suggests that environmental factors such as forestation and biodiversity are unlikely to be general levers for disease control. Although there were substantial differences among diseases in the putative drivers associated with the spatial distribution of disease burden and change in burden over time, there were some associations that held across many diseases. We consider the advantages and disadvantages of country-level analyses, the relationships among the putative drivers, and the importance of each driver in determining disease burden, and end by identifying commonalities among disease agents.

Country-level analyses facilitate global-scale insight but have important constraints. One limitation is that countries vary by many orders of magnitude in size, and both ecological and economic processes might operate differently in small versus large countries. To avoid mixing scales—and possibly losing the ability to detect important patterns—we limited our analysis to 60 intermediate-sized countries (more than 30% of all countries). This approach sidesteps the 'ecological fallacy', a possible source of error in country-level analyses. However, it also limits scope of inference, making our results

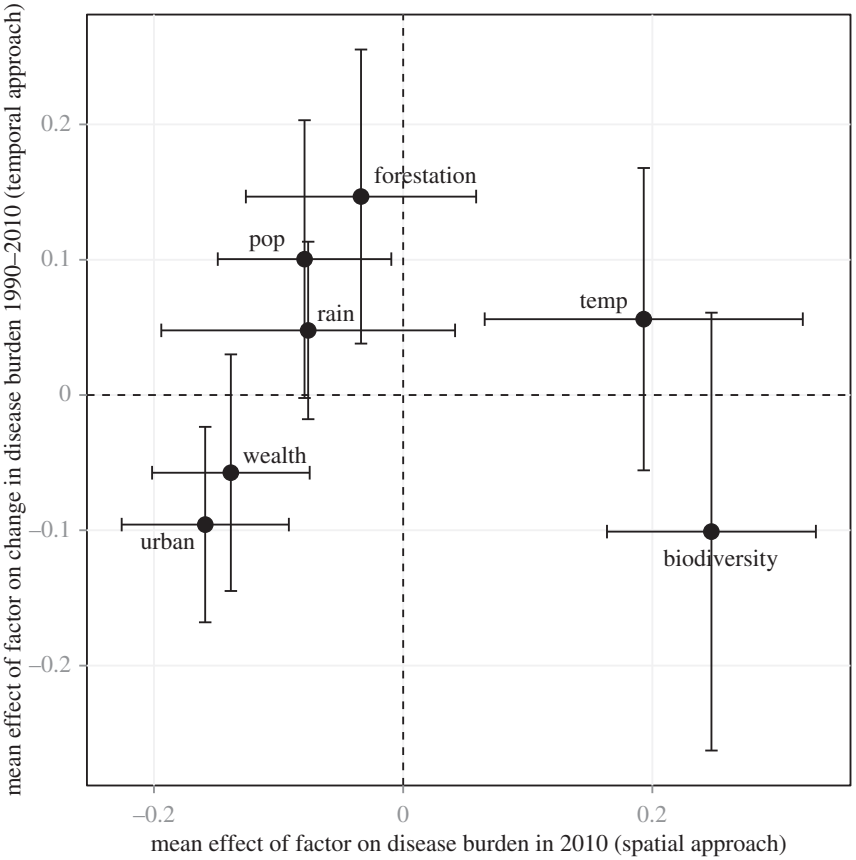

Figure 4. Results of spatial (*x*-axis) and temporal (*y*-axis) meta-analyses, which summarize the results of PLS-SEM across 24 diseases. Points represent mean effect sizes for the effects of wealth, per cent of population living in urban environments, population density, forest cover, temperature, precipitation and biodiversity. Error bars represent 95% confidence intervals.

Table 2. Results of PLS-SEM.

	spatial				temporal			
parameter	estimate	s.e.	*z*	*p*-value	estimate	s.e.	*z*	*p*-value
forestation	−0.0338	0.0471	−0.7177	0.4729	0.1467	0.0554	2.6463	0.0081
biodiversity	0.2474	0.0428	5.7828	<0.0001	−0.1010	0.0826	−1.2227	0.2214
urbanization	−0.1589	0.0342	−4.6391	<0.0001	−0.0958	0.0369	−2.5976	0.0094
population	−0.0791	0.0355	−2.2271	0.0259	0.1004	0.0524	1.9175	0.0552
wealth	−0.1383	0.0322	−4.2889	1.80×10^{-5}	−0.0574	0.0446	−1.2870	0.1981
temperature	0.1931	0.0652	2.9632	0.0030	0.0561	0.05703	0.9830	0.3256
precipitation	−0.0762	0.0601	−1.2674	0.2050	0.0478	0.0335	1.4272	0.1535

