Anticipating infectious disease re-emergence and elimination: a test of early warning signals using empirically based models

2021-12-29

Forecasts of the emergence, re-emergence, and elimination of human infectious diseases would allow for proactive, rather than reactive, decisions that could save lives. Recent theory suggests that a generic feature of dynamical systems approaching a tipping point – critical slowing down – can anticipate disease re-emergence and elimination. Empirical demonstrations of critical slowing down in real disease dynamics are scarce, but are essential before we can implement model-independent outbreak detection systems. Here, we use fitted, mechanistic models of measles transmission in four Nigerien cities to detect critical slowing down through statistical early warning signals. We find that several early warning signals accurately anticipate measles re-emergence and elimination, suggesting that critical slowing down can be detected before tipping points in real disease dynamics. Broadly, our findings suggest that early warning signals, coupled with decision-support algorithms and expert judgment, could provide the basis for outbreak early detection systems.

# Introduction

Forecasts of the emergence and re-emergence of infectious diseases have the potential to save lives, money, and human productivity by allowing for proactive, rather than reactive, preparedness measures [1]. Similarly, indicators of the elimination of infectious diseases can signal the effectiveness of ``end game’’ strategies aimed at disease eradication [2]. Predicting (re)emergence and elimination is possible with parametric mathematical models of disease transmission, but their success relies on detailed understanding of the underlying transmission dynamics and adequate data [3]. We often do not have enough information (or time) to parameterize such models. An alternative approach is to use model-independent statistical signals that portend infectious disease (re)emergence and elimination by detecting critical slowing down as the system approaches a critical transition [4,5].

Emergence and elimination of an infectious disease both involve a critical transition, reflected in deterministic models by a *transcritical bifurcation*, that occurs at the critical point where the effective reproduction number (, corresponding to the number of secondary cases that arise from a single infected case in a population) is equal to one [6]. Thus, subcritical () and supercritical () systems represent alternative dynamical regimes [4,7,8].

Critical transitions in stochastic systems, such as systems of disease transmission, are often accompanied by critical slowing down, a reduction in the resilience of a system to perturbations [9,10]. Critical slowing down (CSD) can be measured by changes in the dynamical features of the system: early warning signals (EWS) such as an increase in the variance and autocorrelation [7,11]. Recent theoretical work suggests that CSD occurs as disease dynamics approach from below (emergence) [4,12] and from above (elimination) [2,4,13], and that several EWS can anticipate the critical transition [14–16]. These findings suggest that early warning signals could be operationalized to develop early warning systems of disease emergence and elimination, or re-emergence and outbreaks of endemic diseases.

However, operationalizing EWS, and deploying early warning systems based on them, faces many challenges [1,17]. For example, using EWS in an “online mode” requires choosing temporal windows over which EWS are calculated. These moving windows should be long enough to provide reliable statistics, but short enough to forget the past so as to not overwhelm information contained in new observations. This will be especially important for diseases that fluctuate seasonally, where EWS might always increase and then decrease over the course of the year, requiring the end-user to reset computations each season [16]. Another challenge is defining thresholds for detection of an upcoming tipping point. Detection thresholds can be based on the absolute value of an EWS (e.g., warning if variance exceeds some value), the trend in an EWS over time (e.g., warning if the correlation of variance with time exceeds some value), or an algorithmic combination of many factors (e.g., variance and autocorrelation increases above some value several observation periods in a row). Ref. [18] has pioneered the approach of developing algorithms for combining EWS, their values, and their trends to best detect disease emergence and elimination, but much work remains to be done.

Given the amount of research still required to operationalize EWS, it is imperative that we stress test the hypothesis that EWS anticipate critical transitions before we allocate time and resources to the task of deploying early warning systems. One way to stress test EWS is through empirical case studies: can EWS anticipate re-emergence and elimination in observed time series of disease incidence [19]? However, uncritical application of EWS to observed data could lead to researchers getting the right answer for the wrong reasons. EWS might perform well for a given time series, but for reasons having nothing to do with critical slowing down (DUTTA 2018; Dakos 2015, Kefi 2013–ADD) . Likewise, critical transitions may occur in the absence of early warning signals [20]. Indeed, without knowing the critical point (i.e., when ) it is impossible to know if EWS are in fact sending us the signal we think they are.

Another option is to use fitted models of disease transmission to test EWS. This offers several advantages. First, using a model to simulate time series of cases means we also have access to a time series of , which allows us to know precisely when the critical transition occurs. This means we know whether we are getting the right answer for the right reason. Second, we can simulate replicate time series to account for the inherent stochasticity of disease transmission. This means our conclusions are not based on one-off events that bias conclusions [21]. Third, we can specifically simulate re-emergence and elimination events. This means we can separate the stress testing of EWS from the research necessary to operationalize EWS. Thus, we have all the flexibility of a theoretical model, but we remain tethered to reality because the model parameters are fitted to real data.

