Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data

Amy Wesolowski[a,b,c,1,2], C. J. E. Metcalf[d,e,f,1,2], Nathan Eagle[a,g], Janeth Kombich[h], Bryan T. Grenfell[d,e], Ottar N. Bjørnstad[i], Justin Lessler[j], Andrew J. Tatem[c,f,k], and Caroline O. Buckee[a,b,c]

[a]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115; [b]Center for Communicable Disease Dynamics, Harvard School of Public Health, Boston, MA 02115; [c]Flowminder Foundation, SE-113 55 Stockholm, Sweden; [d]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; [e]Woodrow Wilson School, Princeton University, Princeton, NJ 08544; [f]Fogarty International Center, National Institute of Health, Bethesda, MD 20892; [g]Department of Computer Science, Northeastern University, Boston, MA 02115; [h]Department of Biological Sciences, University of Kabianga, Kericho Country, Kenya; [i]Center for Infectious Disease Dynamics, The Pennsylvania State University, State College, PA 16801; [j]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205; and [k]Department of Geography and Environment, University of Southampton, Southampton SO17 1BJ, United Kingdom

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved July 21, 2015 (received for review December 11, 2014)

Changing patterns of human aggregation are thought to drive annual and multiannual outbreaks of infectious diseases, but the paucity of data about travel behavior and population flux over time has made this idea difficult to test quantitatively. Current measures of human mobility, especially in low-income settings, are often static, relying on approximate travel times, road networks, or cross-sectional surveys. Mobile phone data provide a unique source of information about human travel, but the power of these data to describe epidemiologically relevant changes in population density remains unclear. Here we quantify seasonal travel patterns using mobile phone data from nearly 15 million anonymous subscribers in Kenya. Using a rich data source of rubella incidence, we show that patterns of population travel (fluxes) inferred from mobile phone data are predictive of disease transmission and improve significantly on standard school term time and weather covariates. Further, combining seasonal and spatial data on travel from mobile phone data allows us to characterize seasonal fluctuations in risk across Kenya and produce dynamic importation risk maps for rubella. Mobile phone data therefore offer a valuable previously unidentified source of data for measuring key drivers of seasonal epidemics.

rubella | mobile phones | population mobility | Kenya | seasonality

Seasonal variation in infectious disease incidence is a ubiquitous phenomenon observed for a range of pathogens such as malaria, measles, and influenza (1–7). Understanding and quantifying key mechanisms that drive seasonal variability such as climatic conditions (malaria and influenza) or patterns of human aggregation (measles and influenza) contribute to our fundamental understanding of epidemic dynamics; they also have important implications for evaluating public health measures that may reduce transmission such as vaccination and school closures (8–11).

The effectiveness of any public health measure designed to reduce seasonal transmission by modifying patterns of human aggregation and travel will depend on the degree to which transmission depends on population density and movement. Direct measures of population travel are rare (2, 4, 12, 13). As a result, proxy measures such as school terms and rainfall patterns have been used (1, 9, 13–15). Term time forcing, where school-driven aggregation leads to seasonal peaks of transmission for directly transmitted immunizing infections such as measles, mumps, and rubella, has been observed in many high-income countries (8, 16, 17) [England and Wales (8), Peru (15), and Denmark (17)]. On the other end of the spectrum in the low-income, predominantly agricultural context of Niger (13), analysis of night lights indicates that peaks in transmission reflect population changes resulting from annual mass migrations of individuals between agricultural areas to cities in the dry season (1). School terms and holidays versus agricultural movements likely represent the extremes in terms of density-related drivers of transmission in high- and low-income settings, respectively. Identifying where other low-income countries lie along this continuum will be important in order to plan public health measures and understand the dynamics of these infectious diseases.

Recently, the ubiquity of mobile phone ownership and use, particularly in low-income settings, has produced vast datasets on the individual movement patterns of millions of people (18–20). These data represent a potentially valuable previously unidentified resource for the direct observation of seasonal population dynamics on refined spatial scales that underlie disease transmission. Unlike other static measures of travel such as road networks, small-scale surveys, travel time surfaces, or census data, mobile phone data can provide a dynamic picture of mobility and population travel and changes over large geographic scales (1, 5, 12, 13, 21). However, their relevance to infectious disease dynamics has yet to be formally assessed, requiring parallel longitudinal estimates of phone use and transmission rates (19). For example, although we have previously estimated the rate of malaria importation across Kenya using spatially stratified prevalence estimates in combination with mobile phone data (20), because these estimates were static, the interaction between temporal dynamics in human mobility and malaria incidence was impossible to assess.

