Impact of human mobility on the emergence of dengue epidemics in Pakistan

Amy Wesolowski\textsuperscript{a,b}, Taimur Qureshi\textsuperscript{c}, Maciej F. Boni\textsuperscript{id,e}, Pål Roe Sundsø\textsuperscript{f}, Michael A. Johansson\textsuperscript{b,f}, Syed Basit Rasheed\textsuperscript{g}, Kenth Enge-Monsen\textsuperscript{h}, and Caroline O. Buckee\textsuperscript{a,b,1}

\textsuperscript{a}Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA 02115; \textsuperscript{b}Center for Communicable Disease Dynamics, Harvard T. H. Chan School of Public Health, Boston, MA 02115; \textsuperscript{c}Telenor Research, Telenor Group, N-1360 Fornebu, Norway; \textsuperscript{d}Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Ho Chi Minh City, Vietnam; \textsuperscript{e}Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7FZ, United Kingdom; \textsuperscript{f}Division of Vector-Borne Diseases, Centers for Disease Control, San Juan, Puerto Rico 00920; and \textsuperscript{g}Department of Zoology, University of Peshawar, Peshawar 25120, Pakistan

Edited by Burton H. Singer, University of Florida, Gainesville, FL, and approved August 6, 2015 (received for review April 2, 2015)

The recent emergence of dengue viruses into new susceptible human populations throughout Asia and the Middle East, driven in part by human travel on both local and global scales, represents a significant global health risk, particularly in areas with changing climatic suitability for the mosquito vector. In Pakistan, dengue has been endemic for decades in the southern port city of Karachi, but large epidemics in the northeast have emerged only since 2011. Pakistan is therefore representative of many countries on the verge of countrywide endemic dengue transmission, where prevention, surveillance, and preparedness are key priorities in previously dengue-free regions. We analyze spatially explicit dengue case data from a large outbreak in Pakistan in 2013 and compare the dynamics of the epidemic to an epidemiological model of dengue virus transmission based on climate and mobility data from ~40 million mobile phone subscribers. We find that mobile phone-based mobility estimates predict the geographic spread and timing of epidemics in both recently epidemic and emerging locations. We combine transmission suitability maps with estimates of seasonal dengue virus importation to generate fine-scale dynamic risk maps with direct application to dengue containment and epidemic preparedness.

Dengue is the most rapidly spreading mosquito-borne disease worldwide (1, 2). Half the global population now lives in at-risk regions for dengue virus transmission, due to the wide distribution of the mosquito vector, \textit{Aedes aegypti}, which thrives in peri-urban areas and transmits the virus between humans (3). Dengue virus can cause acute febrile illness and carries the risk of severe disease, hospitalization, and shock syndrome, especially in clinical settings with little experience treating dengue patients. There is currently no specific therapeutic protocol for, or vaccine against, infection (1). Current control measures focus on vector control, although these measures are often logistically difficult and have shown varying efficacy in controlling epidemics (4). In the absence of effective prevention and treatment, public health system preparedness remains the single most important tool for minimizing morbidity and mortality as dengue epidemics spread beyond endemic areas (5, 6).

The introduction of dengue into new populations is mediated by travel of infected individuals to areas that can support transmission, because mosquito vectors move only short distances during their lifespans (3, 7–12). International travel to endemic countries has resulted in imported cases and outbreaks in Europe and the Americas (2, 8, 10, 13). Local variation in transmission, within a single city for example, is also driven by mobility patterns of individuals on short timescales (7). Forecasting methods are needed to spatially target interventions and epidemic preparedness measures that reflect both the changing temporal risks of importation and environmental suitability that go beyond solely climate-based methods (14). Dengue has long been endemic in most Southeast Asian countries (1), but has more recently emerged in parts of the Middle East and South Asia, including Pakistan (15, 16). In Pakistan, the transmission of dengue viruses was largely confined to the southern city of Karachi until 2011 when a large dengue epidemic with over 20,000 cases occurred in the northeastern city of Lahore (16), causing significant morbidity and mortality. In 2013, a second large epidemic occurred in northeastern Pakistan in Punjab and Khyber-Pakhtunkhwa (KP) provinces, establishing the region as an emerging focus of seasonal dengue epidemics. It has been hypothesized that the recent geographic expansion of \textit{A. aegypti} mosquito vectors, changing environmental suitability, and human importation of dengue from endemic regions all contributed to the emergence of dengue in northern areas (17). Pakistan is therefore representative of many countries that are on the verge of countrywide endemic dengue transmission and are struggling to contain its emergence into previously dengue-free regions.

