

Research



Cite this article: Harris MJ, Hay SI, Drake JM. 2020 Early warning signals of malaria resurgence in Kericho, Kenya. *Biol. Lett.* **16**: 20190713.
<http://dx.doi.org/10.1098/rsbl.2019.0713>

Received: 11 October 2019
Accepted: 19 February 2020

Subject Areas:
health and disease and epidemiology

Keywords:
malaria, forecasting, infectious disease, resurgence, statistics

Author for correspondence:
Mallory J. Harris
e-mail: mharris9@stanford.edu

Electronic supplementary material is available online at <http://dx.doi.org/10.6084/m9.figshare.c.4880289>.

Early warning signals of malaria resurgence in Kericho, Kenya

Mallory J. Harris^{1,3}, Simon I. Hay^{4,5} and John M. Drake^{1,2}

¹Odum School of Ecology, and ²Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA 30602, USA

³Biology Department, Stanford University, 371 Serra Mall, Stanford, CA, USA

⁴Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA 98121, USA

⁵Department of Health Metrics Sciences, School of Medicine, University of Washington, Seattle, WA 98121, USA

id MJH, 0000-0001-8196-0430; SIH, 0000-0002-0611-7272

Campaigns to eliminate infectious diseases could be greatly aided by methods for providing early warning signals of resurgence. Theory predicts that as a disease transmission system undergoes a transition from stability at the disease-free equilibrium to sustained transmission, it will exhibit characteristic behaviours known as critical slowing down, referring to the speed at which fluctuations in the number of cases are dampened, for instance the extinction of a local transmission chain after infection from an imported case. These phenomena include increases in several summary statistics, including lag-1 autocorrelation, variance and the first difference of variance. Here, we report the first empirical test of this prediction during the resurgence of malaria in Kericho, Kenya. For 10 summary statistics, we measured the approach to criticality in a rolling window to quantify the size of effect and directions. Nine of the statistics increased as predicted and variance, the first difference of variance, autocovariance, lag-1 autocorrelation and decay time returned early warning signals of critical slowing down based on permutation tests. These results show that time series of disease incidence collected through ordinary surveillance activities may exhibit characteristic signatures prior to an outbreak, a phenomenon that may be quite general among infectious disease systems.

1. Introduction

Despite modern advances in disease control, the World Health Organization reports that nearly one-third of deaths in developing countries are due to infectious diseases [1]. So far, only one human global disease eradication campaign has been successful (smallpox). Especially in low-resource settings, premature abandonment of disease elimination programmes as other health priorities compete for limited resources may result in reestablishment, even when the disease is on the brink of elimination. For instance, in Sri Lanka, malaria incidence climbed from 17 to 440 644 cases between 1963 and 1968 [2,3]. Indeed, malaria resurgence occurred repeatedly throughout the last century. A systematic review identified 75 resurgence events worldwide between 1930 and 2000, attributed largely to the weakening of malaria control programmes, increasing the potential for transmission and evolution of drug and insecticide resistance [2]. Exactly those conditions are found for other diseases now on the cusp of global eradication (e.g. polio [4], Dracunculiasis [5] and human Africa trypanosomiasis [6]). It follows that effective disease elimination might be guided by techniques for differentiating between efforts that will lead to stable disease elimination and scenarios where cases are diminished without a corresponding shift in the system's stability. Few methods are available that are appropriate to

Table 1. Names and equations for 10 summary statistics used in these analyses. The correlation coefficient and p -value corresponding to testing between December 1981 and April 1993 is also given for each indicator.

statistic	formula	correlation coefficient (τ)	p -value
mean	$m_i(x_j)$	1.000	0.157
variance	$m_i((x_j - \text{mean}_i)^2)$	0.791	0.039
(variance first difference)	$\text{variance}_i - \text{variance}_{i-1}$	0.461	0.153
autocovariance	$m_i((x_j - \text{mean}_i)(x_{j-1} - \text{mean}_{j-1}))$	0.890	0.008
(lag-1 autocorrelation)	$\frac{\text{autocovariance}_i}{(\text{variance}_i * \text{variance}_{i-1})^2}$	0.739	0.064
(decay time)	$\frac{-1}{\log(\min(\max(\text{autocorrelation}_i, 0), 1))}$	0.739	0.033
(index of dispersion)	$\frac{\text{variance}_i}{\text{mean}_i}$	0.709	0.063
(coefficient of variation)	$\frac{(\text{variance}_i)_0^5}{\text{mean}_i}$	0.537	0.168
skewness	$\frac{m_i((x_j - \text{mean}_i)^3)}{(\text{variance}_i)^{1.5}}$	0.145	0.415
kurtosis	$\frac{m_i((x_j - \text{mean}_i)^4)}{(\text{variance}_i)^2}$	-0.017	0.505

the relevant problems of limited data and non-stationary disease dynamics at these transition points.