relevant only for intermediate-sized countries. Although it is unfortunate that we could not include large countries in our analysis, especially considering that they contain the bulk of the world's people, biodiversity and forest, we believe that including large countries might have obscured patterns and biased results (see above). Countries that were excluded for reasons other than size were those where data were sparse—including many Soviet bloc (Belarus, Estonia, Georgia, Hungary, Latvia, Lithuania, Poland, Russian Federation, Serbia, Slovakia and Slovenia) and Eastern European nations (Bosnia/Herzegovina, Croatia, Czech Republic, Macedonia, Montenegro and Serbia), along with countries experiencing political upheaval or repression around the year 1990 (Cambodia, Democratic Republic of the Congo, Liberia, Myanmar, North Korea and Timor-Leste). Because these countries were not included in our analysis, our conclusions cannot be extrapolated to Eastern Europe or to countries experiencing active conflict. Our approach does, however, allow us to examine the relationships between putative drivers and disease outcomes for other nations at the scale most relevant for policy-making: the country level.

Structural equation models allowed us to account for collinearity and investigate relationships among drivers. In short, many factors were associated with the high disease burden observed in tropical countries. Low-latitude nations were poorer, warmer and possessed greater biodiversity than high-latitude nations, consistent with well-known epidemiological and biogeographic patterns. People in wealthy

nations were also more likely to live in cities, an effect that appears to reduce disease, contrary to the intuition that increasing human density driven by urbanization should increase disease transmission (see below). Furthermore, through the direct effects indicated by the coefficients in figure 1, we can also estimate indirect and total effects for each driver. For instance, although there was no significant direct association between precipitation and leishmaniasis, an indirect negative association between precipitation and leishmaniasis (−0.18) can be estimated by multiplying the direct association between precipitation on forestation (+0.53) and the direct association between forestation on leishmaniasis (−0.34).

Both the spatial and temporal analyses identified a negative association between urbanization and disease (figure 4), as did our parallel generalized linear mixed model panel analysis (electronic supplementary material, S4). This result contradicts the common expectation that urbanization should increase the risk of disease outbreaks (e.g. [15,71]). Urbanization might reduce disease by increasing sanitation through enhanced availability of sewers and water distribution systems, reducing distance to medical facilities, allowing public health campaigns (e.g. spraying for mosquitoes, mass drug administration, vaccination) to be carried out more efficiently and effectively, facilitating the transfer of public health information or separating human populations from disease vectors and reservoirs living in forested areas. That urbanization might reduce disease is good news for a world that is rapidly urbanizing [72].

Population density had a less straightforward association with disease burden. In the spatial analysis, more densely populated countries had lower disease burdens, but the temporal analysis suggested a non-significant increase in disease with increasing population density. Within the temporal analysis, eight diseases increased with increasing population density, but two highly infectious directly transmitted diseases (measles and varicella) decreased in burden with increasing population density. This finding runs counter to predictions based on density-dependent disease transmission [17]. For some diseases, this could arise due to an encounter-dilution effect; that is, if the number of infectious propagules is held constant or increases less rapidly than does host population density, then increasing population density can dilute *per capita* risk [73]. This mechanism would be most likely to operate for zoonoses or vector-borne diseases because zoonotic propagules or infected vectors are at least to some extent independent from human density. In our study, this mechanism could explain the negative association between burden of Chagas disease and population density. For the directly transmitted diseases such as measles and varicella, contacts might not increase linearly with host density because contact rates with new susceptibles saturate as density increases; this pattern has been posited on theoretical grounds [20] and supported by experiments [74] and the observation that the number of zoonotic disease outbreaks *per capita* declines with increasing population density in Asia [7] (see our re-analysis in the electronic supplementary material, S6). Furthermore, DALYs measure health impacts, which could also vary with population density. For instance, healthcare might be more effective where people are concentrated, reducing disease burden due to complications from infection [75]. In sum, our results for the demographic drivers suggest that global trends toward increasing urbanization might, counterintuitively, bring reductions in disease burden, at least for long-established human diseases.