Measuring changing risks of importation events that spark epidemics has been extremely challenging on the refined temporal and spatial scales necessary to inform local policies (18). Being able to predict when to prepare surveillance systems and health facilities for dengue outbreaks could dramatically reduce the morbidity and mortality associated with epidemics and would allow policy makers to pinpoint regions that are particularly vulnerable to imported cases, for vector control. Mobile phone data offer direct measures of human aggregation and movement.

Significance

Dengue virus has rapidly spread into new human populations due to human travel and changing suitability for the mosquito vector, causing severe febrile illness and significant mortality. Accurate predictive models identifying changing vulnerability to dengue outbreaks are necessary for epidemic preparedness and containment of the virus. Here we show that an epidemiological model of dengue transmission in travelers, based on mobility data from ~40 million mobile phone subscribers and climatic information, predicts the geographic spread and timing of epidemics throughout the country. We generate fine-scale dynamic risk maps with direct application to dengue containment and epidemic preparedness.


Conflict of interest statement: M.F.B. has worked as a paid consultant to Visterra, Inc. in Cambridge, MA.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

\textsuperscript{1}To whom correspondence should be addressed. Email: cbuckee@hsph.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1504964112/-/DCSupplemental.
and represent a unique source of information on the human determinants of the geographic expansion of emerging epidemic diseases like dengue. Here, we conduct a retrospective epidemiological analysis of large dengue outbreaks in Pakistan in 2013, to examine the predictive ability of an epidemiological model that integrates human mobility from the largest mobile phone dataset analyzed to date with climate information. We show that within-country human mobility predicts emerging epidemics in Pakistan, and epidemiological models incorporating this type of data can predict the spatial extent and timing of outbreaks, providing a new approach to forecasting.

Results
Human Mobility in Pakistan Does Not Conform to Standard Model Predictions. To measure human travel patterns underlying the spread of dengue virus across Pakistan in 2013, we estimated the mobility of 39,785,786 mobile phone subscribers between geolocated mobile phone towers in Pakistan between June 1 and December 31 of 2013 (representing ~22% of the population; Materials and Methods). Daily locations and movements were aggregated to measure travel between 356 small, politically defined areas called tehsils (Fig. 1A and Materials and Methods). We compared our data to gravity models of mobility, developed from transportation theory and commonly used to parameterize infectious disease frameworks, to assess whether observed mobility measured using mobile phone data significantly improves upon this standard approach. We have focused on intracountry mobility patterns (Discussion).

The sampled population was extremely mobile. We estimated that between 2.4 million and 4.8 million subscribers traveled between tehsils each day (95% quantile interval: 3.1–4.6 million; details in Materials and Methods). Most travel followed a NW–SE corridor along the major highways (Fig. 1B). Large volumes of travel occurred to and from Karachi, a major population and economic hub of Pakistan, with ~710,000 subscribers traveling to or from the city each day on average (95% quantile interval: 570,000–813,000; Fig. S1). In contrast to expectations of standard mobility models, there was almost no decay in travel with increasing distance (correlation coefficient: −0.064, P < 0.001), although the most frequent destinations for travel were often in a nearby tehsil (Fig. 1C). This pattern reflects the topography, road infrastructure, and population distribution in Pakistan, with the largest cities outside Karachi being located in the northern part of Punjab province that includes the Rawalpindi/Islamabad metropolitan area, Lahore, and Faisalabad. Although there was a decrease in overall movement during Ramadan, we did not observe a systematic difference in the amount or direction of travel patterns in the weeks before and after the holiday (Fig. S1).

Epidemiological Modeling of Dengue Epidemics in 2013. There were 15,535 reported dengue cases in 82 tehsils over 7 mo in Pakistan in 2013 (Fig. 2A and Materials and Methods). Peak timing of the epidemic varied by location, and the majority of cases occurred in and around Karachi, in the northern district of Swat, and in the cities of Lahore and Rawalpindi (Fig. 2A and Table S1). About half of the dengue cases reported occurred in the Mingora area of Swat (KP province), marking the first major outbreak in the region (2011 vs. 2011). We first fitted an ento-epidemiological model to the reported dengue cases in Southern Pakistan, where transmission occurs year round, and to case data in Lahore and Swat in northern Pakistan, where transmission is seasonal due to climatic variation. Southern Pakistan has year-round climatic suitability for dengue vectors and is therefore the most likely source of exported cases to other parts of the country, where greater seasonal temperature extremes limit suitability (16) (SI Text). Although importation from international regions is technically a possibility, given the fairly restrictive political borders we assume here that importation will be negligible compared with the within-country importation rate. We fitted an ento-epidemiological model of dengue dynamics in Karachi (Fig. 2B), with vector dynamics being determined by temperature (19). We fitted the mosquito-biting rate (a = 0.66) and reporting rate (SI Text), yielding parameter values that are consistent with endemic dengue virus transmission (Materials and Methods and SI Text).

