PROCEEDINGS B

Review

Author for correspondence:
C. Jessica E. Metcalf
e-mail: cmetcalf@princeton.edu

THE ROYAL SOCIETY
PUBLISHING

Identifying climate drivers of infectious disease dynamics: recent advances and challenges ahead

C. Jessica E. Metcalf[1,2], Katharine S. Walter[3], Amy Wesolowski[5], Caroline O. Buckee[6,7], Elena Shevliakova[8], Andrew J. Tatem[9,10], William R. Boos[4], Daniel M. Weinberger[3] and Virginia E. Pitzer[3]

[1]Department of Ecology and Evolutionary Biology, and [2]Office of Population Research, Woodrow Wilson School, Princeton University, Princeton, NJ, USA
[3]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, and [4]Department of Geology and Geophysics, Yale University, New Haven, CT, USA
[5]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Helath, Baltimore, MD, USA
[6]Department of Epidemiology, and [7]Center for Communicable Disease Dynamics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[8]NOAA/GFDL, Princeton, NJ, USA
[9]Flowminder Foundation, Stockholm, Sweden
[10]WorldPop project, Department of Geography and Environment, University of Southampton, Southampton, UK

CJEM, 0000-0003-3166-7521

Climate change is likely to profoundly modulate the burden of infectious diseases. However, attributing health impacts to a changing climate requires being able to associate changes in infectious disease incidence with the potentially complex influences of climate. This aim is further complicated by nonlinear feedbacks inherent in the dynamics of many infections, driven by the processes of immunity and transmission. Here, we detail the mechanisms by which climate drivers can shape infectious disease incidence, from direct effects on vector life history to indirect effects on human susceptibility, and detail the scope of variation available with which to probe these mechanisms. We review approaches used to evaluate and quantify associations between climate and infectious disease incidence, discuss the array of data available to tackle this question, and detail remaining challenges in understanding the implications of climate change for infectious disease incidence. We point to areas where synthesis between approaches used in climate science and infectious disease biology provide potential for progress.

1. Introduction

The role of environmental variables and climatic conditions in shaping human health has been recognized for centuries. Infectious diseases, in particular, may be sensitive to climatic conditions through their effects on abundance of vectors such as mosquitoes and ticks [1], pathogen survival outside of the host [2], environmental contamination and exposure to water-borne infections [3], dampening of host immunity [4], disruptions of health status associated with malnutrition linked to droughts or floods, and disruption of health systems by disasters such as floods or hurricanes. Consequently, shifts in climate—the average state of the atmosphere–ocean–land system over time, as well as the day-to-day variability of weather—may affect the burden of infectious diseases now and in the future [5].

Appropriately attributing changes in the burden of infectious diseases to climatic variables, and quantifying this relationship, is a necessary step in evaluating the potential impact of climate change [6]. However, this is complicated by a number of factors. One issue is that the available data for many pathogens often consist of only human cases of incident disease (rather than prevalence of infection in vectors and/or mild or asymptomatic hosts) and rarely span

time horizons reflective of major climatic shifts. As a result, inference into the consequences of climate change requires building on core biological knowledge (e.g. experiments indicating vector or pathogen survival across the range of a specific climatic variable [7]), leveraging spatial variation and range limits associated with climatic conditions [8], or building mechanistic understanding from shorter time-series of climatic variables and disease incidence—a process which itself often builds on biological knowledge and/or spatial variation. Another issue is that while climate acts as an extrinsic driver, infectious disease dynamics also have intrinsic drivers, particularly fluctuations in population-level immunity and susceptibility, as well as the dynamics of human behaviour (e.g. increasing population size and mobility can modify previous geographical limits for vector-borne infections [9]). Disentangling these drivers requires careful statistical and model-based partitioning of possible links between climate and infectious disease incidence while accounting for features of host–pathogen biology such as asymptomatic carriage [10].

In order to attribute a change in infectious disease incidence to climate, Rogers & Randolph [1] specify three criteria: the change in infectious disease incidence must have occurred at the right time, in the right place and in the right direction (consistent with the hypothesized climate–disease relationship). The third criterion requires understanding the mechanism(s) by which climate may affect infectious disease incidence, while the former two criteria require careful analysis of spatio-temporal data. Here, we review the mechanisms by which climatic variables might affect infectious disease transmission, discuss challenges involved in linking climate drivers to infectious disease transmission, provide an overview of statistical and mechanistic models that can be used to quantify these connections, and discuss how these might contribute to generating future projections of the effects of climate on health. Previous reviews have outlined mechanisms underlying climate and infectious disease associations in non-human pathogens [11,12], reviewed conceptual challenges in detection and attribution [6], and provided an overview of core knowledge gaps [13]. Our focus is on infectious diseases in humans, and we draw from examples across a range of pathogens to concretely illustrate the methodological challenges and approaches that have been developed to grapple with projecting future infectious disease incidence under climate change.

2. Potential mechanisms linking climate and infectious diseases

Associations between climatic conditions and infectious disease incidence may be observed at a range of spatial and temporal scales, but associations alone do not indicate causal links. Establishing causality requires identifying whether the association is consistent with a hypothesized mechanism. There are a wealth of potential mechanisms linking climate and infectious diseases, which differ across ecological aspects of the human–pathogen interaction, including route of transmission.

Infectious diseases may be directly transmitted via airborne particles or fomites (e.g. influenza), or indirectly transmitted via food, water (e.g. cholera) or a vector (e.g. malaria, dengue), and could potentially involve non-human reservoir species (zoonotic pathogens, e.g. hantavirus). Each transmission route may be associated with different climatic drivers, outlined in the electronic supplementary material,

table S1. The consequences of climatic variation can range from complete prevention of transmission (across geographical ranges or at certain times of the year) to shifting the magnitude of transmission.

The most direct approach to identifying a mechanistic impact of climate on infectious diseases is experimentation. However, pathogens for which good animal models of transmission exist (e.g. guinea pigs for influenza [2]), or for which experimental studies of vector dynamics can be deployed (e.g. *Aedes aegypti* for dengue [7]), remain a minority. Sometimes, general aspects of the thermal or broader environmental niche are known and can be used to inform the expected direction of the climate–disease relationship.

To appropriately attribute the impact of climate change on infectious disease incidence, we must move beyond simply identifying the direction of the climate–disease relationship to actually quantifying these relationships under natural conditions. Model-based approaches can link changes in climate variables to changes in disease incidence, ideally following intermediate steps on the hypothesized causal pathway. For instance, if rainfall and flooding are thought to influence the risk of exposure for a water-borne disease (e.g. typhoid fever) through increased contamination of drinking water, ideally one would link the hydrologically relevant climate variables to the prevalence of bacterial contamination of water supplies [14] prior to examining correlations with human incidence data.

Importantly, climate probably mediates infectious disease risk via multiple mechanisms, and variability in climatic variables may be even more important than mean levels [15–17]. These factors complicate identifying the role played by climatic variables in shaping the burden of infectious diseases. The scope of climatic variation that can be used to drive inference around the possible impacts of climate is an essential component of meeting this challenge, which we detail next, before outlining available methodological approaches.