Economic factors were also important in determining disease burden. Wealth, which reflects *per capita* GNI, was negatively associated with seven diseases in the spatial analysis and six diseases in the temporal analysis. This pattern might arise from the improvements in sanitation, healthcare infrastructure and transportation that accompany economic growth, and it might be reinforced where disease burden impairs economic growth [11]. Only one disease diverged from this pattern: although HIV/AIDS was less burdensome in wealthier countries, it increased in association with increasing wealth in the temporal analysis. This might be due to reversed causation, where increasing HIV/AIDS triggers increased foreign aid to healthcare programmes designed to reduce disease transmission. This change could also be due to increased surveillance and reporting made possible by increased wealth. For instance, studies on disease outbreaks per country (from the GIDEON database) often find a positive relationship with wealth, presumably due to better reporting and surveillance with increasing wealth [7,8]. Another possible interpretation is that a time lag between economic growth and disease burden obscures their relationship. However, the most plausible explanation for this unexpected pattern is that the poorest countries simultaneously possess the heaviest HIV/AIDS disease burden and the greatest potential proportional change in wealth; that is, for the poorest countries, even a small absolute change in wealth is a large proportional change.

Hotter countries had a greater disease burden than cooler countries (figure 1*a*), but change in temperature was not significantly associated with change in disease burden. In the temporal analysis, increasing temperature had significant positive associations with six diseases, and significant negative associations with three diseases (figure 1*b*). This is consistent with the expectation that changing climate might either increase or decrease disease burden, depending on the disease agent: its transmission strategy, life cycle requirements, host distributions and nonlinear effects of temperature on physiological performance [28]. For example, increasing temperature might create benign environmental conditions that increase vector abundance (e.g. [76]), or it might create harsh environmental conditions that exceed the thermal optimum of vectors or parasite infectious stages, or increase evaporation and therefore desiccation. Although our data do not support a strong role for climate change in infectious disease burden, relationships between disease and climate are best assessed at local scales, due to the nonlinear responses of disease to temperature and because climate can vary substantially within a country [77]. In other words, our analysis should not imply that diseases do not respond to climate, only that we have not seen a general global climate change effect on disease between 1990 and 2010.

Rainfall showed little net association with overall disease burden (figure 4), largely because it had variable associations across disease agents, with some negatively associated and others positively associated. In the spatial analysis, wetter countries had higher burdens of the three geo-helminth diseases (hookworm, whipworm and ascariasis; figure 1*a*), perhaps because geo-helminth eggs must embryonate in moist soil before they are competent to infect another human host. Increasing precipitation over time increased the burden of ascariasis and whipworm, as well as schistosomiasis, possibly because increased rainfall increases run-off of human waste carrying schistosome eggs, delivering eggs to freshwaters where they can infect intermediate host snails (figure 1*b*). On the other

hand, rain had strong negative effects on diseases such as echinococcosis (figure 1*a*,*b*), perhaps because run-off removes *Echinococcus* eggs from the terrestrial habitats where they might be encountered by intermediate hosts.

As we observed with rainfall, associations between diseases and biodiversity were mixed. Countries with more biodiversity per unit area also had greater disease burdens, and increasing biodiversity over time was associated with a non-significant decrease in disease across the disease agents (figure 4), primarily driven by the strong negative response of lymphatic filariasis to increasing biodiversity. The finding that disease is concentrated in biodiverse countries is not new (e.g. [6]), but our analysis is the first to assess the association between biodiversity and disease over time, and it fills an important gap: of all the studies to date that have tested the effects of conservation on human well-being, only 2% address infectious disease [44]. This part of the analysis also reveals a previously hidden effect of biodiversity. To date, all analyses relating biodiversity to human disease burden have used data from only a single time point; our analysis shows that this could be misleading (figure 4), because spatial and temporal patterns diverge strongly for biodiversity. This could suggest that the degree of 'native' biodiversity in a country determines the baseline burden of infection (with high-biodiversity areas experiencing higher burdens than low-biodiversity areas), and biodiversity loss modulates that baseline level (via dilution or amplification effects). However, our results find no general, large-scale dilution effect across disease agents. Instead, the data suggest that disease agents are generally unresponsive to changes in biodiversity: only two diseases of 24 retained biodiversity as a significant driver of disease burden in temporal models, and the effect was positive for one disease (food-borne trematodiases) and negative for the other (lymphatic filariasis; figure 1*b*). Just as for climate effects, relationships between biodiversity and disease transmission are local-scale phenomena that can be swamped, or even reversed, by other processes at larger spatial scales [78]. Although we cannot assess local-scale relationships, it seems clear (and in contrast to past studies [5,8]) that the hypothesized negative link between biodiversity and disease does not apply generally at the country scale.