Here we analyze anonymous mobile phone call records including the location of the routing tower and timing of each call and short message service (SMS) communication between

Significance

Changing patterns of human mobility can drive seasonal outbreaks of infectious diseases, but limited data about travel behavior and population flux over time have made this idea difficult to quantify. Mobile phone data provide a unique source of information about human travel. Here we quantify seasonal travel patterns using mobile phone data from nearly 15 million anonymous subscribers in Kenya. Using a rich data source of rubella incidence, we show that patterns of population fluxes inferred from mobile phone data are predictive of disease transmission and improve significantly on standard school term time and weather covariates, showing for the first time to our knowledge that mobile phone data capture epidemiologically relevant patterns of movement.

Author contributions: A.W., C.J.E.M., B.T.G., and C.O.B. designed research; A.W., C.J.E.M., and C.O.B. performed research; A.W., C.J.E.M., N.E., J.K., B.T.G., O.N.B., J.L., and A.J.T. contributed new reagents/analytic tools; A.W., C.J.E.M., and A.J.T. analyzed data; and A.W., C.J.E.M., B.T.G., A.J.T., and C.O.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]A.W. and C.J.E.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: awesolow@hsph.harvard.edu or cmetcalf@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1423542112/-/DCSupplemental.

14,816,512 subscribers during June 2008 to June 2009 (with February 2009 missing from the dataset) to quantify daily travel patterns (referred to as population fluxes) between provinces (*Materials and Methods* and *Supporting Information*). We combine this with a long-term and highly detailed dataset on rubella incidence to measure how key drivers of human aggregation are associated with disease transmission in this setting.

Results

Transmission and Mobile Phone-Derived Population Fluxes. Unlike the classical two-peak seasonal pattern of many immunizing childhood infections, including rubella (e.g., in Peru, Mexico, Japan, South Africa, and Canada), rubella in Kenya exhibits three peaks of incidence each year (Fig. 1 *A* and *B* and Figs. S1 and S2) (15, 22–25). Similar patterns have been reported in other countries in East Africa, suggesting regional consistency (26). In the majority of provinces, transmission peaks in September and February (except Northeastern province, where inference is weak given paucity of data; Fig. S1). Epidemiological mechanisms driving this within-year triple peak in incidence are the occurrence of three yearly troughs in transmission (Fig. 1*C*) and the relatively low R_0 of this infection (*Materials and Methods* and *Supporting Information*).

To measure the ecological and spatial mechanisms contributing to the three troughs in transmission, we turned to both previous proxy measures of movement and a previously unidentified measure extracted from anonymized mobile phone data. For the latter case, we constructed a daily location for each subscriber and calculated the number of trips from all subscribers between provinces (eight total provinces) per day (*Materials and Methods*, Fig. 2*A*, *Supporting Information*, and Fig. S3) (20). These subscriber locational time series were aggregated to create a province-level daily population flux that described the number of journeys originating from each province per day. This province-level time series was decomposed using a moving average, and the trend was analyzed (*Materials and Methods* and *Supporting Information*).