3. Spatio-temporal scales of variation and confounding factors

Once a hypothesized mechanism is established, the next two criteria require determining whether changes in climate variables can be linked to changes in disease incidence in space or time. Several scales and types of spatio-temporal variation can be leveraged.

(a) Spatial variation

The geographical range limits of infections can powerfully indicate the effects of climate. For many vector-borne infections, climatic changes (particularly increasing temperatures) could increase the geographical range of vectors and thereby increase the size of the at-risk population; conversely, decreases in geographical range could be projected if conditions become too dry [18]. Such changes have been modelled for infections like malaria [19,20], dengue [21], onchocerciasis [22], Chagas disease [23], West Nile virus [24], chikungunya [25] and Rift Valley fever virus [26], and vectors like sand flies [27] and black-legged ticks [28]. However, it is important (and often difficult) to differentiate between places that cannot support ongoing transmission of the pathogen due to suboptimal climatic conditions versus locations where the pathogen has yet

to be introduced or has recently been eliminated by anthropogenic rather than climatic influences [29,30]. Control efforts are likely to be focused in areas of intense transmission, potentially obscuring key climatic conditions.

(b) Seasonality

Seasonal variation is another powerful focus for disentangling climate–infectious disease relationships, providing a repeatable probe of the association between climatic drivers and transmission. Between-year variation, from small deviations [31] through to unusual climatic conditions [32], can be used to identify core climatic drivers if appropriate methods are deployed (see below). However, given the large number of candidate climatic variables, many of which also exhibit seasonal increases and decreases, spurious associations between climatic drivers and a focal infection are inevitable. Confounders linked to human behaviour or demography [33], or seasonality in immune function (e.g. associated with vitamin D metabolism [34]), further complicate inference (electronic supplementary material, S1). Combining seasonal variations with biological knowledge (e.g. [2]) and/or variation through space [35] can help to identify a signature of a driving covariate among multiple covarying climatic variables.

(c) Multi-annual variation

Over longer time-courses, multi-annual fluctuations in climate, such as the El Niño Southern Oscillation (ENSO), provide strong signatures against which to test associations between climate and infectious diseases [36], and can help tease apart the effect of different climatic factors that do not necessarily covary on longer than seasonal time scales [37]. Non-stationary changes in climate (e.g. warming, multi-decadal drought) can also be examined, but establishing robust links with these longer-time-scale variations in climate requires amassing infectious disease time series of sufficiently long duration, which is challenging [38]. Even if data are available, long-term signatures of associations between climate and infectious disease dynamics may be overwhelmed by changes in surveillance practices and/or the introduction of interventions, as well as non-stationary changes in human ecology (e.g. increasing density, urbanization [29]) and evolution in pathogen (e.g. emergence of drug resistance [39]) or vector populations (e.g. climate-imposed selection may alter mosquito diapause, shifting the range of mosquito-borne pathogens [40]). Thus, observed associations between long-term changes in climate and infectious disease burden must be interpreted with caution, taking into account the underlying changes in the human and pathogen populations and their interface [30].

(d) Combining spatial and temporal scales

'Location-specific' associations, which cannot be extrapolated to regions where the climate may differ, or temporal associations that do not align with range limits linked to climate, pose a serious challenge to characterizing the effects of climate change. In many instances, these may be indicative of non-mechanistic associations. However, it is also possible that the important drivers of transmission vary depending on the range of climatic conditions in a given location. For example, the climatic drivers of influenza transmission, or the shape of the climate–disease relationship, may vary between temperate and tropical regions [41,42]. Understanding how climate change might affect this pathogen will require accounting for these differences.

4. Estimating climate effects on infectious diseases

The scales of variation delineated have the potential to inform our estimates of the association between climate drivers and disease incidence. However, for communicable diseases (pathogens transmitted from person to person, either directly or indirectly), several factors make this challenging. First, the infectious process is typically unobserved. Second, many infections are to some degree immunizing, meaning that one new infection both magnifies transmission, but also depletes the pool of susceptible individuals, resulting in feedbacks in the transmission dynamics which can obscure climate signatures (figure 1). Attributing changes in infectious diseases to specific climate drivers thus requires accounting for nonlinearities in the risk of exposure and susceptible depletion, a fact recognized since at least the 1950s, when Hope Simpson quantified 'infectiousness' for measles, chickenpox and mumps by carefully accounting for individuals in households who had previously been exposed and were thus no longer susceptible. He found no seasonal variation in 'infectiousness', despite seasonality in incidence [44]. The core of his analysis underscores the essential limitation of using traditional time-series approaches for infectious disease dynamics—the population exposed and at risk changes rapidly as a result of the very dynamics of infection. Autocorrelations and lags that do not take into account the underlying biology can result in biases [45] and more complex inferential frameworks may be required.

An array of methods is available for disentangling links between climate and infectious diseases, ranging from statistical to mechanistic (i.e. dynamic) modelling approaches (figure 2). Statistical models focus on matching patterns of variation in climatic variables with the distribution of observed cases of disease in time and/or space; mechanistic models seek to dissect these relationships by explicitly accounting for the processes of transmission of infection and observation of disease [1]. Both approaches are useful, but their applicability depends on characteristics of the pathogen and the host–pathogen relationship (see the electronic supplementary material, table S2.)