Here, we use fitted model simulations of measles dynamics to test whether CSD anticipates critical transitions in real disease dynamics. We focus on two scenarios: the re-emergence of measles following a large outbreak, a situation typical of measles dynamics in the Sahel [22], and the elimination of measles by gradual improvement of routine vaccination. We seek to answer two related questions. First, can CSD distinguish between time series of disease incidence when the underlying dynamics are far from and near to a critical transition? Second, can CSD anticipate disease re-emergence and elimination?

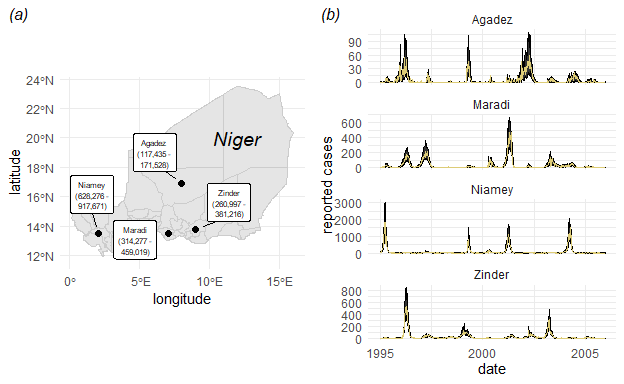
To answer these questions, we fit mechanistic models of disease transmission to time series of measles incidence in four Nigerien cities [22,23]. We then use the fitted models to perform model experiments designed to test the performance of several EWS, which signify CSD, at anticipating re-emergence and elimination. Our results confirm theoretical expectations about several EWS and associated CSD. In particular, we show that CSD before a critical transition is detectable by several EWS in realistic scenarios. However, our study highlights the limitations of EWS in situations where disease re-emergence and elimination occurs rapidly. Moreover, we find that EWS perform better at detecting CSD before re-emergence than before elimination.

# Materials and methods

## Data

We used weekly measles case report data from four Nigerien cities: Agadez, Maradi, Niamey, and Zinder (figure *a*). The data were collected over an 11 year period from 1995-2005 (figure *b*).

These data are ideal for testing theory on CSD in disease dynamics because each city has a different population size (spanning means of about 150 to 750 thousand during this time period), different dynamics in terms of outbreak sizes (maximum weekly incidence from 60 to 1845 cases) and length of inter-epidemic periods (2-5 years), and has different amounts of demographic stochasticity due to differences in population size. Such differences provide a natural gradient of extra-demographic and environmental stochasticity that may influence CSD [24–27]. The data were obtained from the Niger Ministry of Health [28].



Locations of data sources and observed and predicted measles dynamics. (*a*) Locations and 1995–2005 population-size–ranges (in parantheses) of our four focal cities in Niger. (*b*) Time series of weekly reported cases (incidence data; yellow solid lines) and the 68% prediction intervals (black ribbons) for one-week-ahead predictions from our fitted SEIR models for each city.

## Stochastic *SEIR* model

The model is a discrete-time approximation of a continuous-time SEIR model with limited demography, specified as a set of difference equations,

where is a time step of one day, and are the number of times different transition events occur in the time interval of . is the number of births, is the number of newly infected individuals that have the disease but are not infectious, is the number of newly infectious individuals, is the number of imported infections, and is the number of newly recovered individuals who are no longer infectious and have life-long immunity. The stochastic random variables are distributed as:

where is the *per capita* birth rate at time , is total population size in each city at time , is the vaccination probability, is the rate of imported infections, and , , and are the probabilities of exposure, becoming infectious, and recovery, respectively. These probabilities reflect the processes of transmission, transition from the latent period to the infectious period, and recovery, which were modeled as:

where is a time-varying rate of transmission, is a time-invariant rate of progression from the exposed class to the infectious class, and is a time-invariant recovery rate. Rate of transmission was modeled as,

where is the minimum transmission rate over the season, and the term is a B-spline to model seasonality in transmission. The B-spline bases () are periodic with a 1 year period. The transmission rate () is also subject to stochastic process noise at each time step, , which we model as gamma-distributed white (temporally uncorrelated) noise with mean 1 and intensity [29].

Death is not included in the model because the rate of infection is much faster than the rate of death. Excluding deaths means we can avoid making further assumptions about demographic rates – we are already making assumptions about birth rates (e.g., the rate is the same across cities, but with city-specific population size). Demographic stochasticity in births and imported infections was inlcuded by drawing time-specific values from Poisson distributions. In this model, the effective reproduction number at time may be approximated as: .

Observed case reports () were drawn from a negative binomial distribution subject to a constant reporting fraction () and dispersion parameter ,

where are the accumulated cases that transition from the infected class to the recovered class in a one week period. In this parameterization of the negative binomial, the mean is equal to and the the variance is equal to .