Disease burden was similar in heavily forested countries when compared with lightly forested countries (figure 1*a*), but increasing forestation over time was correlated with increases in disease burden (figure 1*b*), driven primarily by lymphatic filariasis and other zoonoses, but also by ascariasis (a geo-helminth) and leprosy (a directly transmitted disease). Forest may provide habitat for vectors (e.g. mosquitoes that carry lymphatic filariasis), create environmental conditions amenable for parasite development (e.g. moist soil for *Ascaris* eggs) or be associated with living conditions that facilitate close contact (e.g. small communities affected by leprosy). The fact that lymphatic filariasis is positively associated with forest and negatively associated with biodiversity suggests a tension between two forces influencing disease burden: increasing forest habitat might facilitate vector populations, while increasing biodiversity might reduce the proportion of reservoir hosts in the vertebrate population. This tension is well recognized in the disease ecology literature (e.g. [21,79,80]), and highlights the difficulty of using conservation as a public health tool. Given that increasing forestation tends to increase zoonotic disease burden, forest conservation would seem to be a win–lose approach to conservation and public health. As deforestation and forest conservation projects proceed, potential disease-related advantages or collateral impacts should be incorporated into cost–benefit analyses [81]; where collateral impacts are expected, projects should plan for increased surveillance, prophylaxis and treatment [17], or better separation between forested areas and human settlements.

It has been posited that parasite life cycles can be used to predict their response to anthropogenic environmental changes such as pollution [82], resource extraction [83–86], global warming [28] and biodiversity loss [22]. In the temporal dendrogram (figure 2*a*), several diseases (typhoid, leprosy, HIV/AIDS, pertussis, hookworm, measles, diphtheria and varicella) clustered due to similarities in their responsiveness to urbanization, and the geo-helminths clustered due to shared association with precipitation. These diseases might converge in their response to drivers due to their shared transmission strategies, but this is the exception rather than the rule: most disease responses were idiosyncratic (*sensu* [34]), and seemingly similar disease agents often did not respond to the same drivers. This variability suggests that, although some general drivers (e.g. urbanization) tend to reduce infectious disease, we should expect no silver bullet for infectious disease control; each disease poses a unique challenge to public health.

5. Conclusion

Despite effective medical treatments for most infectious diseases, infection drives global morbidity and death. For the countries we analysed, infectious disease burdens correlate with human demography and economics, whereas associations with environmental and biotic factors are less clear-cut. The oft-hypothesized negative relationship between biodiversity and disease was not supported by our data; where biodiversity change did associate with disease burden change, it had mixed effects. Most policy decisions are made at a country level, and our analysis shows that—at this scale—conserving biodiversity may not be an effective lever for improving public health. If there is any broad-brush approach that would simultaneously reduce the burden of multiple infectious diseases, it appears to be the promotion of urbanization and economic development.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material. Software used for PLS-SEM analysis SmartPLS (v. 3.2.4) can be obtained at http://www.smartpls.de/.

Authors' contributions. C.L.W., A.M., H.S.Y. and K.D.L. conceived the research question and designed the approach, gathered the data, analysed the data, and wrote the paper. D.K. provided data and revised the manuscript draft. All authors approved the version to be published.

Competing interests. We have no competing interests.

Funding. C.L.W. was supported by a Michigan Fellowship from the Michigan Society of Fellows and the Department of Ecology and Evolutionary Biology at the University of Michigan. A.M. was supported by the Department of Environmental Science, Policy, and Management at the University of California Berkeley. H.S.Y. was supported by NSF DEB 1556786. D.K. was supported by the Department of Geographical Sciences at the University of Maryland. K.D.L. was supported by the US Geological Survey.

Acknowledgements. The authors acknowledge Ralf Meisenzahl (choosing economic indicators), Jeff Vincent (parsing forest cover from plantation cover), Ian Wang (developing statistical approaches to spatial data) and Maggie Klope (figure in the electronic supplementary material, S3).