(a) Traditional time-series modelling approaches

Statistical modelling approaches for climate–disease associations include regression approaches, such as generalized linear models or generalized additive models (which can also include time-dependent components), and conventional time-series models, such as autoregressive integrated moving average (ARIMA) models. Their appropriateness is shaped by the degree to which hidden processes linked to transmission or immunity result in confounding or non-independence (electronic supplementary material, table S2). Where climate affects susceptibility to symptomatic disease (in turn affecting detectability by the surveillance system) rather than the infection process, as reported for *Streptococcus pyogenes* [46], nonlinear feedbacks between climate and disease incidence should be minimal, and statistical approaches appropriate. Likewise, if the primary reservoir for infection is a non-human host
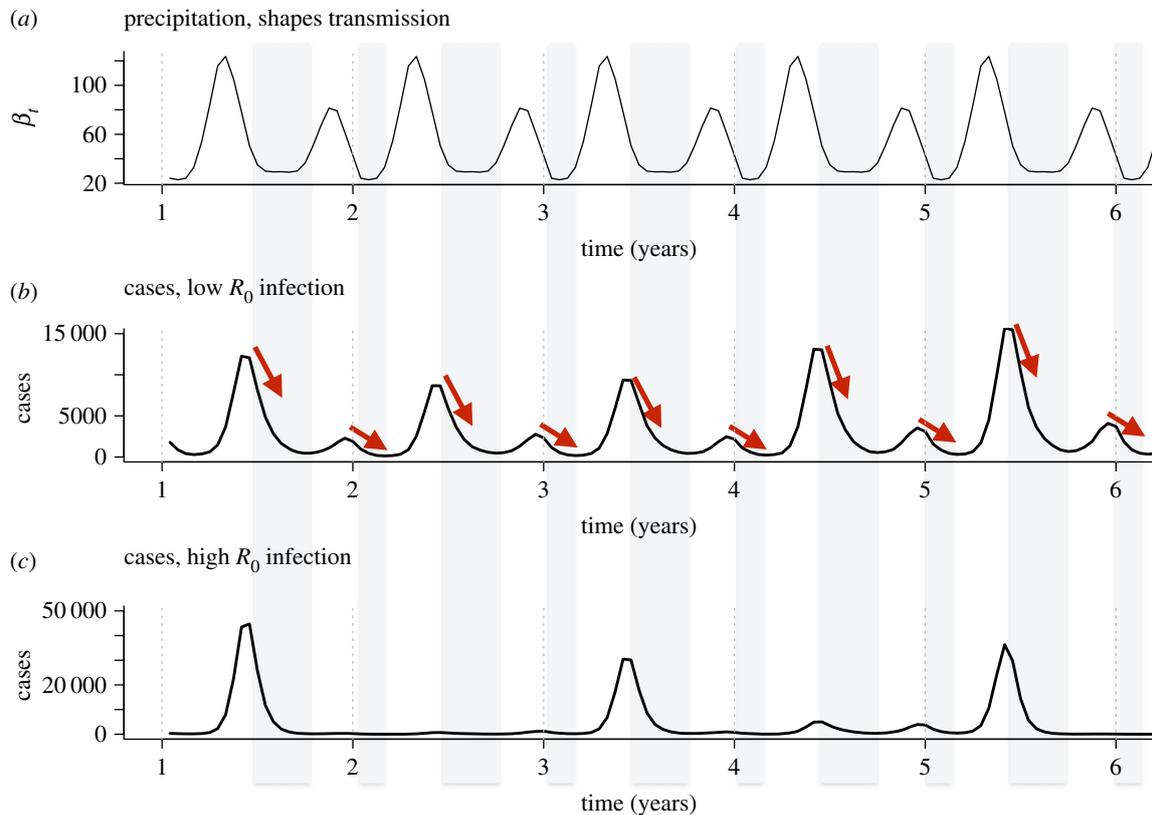
Figure 1. Issues with linking incidence and climate drivers for immunizing infections (a) Simulated precipitation over 6 years, set to reflect bimodal peaks in each year as observed (e.g. in Kenya), with periods of low precipitation shown in grey. (b) Resulting cases for a completely immunizing infection with a two-week generation time and a low basic reproductive number (R_0) (set to $R_0 = 5$) simulated using an SIR transmission dynamic model as in [43]. The resulting time series shows a clear footprint where periods of low precipitation and thus low transmission correspond to outbreaks turning over (red arrows), where conversely high precipitation is associated with increases in incidence. (c) Resulting cases for an identical infection simulated using an SIR model, but with a high basic reproductive number ($R_0 = 32$). In this case, the resulting time series shows a much more erratic picture, with little direct indication of the impact of precipitation on cases, especially in low incidence years (i.e. no increase in cases with increases in precipitation in year 2), as a result of the dominant multi-annual period resulting from the intrinsic dynamics.
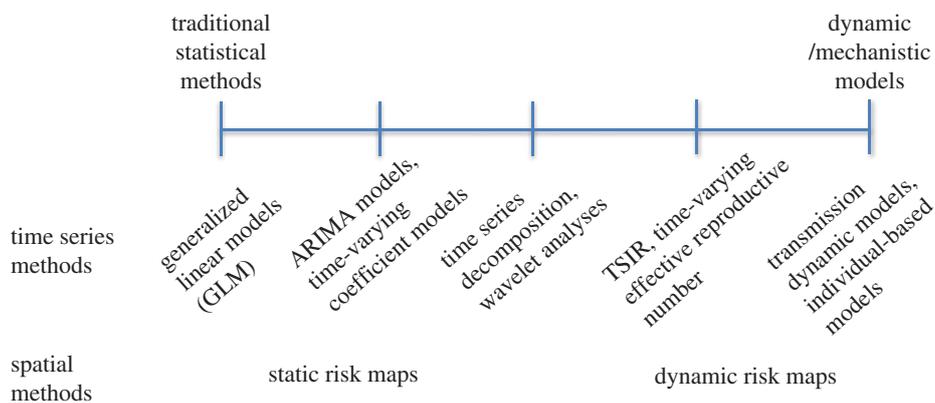

Figure 2. Schematic of the range of modelling approaches ordered across a spectrum of increasing incorporation of mechanism. (Online version in colour.)

and the impact of climate is to mediate pathogen exposure, possibly by affecting vectors, then nonlinear modelling approaches are not required unless relevant dynamic feedbacks are identified. Accordingly, traditional time-series modelling approaches have been applied to examine associations between extreme precipitation and water-borne diseases in the USA [47], and rainfall, humidity and cutaneous leishmaniasis in Brazil [48]. For infections with relatively low transmission rates, and a long generation interval, traditional time-series approaches could also be applied, because there will only be weak dependence between consecutive

observations (i.e. incident cases of disease; figure 1). Finally, conventional time-series approaches are appropriate for exploratory analyses to identify potentially relevant variables for further inquiry using mechanistic approaches [35].

When carefully trained, evaluated and tested (e.g. via out-of-sample prediction), non-mechanistic autoregressive models may have good predictive power for forecasting burden over shorter time horizons [49], even outperforming models that incorporate the 'true' underlying drivers; but over the longer term, non-stationary changes in climate, or variables like demography and susceptible depletion, may complicate this picture.

(b) Time-series decomposition, wavelet analysis and synchrony

The periodicity in the dynamics of some infectious diseases as well as climate (including both seasonal and multi-annual cycles) lends itself to a range of methods specifically designed to quantify non-stationary associations between time series, such as changes in mean, variance and/or period of oscillations over time. Methods for time-series decomposition (including moving averages, seasonal and trend decomposition using Loess and Bayesian structural time-series models [50]) seek to separate out long-term trends in incidence from seasonal and multi-annual cycles and random 'noise', bridging the gap between traditional statistical approaches and mechanistic models. Wavelet analysis is especially appropriate for identifying changes in the periodicity and/or phase of incident cases and correlating them with exogenous variables [51]. Wavelets have been used to evaluate the synchrony between rainfall and cholera across Africa [52] and Haiti [53], and to suggest that warmer locations serve as sources of out-going waves of dengue across Southeast Asia [54]. Such approaches may be particularly powerful in settings where nonlinear aspects of the dynamics, such as susceptible feedbacks, are expected, but are hard to model explicitly given complexities in the biology and/or a paucity of data available to reflect these (e.g. multiple strains, asymptomatic carriage). The wavelet spectra allow partitioning of frequencies reflecting such biologically driven resonances (consistent with the intrinsic disease dynamics), allowing other frequencies and their associations with extrinsic variables (e.g. climate variables) to be evaluated. However, the inference remains correlative rather than causative, because climate drivers such as ENSO may have multi-annual cycles that align by chance with multi-annual dynamics of an infectious disease. But if multiple lines of evidence (e.g. average age of infection, exponential growth rates) suggest that the transmission rate of an infection and associated intrinsic dynamics are inconsistent with the observed multi-annual cycles, yet these cycles align with a climatic variable such as ENSO, this provides supporting evidence for a role of this climate variable on transmission.