Forecasting methods that depend directly on raw case incidence have been proposed for infectious diseases, but techniques such as the moving epidemic method, which establishes thresholds for identifying the start of an epidemic, rely on regular, seasonal trends and are therefore not applicable for predicting the resurgence of vector-borne diseases [7]. Recent theory suggests that critical transitions in infectious disease systems may be detected using non-parametric statistical methods [8,9]. Resurgence is a type of critical transition, wherein the stable state of the disease system shifts from the disease-free equilibrium (where cases are due to imported infections and subcritical stuttering chains of local transmission) to self-sustained transmission. In systems where the critical transition is driven by a gradual shift in underlying disease parameters toward conditions favouring disease transmission, any outbreak is preceded both by an approach to criticality, which ends when transmission is exactly maintained by an effective reproduction number of one, and a delay as susceptible individuals accumulate in the population [10]. During the approach to criticality, the transmission system is expected to exhibit slower returns to mean (i.e. longer and longer chains of subcritical transmission) following the introduction of cases, which is a manifestation of critical slowing down [11]. The mathematical theory of critical slowing down has recently been worked out in detail for a number of specific disease systems [8–10,12]. General predictions of the theory are that cases will exhibit an increase in several summary statistics (table 1) as the critical point is approached [8,13,14]. These early warning signals would precede the increase in cases at the beginning of the supercritical period that triggers an alert in the moving epidemic method [7].

This paper reports the first test of this theory on real rather than simulated data for a vector-borne disease. Here,

we show the evidence of critical slowing down in a time series of monthly *Plasmodium falciparum* malaria case incidence leading up to a resurgence in Kericho, Kenya. The April 1993 resurgence event has been attributed to the parasite's development of resistance to chloroquine, the anti-malarial drug administered to workers as a control strategy on these plantations [15]. Thus, the proposed mechanism of the disease's return is a gradual underlying process—the increasing frequency of resistance in the population—rather than a sudden event like importation. In five of the 10 indicators, we detect increases that serve as early warning signals of malaria resurgence during the approach to criticality. Theory suggests this phenomenon may be a broadly applicable one.

2. Methods

Data were obtained from two relatively isolated tea plantations for 1965 through 2002. The company-operated hospital provided case numbers based on stained blood smears from patients with suspected malaria [16,17]. Malaria reached epidemic levels in these areas during World War II, carried by soldiers returning from Ethiopia. In March 1948, a combination of proguanil prophylaxis and DDT house spraying brought malaria incidence down, and transmission was mostly controlled via the anti-malarial medication chloroquine until the April 1993 outbreak, when a spike in the number of cases exceeded all prior months by two and a half fold (figure 1) [15]. Operationally, we therefore define April 1993 as the boundary between subcritical and supercritical transmission. It is known that chloroquine remained fully effective against *P. falciparum* through December 1981 [16]. Thus, we considered December 1981 to April 1993 to be the period during which the system approached criticality. We performed a sensitivity analysis to determine the effect of alternative endpoints for this window (electronic supplementary material, files S1 and S2). The theory of critical slowing down developed by O'Regan *et al.* [8] predicts that the indicators of critical slowing down should have increased during this time (table 1).