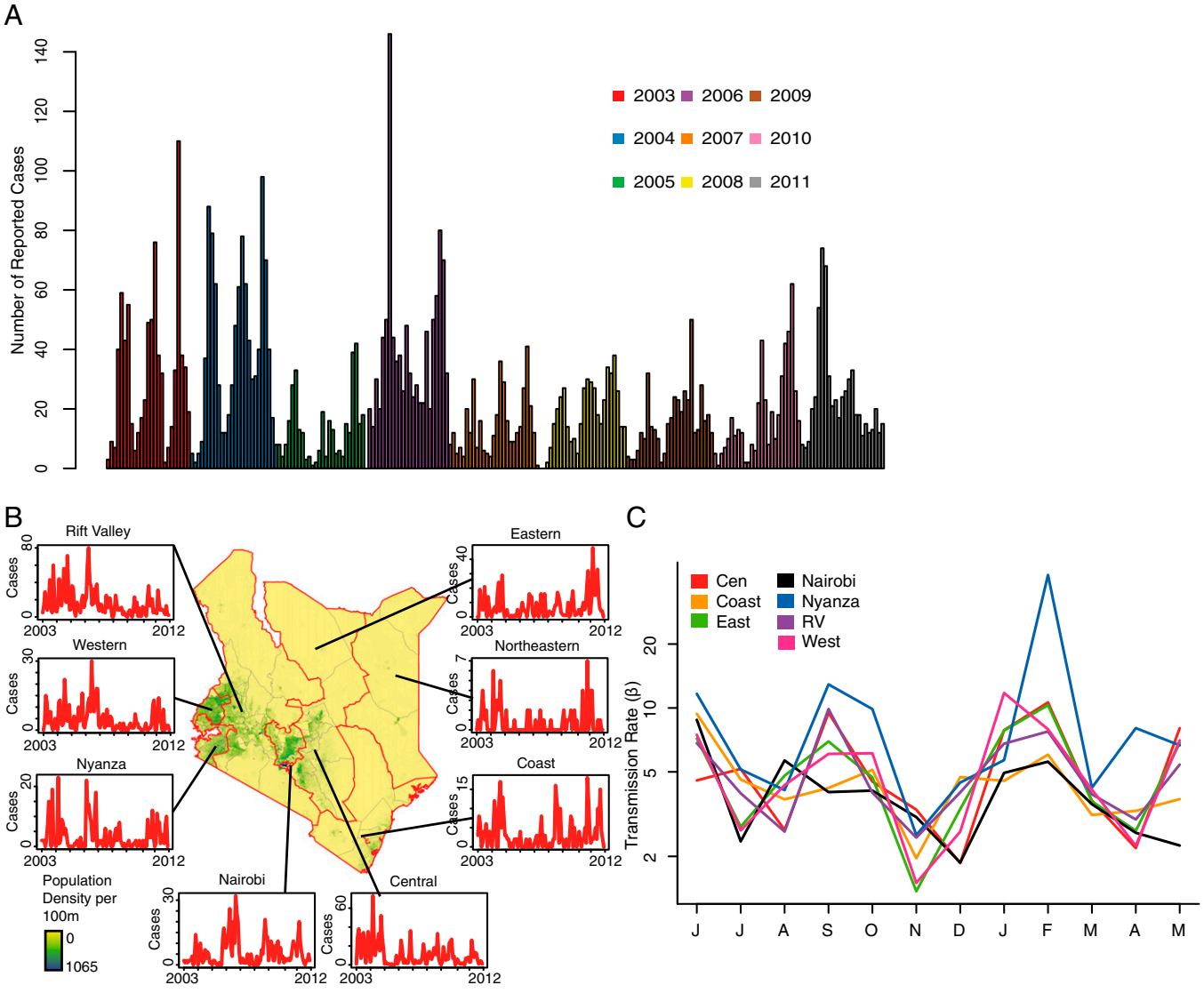

Fig. 1. The rubella patterns in Kenya. (*A*) Over the course of the rubella dataset (January 2008 to January 2012), the biweekly number of reported rubella cases is shown. (*B*) A map of Kenya with provinces outlined in red along with the rubella case data per province. (*C*) The monthly transmission (beta) estimates per province. In the majority of provinces, there are two pronounced peaks in transmission during September and January–March with a number of locations also peaking in May–June (see Central province in red, for example). The peaks in transmission varied geographically with the Nyanza province peaking in February, for example.