## Model fitting and inference

The model described above was fitte to the time series of case reports (incidence data) from each of our focal cities using Maximization by Iterated particle Filtering (MIF) [30]. We estimated 14 parameters for each city: six seasonal transmission parameters (), minimum transmission rate (), three initial conditions , the importation rate (), reporting fraction (), one parameter accounting for process noise (), and one parameter accounting for measurement noise (). To help parameter identifiability, mean incubation period was set to days and mean infectious period was set to days. The vaccination probability () was set to 70% for all times , consistent with reported vaccination coverage [22].

MIF relies on particle filtering, which estimates the likelihood of fixed parameters by integrating state variables of a stochastic system. To identify the maximum likelihood estimates, MIF lets parameters take a random walk during the filtering process and selectively propagates forward parameter sets (i.e., particles) with the highest likelihood. The variance of the random walk decreases at each iteration of MIF, where a MIF iteration means one filtering pass through the time series. This procedure converges toward the maximimum likelihood estimates (MLEs), in theory.

We used the IF2 algorithm [30] implemented in the R [31] package pomp version 1.18 [32,33] to conduct MIF. To initialize MIF, we generated 5000 parameter sets using Latin Hypercube Sampling over large ranges of the parameter values. We then performed two rounds of MIF, each for 100 iterations, with 10,000 particles, and geometric cooling. For the first round of MIF we set cooling.factor = 1. For the second round, which was initialized using the collection of parameter sets from the end of the first round, we set cooling.factor = 0.9. We computed the log likelihood of 5000 final MIF parameter sets (i.e., parameter sets collected after 200 MIF iterations) as the log of the mean likelihoods of 50 replicate particle filters with 10,000 particles each. At this stage, we assume the parameter set with highest log likelihood is the MLE.

We used bootstrapping to estimate approximate 95% confidence intervals for all parameters, conducted for each city independently, as follows. First, we simulated 100 realizations from the fitted model using the MLE parameters. Second, we fitted the SEIR model to each of the 100 bootstrap simulations using the same MIF procedure described above, except we initiated the parameter search from 50 parameter sets rather then 5000. We reduced the number of parameter sets due to the computational constraints of fitting 100 simulated data sets for each of the four cities. Third, we identified the MLE parameter set for each of the 100 bootstrap simulations from among the 50 MIF parameter sets. Last, we calculated summary statistics (mean, median, quantiles) from the distribution of 100 MLE parameters (electronic supplement).

## Model assessment

We used the MLE parameter sets to make one-week-ahead predictions and compared observed and expected case counts. To make one-week-ahead predictions, we used particle filtering with 50,000 particles and retained the mean and standard deviation of all latent states across all particles before they were filtered at each time step. We used the mean predictions () to assess model fit using a generalized coefficient of determination, calculated as: [34]. In addition to comparing model expectations to in-sample observed data, we also compared our fitted SEIR models to two benchmarks: a negative binomial sampling model that assumes independent and identically distributed observations and a SARIMA model (electronic supplement). We fit the benchmark models to the observed data and then compared models using Akaike’s Information Criterion (AIC) to account for differences in the number of model parameters [35].

## Model simulations

For fitting the SEIR model, we used known population size interpolated between years. This meant we were able to ignore certain demographic processes. For example, from the susceptible pool were ignored under the assumption that the infection rate was much faster than the death rate. We also ignored the recovered class because their dynamics, outside of the contribution to population size (which was assumed known), do not impact the compartments. However, births and deaths from all compartments, including , were needed when simulating the model over arbitrarily long time periods that do not necessarily represent real times for which we would have information on population size. Details on the *simulating model* are in the Online Supporting Information.

### Simulating re-emergence

To simulate re-emergence of measles, we manipulated the initial size of the susceptible pool to simulate an increase from low to high . Doing so allowed us to test whether EWS can distinguish between windows of time when is far from a critical transition and when is near a critical transition. We reduced the initial fraction of susceptible individuals by multiplying the MLE for by six discounting factors: 1e-4, 0.1, 0.2, 0.3, 0.4, and 0.5. These discounting factors represent situations of susceptible depletion after outbreaks of various size. After defining based on the discounting factor, we then set the initial number of recovered individuals to and set the initial number of exposed and infected individuals to zero. Initial population size for simulation scenarios, , was set to the mean population size for each city over the 1995-2005 time period. We then simulated the model forward for forty years using the mean birth rate for the entire country () and setting the death rate equal to the birth rate () to achieve a constant equilibrium total population size over the course of the simulation (total population size does vary, though, because of stochasticity in the model). Forty years was long enough for to reach or exceed 1 for each simulation replicate. Several outbreaks are seen within 11 years in the data (figure 1b), but times to outbreaks were longer in some scenarios where the susceptible population was initialized as much smaller than ever observed in reality. Because the model is stochastic, we repeated these simulations 500 times for each city–susceptible-discount combination.

Next, we split each simulated time series into null and test intervals. First, averaging across all simulations for a city-susceptible discount combination, we found the simulation year in which reaches or exceeds 1 and excluded years past that year (electronic supplement). We split the remaining time series into two windows of equal length (figure *a*). The null interval is the first window, where is increasing but far from 1. The test interval is the second window, where is increasing and approaching 1. We did this for each city and for each level of susceptible depletion. We calculated EWS over null and test intervals separately.