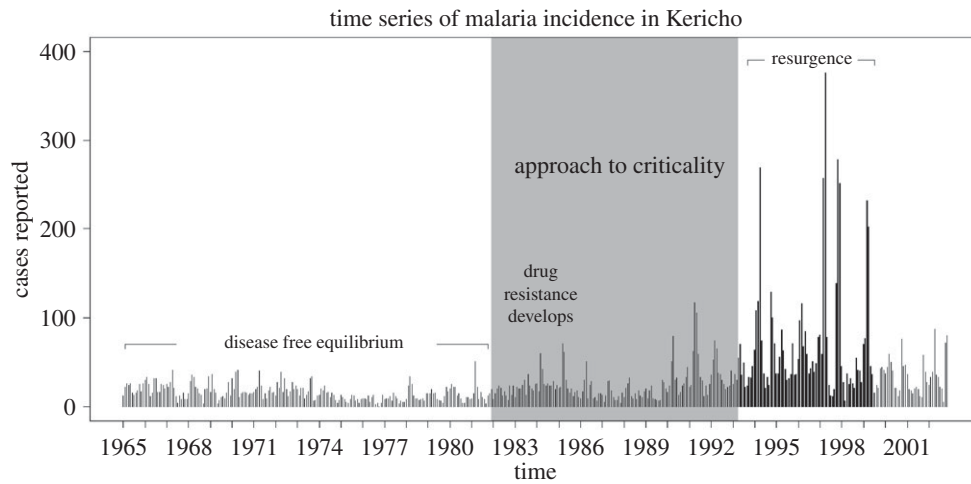


Figure 1. Monthly cases reported over the course of the time series. The approach to criticality preceding the 1993 outbreak is indicated by grey shading.

To test this prediction, from December 1981 to April 1993, we calculated each of these 10 leading indicators over a rolling window (i.e. the statistics were repeatedly calculated within a fixed size temporal subsample of the data, shifting the window by one month for each computation) (table 1). To better distinguish the signal of critical slowing down from background noise and periodic trends, we preprocessed the raw case data by detrending with a Gaussian kernel. The early warning signal itself was calculated so that for a moving window centred at time index i , a corresponding kernel weight w_{ij} was assigned to the value of the statistic (mean, variance, etc.) at index j and then normalized so that $N_i = \sum_j w_{ij}$. Thus, for the time i and statistic f , the rolling window estimator m_i was defined as

$$m_i(f_j(x)) = \sum_{j=i-b+1}^i \frac{w_{ij}f_j(x)}{N_i},$$

where the only tuning parameter is the bandwidth $b=40$ (here set to 40 months, exceeding the period of malaria's seasonality to reduce any influence of seasonality on the overall trend). A sensitivity analysis shows that statistical results are relatively insensitive to the choice of bandwidth (electronic supplementary material, file S2). All analyses were conducted in R using the *spaero* package [18], a freely available package that provides a wide range of methods for investigating critical slowing down in time series. Data and code for reproducing analyses reported in this paper are available online at <https://github.com/mjharris95/Kericho-EWS.git>.

The performance of each indicator as an early warning signal was then assessed by performing significance tests on the rolling window statistics as follows. First, we determined the trend in each indicator over time by calculating Kendall's correlation coefficient, τ , between the calculated rolling window statistic at each time and the time index i . There are other statistics one could use (e.g. Spearman's correlation coefficient), but the use of Kendall's τ has become standard in the literature on anticipating critical transitions and our results are robust to choice of correlation coefficient (electronic supplementary material, file S3) [19–22]. The sign of τ indicates the direction of change, while the magnitude corresponds to the strength of the temporal trend. Unlike simulation studies, where an unlimited number of examples are available to investigate detectability, analysis based on data requires a means to assign statistical significance without access to replication. We used the following permutation approach to assess the significance of the estimated correlations. We first generated 10 000 permutations of case incidence during the approach to criticality (December 1981 to April 1993), thus preserving the general features of the original data (e.g. amplitudes of fluctuations) while eliminating underlying patterns of serial

dependence. We then performed rolling window analyses on the permuted data to generate a null distribution of correlation coefficients for each indicator. The false positive rates of the 10 tests were estimated by finding the proportion of these permutations with a value for τ more extreme than that calculated for the Kericho data. The resulting p -values quantify the reliability of their corresponding statistics as early warning signals. In an applied surveillance setting, time series data would be updated regularly. To reflect this, we first included only data collected at least 96 months prior to the critical transition (April 1985) and repeatedly added an additional month of data, recalculating p -values for each month.