In northern Pakistan, we used our epidemiological framework to estimate the timing of the introduction of the first case that sparked the epidemics in each area, to compare these estimates with our mechanistic model of imported infections using mobile phone data. We performed a sensitivity analysis to estimate an expected range of dates for the introduction of dengue to Lahore and Mingora, where most cases were concentrated (Materials and Methods and Fig. 3A and B). Interestingly, Lahore had suffered its first major outbreak 3 y before this epidemic, and we expect that immunity may have played a significant role in determining transmission dynamics (16). Mingora, on the other hand, represents an effectively naïve population. In the absence of serotype information, we took the simplest approach and assumed each population was immunologically naïve; however, we
hypothesize that the delay between the first cases and the peak of the epidemic in Lahore may have been caused by immunity from the 2011 outbreak. We estimated that the first case was introduced to Lahore during the second week of May (between days 124 and 130; SI Text and Fig. 3A) a few days earlier than the first reported case (day 133). In Mingora, on the other hand, we estimated that the first introductions likely occurred in August (between days 202 and 231) (Fig. 3B), a few weeks before the first reported case in the city.

**Models of Dengue Virus Importation Based on Mobile Phone Data Accurately Predict the Spatial Extent and Timing of Epidemics.** We next modeled the number of infected individuals traveling from the endemic areas in southern Pakistan to all other tehsils, using different approaches to characterize mobility: direct observations from the mobile phone data and various modified gravity models of travel (Materials and Methods) (10, 19). To compare the performance of the mobile phone data against the next best alternative, we used a parameter-free gravity model (referred to as the diffusion model) that is equivalent to a population-weighted spatial diffusion model (Materials and Methods) based on the travel time distance between the origin and the destination. In addition to these models, we fitted a gravity model to the mobile phone data, to determine whether simple adjustments would significantly improve our predictions (SI Text and Fig. S2).

We compared the timing of predicted importation events from endemic areas in southern Pakistan, based on our mobility model, to the estimated first dengue case inferred from case report data from Lahore and Mingora (SI Text). The timing of the first dengue case estimated from the epidemiological data overlapped well with the predicted introductions from southern Pakistan to Lahore from the importation model (Fig. 3A), with the first introductions occurring approximately 1 mo earlier. In Mingora, the predicted timing of imported infections using our model occurred 2 wk before the first reported case, consistent with the serial interval for dengue and the estimates of the first case from epidemiological data (Fig. 3B). Crucially, the diffusion model does not predict any introductions from endemic areas in southern Pakistan to Mingora. Thus, travel patterns measured using mobile phone data predict introduction events consistent with outbreaks in both emerging (Lahore) and previously dengue-free (Mingora) regions. Our ability to measure these importations in more remote places like Mingora was somewhat sensitive to the Karachi model fit, in particular the reporting rate, although the mobile phone data are always able to predict earlier, more frequent, and more accurate introductions than the diffusion model (SI Text). The modeled interaction between seasonal variations in vectorial capacity and the dynamics of importation events provide accurate predictions about the location and timing of epidemics in different epidemiological settings and regions of the country.

We combined our estimates of imported cases with an index of climatic suitability for dengue vectors (19) across Pakistan to predict the potential for spread of the virus in areas with introduced cases and assessed how well the spatiotemporal dynamics of the epidemic were captured. In general, the spatial extent and epidemic timing of cases were predicted more accurately using mobile phone data (Fig. 4) compared with the diffusion models, which both predicted early spread along the major highway connecting Karachi to other highly populated areas.

**A New Approach to Dynamic Risk Mapping for Epidemic Preparedness.** We next constructed risk maps to identify areas of the country that were vulnerable to epidemics due to a combination of seasonally varying climatic suitability based on temperature and relative humidity (vectorial capacity index) for dengue transmission and
seasonal patterns of importation of dengue virus from southern Pakistan. By combining the daily suitability index (Fig. 3A) with estimated introductions of infected travelers from endemic areas we created a composite epidemic risk measure that varied in space and time and showed a distinctively different pattern of risk than a climate-based suitability map alone (Fig. 3B). In southern Pakistan outside of Karachi, for example, a low risk of epidemic dengue results from low numbers of infected travelers visiting the area. Along the Afghanistan border, on the other hand, low risk of epidemics is due to temperature extremes that prevent continuous dengue transmission. The highest-risk regions are in Punjab and KP provinces in the northeast of the country, where a combination of frequent introductions from Karachi and long periods of sustained climatic suitability for transmission create the necessary conditions for outbreaks.