(c) Regression-based TSIR models to extract transmission

A potentially powerful approach is to harness statistical methods within a mechanistic framing. One such semi-mechanistic approach examines associations between climate and the *transmission rate* by reconstructing the underlying susceptible and infectious populations. As with Hope Simpson's approach [44], the idea is to condition on susceptibility and exposure to the pathogen in order to extract a measure of 'infectiousness', although here the rate of exposure must be inferred from the underlying time series. This concept underlies time-series susceptible–infected–recovered regression models [43], which build on estimating the size of the susceptible population (S_t) by taking into account depletion of susceptibles via infection or vaccination, and replenishment via births. The number of incident infections at time $t + 1$ is then modelled as $I_{t+1} = \beta_t I_t S_t$, where β_t is the rate of transmission from an infected to susceptible individual; the time-step should approximately equal the serial interval of the infection. This equation can be linearized by taking the log of both sides, then fitted using a regression with I_t (the observed incidence,

corrected for underreporting) and S_t as offsets:

$$\log(I_{t+1}) = \log(\beta_t) + \log(I_t) + \log(S_t).$$

Seasonal (and other temporal) fluctuations in the transmission rate β_t can then be estimated by choosing an appropriate focal time-scale. The resulting estimates of transmission can be analysed to evaluate their association with seasonal climatic (and non-climatic) drivers. Alternatively, β_t can be framed as a function of climatic variables or other intermediate variables (e.g. environmentally sampled viral or bacterial particles) and these associations can be directly evaluated. For pathogens that are not completely immunizing, waning of immunity can be captured (e.g. using Taylor expansions [37]). Such approaches have been used to suggest that seasonal migration patterns linked to agriculture shape measles dynamics in Niger [55], and reveal temperature-dependence of rotavirus transmission in England and The Netherlands [56].

(d) Partitioning drivers of R_E

A related approach is to extract the time-varying effective reproductive number $R_{E,t}$ from time series of incidence [57], defined at time t as $R_{E,t} = R_0 S_t$, where R_0 is the constant basic reproductive number (i.e. the average number of secondary infections expected from one infectious individual in a fully susceptible population), and S_t is the number of susceptibles at time t; $R_{E,t}$ is thus the expected number of secondary infections accounting for population immunity. This relationship can be expanded to encompass the influence of potential drivers [31]:

$$R_{E,t} = R_0 S_t \prod_k \delta_{t,k}^{\beta_k}.$$

Variables associated with hypothesized mechanisms by which climatic variables shape transmission can be incorporated via the k potential drivers, $\delta_{t,k}$. Following linearization (taking logs, as above), the coefficients β_k can be estimated via regression (e.g. revealing how absolute humidity and school holidays shape influenza-like illness in The Netherlands [31]).

(e) Dynamic models

The next extension is to fit a full transmission dynamic model to incidence data. Such models explicitly attempt to represent reasonably complex life cycles, and differentiate between infection and disease. Again, there are two possible approaches. First, data from climate drivers can be directly used within the dynamic model. For example, Shaman *et al.* [58] developed a dynamic model for influenza encompassing a parametric relationship between absolute humanity and R_0. A similar approach was used to examine associations between rainfall and cholera incidence in Haiti [59]. Experimental data on the relationship between temperature and mosquito demography were incorporated into a model that revealed how climate shaped the first reported dengue outbreak in Europe [60]; similar models have been used to hindcast seasonal outbreaks of malaria in India to evaluate an early warning system [61]. Full dynamical models can also be framed to allow formal statistical comparisons between hypotheses (e.g. to evaluate the role of rainfall versus feedbacks due to immunity in driving malaria cycles [62]). Alternatively, seasonality in transmission can be approximated (e.g. using a sinusoidal forcing function, for which the amplitude and phase are directly estimated), and association with climate drivers evaluated post-hoc. This

approach has been used to examine the climatic drivers underlying the distinct seasonal and spatial patterns of RSV in the USA [35] and the Philippines [63]. Out-of-sample model validation is always essential, the impact of interventions must be considered [64] and care must be taken to avoid common inferential pitfalls (below).

(f) Risk maps

For vector-borne and zoonotic diseases where endemicity is shaped by the local environment, risk maps estimate the spatial distribution of human risk of infection, capturing heterogeneity at local [65], regional, country-wide [66,67] and global scales [68–71]. Risk-mapping approaches integrate observed occurrences (i.e. prevalence or incidence data) at discrete locations to generate an interpolated map of estimated risk of human infection and/or disease, generally building on environmental, ecological, socio-demographic and/or human susceptibility covariates [69,72], and sometimes known biological dependencies (e.g. of vectors on climate [73]). Underlying methods range from model-based geostatistics (especially where data are rich), to machine-learning approaches (e.g. boosted regression trees [71]), and generally encompass formal quantification of uncertainty [74,75].

Increasingly, methods underlying risk maps encompass core biological mechanisms via transmission models, allowing attribution of the impact of interventions [76] through to projection of burden, for example, by inferring attack rates and modelling demographic processes for Zika virus [77]. Dynamic spatial risk maps based on mechanistic models enable hypothesis-testing of the climatic drivers of infectious disease risk, and attribution of specific events, such as the role of the 2015 El Niño in the emergence of Zika virus [78]. However, parametrizing such risk-map models requires detailed biological and epidemiological knowledge and may be difficult to develop for broad spatio-temporal scales. Model validation against observed changes in the spatial distribution of disease is essential.

5. Methodological challenges for climate–disease models

The analysis of links between climate and infectious disease incidence using population-level data raises a series of statistical challenges (beyond the broader epistemological challenges that arise from attempting to indirectly infer mechanism, and process-based challenges; figure 1). First, for any pathogen, there are many possible extractions of climate covariates, which are likely to be highly collinear. In the best-case scenario, clear candidate predictors present themselves through knowledge of biological mechanisms, usually from experimental data (electronic supplementary material, table S1). The most compelling evidence for quantifying the impacts of climate on infectious diseases comes when experimental evidence for the climate–pathogen relationship is used to parametrize models describing population-level patterns at broad spatial and temporal scales (e.g. [79,80]). Where no information is available to guide covariate choice, care should be taken to avoid pitfalls commonly encountered in regression featuring multiple comparisons [81]. Approaches such as principal component analysis and Lasso might be used to reduce the dimensionality of the system, while Bayesian model averaging can be used to identify combinations of potential predictors that provide the best fit to

the data while accounting for model uncertainty. If strong relationships between principal components and disease incidence or transmission emerge, the climatic variables captured by the focal principal component can be further explored in laboratory or field experiments to pinpoint potential individual-level mechanisms.