3. Results

The test values for all indicators except for kurtosis and skewness lie in the right tails of their null distributions, consistent with the prediction that critical slowing down results in an increase in these statistics (electronic supplementary material, file S4). Visually, it appears that most of the indicators exhibit intermittent periods of both upward and downward trend. The first difference of variance was the first indicator to return a low p -value (about 65 months prior to the critical transition), although the p -values for this indicator increased after a few months, remaining high for the remainder of the study (figure 2). The p -values for increases in decay time and lag-1 autocorrelation both dropped rapidly 24 months prior to the critical transition. The p -values for increases in autocovariance and variance also rapidly decreased 24 months prior to the critical transition and dropped below 0.05 fifteen and four months prior to the critical transition, respectively.

4. Discussion

These results show that simple summary statistics may serve as early warning signals of disease resurgence. Nine of the 10 statistical indicators increased during the approach to criticality, as predicted by the theory of critical slowing down, although only the increases in half of these indicators corresponded to low p -values (below 0.05) at any point in the 96 months preceding the critical transition. The first difference of variance was the first statistic to return early warning signals, followed by autocorrelation and decay time, then autocovariance, then variance. In the months immediately preceding the critical transition, the p -values for increases in the latter four declined sharply. This suggests that certain

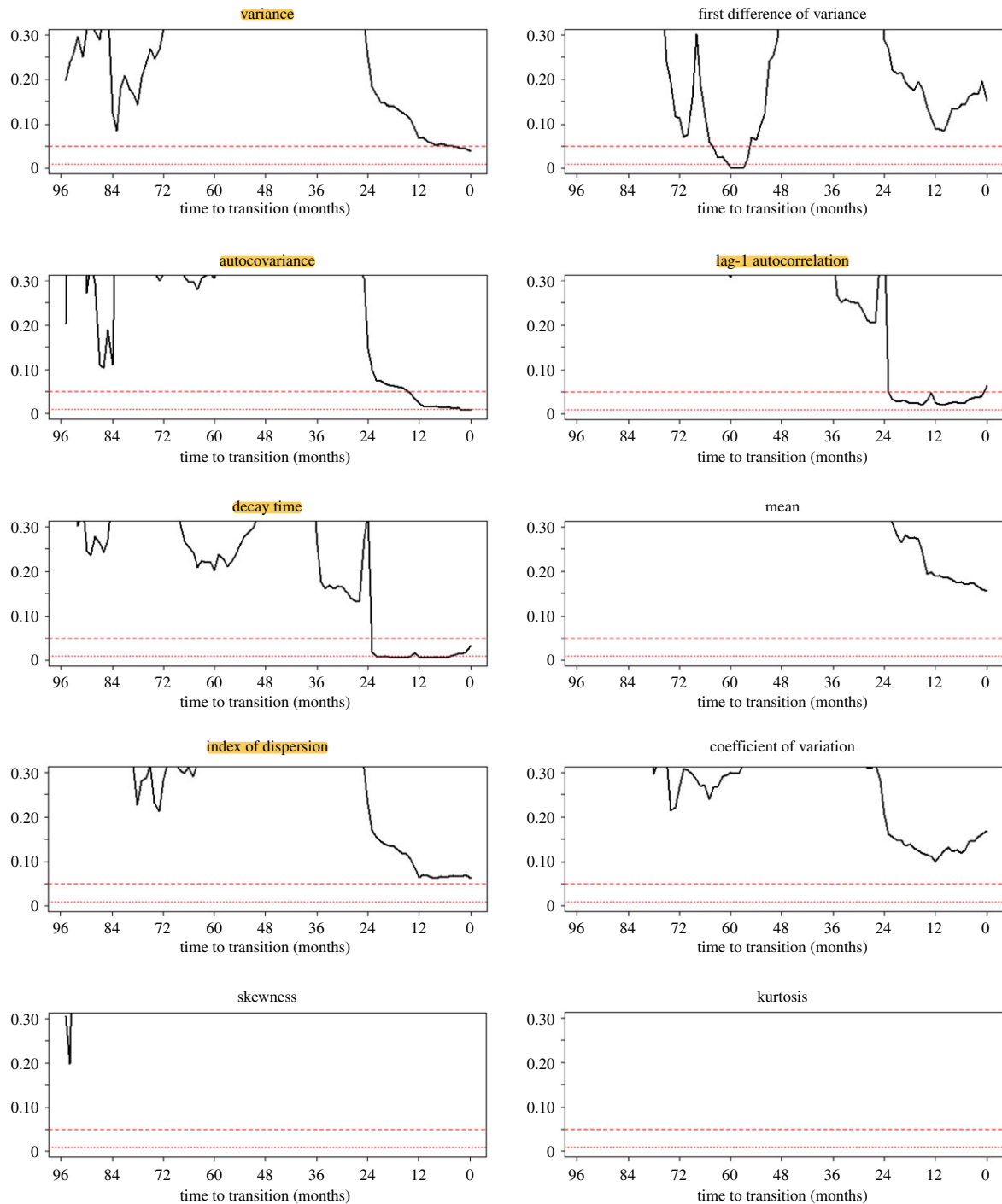


Figure 2. Time series. These plots give the p -value of the signal from each indicator, computed monthly starting in April 1985, 96 months prior to the notional month of critical transition, to April 1993. Red horizontal lines indicate p -values of 0.05 and 0.01 for reference.

indicators may be more reliable early warning signals of critical slowing down at different times during the approach to criticality. The finding that variance and lag-1 autocorrelation are relatively robust early warning indicators is consistent with findings in simulation studies [8,12,19] and in other (non-disease) natural systems [11,20]. Our conclusion is that the statistical signatures of critical slowing down may be detected in ordinary surveillance data prior to disease resurgence and therefore should be incorporated into monitoring programmes and decision support for proactive response. The statistical features of the lag between the critical transition and disease resurgence, particularly in vector-borne disease systems, are not yet well understood [10].

There are qualifications to this conclusion. Our methods assume that the system is forced through a critical transition due to some slow-moving underlying mechanism, causing one or more disease transmission parameters to vary over time. The Kericho example meets these conditions, drawing from a nearly closed system without major human immigration, and undergoing a critical transition due to the gradual development of parasite resistance to chloroquine, with drug-resistant parasites becoming prevalent in the areas of Kenya starting in the early 1980s [23]. Vector-borne disease resurgence provoked by imported cases or a sudden change in the efficacy or administration of control measures is not expected to be detected in this manner. However, the accumulation of insecticide and pesticide resistance remains