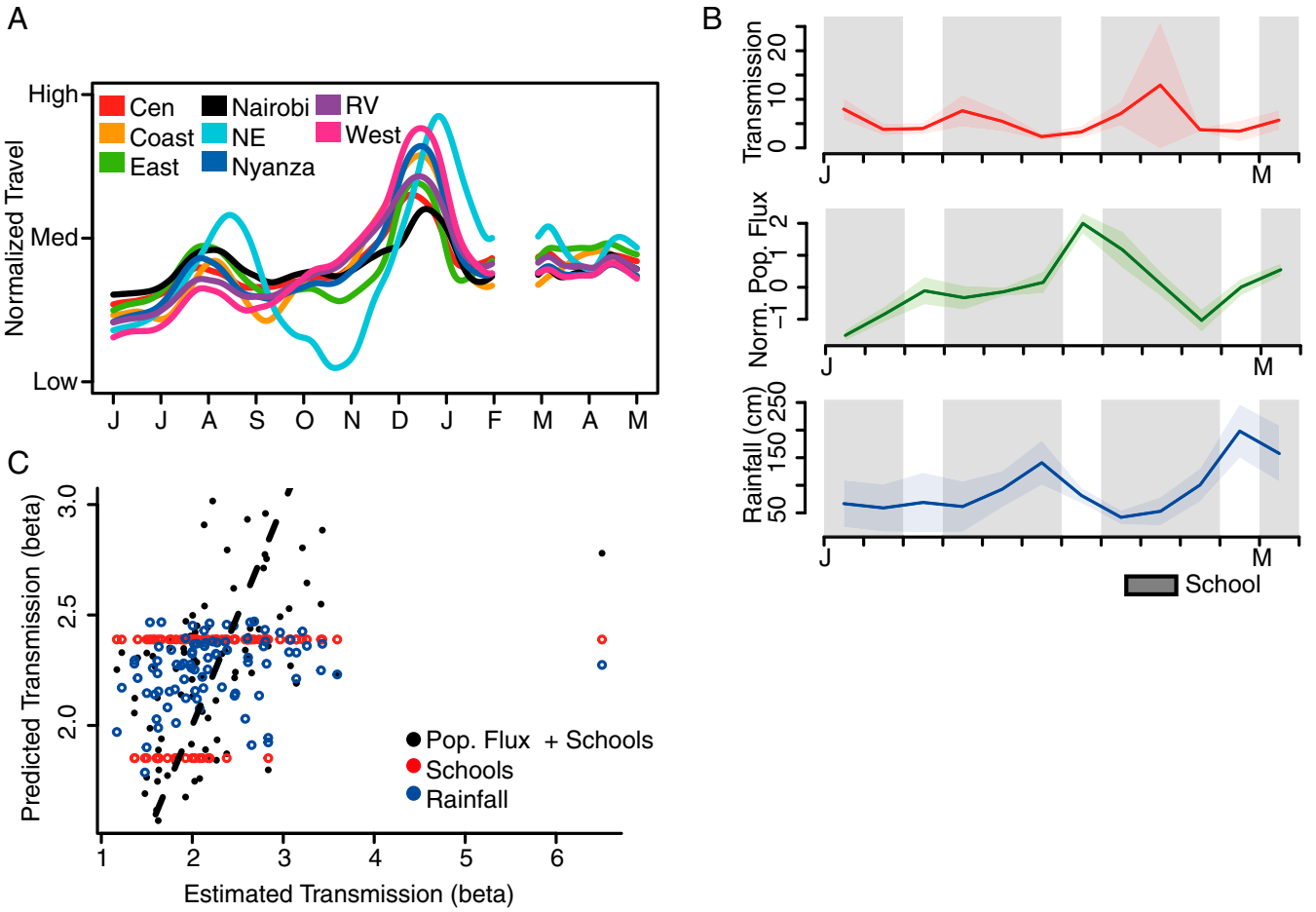

Fig. 2. The relationship between population flux, rainfall, school terms, and rubella transmission per province. (*A*) The normalized population fluxes measured from the mobile phone data per month in each province with months ordered from the start of the dataset (June) until the following May. In nearly all provinces, travel peaked in December and July/August, although this varied by location. (*B*) Monthly mean values from provinces for transmission and the three covariates (rainfall, school terms, and population flux). The time series is shown highlighting the variability in these estimates over the course of the year. We would expect that transmission would decrease during school breaks (white area) and population fluxes would increase during these times (white area). (*C*) Using the province level covariates shown in *B*, we constructed a number of models using various subsets of the covariates to predict transmission values. The estimated (from the TSIR model) versus predicted values for transmission (to aid with comparison a black *xy* line is drawn) from the minimum adequate model (black, includes population flux and school terms as covariates) and a model based on only school terms (red) or rainfall (blue). The model based on movement and school terms (adjusted $R^2 = 0.19$, $P < 0.001$) provides the best estimates (only rainfall, adjusted $R^2 = 0.03$, $P = 0.056$; only school terms, adjusted $R^2 = 0.09$, $P = 0.003$). In comparison, the model based only on school terms only produces two transmission values, whereas the model based only on rainfall produces estimates over a smaller range than the model including population flux.

As has been found in high-income settings, but in contrast to other African countries, rubella transmission was strongly positively related to school term times (*t* test, $t = -3.6681$, df = 80.505, $P < 1e\text{-}3$ for provinces excluding the Northeastern province where power for analysis was weak); transmission was not significantly correlated with rainfall (mean correlation coefficient, −0.13; 95% quantile interval, −0.34, 0.07, $P > 0.2$, excluding the Northeastern province; Fig. S3 and Table S1) (1, 13, 27, 28). There was also a strong positive relationship with province-level population flux in the previous month normalized for each province (mean correlation coefficient, 0.38; 95% quantile interval, 0.18, 0.55, $P < 1e\text{-}3$, excluding the Northeastern province). We further analyzed the relationship between these variables for districts (68 districts in Kenya), and although there exists heterogeneity between locations, we found the same broad general relationships (Fig. S4).

To explore the relative importance of the three drivers, we fit all three (rainfall, school terms, and population flux; Fig. 2*B*) as main effects in a linear regression with seasonal transmission as the response variable (excluding the Northeastern province and taking the square root of the response variable to normalize the errors). Including population flux significantly improved the model ($F_{1,82} = 19.59$, $R^2 = 0.19$, $P < 0.001$; Table S2 and Fig. 2*C*), and further including either rainfall ($F_{1,81} = 0.40$, $P > 0.5$) or school terms ($F_{1,81} = 3.41$, $P > 0.05$) did not significantly improve the model fit (see *Supporting Information* for models including site effects; results remain qualitatively the same). To highlight the improved ability of population fluxes to predict transmission we also constructed two comparison models using either only rainfall or only school terms (Fig. 2*C*; also see Tables S2 and S3 and Fig. S5); these two models explained considerably less of the variance, with adjusted R^2 values of adjusted $R^2 = 0.03$ ($P = 0.056$) and adjusted $R^2 = 0.09$ ($P = 0.003$), respectively.