### Simulating elimination

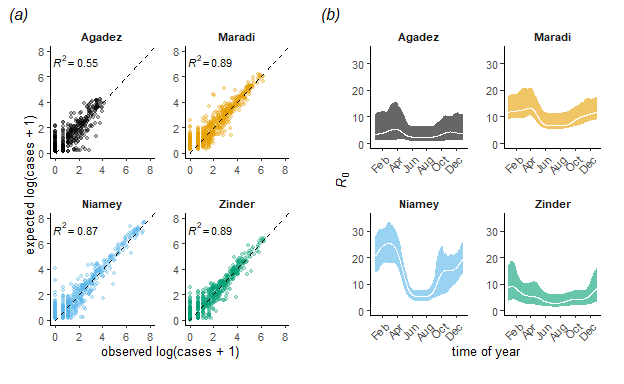
To simulate elimination, we simulated an improvement in routine vaccination in which the vaccination probability linearly increased over time to eventually reach 100%, i.e. eradication (figure S3). We ran simulations for 100 years, starting with 50 years of dynamics at the baseline vaccine coverage reported for Niger of 70%, [22]. Note that vaccination is included in our model by multiplying the birth rate of susceptibles by . At year 50, we initiated the improvement in routine vaccination and let the model run for another 50 years. We ran simulations across six vaccination ``speeds’’ (the rate at which ; electronic supplement), simulating situations of slow and fast approaches to elimination. As in the re-emergence simulations, we set the birth rate equal to the death rate to achieve a constant equilibrium population size.

We then split each time series into null and test intervals for calculating EWS. The test interval was defined as the window of time between the start of the vaccination improvements (year 50) and the time at which the vaccination probability reached the vaccination threshold of (figure *a*); was approximated for each city using the MLE parameters as: , where is the maximum value of over its seasonal variation. The null interval was defined as the window of time that ends at the start of the vaccination improvements (end of year 49) and starts at a time that results in an interval equal in length to the test interval (figure *a*). EWS were then calculated for each interval.

## Calculating early warning signals

We considered nine candidate early warning signals (table S3). We used the spaero::get\_stats() function [36] in R [31] to calculate EWS according to the formulas in table S3. All EWS except the coefficient of variation are expected to increase as approaches 1 from below [4,13,15]. We are not aware of theoretical results for the behavior of these EWS as approaches 1 from above that are applicable to our fitted model’s dynamics, which are highly non-linear (figure S2). But a natural expectation is that the mean should decrease as the endemic equilibrium of our model’s deterministic skeleton moves toward zero.

For each simulation of re-emergence and elimination, we calculated EWS for the time series of expected cases in the null and test intervals. This yielded a distribution of EWS over the 500 null and test intervals. We assessed the performance of each EWS using the Area Under the Curve (AUC) statistic. Specifically, we use AUC to calculate the amount of overlap between the distributions of each EWS from the null and test intervals. Values of AUC far from 0.5 (i.e., close to 0 or 1) indicate a greater degree of separation and thus better performance of a particular EWS in terms of classifying whether is close to a critical transition. We calculated AUC as: where is the sum of the ranks of test set EWS statistics in a combined set of null and test statistics (lower numbers have lower ranks), is the number of test of statistics and is the number of null statistics. The AUC of an EWS is the probability that a randomly chosen EWS value from the test set is higher than an EWS value randomly chosen from the null set [37]. Therefore, AUC should be high (closer to 1) when an EWS is expected to increase as a critical transition is approached, whereas AUC should be low (closer to 0) when an EWS is expected to decrease.



Accuracy of the fitted *SEIR* models and estimated seasonality. (*a*) Comparison of in-sample model predictions and observations for each city. Expected cases are one-week-ahead predictions from the fitted models. The dashed line shows 1:1. Coefficients of determination () were calculated as the reduction in the sum-of-squared errors from model predictions relative to a null model of the mean number of cases (Methods). (*b*) The estimated seasonality of the basic reproductive ratio () for each city. was approximated as: , where is the incubation period, is the infectious period, is the time-specific rate of transmission, and is the death rate. Only is estimated by our model. We set = 8 days, = 5 days, and = 0.05 for calculating as shown in this figure. The white line is calculated using the MLE parameters; shaded regions are the bootstrapped 95% confidence intervals.