leading causes of malaria resurgence [24]. Moreover, in many areas, slow environmental forcing may drive additional changes in transmission, for instance as climate change shifts the geographic region suitable for malaria transmission [25]. Furthermore, this framework could be used to predict a critical transition toward elimination, provided that the system undergoes a gradual approach to criticality, for instance through increasing insecticide and bed net usage [8,9]. Thus, we think monitoring programmes based on critical slowing down might be used to inform intervention strategies by helping to distinguish between true changes in system stability and a temporary decrease in cases.

Although this study provides evidence that rolling window statistics may be used to anticipate disease resurgence, applications may be hampered by data limitations [12]. For instance, under-reporting, especially failure to detect asymptomatic carriers, may corrupt disease surveillance data. In our study, since the population eligible for healthcare remained constant over the course of the time series, inaccurate case reports are unlikely to vary in any systematic way, justifying our decision to use case reporting as a measure of incidence [16]. Moreover, theoretical studies suggest that early warning indicators based on critical slowing down are surprisingly robust to imperfect reporting [9,12,26]. Under intervention strategies that change case recovery rate, variance is expected to be robust to under-reporting, whereas autocorrelation is expected to be sensitive to under-reporting [8]. Further theoretical studies are needed on the indicators of critical slowing down, particularly decay time, in the context of vector-borne disease systems. Investment in improved surveillance could alleviate the risk of costly resurgence responses. A related problem is that the models of the disease system approaching a critical transition used to identify potential early warning signals consider the total number of infectious individuals, whereas reported data (such as studied here) reflect the number of cases reported in a sampling interval (here, monthly hospital reports) [16,26]. In our case, since the resolution of the time series is a month, exceeding the initial infectious period of malaria, and the system slowly approaches resurgence, new cases are approximately equivalent to the total infectious individuals at a time point. But, such convenient alignment of case reporting and disease transmission may not be obtained in other situations. We recommend more extensive investigation of critical slowing down across infectious disease systems, particularly to identify which indicators most reliably predict critical slowing down and to analyse the effects of bandwidth and notional month of transition on early warning signal detection. Additional studies could also inform the selection of detection thresholds based on p -value and correction for testing based on multiple statistics

that are not independent. Data requirements for such testing include regularly reported measures of incidence for tens to hundreds of infection generations, occurrence of outbreaks and a method for determining when the system reached the critical point. This last criterion could be satisfied either through parametric modelling of the data themselves or from independent information.