The repeatability of seasonality in both climate variables and disease incidence brings another challenge. For example, a low absolute humidity and low influenza incidence occur at approximately the same time of the year, spurious associations are possible. A permutation test can address this— the time course of drivers within each year is randomly redistributed across years and the analyses re-run [31]. Only parameters whose association is significantly stronger than in the randomized time-series are likely to reflect a true association. Alternatively, associations between the *residuals* of climatic factors and model-predicted incidence can be explored to determine whether *unusual* climatic conditions (that differ from the expected conditions given the time of year) are also associated with higher or lower disease incidence [35,58]. There is promise but also challenge in harnessing the impact of extreme events to strengthen inference [82]. 'Model-free methods', such as convergent cross-mapping have been used to infer a role of absolute humidity on influenza [42]; however, the scope of inference with such methods may be limited in settings of periodic fluctuations [83].

A final statistical challenge is the potential lack of congruence between the spatial and temporal scale of available data and the key scale of mechanism [84]. For example, if one day's strong rainfall results in a pulse in transmission of an enteric infection, but disease incidence data are only available monthly, or at a larger spatial scale, this detail will be missed. It has been suggested that this type of mismatch may result in stronger inference with broad-scale trends like El Niño rather than specific climate drivers like temperature [85].

6. Methodological challenges related to climate data and climate change models

Observations of atmospheric and hydrological variables will be key to deploying the methods described above. Routine observations are made from surface stations [86], weather balloons [87], satellite radiometers [88], etc. These are used with fluid dynamical constraints to produce estimates of the global atmospheric state at resolutions of tens of kilometres [89]. All of these data sources have uncertainties and biases; consequently, the appropriate source depends on the question of interest, pathogen biology and the spatio-temporal scales of the infectious disease data. Potential variables of interest include directly observed quantities such as temperature or precipitation, or derived measures of heat stress or drought.

Disease incidence in future decades would ideally be predicted using output from climate models, but this effort is hampered by poor climate model resolution and high model uncertainty and bias. Infectious diseases are often sensitive to variability on short spatial scales that is not well-represented by global climate models, with typical horizontal resolutions on the order of 100 km. Statistical or dynamical downscaling—the latter employing higher-resolution climate models integrated only in a particular region—can provide data at the requisite scales, but rely on potentially biased boundary conditions obtained from global models. Global models also

typically produce biased estimates of extreme events, such as tropical precipitation extremes [90]; this is problematic as the dynamics of infectious diseases may be more sensitive to extremes or variability than to time-mean properties, and associations may be strongly nonlinear.

However, there is opportunity to identify disease–climate associations that rely on climatic variables with more certain projected changes. For example, although predictions of next-century regional precipitation change are highly uncertain, there is evidence that aridity over many continental regions has increased and will continue to do so in coming decades [91]. The seasonal cycle of rainfall is biased in models of some tropical regions [90], with, for example, potentially serious consequences for malaria burden projections. However, projected delays in the seasonal onset of monsoon rainfall as climate warms are robust across models and stem from well-understood atmospheric thermodynamics [92]. Tight interdisciplinary work, in which climate scientists work closely with infectious disease modellers, thus has the potential to lead to discovery of emergent constraints on future disease incidence.

7. Linking climate change models to infectious disease models

Forecasting future infectious disease incidence under climate change requires jointly modelling both changes to future climate and the climate–disease relationship, accounting for uncertainties in both. On short time scales, data-assimilation approaches used in weather forecasting have been applied to influenza [93], taking as a premise noisy data and using statistical filtering to reduce forward error propagation [94] to provide reliable real-time forecasts of influenza incidence across a large number of US cities [93,95], as well as allowing nuanced evaluation of different data-assimilation and filtering approaches [94]. Across longer time scales, an array of possible climate and infectious disease models can be combined in a multimodel ensemble framework, allowing formal assessment of the range of possible future outcomes and identification of the origin of discrepancies between models [96]. This approach has been used to forecast future malaria distributions under climate change—one of the few pathogens for which the necessary diversity of models for the climate–disease relationship exists [80].

For infectious diseases, as for other anticipated impacts of climate change [96], the footprint of human adaptation, ranging from interventions such as vector control to vaccination strategies altering the landscape of immunity, or human movement shifting diseases, complicate forecasting. But near-term forecasts [93] or longer-term projections that explicitly do not account for these subtleties [80] provide a powerful starting point, from which refinements can subsequently be made. In both instances, model projections must be continually tested by comparing with out-of-sample data, and updated when the models fail to capture relevant patterns in data. One potentially fruitful direction is to build on methods used in the climate community to increase skill via retrospective analyses of weather and climate forecast models [93]. Retrospective forecasting of infectious disease outcomes on time scales of months to years would allow the impact of the seasonal cycle, ENSO and other interannual climate variability to be assessed [97]. Importantly, even if the climate prediction itself is wrong, an ensemble approach could assess whether infectious disease projections conditioned on the actual climate occurrence exhibit reasonable skill.

8. Conclusion

To evaluate the impact of climate change on the incidence of infectious diseases, we must move beyond identifying simple correlations and statistical associations between climatic variables and incident cases of disease to identifying the underlying causal mechanisms. Just as the methods used for weather forecasting differ from those used for long-range climate projections, different methods are needed to make out-of-sample predictions about future trends in incidence under climate change. Traditional statistical approaches are useful under certain conditions (electronic supplementary material, table S2), for exploratory analyses, and even for short- to medium-term forecasting (e.g. [98]). However, more mechanistic modelling approaches are needed if climatic drivers impact the *transmission* of infection from infected to susceptible hosts. Epidemiologists can take cues from the climate modelling community by seeking to better understand and incorporate the underlying properties that influence the observed behaviour of the climate–disease system, and by routinely testing an ensemble of models by retrospectively comparing and validating models against data. From forecasting influenza epidemics over the short term [94] to disentangling the impact of climate change on malaria [80], great strides are being made, but much remains to be done.

References

1. Rogers D, Randolph S. 2006 Climate change and vector-borne diseases. *Adv. Parasit.* **62**, 345–381. (doi:10.1016/S0065-308X(05)62010-6)

2. Lowen AC, Mubareka S, Steel J, Palese P. 2007 Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* **3**, e151. (doi:10.1371/journal.ppat.0030151)

3. Reiner RC, King AA, Emch M, Yunus M, Faruque A, Pascual M. 2012 Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. *Proc. Natl*

7

rspb.royalsocietypublishing.org Proc. R. Soc. B **284**: 20170901

8

rspb.royalsocietypublishing.org Proc. R. Soc. B 284: 20170901

Acad. Sci. USA **109**, 2033–2036. (doi:10.1073/pnas.1108438109)

4. Foxman EF, Storer JA, Vanaja K, Levchenko A, Iwasaki A. 2016 Two interferon-independent double-stranded RNA-induced host defense strategies suppress the common cold virus at warm temperature. *Proc. Natl Acad. Sci. USA* **113**, 8496–8501. (doi:10.1073/pnas.1601942113)

5. McMichael AJ, Woodruff RE, Hales S. 2006 Climate change and human health: present and future risks. *Lancet* **367**, 859–869. (doi:10.1016/S0140-6736(06)68079-3)

6. Stone D *et al.* 2013 The challenge to detect and attribute effects of climate change on human and natural systems. *Clim. Change* **121**, 381–395. (doi:10.1007/s10584-013-0873-6)