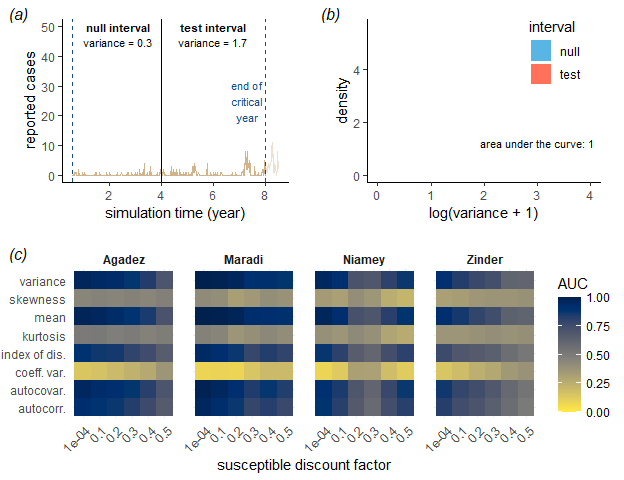
# Results

The fitted models adequately reproduce observed dynamics (figure *b*), with in-sample s from one-week-ahead predictions ranging from 0.55 for Agadez to 0.89 for Maradi (figure *a*). The fitted models also had lower AIC values than two benchmarking models (table S1). The estimated seasonality is consistent with previously estimated patterns (figure S2), including the decline in seasonality amplitude as population size decreases (figure *b*) [22].

Our model for Agadez performed poorly relative to the other cities, but still did better than non-mechanistic models (table S1). Maximum likelihood estimates and bootstrapped 95% confidence intervals for all parameters are in the electronic supplement (tables S4-S7).

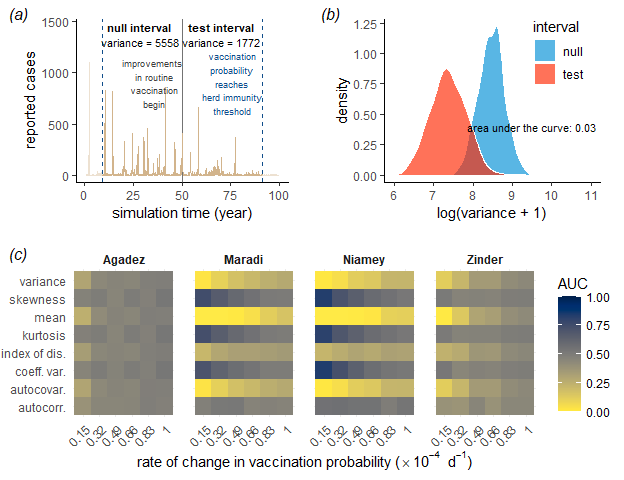
The EWS generally performed as expected from theory on the approach to re-emergence. Most EWS increased as the critical transition is approached, resulting in AUC values above 0.5 and near 1 (figure ). Skewness, kurtosis, and coefficient of variation performed poorly across all levels of susceptible depletion in all cities.

Variance, mean, index of dispersion, autocovariance, and autocorrelation all performed equally well at predicting re-emergence (figure *c*). Their performance declined as the size of the susceptible pool increased (i.e., a larger susceptbile discounting factor). This is expected because a larger susceptible pool results in more rapid returns to , which, in turn, results in shorter null and test intervals, making estimates of EWS less precise [14]. Thus, re-emergence may prove difficult to anticipate in “fast” transmission systems, as observed in simulations by ref. [14] and seen here when susceptible depletion was relatively small (figure *c*).



Performance of early warning signals (EWS) over fixed windows on the approach to emergence. (*a*) A typical example of an emergence simulation for Maradi. The two vertical blue lines indicate the start (left-most line) and end (line for critical year) of the full window. The black line demarcates the division between the equal-length null and test intervals, in which we show the calculated variance. (*b*) Empirical densities of variance in the null and test intervals across 500 simulations and the associated area under the curve (AUC) statistic. (*c*) Heatmap of AUC statistics for each EWS at each level of susceptible discount factor. AUC values closer to 0 or 1 indicate higher ability to distinguish among time series near and far from a critical transition. See figure S8 for a visualization of how susceptible discounting factor maps to number of weeks in the null and test intervals.

The EWS performed less well when anticipating elimination, relative to emergence (figure ). Only three metrics were reliable: mean, autocovariance, and variance. All three metrics decreased as approached the critical transition (figure S6). As in the case of anticipating elimination, AUC values moved closer to 0.5 as the rate of improvement in routine vaccination increased (figure *c*).



Performance of early warning signals (EWS) over fixed windows on the approach to elimination. (*a*) A typical example of an elimination simulation for Maradi. The two vertical blue lines indicate the start (left-most line) and end (line for critical year) of the full window. The black line demarcates the division between the equal-length null and test intervals, in which we show the calculated variance. (*b*) Empirical densities of variance in the null and test intervals across 500 simulations and the associated area under the curve (AUC) statistic. (*c*) Heatmap of AUC statistics for each EWS at each speed of approach to herd immunity. AUC values closer to 0 or 1 indicate higher ability to distinguish among time series near and far from a critical transition. See figure S8 for a visualization of how vaccination speed maps to number of weeks in the null and test intervals.

In all, the suite of EWS suggested that indicators of critical slowing down have some predictive value for emergence and elimination events. We found similar results for the approach to elimination when calculating EWS over a moving window of 35 weeks in the null and test intervals (figure S9b). But, all EWS performed worse when predicting the approach to emergence over the moving window (figure S9).