In conclusion, the results of this study support the use of rolling window calculations of indicators of critical slowing down in disease surveillance [8,9,12,14,26]. This study also demonstrates that early warning signals can be detected in monthly incidence data several years prior to a critical transition. The algorithmic nature of these methods points to the possibility that automated, model-independent disease forecasting systems could be developed for application directly to clinical data to assist in the detection of trends and predict disease resurgence, improving the efficiency of disease control campaigns and contributing to a welcome shift from reactive to preemptive response. For instance, we envision early warning systems that routinely monitor not only changes in the number of reported cases but also changes in these summary statistics and trigger an alert when a threshold is reached [8,9,12,14,26]. The detection of early warning signals could be followed by efforts to more specifically identify whether underlying features of the system are shifting (e.g. the development of drug resistance). In sum, being able to predict critical transitions might make malaria control (and other vertical eradication programmes) more efficient by prompting preemptive action prior to resurgence.

Data accessibility. Code used in the analyses and to generate figures and the corresponding data are available at <https://github.com/mjharris95/Kericho-EWS.git>

Authors' contributions. J.M.D., M.J.H. and S.I.H. designed and completed the study. J.M.D. and M.J.H. analysed the data. All authors contributed to the drafting of the manuscript. All authors agree to be held accountable for the content therein and approve the final version of the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported through the University of Georgia's Odum School of Ecology, the UGA Center for Undergraduate Research Opportunities and National Institute of General Medical Sciences of the National Institutes of Health under Award Number U01GM110744. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

Acknowledgements. The malaria inpatient data were supplied by the chief physician Dr Walter Odonde, Unilever Tea Kenya Ltd and updated to 2010 by Mr Geoffrey Kores, the Records Officer. Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results.

References

1. Mohsen N *et al.* 2017 Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet (London, England)* **390**, 1151–1210. (doi:10.1016/S0140-6736(17)32152-9)
2. Cohen JM, Smith DL, Cotter C, Ward A, Yamey G, Sabot OJ, Moonen B. 2012 Malaria resurgence: a systematic review and assessment of its causes. *Malar. J.* **11**, 122. (doi:10.1186/1475-2875-11-122)
3. Gubler DJ. 1998 Resurgent vector-borne diseases as a global health problem. *Emerg. Infect. Dis.* **4**, 442–450. (doi:10.3201/eid0403.980326)
4. Khan F, Datta SD, Quddus A, Vertefeuille JF, Burns CC, Jorba J, Wassilak SGF. 2018 Progress toward polio eradication—worldwide, January 2016–March 2018. *MMWR. Morb. Mortal. Wkly. Rep.* **67**, 524–528. (doi:10.15585/mmwr.mm6718a4)
5. Cromwell EA *et al.* 2018 Slaying little dragons: the impact of the Guinea Worm Eradication Program on