7. Yang H, Macoris M, Galvani K, Andrighetti M, Wanderley D. 2009 Assessing the effects of temperature on the population of *Aedes aegypti*, the vector of dengue. *Epidemiol. Infect.* **137**, 1188–1202. (doi:10.1017/S0950268809002040)

8. Hay SI *et al.* 2013 Global mapping of infectious disease. *Phil. Trans. R. Soc. B* **368**, 20120250. (doi:10.1098/rstb.2012.0250)

9. Tatem AJ, Smith DL. 2010 International population movements and regional *Plasmodium falciparum* malaria elimination strategies. *Proc. Natl Acad. Sci. USA* **107**, 12 222–12 227. (doi:10.1073/pnas.1002971107)

10. Duong V *et al.* 2015 Asymptomatic humans transmit dengue virus to mosquitoes. *Proc. Natl Acad. Sci. USA* **112**, 14 688–14 693. (doi:10.1073/pnas.1508114112)

11. Altizer S, Ostfeld RS, Johnson PT, Kutz S, Harvell CD. 2013 Climate change and infectious diseases: from evidence to a predictive framework. *Science* **341**, 514–519. (doi:10.1126/science.1239401)

12. Dobson A, Molnár PK, Kutz S. 2015 Climate change and Arctic parasites. *Trends Parasitol.* **31**, 181–188. (doi:10.1016/j.pt.2015.03.006)

13. Rohr JR, Dobson AP, Johnson PT, Kilpatrick AM, Paull SH, Raffel TR, Ruiz-Moreno D, Thomas MB. 2011 Frontiers in climate change–disease research. *Trends Ecol. Evol.* **26**, 270–277. (doi:10.1016/j.tree.2011.03.002)

14. Karkey A *et al.* 2016 The ecological dynamics of fecal contamination and Salmonella Typhi and Salmonella Paratyphi A in municipal Kathmandu drinking water. *PLoS Negl. Trop. Dis.* **10**, e0004346. (doi:10.1371/journal.pntd.0004346)

15. Paaijmans KP, Blanford S, Bell AS, Blanford JI, Read AF, Thomas MB. 2010 Influence of climate on malaria transmission depends on daily temperature variation. *Proc. Natl Acad. Sci. USA* **107**, 15 135–15 139. (doi:10.1073/pnas.1006422107)

16. Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW. 2011 Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **108**, 7460–7465. (doi:10.1073/pnas.1101377108)

17. Carrington LB, Seifert SN, Armijos MV, Lambrechts L, Scott TW. 2013 Reduction of *Aedes aegypti* vector competence for dengue virus under large temperature fluctuations. *Am. J. Trop. Med. Hyg.* **88**, 689–697. (doi:10.4269/ajtmh.12-0488)

18. Yamana TK, Eltahir EA. 2013 Projected impacts of climate change on environmental suitability for malaria transmission in West Africa. *Environ. Health Persp.* **121**, 1179. (doi:10.1289/ehp.1206174)

19. Baeza A, Bouma MJ, Dobson AP, Dhiman R, Srivastava HC, Pascual M. 2011 Climate forcing and desert malaria: the effect of irrigation. *Malaria J.* **10**, 190. (doi:10.1186/1475-2875-10-190)

20. Patz JA, Olson SH. 2006 Malaria risk and temperature: influences from global climate change and local land use practices. *Proc. Natl Acad. Sci. USA* **103**, 5635–5636. (doi:10.1073/pnas.0601493103)

21. Rogers DJ. 2015 Dengue: recent past and future threats. *Phil. Trans. R. Soc. B* **370**, 20130562. (doi:10.1098/rstb.2013.0562)

22. Cheke RA *et al.* 2015 Potential effects of warmer worms and vectors on onchocerciasis transmission in West Africa. *Phil. Trans. R. Soc. B* **370**, 20130559. (doi:10.1098/rstb.2013.0559)

23. Medone P, Ceccarelli S, Parham PE, Figuera A, Rabinovich JE. 2015 The impact of climate change on the geographical distribution of two vectors of Chagas disease: implications for the force of infection. *Phil. Trans. R. Soc. B* **370**, 20130560. (doi:10.1098/rstb.2013.0560)

24. Paz S. 2015 Climate change impacts on West Nile virus transmission in a global context. *Phil. Trans. R. Soc. B* **370**, 20130561. (doi:10.1098/rstb.2013.0561)

25. Campbell LP, Luther C, Moo-Llanes D, Ramsey JM, Danis-Lozano R, Peterson AT. 2015 Climate change influences on global distributions of dengue and chikungunya virus vectors. *Phil. Trans. R. Soc. B* **370**, 20140135. (doi:10.1098/rstb.2014.0135)

26. Barker CM, Niu T, Reisen WK, Hartley DM. 2013 Data-driven modeling to assess receptivity for Rift Valley fever virus. *PLoS Neglect. Trop. D* **7**, e2515. (doi:10.1371/journal.pntd.0002515)

27. Fischer D, Moeller P, Thomas SM, Naucke TJ, Beierkuhnlein C. 2011 Combining climatic projections and dispersal ability: a method for estimating the responses of sandfly vector species to climate change. *PLoS Neglec. Trop. D* **5**, e1407. (doi:10.1371/journal.pntd.0001407)

28. Brownstein JS, Holford TR, Fish D. 2003 A climate-based model predicts the spatial distribution of the Lyme disease vector *Ixodes scapularis* in the United States. *Environ. Health Persp.* **111**, 1152. (doi:10.1289/ehp.6052)

29. Tatem AJ, Gething PW, Smith DL, Hay SI. 2013 Urbanization and the global malaria recession. *Malaria J.* **12**, 133. (doi:10.1186/1475-2875-12-133)

30. Gething PW, Smith DL, Patil AP, Tatem AJ, Snow RW, Hay SI. 2010 Climate change and the global malaria recession. *Nature* **465**, 342–345. (doi:10.1038/nature09098)

31. te Beest DE, van Boven M, Hooiveld M, van den Dool C, Wallinga J. 2013 Driving factors of influenza transmission in the Netherlands. *Am. J. Epidemiol.* **178**, 1469–1477. (doi:10.1093/aje/kwt132)

32. Pascual M, Rodó X, Ellner SP, Colwell R, Bouma MJ. 2000 Cholera dynamics and El Nino-southern oscillation. *Science* **289**, 1766–1769. (doi:10.1126/science.289.5485.1766)

33. Martinez PP, King AA, Yunus M, Faruque A, Pascual M. 2016 Differential and enhanced response to climate forcing in diarrheal disease due to rotavirus across a megacity of the developing world. *Proc. Natl Acad. Sci. USA* **113**, 4092–4097. (doi:10.1073/pnas.1518977113)

34. Stevenson T *et al.* 2015 Disrupted seasonal biology impacts health, food security and ecosystems. *Proc. R. Soc. B* **282**, 20151453. (doi:10.1098/rspb.2015.1453)

35. Pitzer VE, Viboud C, Alonso WJ, Wilcox T, Metcalf CJ, Steiner CA, Haynes AK, Grenfell BT. 2015 Environmental drivers of the spatiotemporal dynamics of respiratory syncytial virus in the United States. *PLoS Pathog.* **11**, e1004591. (doi:10.1371/journal.ppat.1004591)