# Discussion

The ability to detect early warnings of outbreaks of potentially fatal diseases such as measles during non-epidemic periods of unpredictable duration may facilitate planning such as enhanced surveillance to expedite outbreak detection, implementation of serological surveys to identify immunity gaps, and initiation of targeted supplemental vaccination [38,39]. Further, it has been argued that measles eradication requires consideration of local demographic factors [40] and that regional outbreak response vaccination (ORV) can be effective when initiated early, even if coverage is suboptimal [38]. Consequently indicators that a region is on a path to outbreak or elimination can help to prioritize the timing and distribution of limited resources. Using empirically-based disease transmission models, we found that generic indicators of critical slowing down were informative regarding simulated re-emergence and elimination of measles.

The generic indicators performed better in scenarios of re-emergence. Moreover, these indicators behaved as expected based on a simple one-dimensional model of fluctuations around an equilibrium with gradually declining stability – the prototypical model of critical slowing down. This is because the autocorrelation and variance increased on the approach to (figure *c*). If, alternatively, the variance had increased but the autocorrelation had not, a better supported model would be an equilibrium with fixed stability but subject to increasing intensity of perturbations. From a model selection point of view, our simulations of re-emergence can be said to support the prototypical critical slowing down model. And since our simulations were fit to measles data and had high one-step-ahead predictive accuracy, it suggests that the prototypical model of critical slowing down may be supported by data sets similar to the ones we fit and can provide predictive value for them. The definitive test of this approach, or course, would be to record prospective predictions of re-emergence based on indicators of critical slowing down and evaluate their accuracy.

The lag-1 autocorrelation was not a strong indicator of the transition to elimination (figure *c*). Thus, the prototypical model of critical slowing down appears less useful in this scenario. Other EWS (variance, autocovariance, and mean) did show altered behavior as the transition to elimination was approached, but these EWS were less sensitive under elimination scenarios compared to re-emergence scenarios. These results suggest that although distributional changes in indicators occur prior to disease elimination, interpreting these changes in terms of loss of stability of an endemic equilibrium, or declining , requires a model which accounts for complicating system features such as seasonality, non-linearity, damped oscillations, and local extinction.

A potential limitation of our findings is that the susceptible discount factors in our simulation study (figure *c*) might be smaller than the factors that occur in reality. A smaller discount factor corresponds to a larger level of susceptible depletion. To check the relevance of this limitation, we calculated the level of susceptible depletion after outbreaks (defined as years where the total number of cases reached 80% of the maximum observed) across one hundred replicate simulations (electronic supplement). We found that susceptible depletion was less than 0.5, the smallest susceptible depletion level we tested, for 0.9% of outbreaks in Agadez, 21% of outbreaks in Maradi, 100% of outbreaks in Niamey, and 26% of outbreaks in Zinder. These statistics do not detract from our main findings of CSD in measles dynamics, but they do suggest that EWS might be less useful in some cases than in others. For example, AUC values for emergence at the 0.5 level of susceptible depletion are already low for most cities (figure *c*). Thus, our analysis methods are not practical for cities that rarely experience levels of susceptible depletion below 0.5 (e.g., Agadez).

Unpredictable, recurring outbreaks with seasonality in transmission such as those observed for measles in Niger during 1995–2005 are challenging settings for the application of EWS in part because theoretical predictions of EWS behavior [14,27] are typically restricted to models with simpler dynamics. Consequently, the development of robust, model-independent early warning systems for infectious diseases [1] likely will benefit from study of the behavior of EWS in more complex models. Also, although we have shown that critical slowing down precedes tipping points in stochastic models that were fit to data with state-of-the-art methods, how to operationalize the phenomenon of critical slowing down remains an open research area [41]. Emerging technologies like artificial intelligence might offer new ways to find optimal detection thresholds for early warning signals. But there will always be a role for expert judgment. Early warning signals, although powerful and now accompanied with additional support, will likely be just one part of a decision-support toolkit.

# Acknowledgements

We thank Tobias Brett for comments on the modeling approach and Paige Miller and Éric Marty for helpful comments on early versions of this work. We thank Epicentre/MSF-OCP and the Niger Ministry of Health for making the data available.

# Funding

This research was funded by the National Institute of General Medical Sciences of the National Institutes of Health (Award Number U01GM110744). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This work was done on the Olympus High Performance Compute Cluster located at the Pittsburgh Supercomputing Center at Carnegie Mellon University, which is supported by National Institute of General Medical Sciences Modeling Infectious Disease Agent Study (MIDAS) Informatics Services Group grant 1U24GM110707.