Disclaimer. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

References

1. World Health Organization 2016 *World Health Statistics 2016: Monitoring Health for the SDGs* (*Sustainable Development Goals*). Geneva, Switzerland: World Health Organization.
2. World Health Organization 2008 *The global burden of disease: 2004 update*. Geneva, Switzerland: World Health Organization.
3. Guernier V, Hochberg ME, Guégan J-F. 2004 Ecology drives the worldwide distribution of human diseases. *PLoS Biol*. **2**, e141. (doi:10.1371/journal.pbio.0020141
4. Bonds MH, Dobson AP, Keenan DC. 2012 Disease ecology, biodiversity, and the latitudinal gradient in income. *PLoS Biol*. **10**, e1001456. (doi:10.1371/journal.pbio.1001456)
5. Hay SI *et al.* 2013 Global mapping of infectious disease. *Phil. Trans. R. Soc. B* **368**, 20120250. (doi:10.1098/rstb.2012.0250)
6. Dunn RR, Davies TJ, Harris NC, Gavin MC. 2010 Global drivers of human pathogen richness and prevalence. *Proc. R. Soc. B* **277**, 2587–2595. (doi:10.1098/rspb.2010.0340)
7. Morand S, Jittapalapong S, Suputtamongkol Y, Abdullah M, Huan T. 2014 Infectious diseases and their outbreaks in Asia-Pacific: biodiversity and its regulation loss matter. *PLoS ONE* **9**, e90032. (doi:10.1371/journal.pone.0090032)
8. Jones KE, Patel NG, Levy M, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008 Global trends in emerging infectious diseases. *Nature* **451**, 990–993. (doi:10.1038/nature06536)
9. Morand S, Waret-Szkuta A. 2012 Determinants of human infectious diseases in Europe: biodiversity and climate variability influences. *Bull. Epidemiol. Hebdomadaire* **12**, 156–159.
10. Murray KA, Preston N, Allen T, Zambrana-Torrelio C, Hosseini PR, Daszak P. 2015 Global biogeography of human infectious diseases. *Proc. Natl. Acad. Sci. USA* **112**, 12 746–12 751. (doi:10.1073/pnas.1507442112)
11. Bonds MH, Keenan DC, Rohani P, Sachs JD. 2010 Poverty trap formed by the ecology of infectious diseases. *Proc. R. Soc. B* **277**, 1185–1192. (doi:10.1098/rspb.2009.1778)
12. Ngonghala CN, Plucinski MM, Murray MB, Farmer PE, Barrett CB, Keenan DC, Bonds MH. 2014 Poverty, disease, and the ecology of complex systems. *PLoS Biol*. **12**, e1001827. (doi:10.1371/journal.pbio.1001827)
13. Ezzati M *et al.* 2003 Estimates of global and regional potential health gains from reducing muliple major risk factors. *Lancet* **362**, 271–280.
14. Smith KR, Ezzati M. 2005 How environmental health risks change with development: The epidemiologic and environmental health risk transitions revisited. *Annu. Rev. Environ. Resour*. **30**, 291–333. (doi:10.1146/annurev.energy.30.050504.144424)
15. Foley JA *et al.* 2005 Global consequences of land use. *Science* **309**, 570–574. (doi:10.1126/science.1111772)
16. Garchitorena A *et al.* 2017 Disease ecology, health and the environment: a framework to account for ecological and socio-economic drivers in the control of neglected tropical diseases. *Phil. Trans. R. Soc. B* **372**, 20160128. (doi:10.1098/rstb.2016.0128)
17. Anderson RM, May RM. 1978 Regulation and stability of host–parasite population interactions: I. Regulatory processes. *J. Anim. Ecol*. **47**, 219–247. (doi:10.2307/3933)
18. Patterson JH, Ruckstuhl KE. 2013 Parasite infection and host group size: a meta-analytical review. *Parasitology* **140**, 803–813. (doi:10.1017/S0031182012002259)
19. Rifkin JL, Nunn CL, Garamszegi LZ. 2012 Do animals living in larger groups experience greater parasitism? A meta-analysis. *Am. Nat*. **180**, 70–82. (doi:10.1086/666081)
20. Heesterbeek J, Metz J. 1993 The saturating contact rate in marriage- and epidemic models. *J. Math. Biol*. **31**, 529–539. (doi:10.1007/BF00173891)
21. Valle D, Clark J. 2013 Conservation efforts may increase malaria burden in the Brazilian Amazon. *PLoS ONE* **8**, e57519. (doi:10.1371/journal.pone.0057519)
22. Wood CL, Lafferty KD, DeLeo GA, Young HS, Hudson P, Kuris AM. 2014 Does biodiversity protect humans against infectious disease? *Ecology* **95**, 817–832. (doi:10.1890/13-1041.1)