36. Fisman DN, Tuite AR, Brown KA. 2016 Impact of El Niño Southern Oscillation on infectious disease hospitalization risk in the United States. *Proc. Natl Acad. Sci. USA* **113**, 14 589–14 594. (doi:10.1073/pnas.1604980113)

37. Koelle K, Rodo X, Pascual M, Yunus M, Mostafa G. 2005 Refractory periods to climate forcing in cholera dynamics. *Nature* **436**, 696–700. (doi:10.1038/nature03820)

38. Rohani P, King AA. 2010 Never mind the length, feel the quality: the impact of long-term epidemiological data sets on theory, application and policy. *Trends Ecol. Evol.* **25**, 611–618. (doi:10.1016/j.tree.2010.07.010)

39. Patz JA, Campbell-Lendrum D, Holloway T, Foley JA. 2005 Impact of regional climate change on human health. *Nature* **438**, 310–317. (doi:10.1038/nature04188)

40. Denlinger DL, Armbruster PA. 2014 Mosquito diapause. *Ann. Rev. Entomol.* **59**, 73–93. (doi:10.1146/annurev-ento-011613-162023)

41. Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, Comrie A, Viboud C. 2013 Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS Pathog.* **9**, e1003194. (doi:10.1371/journal.ppat.1003194)

42. Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. 2016 Global environmental drivers of influenza. *Proc. Natl Acad. Sci. USA* **113**, 13 081–13 086. (doi:10.1073/pnas.1607747113)

43. Bjørnstad ON, Finkenstadt B, Grenfell BT. 2002 Endemic and epidemic dynamics of measles: estimating epidemiological scaling with a time series SIR model. *Ecol. Monog.* **72**, 169–184. (doi:10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2)

44. Hope Simpson R. 1952 Infectiousness of communicable diseases in the household: (measles, chickenpox, and mumps). *Lancet* **260**, 549–554. (doi:10.1016/S0140-6736(52)91357-3)

45. Imai C, Hashizume M. 2015 A systematic review of methodology: time series regression analysis for

environmental factors and infectious diseases. *Trop. Med. Health* **43**, 1–9. (doi:10.2149/tmh.2014-21)

46. Olafsdottir L, Erlendsdóttir H, Melo-Cristino J, Weinberger D, Ramirez M, Kristinsson K, Gottfredsson M. 2014 Invasive infections due to *Streptococcus pyogenes*: seasonal variation of severity and clinical characteristics, Iceland, 1975 to 2012. *Euro Surveill.* **19**, 5–14. (doi:10.2807/1560-7917.ES2014.19.17.20784)

47. Curriero FC, Patz JA, Rose JB, Lele S. 2001 The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Am. J. Publ. Health.* **91**, 1194–1199. (doi:10.2105/AJPH.91.8.1194)

48. Lewnard JA *et al.* 2014 Forecasting temporal dynamics of cutaneous Leishmaniasis in Northeast Brazil. *PLoS Neglect. Trop. D* **8**, e3283. (doi:10.1371/journal.pntd.0003283)

49. Johansson MA, Reich NG, Hota A, Brownstein JS, Santillana M. 2016 Evaluating the performance of infectious disease forecasts: a comparison of climate-driven and seasonal dengue forecasts for Mexico. *Sci. Rep.* **6**, 33707. (doi:10.1038/srep33707)

50. Harvey AC. 1990 *Forecasting, structural time series models and the kalman filter*. Cambridge, UK: Cambridge University Press.

51. Cazelles B, Chavez M, Berteaux D, Ménard F, Vik JO, Jenouvrier S, Stenseth NC. 2008 Wavelet analysis of ecological time series. *Oecologia* **156**, 287–304. (doi:10.1007/s00442-008-0993-2)

52. De Magny GC, Guégan J-F, Petit M, Cazelles B. 2007 Regional-scale climate-variability synchrony of cholera epidemics in West Africa. *BMC Infect. Dis.* **7**, 242. (doi:10.1186/1471-2334-7-20)

53. Gaudart J *et al.* 2013 Spatio-temporal dynamics of cholera during the first year of the epidemic in Haiti. *PLoS Neglect. Trop Dis.* **7**, e2145. (doi:10.1371/journal.pntd.0002145)

54. van Panhuis WG *et al.* 2015 Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. *Proc. Natl Acad. Sci. USA* **112**, 13 069–13 074. (doi:10.1073/pnas.1501375112)

55. Ferrari MJ, Grais RF, Bharti N, Conlan AJK, Bjørnstad ON, Wolfson LJ, Guerin PJ, Djibo A, Grenfell BT. 2008 The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684. (doi:10.1038/nature06509)

56. Atchison C, Tam C, Hajat S, Van Pelt W, Cowden J, Lopman B. 2010 Temperature-dependent transmission of rotavirus in Great Britain and The Netherlands. *Proc. R. Soc. B* **277**, 933–942. (doi:10.1098/rspb.2009.1755)

57. Wallinga J, Lipsitch M. 2007 How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* **274**, 599–604. (doi:10.1098/rspb.2006.3754)

58. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. 2010 Absolute humidity and the seasonal onset of influenza in the Continental US. *PLoS Biol.* **8**, e1000316. (doi:10.1371/journal.pbio.1000316)

59. Eisenberg MC, Kujbida G, Tuite AR, Fisman DN, Tien JH. 2013 Examining rainfall and cholera dynamics in Haiti using statistical and dynamic modeling approaches. *Epidemics* **5**, 197–207. (doi:10.1016/j.epidem.2013.09.004)

60. Lourenço J, Recker M. 2014 The 2012 Madeira dengue outbreak: epidemiological determinants and future epidemic potential. *PLoS Negl. Trop. Dis.* **8**, e3083. (doi:10.1371/journal.pntd.0003083)

61. Lauderdale JM *et al.* 2014 Towards seasonal forecasting of malaria in India. *Malar. J.* **13**, 310. (doi:10.1186/1475-2875-13-310)

62. Laneri K, Bhadra A, Ionides EL, Bouma M, Dhiman RC, Yadav RS, Pascual M, Ferguson NM. 2010 Forcing versus feedback: epidemic malaria and monsoon rains in northwest India. *PLoS Comput. Biol.* **6**, e1000898. (doi:10.1371/journal.pcbi.1000898)

63. Paynter S *et al.* 2014 Using mathematical transmission modelling to investigate drivers of respiratory syncytial virus seasonality in children in the Philippines. *PLoS ONE* **9**, e90094. (doi:10.1371/journal.pone.0090094)

64. Roy M, Bouma M, Dhiman RC, Pascual M. 2015 Predictability of epidemic malaria under non-stationary conditions with process-based models combining epidemiological updates and climate variability. *Malar. J.* **14**, 419. (doi:10.1186/s12936-015-0937-3)

65. Bejon P *et al.* 2014 A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots. *Elife* **3**, e02130. (doi:10.7554/eLife.02130)

66. Diuk-Wasser MA *et al.* 2010 Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States. *Glob. Ecol. Biogeogr.* **19**, 504–514.