# Authors’ contributions

ATT conceived of the study, designed the study, carried out the statistical and modeling analysis, and drafted the manuscript; EBO participated in the design of the study, participated in statistical and modeling analysis, and critically revised the manuscript; MJF and AWP helped revise the manuscript; PR participated in the conception and design of the study, guided the statistical and modeling analysis, and helped draft and revise the manuscript; JMD coordinated the study, participated in the conception and design of the study, guided the statistical and modeling analysis, and helped draft and revise the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

# References

1. Han BA, Drake JM. 2016 Future directions in analytics for infectious disease intelligence. *EMBO reports*, e201642534. (doi:[10.15252/embr.201642534](https://doi.org/10.15252/embr.201642534))

2. Drake JM, Hay SI. 2017 Monitoring the path to the elimination of infectious diseaes. *Tropical Medicine and Infectious Disease* **2**, 20.

3. Metcalf CJE, Lessler J. 2017 Opportunities and challenges in modeling emerging infectious diseases. *Science* **357**, 149–152. (doi:[10.1126/science.aam8335](https://doi.org/10.1126/science.aam8335))

4. O’Regan SM, Drake JM. 2013 Theory of early warning signals of disease emergenceand leading indicators of elimination. *Theoretical Ecology* **6**, 333–357. (doi:[10.1007/s12080-013-0185-5](https://doi.org/10.1007/s12080-013-0185-5))

5. Drake JM *et al.* 2019 The statistics of epidemic transitions. *PLoS Computational Biology* **15**, e1006917. (doi:[10.1371/journal.pcbi.1006917](https://doi.org/10.1371/journal.pcbi.1006917))

6. Heffernan JM, Smith RJ, Wahl LM. 2005 Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface* **2**, 281–293. (doi:[10.1098/rsif.2005.0042](https://doi.org/10.1098/rsif.2005.0042))

7. Scheffer M *et al.* 2009 Early-warning signals for critical transitions. *Nature* **461**, 53–59. (doi:[10.1038/nature08227](https://doi.org/10.1038/nature08227))

8. Scheffer M *et al.* 2012 Anticipating critical transitions. *Science* **338**, 344–348. (doi:[10.1126/science.1225244](https://doi.org/10.1126/science.1225244))

9. Wissel C. 1984 A universal law of the characteristic return time near thresholds. *Oecologia* **65**, 101–107. (doi:[10.1007/BF00384470](https://doi.org/10.1007/BF00384470))

10. Nes EH van, Scheffer M. 2007 Slow Recovery from Perturbations as a Generic Indicator of a Nearby Catastrophic Shift. *The American Naturalist* **169**, 738–747. (doi:[10.1086/516845](https://doi.org/10.1086/516845))

11. Carpenter SR, Brock WA. 2006 Rising variance: A leading indicator of ecological transition. *Ecology Letters* **9**, 311–318. (doi:[10.1111/j.1461-0248.2005.00877.x](https://doi.org/10.1111/j.1461-0248.2005.00877.x))

12. Dibble CJ, O’Dea EB, Park AW, Drake JM. 2016 Waiting time to infectious disease emergence. *Journal of the Royal Society Interface* **13**, 20160540. (doi:[10.1098/rsif.2016.0540](https://doi.org/10.1098/rsif.2016.0540))

13. O’Regan SM, Lillie JW, Drake JM. 2016 Leading indicators of mosquito-borne disease elimination. *Theoretical Ecology* **9**, 269–286. (doi:[10.1007/s12080-015-0285-5](https://doi.org/10.1007/s12080-015-0285-5))

14. Brett TS, Drake JM, Rohani P. 2017 Anticipating the emergence of infectious diseases. *Journal of the Royal Society Interface* **14**, 20170115. (doi:[10.1098/rsif.2017.0115](https://doi.org/10.1098/rsif.2017.0115))

15. Brett TS, O’Dea EB, Marty É, Miller PB, Park AW, Drake JM, Rohani P. 2018 Anticipating epidemic transitions with imperfect data. *PLoS Computational Biology* **14**, e1006204. (doi:[10.1371/journal.pcbi.1006204](https://doi.org/10.1371/journal.pcbi.1006204))

16. Miller PB, O’Dea EB, Rohani P, Drake JM. 2017 Forecasting infectious disease emergence subject to seasonal forcing. *Theoretical Biology and Medical Modelling* **14**, 17. (doi:[10.1186/s12976-017-0063-8](https://doi.org/10.1186/s12976-017-0063-8))

17. Boettiger C, Hastings A. 2013 Tipping points: From patterns to predictions. *Nature* **493**, 157–158.

18. Brett TS, Rohani P. 2020 Dynamical footprints enable detection of disease emergence. *PLoS Biology* **18**, e3000697. (doi:[10.1371/journal.pbio.3000697](https://doi.org/10.1371/journal.pbio.3000697))

19. Harris MJ, Hay SI, Drake JM. 2020 Early warning signals of malaria resurgence in Kericho, Kenya. *Biology Letters* **16**, 20190713. (doi:[10.1098/rsbl.2019.0713](https://doi.org/10.1098/rsbl.2019.0713))

20. Boettiger C, Hastings A. 2012 Early warning signals and the prosecutor’s fallacy. *Proceedings of the Royal Society B: Biological Sciences* **279**, 4734–4739. (doi:[10.1098/rspb.2012.2085](https://doi.org/10.1098/rspb.2012.2085))