- dracunculiasis disability averted from 1990 to 2016. *Gates Open Res.* **2**, 30. (doi:10.12688/gatesopenres.12827.1)
6. Rock KS *et al.* 2018 Assessing strategies against gambiense sleeping sickness through mathematical modeling. *Clin. Infect. Dis.* **66**, S286–S292. (doi:10.1093/cid/ciy018)
 7. Vega T, Lozano JE, Meerhoff T, Snacken R, Mott J, Ortiz de Lejarazu R, Nunes B. 2013 Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method. *Influenza Other Respi. Viruses* **7**, 546–558. (doi:10.1111/j.1750-2659.2012.00422.x)
 8. O'Regan SM, Lillie JW, Drake JM. 2016 Leading indicators of mosquito-borne disease elimination. *Theor. Ecol.* **9**, 269–286. (doi:10.1007/s12080-015-0285-5)
 9. Drake JM, Hay SI. 2017 Monitoring the path to the elimination of infectious diseases. *Trop. Med. Infect. Dis.* **2**, 20. (doi:10.3390/tropicalmed2030020)
 10. Dibble CJ, O'Dea EB, Park AW, Drake JM. 2016 Waiting time to infectious disease emergence. *J. R. Soc. Interface* **13**, 20160540. (doi:10.1098/rsif.2016.0540)
 11. Scheffer M. 2009 *Critical transitions in nature and society*. Princeton, NJ: Princeton University Press.
 12. Brett TS, O'Dea EB, Marty É, Miller PB, Park AW, Drake JM, Rohani P. 2018 Anticipating epidemic transitions with imperfect data. *PLoS Comput. Biol.* **14**, e1006204. (doi:10.1371/journal.pcbi.1006204)
 13. Nesbit RM, Gurney WSC. 1982 *Modelling fluctuation populations*. Hoboken, NJ: Wiley.
 14. Miller PB, O'Dea EB, Rohani P, Drake JM. 2017 Forecasting infectious disease emergence subject to seasonal forcing. *Theor. Biol. Med. Model.* **14**, 17. (doi:10.1186/s12976-017-0063-8)
 15. Shanks GD, Hay SI, Omumbo JA, Snow RW. 2005 Malaria in Kenya's Western Highlands. *Emerg. Infect. Dis.* **11**, 1425–1432. (doi:10.3201/eid1109.041131)
 16. Shanks GD, Hay SI, Stern DI, Biomndo K, Snow RW. 2002 Meteorologic influences on *Plasmodium falciparum* malaria in the highland tea estates of Kericho, Western Kenya. *Emerg. Infect. Dis.* **8**, 1404–1408. (doi:10.3201/eid0812.020077)
 17. Harris MJ, Hay SI, Drake JM. 2020 Data from: code to accompany evidence of critical slowing down prior to malaria resurgence in Kericho, Kenya. *Github*. See <https://github.com/mjharris95/Kericho-EWS>.
 18. O'Dea E. 2016 spaero: software for project AERO.
 19. Pananos AD, Bury TM, Wang C, Schonfeld J, Mohanty SP, Nyhan B, Salathé M, Bauch CT. 2017 Critical dynamics in population vaccinating behavior. *Proc. Natl Acad. Sci. USA* **114**, 13 762–13 767. (doi:10.1073/pnas.1704093114)
 20. Lenton TM. 2011 Early warning of climate tipping points. *Nat. Clim. Chang.* **1**, 201–209. (doi:10.1038/nclimate1143)
 21. Taranu ZE, Carpenter SR, Frossard V, Jenny J-P, Thomas Z, Vermaire JC, Perga M-E. 2018 Can we detect ecosystem critical transitions and signals of changing resilience from paleo-ecological records? *Ecosphere* **9**, e02438. (doi:10.1002/ecs2.2438)
 22. Kéfi S *et al.* 2014 Early warning signals of ecological transitions: methods for spatial patterns. *PLoS ONE* **9**, e92097. (doi:10.1371/journal.pone.0092097)
 23. Spencer HC, Kaseje DCO, Brandling-Bennett AD, Oloo AJ, Churchill FC, Koech DK. 1987 Changing response to chloroquine of *Plasmodium falciparum* in Saradidi, Kenya, from 1981 to 1984. *Ann. Trop. Med. Parasitol.* **81**, 98–104. (doi:10.1080/00034983.1987.11812194)
 24. Bødker R, Kisinza W, Malima R, Msangeni H, Lindsay S. 2000 Resurgence of malaria in the Usambara mountains, Tanzania, an epidemic of drug-resistant parasites. *Glob. Chang. Hum. Heal.* **1**, 134–153. (doi:10.1023/A:1010077105257)
 25. Ryan SJ, McNally A, Johnson LR, Mordecai EA, Ben-Horin T, Paaijmans KP, Lafferty KD. 2015 Mapping physiological suitability limits for malaria in Africa under climate change. *Vector Borne Zoonotic Dis.* **15**, 718–725. (doi:10.1089/vbz.2015.1822)
 26. O'Dea EB, Drake JM. 2018 Disentangling reporting and disease transmission. *Theor. Ecol.* **12**, 89–98. (doi:10.1007/s12080-018-0390-3)