23. Whittles L, Didelot X. 2016 Epidemiological analysis of the Eyam plague outbreak of 1665–1666. *Proc. R. Soc. B* **283**, 20160618. (doi:10.1098/rspb.2016.0618)
24. Murray CJ, Lopez AD, Chin B, Feehan D, Hill KH. 2007 Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *Lancet* **368**, 2211–2218. (doi:10.1016/S0140-6736(06)69895-4)
25. Phillips D. 1993 Urbanization and human health. *Parasitology* **106**, S93–S107. (doi:10.1017/S0031182000086145)
26. Neiderud C-J. 2015 How urbanization affects the epidemiology of emerging infectious diseases. *Infect. Ecol. Epidemiol One Health J*. **5**, 27060. (doi:10.3402/iee.v5.27060)
27. Gualdieri L, Piemonte M, Alfano S, Maffei R, Della Pepa ME, Rinaldi L, Galdiero M, Galdiero M, Cringoli G. 2016 Immigrants living in an urban *milieu* with sanitation in Southern Italy: persistence and transmission of intestinal parasites. *Parasitol. Res*. **115**, 1315–1323. (doi:10.1007/s00436-015-4868-2)
28. Lafferty KD. 2009 The ecology of climate change and infectious diseases. *Ecology* **90**, 888–900. (doi:10.1890/08-0079.1)
29. Estrada-Pena A, Ostfeld RS, Peterson AT, Poulin R, de la Fuente J. 2014 Effects of environmental change on zoonotic disease risk: an ecological primer. *Trends Parasitol*. **30**, 205. (doi:10.1016/j.pt.2014.02.003)
30. Wood CL, Lafferty KD. 2013 Biodiversity and disease: a synthesis of opposing ecological perspectives on Lyme disease transmission. *Trends Ecol. Evol*. **28**, 239–247. (doi:10.1016/j.tree.2012.10.011)
31. Ostfeld RS, Keesing F. 2013 Straw men don't get Lyme disease: response to Wood and Lafferty. *Trends Ecol. Evol*. **28**, 502–503. (doi:10.1016/j.tree.2013.05.009)
32. Lafferty K, Wood C. 2013 It's a myth that protection against disease is a strong and general service of biodiversity conservation: response to Ostfeld and Keesing. *Trends Ecol. Evol*. **28**, 503–504. (doi:10.1016/j.tree.2013.06.012)
33. Salkeld DJ, Padgett K, Jones J. 2013 A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. *Ecol. Lett*. **16**, 679–686. (doi:10.1111/ele.12101)
34. Randolph S, Dobson A. 2012 Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology* **139**, 847–863. (doi:10.1017/S0031182012000200)
35. Ostfeld RS. 2013 A Candide response to Panglossian accusations by Randolph and Dobson: biodiversity buffers disease. *Parasitology* **140**, 1196–1198. (doi:10.1017/S0031182013000541)
36. Civitello DJ *et al.* 2015 Biodiversity inhibits parasites: broad evidence for the dilution effect. *Proc. Natl. Acad. Sci. USA* **112**, 8667–8671. (doi:10.1073/pnas.1506279112)
37. Salkeld DJ, Padgett K, Jones JH, Antolin MF. 2015 Public health perspective on patterns of biodiversity and zoonotic disease. *Proc. Natl. Acad. Sci. USA* **112**, E6261. (doi:10.1073/pnas.1517640112)
38. Wood CL, Lafferty KD, DeLeo GA, Young HS, Hudson PJ, Kuris AM. 2016 Does biodiversity protect humans against infectious disease? Reply. *Ecology* **97**, 542–546.
39. Levi T, Massey AL, Holt RD, Keesing F, Ostfeld RS, Peres CA. 2016 Does biodiversity protect humans against infectious disease? Comment. *Ecology* **97**, 536–542. (doi:10.1890/15-354.1)
40. Myers SS, Gaffikin L, Golden CD, Ostfeld RS, Redford KH, Ricketts TH, Turner WR, Osofsky SA. 2013 Human health impacts of ecosystem alteration. *Proc. Natl Acad. Sci. USA* **110**, 18 753–18 760. (doi:10.1073/pnas.1218656110)
41. Keesing F, Holt RD, Ostfeld RS. 2006 Effects of species diversity on disease risk. *Ecol. Lett*. **9**, 485–498. (doi:10.1111/j.1461-0248.2006.00885.x)
42. Hough RL. 2014 Biodiversity and human health: evidence for causality? *Biodivers. Conserv*. **23**, 267–288. (doi:10.1007/s10531-013-0614-1)
43. McKinnon MC *et al.* 2016 What are the effects of nature conservation on human well-being? A systematic map of empirical evidence from developing countries. *Environ. Evid*. **5**, 8. (doi:10.1186/s13750-016-0058-7)
44. Pattanayak SK, Wendland KJ. 2007 Nature's care: diarrhea, watershed protection, and biodiversity conservation in Flores, Indonesia. *Biodivers. Conserv*. **16**, 2801–2819. (doi:10.1007/s10531-007-9215-1)