Population-Scale Consequences of Seasonal and Spatial Population Fluxes. With evidence that the mobile phone data captures epidemiologically relevant human movements for rubella in Kenya, we next used the mobile phone data to construct maps to characterize spatial and temporal patterns of risk of rubella introduction based on population fluxes across the country on scales finer than possible using only the rubella data (Fig. 3). We recalculated the

population flux per district (*Materials and Methods*) and analyzed these finer spatial data. We identified both areas at a high risk for rubella introductions and how these locations and assessment of risk varied over the year. As suggested in Fig. 1, areas most at risk vary over the course of the year, although Nairobi District and districts in the Central province remain at high risk throughout the year (Fig. 3). During school terms, the risk decreases across the entire country, although the most noticeable differences are in western Kenya, where the risk during school breaks is relatively much higher than during the term. Given the variability in the amount of risk and the location of these risks over the course the year, spatially and temporally targeted control will be essential to maintaining gains that follow introduction of vaccination.

Discussion

Our analysis of population fluxes from mobile phone data and rubella transmission dynamics shows that mobile phone data may be used to capture seasonal human movement patterns relevant for understanding childhood infection dynamics and significantly outperform previous proxies used in this context (1, 13). This suggests that mobile phone-derived data may represent an improvement on broadly used qualitative correlates of temporal variations in transmission (schools terms, measured as "on" or "off"; Fig. S5) or suspected phenomenological drivers (rainfall, where the strength of the relationship is likely to be highly spatially variable).

The power of population fluxes quantified by mobile phone data to describe patterns of transmission linked to childhood infections is somewhat surprising because previous work suggests that mobile phone ownerships is concentrated in adults and toward urban, more educated males (29, 30). However, this supports previous evidence that children rarely travel without adults and also underscores the importance of these population-scale movements for dynamics of infections like rubella (31). Although mobile phone data are inherently biased by ownership, there exist few data sources that can directly record daily movement patterns over the range of spatial (the dynamics of an entire country) and temporal (12 months of data) scale. We also analyzed the role of birth seasonality on rubella dynamics and did not find a strong association between transmission and seasonal birth rates (*Supporting Information* and Fig. S6). Moreover, a detailed age-structured simulation confirms that the average age of infection for rubella in Kenya makes the effect of susceptible births on the dynamics negligible (*Supporting Information*) (32).

Rubella incidence is frequently underreported, as a result of the mild, even asymptomatic manifestation of the infection (15). Evidence that mobile phone data can capture population fluctuations that relate to rubella transmission provides grounds for optimism that patterns of incidence might be inferred indirectly. This is relevant not only to projecting timing of outbreaks but also, if epidemiologically relevant movement can be measured, to predicting patterns of increased congenital rubella syndrome (CRS) incidence stemming from metapopulation dynamics [local extinction of rubella followed by aging of susceptibles into childbearing ages means that subsequent introduction of rubella can result in CRS cases (15)]. Because introduction of rubella-containing vaccine is increasingly being considered across low-income countries (33), this could be used to help guide targeted vaccination efforts spatially and also to infer effects of geographic heterogeneities in vaccination coverage on the CRS burden (22), although more work would need to be conducted to fully understand the relationship between mobility measured via mobile phones and the populations of interest. Many of the same principles that apply to rubella also apply to measles, which is of increasing relevance because all WHO regions currently have measles elimination targets. As we move toward elimination goals for measles or other vaccine-preventable infections, mobile phone data offers enormous potential for quantifying daily movement patterns at particular spatial scales, which will be important in order to maintain elimination gains by indicating key areas to target to minimize reintroductions and ongoing spread. The resolutions of these data are variable, however, so the utility of this approach will need to be considered separately for different policy questions. In particular, mobile phone data can primarily describe within-country movement patterns on a relatively large spatial scale and so will be less relevant for public health planning when either international travel has large effects on disease dynamics or transmission is highly spatially localized.