67. Hahn MB, Jarnevich CS, Monaghan AJ, Eisen RJ. 2016 Modeling the Geographic Distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the Contiguous United States. *J. Med. Entomol.* **53**, 1176–1191. (doi:10.1093/jme/tjw076)

68. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, Tatem AJ, Hay SI. 2011 A new world malaria map: *Plasmodium falciparum* endemicity in 2010. *Malar. J.* **10**, 1. (doi:10.1186/1475-2875-10-378)

69. Gething PW *et al.* 2012 A long neglected world malaria map: *Plasmodium vivax* endemicity in 2010. *PLoS Negl. Trop. Dis.* **6**, e1814. (doi:10.1371/journal.pntd.0001814)

70. Bhatt S *et al.* 2013 The global distribution and burden of dengue. *Nature* **496**, 504–507. (doi:10.1038/nature12060)

71. Messina JP *et al.* 2016 Mapping global environmental suitability for Zika virus. *Elife* **5**, e15272. (doi:10.7554/eLife.15272)

72. Howes RE *et al.* 2011 The global distribution of the Duffy blood group. *Nat. Commun.* **2**, 266. (doi:10.1038/ncomms1265)

73. Ryan SJ, McNally A, Johnson LR, Mordecai EA, Ben-Horin T, Paaijmans K, Lafferty KD. 2015 Mapping physiological suitability limits for malaria in Africa under climate change. *Vector-Borne Zoonot.* **15**, 718–725. (doi:10.1089/vbz.2015.1822)

74. Diggle P, Moyeed R, Rowlingson B, Thomson M. 2002 Childhood malaria in the Gambia: a case-study in model-based geostatistics. *J. R. Stat. Soc. C-Appl.* **51**, 493–506. (doi:10.1111/1467-9876.00283)

75. Clements AC, Reid HL, Kelly GC, Hay SI. 2013 Further shrinking the malaria map: how can geospatial science help to achieve malaria elimination? *Lancet Infect. Dis.* **13**, 709–718. (doi:10.1016/S1473-3099(13)70140-3)

76. Bhatt S *et al.* 2015 The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* **526**, 207–211. (doi:10.1038/nature15535)

77. Perkins A, Siraj A, Ruktanonchai CW, Kraemer M, Tatem A. 2016 Model-based projections of Zika virus infections in childbearing women in the Americas. *Nat. Microbiol.* **1**, 1626. (doi:10.1101/039610)

78. Caminade C, Turner J, Metelmann S, Hesson JC, Blagrove MS, Solomon T, Morse AP, Baylis M. 2017 Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proc. Natl Acad. Sci. USA* **114**, 119–124. (doi:10.1073/pnas.1614303114)

79. Shaman J, Kohn M. 2009 Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc. Natl Acad. Sci. USA* **106**, 3243–3248. (doi:10.1073/pnas.0806852106)

80. Caminade C *et al.* 2014 Impact of climate change on global malaria distribution. *Proc. Natl Acad. Sci. USA* **111**, 3286–3291. (doi:10.1073/pnas.1302089111)

81. Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP. 2006 Why do we still use stepwise modelling in ecology and behaviour? *J. Anim. Ecol.* **75**, 1182–1189. (doi:10.1111/j.1365-2656.2006.01141.x)

82. Rodó X *et al.* 2013 Climate change and infectious diseases: can we meet the needs for better prediction? *Clim. Change* **118**, 625–640. (doi:10.1007/s10584-013-0744-1)

83. Cobey S, Baskerville EB. 2016 Limits to causal inference with state-space reconstruction for infectious disease. *PLoS ONE* **11**, e0169050. (doi:10.1371/journal.pone.0169050)

84. Heaney A, Little E, Ng S, Shaman J. 2016 Meteorological variability and infectious disease in Central Africa: a review of meteorological data quality. *Ann. NY Acad. Sci.* **1382**, 31–43. (doi:10.1111/nyas.13090)

85. Hallett TB, Coulson T, Pilkington JG, Clutton-Brock TH, Pemberton JM, Grenfell BT. 2004 Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature* **430**, 71–75. (doi:10.1038/nature02708)

86. Menne M, Durre I, Vose R, Gleason B, Houston T. 2012 An overview of the global historical climatology network-daily database. *J. Atmos. Ocean Tech.* **29**, 897–910. (doi:10.1175/JTECH-D-11-00103.1)

87. Durre I, Vose RS, Wuertz DB. 2006 Overview of the integrated global radiosonde archive. *J. Climate* **19**, 53–68. (doi:10.1175/JCLI3594.1)

88. Knapp KR et al. 2012 Globally gridded satellite (GridSat) observations for climate studies. *Bull. Am. Meteorol. Soc.* **92**, 893–907. (doi:10.1175/2011BAMS3039.1)

89. Dee D et al. 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597. (doi:10.1002/qj.828)

90. O'Gorman PA. 2015 Precipitation extremes under climate change. *Curr. Clim. Change Rep.* **1**, 49–59. (doi:10.1007/s40641-015-0009-3)

91. Dai A. 2013 Increasing drought under global warming in observations and models. *Nat. Clim. Change* **3**, 52–58. (doi:10.1038/nclimate1633)

92. Seth A, Rauscher SA, Biasutti M, Giannini A, Camargo SJ, Rojas M. 2013 CMIP5 projected changes in the annual cycle of precipitation in monsoon regions. *J. Clim.* **26**, 7328–7351. (doi:10.1175/JCLI-D-12-00726.1)

93. Shaman J, Karspeck A. 2012 Forecasting seasonal outbreaks of influenza. *Proc. Natl Acad. Sci. USA* **109**, 20 425–20 430. (doi:10.1073/pnas.1208772109)

94. Yang W, Karspeck A, Shaman J. 2014 Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLoS Comput. Biol.* **10**, e1003583. (doi:10.1371/journal.pcbi.1003583)

95. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. 2013 Real-time influenza forecasts during the 2012–2013 season. *Nat. Commun.* **4**, 2837. (doi:10.1038/ncomms3837)

96. Schellnhuber HJ, Frieler K, Kabat P. 2014 The elephant, the blind, and the intersectoral intercomparison of climate impacts. *Proc. Natl Acad. Sci. USA* **111**, 3225–3227. (doi:10.1073/pnas.1321791111)

97. Ehsan MA, Tippett MK, Almazroui M, Ismail M, Yousef A, Kucharski F, Omar M, Hussein M, Alkhalaf AA. 2016 Skill and predictability in multimodel ensemble forecasts for Northern Hemisphere regions with dominant winter precipitation. *Clim. Dyn.* **48**, 3309–3324. (doi:10.1007/s00382-016-3267-4)

98. Weinberger D, Baroux N, Grangeon J-P, Ko AI, Goarant C. 2014 El Niño Southern oscillation and leptospirosis outbreaks in New Caledonia. *PLoS Neglect. Trop. D* **8**, e2798. (doi:10.1371/journal.pntd.0002798)

10

rspb.royalsocietypublishing.org   *Proc. R. Soc. B* **284**: 20170901