45. Bauch SC, Birkenbach AM, Pattanayak SK, Sills EO. 2015 Public health impacts of ecosystem change in the Brazilian Amazon. *Proc. Natl. Acad. Sci. USA* **112**, 7414–7419. (doi:10.1073/pnas.1406495111)
46. Brownstein JS, Skelly DK, Holford TR, Fish D. 2005 Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. *Oecologia* **146**, 469–475. (doi:10.1007/s00442-005-0251-9)
47. Fornace KM *et al.* 2016 Association between landscape factors and spatial patterns of *Plasmodium knowlesi* infections in Sabah, Malaysia. *Emerg. Infect. Dis*. **22**, 201–208. (doi:10.3201/eid2202.150656)
48. Kilpatrick AM, Salkeld DJ, Titcomb G, Hahn MB. 2017 Conservation of biodiversity as a strategy for improving human health and well-being. *Phil. Trans. R. Soc. B* **372**, 20160131. (doi:10.1098/rstb.2016.0131)
49. Smith KF, Goldberg M, Rosenthal S, Carlson L, Chen J, Chen C, Ramachandran S. 2014 Global rise in human infectious disease outbreaks. *J. R. Soc. Interface* **11**, 20140950. (doi:10.1098/rsif.2014.0950)
50. Just MG *et al.* 2014 Global biogeographic regions in a human dominated world: the case of human diseases. *Ecosphere* **5**, 143. (doi:10.1890/ES14-00201.1)
51. Murray CJ *et al.* 2012 Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2197–2223. (doi:10.1016/S0140-6736(12)61689-4)
52. Murray CJ. 1994 Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull. World Health Organ*. **72**, 429–445.
53. Freedman DA. 1999 Ecological inference and the ecological fallacy. *Int. Encycl. Soc. Behav. Sci*. **6**, 4027–4030.
54. Pearce N. 2000 The ecological fallacy strikes back. *J. Epidemiol. Community Health* **54**, 326–327. (doi:10.1136/jech.54.5.326)
55. Simpson E. 1951 The interpretation of interaction in contingency tables. *J. R. Stat. Soc. B* **13**, 238–241.
56. Tu Y-K, Gunnell D, Gilthorpe MS. 2008 Simpson's Paradox, Lord's Paradox, and suppression effects are the same phenomenon – the reversal paradox. *Emerg. Themes Epidemiol*. **5**, 2. (doi:10.1186/1742-7622-5-2)
57. Jenkins CN, Pimm SL, Joppa LN. 2013 Global patterns of terrestrial vertebrate diversity and conservation. *Proc. Natl Acad. Sci. USA* **110**, E2602–E2610. (doi:10.1073/pnas.1302251110)
58. Barlow J *et al.* 2016 Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **535**, 144–149. (doi:10.1038/nature18326)
59. Newbold T *et al.* 2015 Global effects of land use on local terrestrial biodiversity. *Nature* **520**, 45–50. (doi:10.1038/nature14324)
60. Newbold T *et al.* 2016 Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science* **353**, 288–291. (doi:10.1126/science.aaf2201)
61. Gray CL *et al.* 2016 Local biodiversity is higher inside than outside terrestrial protected areas worldwide. *Nat. Commun*. **7**, 12306. (doi:10.1038/ncomms12306)
62. Hansen M *et al.* 2013 High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853. (doi:10.1126/science.1244693)
63. Kim D-H, Sexton JO, Noojipady P, Huang C, Anand A, Channan S, Feng M, Townshend JR. 2014 Global, Landsat-based forest-cover change from 1990 to 2000. *Remote Sens. Environ*. **155**, 178–193. (doi:10.1016/j.rse.2014.08.017
64. Tropek R, Sedlacek O, Beck J, Keil P, Musilova Z, Simova I, Storch D. 2014 Comment on 'High-resolution global maps of 21st-century forest cover change'. *Science* **344**, 981. (doi:10.1126/science.1248753)
65. Gaveau DLA *et al.* 2014 Four decades of forest persistence, clearance, and logging in Borneo. *PLoS ONE* **9**, e101654. (doi:10.1371/journal.pone.0101654
66. Savilaakso S *et al.* 2014 Systematic review of effects on biodiversity from oil palm production. *Environ. Evid*. **3**, 4. (doi:10.1186/2047-2382-3-4)
67. Foster WA *et al.* 2011 Establishing the evidence base for maintaining biodiversity and ecosystem function in the oil palm landscapes of South East Asia. *Phil. Trans. R. Soc. B* **366**, 3277–3291. (doi:10.1098/rstb.2011.0041)