Seasonality is a characteristic feature of many pathogens (6), thereby driving the time of interventions (e.g., influenza vaccination and seasonal malaria chemoprophylaxis). For immunizing infections like rubella, interannual fluctuations in transmission can result in complex multiannual dynamics (34), which may affect the outcome of vaccination programs (13). For vector-borne diseases, seasonality can be influenced by both environmental and behavioral conditions, making determining optimal timing for interventions dependent on being able to quantify both of these factors. Using mobile phone data we can measure both broad seasonal patterns (country-wide dynamics) as well as the variability in seasonal travel on smaller spatial scales that are unobtainable from many existing datasets (see *Supporting Information* for district-level analysis). A major result of our analysis, that mobile phone data capture epidemiologically relevant movement, indicates that this approach represents a powerful tool for quantifying critical drivers of epidemics on epidemiologically relevant spatial and temporal scales. Further investigations to calibrate the relevance of this effect across a broad array of systems will be an important direction for future research.

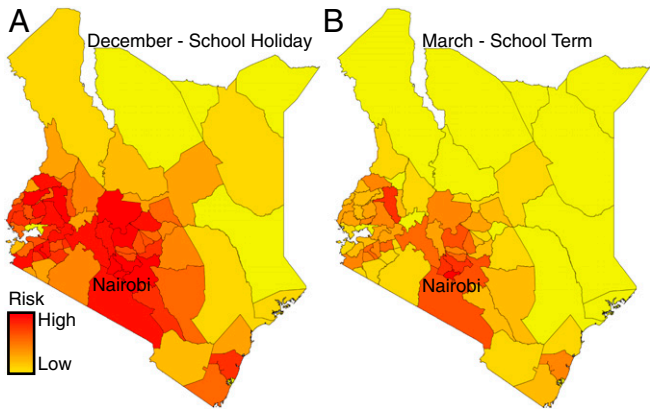

Fig. 3. The seasonal variability in the risk of importation. We analyzed the average amount of population flux per district during (*A*) the major holiday and a school term break (December) and (*B*) during a school term (March). As highlighted in Fig. 1, there are large amounts of population flux and consequently the risk of importation during school breaks (*A* and *B*). In contrast, there is a decrease in the risk of importation during the school term. During the course of the year, the districts with the largest risks vary with higher risks to western Kenya during school breaks. However, Nairobi (shown in red in both maps) consistently remains at a high risk of importation from the large population fluxes.

Materials and Methods

Mobile Phone Data. The leading mobile phone operator in Kenya provided anonymous mobile phone call records that recorded the location of the routing tower and timing of each call and SMS communication between 14,816,512 subscribers during June 2008 to June 2009 (with February 2009 missing from the dataset) (4, 18). In total, over 12 billion mobile phone communications were recorded, including the communication's location at one of 11,920 routing towers. The operator that provided the call data records had ~92% market share at the time of data acquisition. All subscriber data were aggregated to either the province- or district-level scale to further preserve anonymity. These

data were deidentified at the mobile phone operator, contain no personal information, and are considered nonhuman subject research.

Quantifying Population Fluxes. From the mobile phone data the geographic location of the caller and receiver was approximated based on the unique longitude and latitude coordinates for each mobile phone tower. Using the data, a location for each subscriber every time they either made or received a call (or SMS) was obtained. For each day in the dataset, subscribers were assigned a single tower location. If the subscriber made at least one call on that day, then the location of the majority routing tower was assigned. If the subscriber had not made a call on that day, then the location of their most recent routing tower was assigned. This provided a 12-mo time series of tower locations for each subscriber on each day. As in previous studies, trips are calculated by observing when a subscriber's tower location has changed from the previous day (4, 20). For the majority of the analysis, we aggregated towers to the province level based on the tower's location. Thus, only trips between towers in different provinces were considered. We constructed a time series of outgoing travel/population flux for each province (Fig. S3). Incoming and outgoing travel/population fluxes were highly corrected (Pearson's correlation coefficient: 0.997, $P < 0.001$). We decomposed each time series using a moving average and then analyzed the resulting trend. Although mobile phone data ownership is biased toward urban, more educated males, we have previously shown that this does not significantly bias mobility estimates (29, 30). We also aggregated towers to the district level based on the tower's location and recalculated these flux values for travel from each district to all other districts.

Population Data. Estimates of population sizes and numbers of live births for 2009 were obtained from the WorldPop project (www.worldpop.org.uk) using the spatial datasets that were constructed using the methods outlined by Tatem et al. (35) and Linard et al. (5). In brief, 2009 census data at sublocation level were matched to the relevant administrative boundaries and then combined with satellite-derived settlement and land cover data layers to disaggregate the counts to 100-m spatial resolution. Five-year age and sex groupings of population at sublocation level from the census were then used to adjust the output gridded population surface to create a set of 5-y age grouped surfaces that covered all women in the age range 15–49 (women of childbearing age). Subnational urban–rural age-specific fertility rates derived from the most recent demographic and health survey for Kenya (www.dhsprogram.com) were then used to convert these age-structured female count datasets into surfaces of estimates of numbers of live births per grid square.