21. Boettiger C, Hastings A. 2012 Quantifying limits to detection of early warning for critical transitions. *Journal of the Royal Society Interface* **9**, 2527–2539. (doi:[10.1098/rsif.2012.0125](https://doi.org/10.1098/rsif.2012.0125))

22. 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](https://doi.org/10.1038/nature06509))

23. Bharti N, Tatem AJ, Ferrari MJ, Grais RF, Djibo A, Grenfell BT. 2011 Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* (doi:[10.1126/science.1210554](https://doi.org/10.1126/science.1210554))

24. Hastings A, Wysham DB. 2010 Regime shifts in ecological systems can occur with no warning. *Ecology Letters* **13**, 464–472. (doi:[10.1111/j.1461-0248.2010.01439.x](https://doi.org/10.1111/j.1461-0248.2010.01439.x))

25. Dakos V, Van Nes EH, D’Odorico P, Scheffer M. 2012 Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology* **93**, 264–271. (doi:[10.1890/11-0889.1](https://doi.org/10.1890/11-0889.1))

26. O’Dea EB, Park AW, Drake JM. 2018 Estimating the distance to an epidemic threshold. *Journal of the Royal Society Interface* **15**, 20180034. (doi:[10.1098/rsif.2018.0034](https://doi.org/10.1098/rsif.2018.0034))

27. O’Regan SM, Burton DL. 2018 How Stochasticity Influences Leading Indicators of Critical Transitions. *Bulletin of Mathematical Biology* **80**, 1630–1654. (doi:[10.1007/s11538-018-0429-z](https://doi.org/10.1007/s11538-018-0429-z))

28. Niger Ministry of Health. 2008 Weekly measles case reports, 19 95–2004.

29. Bretó C, Ionides EL. 2011 Compound Markov counting processes and their applications to modeling infinitesimally over-dispersed systems. *Stochastic Processes and their Applications* **121**, 2571–2591. (doi:[10.1016/j.spa.2011.07.005](https://doi.org/10.1016/j.spa.2011.07.005))

30. Ionides EL, Nguyen D, Atchadé Y, Stoev S, King AA. 2015 Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences* **112**, 719–724. (doi:[10.1073/pnas.1410597112](https://doi.org/10.1073/pnas.1410597112))

31. R Core Team. 2017 R: A language and environment for statistical computing.

32. King AA, Nguyen D, Ionides EL. 2016 Statistical Inference for Partially Observed Markov Processes via the R Package pomp. *Journal of Statistical Software* **69**, 1–43. (doi:[10.18637/jss.v069.i12](https://doi.org/10.18637/jss.v069.i12))

33. King AA *et al.* 2018 pomp: Statistical Inference for Partially Observed Markov Processes (R package, version 1.18).

34. Martinez-Bakker M, King AA, Rohani P. 2015 Unraveling the transmission ecology of polio. *PLoS Biology* (doi:[10.1371/journal.pbio.1002172](https://doi.org/10.1371/journal.pbio.1002172))

35. Akaike H. 1973 Information theory and an extension of the maximum likelihood princilple. In *Proceedings of the 2nd international symposium on information theory*, pp. 267–281. (doi:[10.1007/978-0-387-98135-2](https://doi.org/10.1007/978-0-387-98135-2))

36. O’Dea EB. 2018 spaero: Software for Project AERO (R package version 0.3.0).

37. Fawcett T. 2006 An introduction to ROC analysis. *Pattern Recognition Letters* **27**, 861–874. (doi:<https://doi.org/10.1016/j.patrec.2005.10.010>)

38. Grais Rf, Conlan Ajk, Ferrari Mj, Djibo A, Le Menach A, Bjørnstad On, Grenfell Bt. 2008 Time is of the essence: Exploring a measles outbreak response vaccination in Niamey, Niger. *Journal of The Royal Society Interface* **5**, 67–74. (doi:[10.1098/rsif.2007.1038](https://doi.org/10.1098/rsif.2007.1038))

39. Weldegebriel GG, Gasasira A, Harvey P, Masresha B, Goodson JL, Pate MA, Abanida E, Chevez A. 2011 Measles Resurgence Following a Nationwide Measles Vaccination Campaign in Nigeria, 2005–2008. *The Journal of Infectious Diseases* **204**, S226–S231. (doi:[10.1093/infdis/jir136](https://doi.org/10.1093/infdis/jir136))

40. Ferrari MJ, Grenfell BT, Strebel PM. 2013 Think globally, act locally: The role of local demographics and vaccination coverage in the dynamic response of measles infection to control. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120141. (doi:[10.1098/rstb.2012.0141](https://doi.org/10.1098/rstb.2012.0141))

41. Shmueli G, Burkom H. 2010 Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics* **52**, 39–51. (doi:[10.1198/TECH.2010.06134](https://doi.org/10.1198/TECH.2010.06134))