68. Fitzherbert EB *et al.* 2008 How will oil palm expansion affect biodiversity? *Trends Ecol. Evol*. **23**, 538–545. (doi:10.1016/j.tree.2008.06.012)
69. Kuris AM. 2012 The global burden of human parasites: who and where are they? How are they transmitted? *J. Parasitol*. **98**, 1056–1064. (doi:10.1645/12-90.1)
70. Hair JF, Hult GTM, Ringle CM, Sarstedt M. 2017 *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: SAGE Publications.
71. Whitmee S. 2017 Safeguarding human health in the Anthropocene epoch: Report of The Rockefeller Foundation–Lancet Commission on planetary health. *Lancet* **386**, 1973–2028. (doi:10.1016/S0140-6736(15)60901-1)
72. United Nations 2014 *World urbanization prospects: 2014 revision*. New York, NY: United Nations.
73. Mooring MS, Hart BL. 1992 Animal grouping for protection from parasites: selfish herd and encounter-dilution effects. *Behaviour* **123**, 173–193. (doi:10.1163/156853992X00011)
74. Johnson MB, Lafferty KD, van Oosterhout C, Cable J. 2011 Parasite transmission in social interacting hosts: monogenean epidemics in guppies. *PLoS ONE* **6**, e22634. (doi:10.1371/journal.pone.0022634)
75. Assaad F. 1982 Measles: summary of worldwide impact. *Rev. Infect. Dis*. **5**, 452–459. (doi:10.1093/clinids/5.3.452)
76. Zhang Y, Bi P, Hiller JE. 2010 Meteorological variables and malaria in a Chinese temperate city: a twenty-year time-series data analysis. *Environ. Int*. **36**, 439–445. (doi:10.1016/j.envint.2010.03.005)
77. Ryan SJ, McNally A, Johnson LR, Mordecai EA, Ben-Horin T, Paaijmans KP, Lafferty KD. 2015 Mapping physiological suitability limits for malaria in Africa under climate change. *Vector-Borne Zoonot. Dis*. **15**, 718–725. (doi:10.1089/vbz.2015.1822)
78. Cohen JM, Civitello DJ, Brace AJ, Feichtinger EM, Ortega CN, Richardson JC, Sauer EL, Liu X, Rohr JR. 2016 Spatial scale modulates the strength of ecological processes driving disease distributions. *Proc. Natl. Acad. Sci. USA* **113**, E3359. (doi:10.1073/pnas.1521657113)
79. Norman R, Bowers RG, Begon M, Hudson PJ. 1999 Persistence of tick-borne virus in the presence of multiple host species: tick reservoirs and parasite mediated competition. *J. Theor. Biol*. **200**, 111–118. (doi:10.1006/jtbi.1999.0982)
80. Swei A, Ostfeld RS, Lane RS, Briggs CJ. 2011 Effects of an invasive forest pathogen on abundance of ticks and their vertebrate hosts in a California Lyme disease focus. *Oecologia* **166**, 91–100. (doi:10.1007/s00442-010-1796-9)
81. Pattanayak SK, Kramer RA, Vincent JR. 2017 Ecosystem change and human health: implementation economics and policy. *Phil. Trans. R. Soc. B* **372**, 20160130. (doi:10.1098/rstb.2016.0130)
82. Lafferty K. 1997 Environmental parasitology: what can parasites tell us about human impacts on the environment? *Parasitol. Today* **13**, 251–255. (doi:10.1016/S0169-4758(97)01072-7)
83. Wood CL, Baum JK, Reddy S, Trebilco R, Sandin S, Zgliczynski B, Briggs AA, Micheli F. 2015 Productivity and fishing pressure drive variability in fish parasite assemblages of the Line Islands, equatorial Pacific. *Ecology* **96**, 1383–1398. (doi:10.1890/13-2154.1)
84. Wood CL, Sandin S, Zgliczynski B, Guerra AS, Micheli F. 2014 Fishing drives declines in fish parasite diversity and has variable effects on parasite abundance. *Ecology* **95**, 1929–1946. (doi:10.1890/13-1270.1)
85. Wood CL, Lafferty KD. 2015 How have fisheries affected parasite communities? *Parasitology* **142**, 134–144. (doi:10.1017/S003118201400002X)
86. Wood CL, Micheli F, Fernandez M, Gelcich S, Castilla JC, Carvajal J. 2013 Marine protected areas facilitate parasite populations among four fished host species of central Chile. *J. Anim. Ecol*. **82**, 1276–1287. (doi:10.1111/1365-2656.12104)