Rubella Data and Dynamics. Concerns about increases in the burden of congenital rubella syndrome have meant that vaccination against rubella has yet to be introduced in many Sub-Saharan African countries including Kenya, and rubella remains prevalent. Using over 9 y of data on rubella cases ($n = 5{,}590$) from each Kenyan province (Fig. 1 *A* and *B*, *Materials and Methods*, Fig. S1, and *Supporting Information*), we quantified seasonal fluctuations of transmission via a time series susceptible–infected–recovered model fitted using trajectory matching with a binomial observation model for case reporting. Seasonal estimates of transmission are shown in Fig. 1*C*; corresponding matched trajectories are in Fig. S1.

Quantifying Seasonal Fluctuations in Transmission for Rubella. Rubella is a completely immunizing infection with a generation time (serial interval; approximately the latent plus infectious period) of ∼18 d (36). We therefore assumed that the time scale of the epidemic process was approximately 2 wk. The number of infected individuals observed after a 2-wk period, I_{t+1}, depends on I_t and the number of susceptible individuals S_t with expectation $\lambda_t = \beta_s S_t I_t^m / N_t$, where β_s is the transmission rate in every biweek in any particular location and the exponent m, usually a little less than 1, captures heterogeneities in mixing not directly modeled by the seasonality (8) and the effects of discretization of the underlying continuous time process. Dividing by N_t captures the fact that social contact networks tend to remain stable with population size. In the same time frame, the number of susceptible individuals will be depleted by this number of new infected individuals and augmented via births. We can therefore generate the full trajectory of susceptible and infected individuals through time as

$$I_{t+1} = \beta_s S_t I_t^m / N_t$$
$$S_{t+1} = S_t + B_t - I_t,$$

where B_t is the number of births, β_s is seasonally varying transmission rates, and initial conditions are set by initial numbers of infected I_0 and susceptibles S_0. To infer the unknown parameters (β_s, I_0 and S_0, and m) we can link this to data by the relationship

$$I^{(r)}{}_T = \text{Binomial}(I_t + I_{t+1}, \rho),$$

where ρ is the reporting rate and $I_T^{(r)}$ is the reported number of infected cases at a monthly time scale; the sum is taken over the previous 2 wk. We have previously found that low reporting rates result in strongly downward biased estimates of m, which result in unrealistic dynamics (15); this also proved to be the case here. For this analysis, we therefore fixed m at a consensus value of 0.97. Previous work (17) indicates that the exact value of m does not affect estimates of seasonal variation in transmission. We fit a different transmission parameter for every site in every month, to quantify how transmission varied throughout the year within each province, and then explored the degree to which the various proxies (including rainfall, school term times, and population flux quantified as above) performed as explanatory variables for fluctuations in transmission.

ACKNOWLEDGMENTS. A.W. was supported by the National Science Foundation (Grant 0750271) and the James S. McDonnell Foundation. A.J.T. acknowledges support from Bill and Melinda Gates Foundation (Grants 49446 and OPP1032350); National Institutes of Health (NIH)/National Institute of Allergy and Infectious Disease (U19AI089674); and the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, NIH. C.O.B. was supported by the Models of Infectious Disease Agent Study program (Cooperative Agreement 1U54GM088558). This work was funded by the Bill and Melinda Gates Foundation; the Science and Technology Directorate, Department of Homeland Security, Contract HSHQDC-12-C-00058 (to B.T.G., C.J.E.M., J.L., and O.N.B.); and the RAPIDD program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, NIH (C.J.E.M., B.T.G., and O.N.B.). C.J.E.M., A.J.T., and C.O.B. acknowledge support from the Wellcome Trust Sustaining Health Grant (106866/Z/15/Z).

1. Bharti N, et al. (2011) Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* 334(6061):1424–1427.
2. Pindolia DK, et al. (2012) Human movement data for malaria control and elimination strategic planning. *Malar J* 11(1):205.
3. Stoddard ST, et al. (2009) The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl Trop Dis* 3(7):e481.
4. Wesolowski A, et al. (2013) The use of census migration data to approximate human movement patterns across temporal scales. *PLoS One* 8(1):e52971.
5. Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ (2012) Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS One* 7(2):e31743.
6. Grassly NC, Fraser C (2006) Seasonal infectious disease epidemiology. *Proc Biol Sci* 273(1600):2541–2550.
7. Mabaso ML, Craig M, Vounatsou P, Smith T (2005) Towards empirical description of malaria seasonality in southern Africa: The example of Zimbabwe. *Trop Med Int Health* 10(9):909–918.
8. Bjørnstad ON, Finkenstadt B, Grenfell BT (2002) Dynamics of measles epidemics: Estimating scaling of transmission rates using a time series SIR model. *Ecol Monogr* 72(2):169–184.
9. Cauchemez S, et al. (2009) Closure of schools during an influenza pandemic. *Lancet Infect Dis* 9(8):473–481.
10. Fine PEM (1993) Herd immunity: History, theory, practice. *Epidemiol Rev* 15(2):265–302.
11. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M (2010) Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol* 8(2): e1000316.
12. Bharti N, et al. (2010) Measles hotspots and epidemiological connectivity. *Epidemiol Infect* 138(9):1308–1316.
13. Ferrari MJ, et al. (2008) The dynamics of measles in sub-Saharan Africa. *Nature* 451(7179):679–684.
14. Fine PE, Clarkson JA (1986) Seasonal influences on pertussis. *Int J Epidemiol* 15(2):237–247.
15. Metcalf CJE, Munayco CV, Chowell G, Grenfell BT, Bjørnstad ON (2011) Rubella metapopulation dynamics and importance of spatial coupling to the risk of congenital rubella syndrome in Peru. *J R Soc Interface* 8(56):369–376.
16. Keeling MJ, Rohani P, Grenfell BT (2001) Seasonally forced disease dynamics explored as switching between attractors. *Physica D* 148(3-4):317–335.
17. Metcalf CJE, Bjørnstad ON, Grenfell BT, Andreasen V (2009) Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proc Biol Sci* 276(1676):4111–4118.
18. González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782.
19. Le Menach A, et al. (2011) Travel risk, malaria importation and malaria transmission in Zanzibar. *Sci Rep* 1:93.
20. Wesolowski A, et al. (2012) Quantifying the impact of human mobility on malaria. *Science* 338(6104):267–270.

21. Macro KNBoSKaI (2010) Kenya Demographic and Health Survey 2008–2009 (Kenya National Bureau of Statistics, Nairobi, Kenya).
22. Metcalf CJE, et al. (2013) Implications of spatially heterogeneous vaccination coverage for the risk of congenital rubella syndrome in South Africa. *J R Soc Interface* 10(78):20120756.
23. Metcalf CJE, et al. (2011) The epidemiology of rubella in Mexico: Seasonality, stochasticity and regional variation. *Epidemiol Infect* 139(7):1029–1038.
24. Terada K (2003) Rubella and congenital rubella syndrome in Japan: Epidemiological problems. *Jpn J Infect Dis* 56(3):81–87.
25. Trottier H, Philippe P (2005) Scaling properties of childhood infectious diseases epidemics before and after mass vaccination in Canada. *J Theor Biol* 235(3): 326–337.
26. Goodson JL, et al. (2011) Rubella epidemiology in Africa in the prevaccine era, 2002-2009. *J Infect Dis* 204(Suppl 1):S215–S225.
27. Martinez-Bakker M, Bakker KM, King AA, Rohani P (2014) Human birth seasonality: Latitudinal gradient and interplay with childhood disease dynamics. *Proc Biol Sci* 281 (1783):20132438.
28. Dorélien AM, Ballesteros S, Grenfell BT (2013) Impact of birth seasonality on dynamics of acute immunizing infections in Sub-Saharan Africa. *PLoS One* 8(10):e75806.
29. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO (2012) Heterogeneous mobile phone ownership and usage patterns in Kenya. *PLoS One* 7(4):e35319.
30. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO (2013) The impact of biases in mobile phone ownership on estimates of human mobility. *J R Soc Interface* 10(81):20120986.
31. Wesolowski A, et al. (2014) Quantifying travel behavior for infectious disease research: A comparison of data from surveys and mobile phones. *Sci Rep* 4:5678.
32. Metcalf CJE, Lessler J, Klepac P, Cutts FT, Grenfell BT (2012) Impact of birth rate, seasonality and transmission rate on minimum levels of coverage needed for rubella vaccination. *Epidemiol Infect* 140(12):2290–2301.
33. Cutts FT, Metcalf CJ, Lessler J, Grenfell BT (2012) Rubella vaccination: Must not be business as usual. *Lancet* 380(9838):217–218.
34. Earn DJD, Rohani P, Bolker BM, Grenfell BT (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287(5453):667–670.
35. Tatem AJ, et al. (2014) Mapping for maternal and newborn health: The distributions of women of childbearing age, pregnancies and births. *Int J Health Geogr* 13:2.
36. Anderson RM, May RM (1991) *Infectious Diseases of Humans* (Oxford Univ Press, Oxford, UK).
37. Mossong J, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 5(3):e74.