The timing of epidemic risk varied considerably in different tehsils (Fig. 3C), with the risk of outbreaks varying by up to 4 mo in different regions. Based on the mobile phone data, epidemic risk occurred much earlier in tehsils in northern regions relatively far from Karachi, in contrast to the diffusion model. Because preparation of local health facilities and surveillance for dengue cases are key components of national response to outbreaks, this variation in timing has important implications for decision making. In particular, given the 2-wk serial interval of dengue (20), mobile phone and disease surveillance data could be used to produce dynamic risk maps and provide location-specific predictions about future risk of epidemics in near real time, giving policy makers time to prepare.

Discussion

For emerging epidemic diseases like dengue, changing environments and dynamic population movements create conditions for geographic expansion of pathogens into naïve populations. Predicting the changing patterns of risk due to these emerging public health threats will be critical to containing their spread and preparing for outbreaks. Mobile phone data provide dynamic population mobility estimates that can be combined with infectious disease surveillance data and seasonally varying environmental data to map these changing patterns of vulnerability in a country where dengue outbreaks are emerging and irregular in many regions. Because these data are continuously being collected by mobile phone operators, these methods could be integrated into national control programs in near real time.

A major limitation of mobile phone data generated by national operators is the difficulty in capturing cross-border travel patterns. International travel may be less relevant for dengue dynamics in Pakistan than in other dengue-endemic countries, however. Less than 2% of the entire Pakistan population are migrants (where their country of origin is not Pakistan) (21). More than half of these migrants are from Afghanistan (2.3 million) (21). Although Pakistan borders four countries (India, Afghanistan, China, and Iran), all aside from India are not endemic for dengue near the Pakistan borders, making importations of dengue cases to Pakistan unlikely (3, 22–24). Estimates of travel between India and Pakistan are difficult to obtain, but travel restrictions between the countries and a limited number of international flights may make travel difficult. Another potential issue with this type of data is the sampling frame, because mobile phone ownership may not represent a representative sample of the population, although we have previously found in an African context that mobility estimates from this kind of data were not significantly affected by ownership bias (25). We expect that the substantial market share and widespread coverage of the mobile phone operator make the possibility of systematic skew in our estimates unlikely, but this cannot be ruled out.

Uncertainties are inherent to any epidemiological model of dengue transmission, in particular regarding the large asymptomatic reservoir and the multiple circulating serotypes of the virus, which impact the ability to estimate both the fraction of the population that is susceptible and the reporting rate accurately (26, 27). In the absence of serotype information, for example, we chose not to include immunological covariates in our model. As Pakistan moves toward an endemic transmission setting in the northeast of the country, however, serotype data will be increasingly important to determine how strain dynamics are likely to affect the likelihood of outbreaks (28). Indeed, we believe that the substantial lag time in Lahore we observed was probably due to the recent large outbreak in 2011, which may have generated substantial immunity and delayed the 2013 epidemic (16, 29).

Fig. 4. The estimated spatial spread of imported dengue. Using the modeled dengue dynamics in Karachi and mobility measured from the mobile phone data or a diffusion model, we estimated the time of the first introduced case to the rest of the country. The mobile phone data predict the earliest introductions in eastern Pakistan near Lahore and inland toward Swat Valley (Mingora). In comparison, the mobility model predicts early introductions in southern Pakistan with few introductions in Mingora. These differences are highlighted in the difference in predictions plot—the number of days earlier (red) from the mobile phone predictions or earlier (yellow) from the diffusion model (without the mobile phone data).

Fig. 5. Dynamic risk mapping for dengue epidemics in Pakistan. (A) Average dengue suitability in Pakistan during 2013 based on temperature and humidity. The southern part of the country is the most suitable for dengue (high values are shown in red). (B) Epidemic risk, measured as a composite of vector suitability and frequency of introductions from endemic areas in southern Pakistan. Places in red are the most at risk whereas those in blue have low risk of epidemics. (C) Variation in the timing of dengue risk by location (per tehsil). A box is drawn if suitability multiplied by estimated introduction events is greater than 0, indicating a nonnegative risk. Each line shown is a tehsil, with the y axis corresponding to the ranked distance from Karachi. The dengue risk is shown using the mobile phone data (blue) or the diffusion model (green) to estimate introduction events from Karachi. In general, the mobile phone data predict earlier risk to tehsils farther away from Karachi, in particular tehsils in Punjab (Faisalabad and Lahore, for example), than the diffusion model data. The diffusion model predicts the earliest risk to nearby tehsils such as Malir.
These uncertainties will continue to make forecasting challenging, although serotype data could improve model accuracy.

Given these difficulties, forecasting dengue epidemics will always need to encompass substantial stochastic variation, despite its importance for targeting limited public health resources. Currently, vector control programs in Pakistan begin during the monsoon season uniformly across the country. We believe that the large estimated lead times this approach offers could aid control managers, providing an early warning system. This approach provides policy-relevant, real-time information about where and when to expect dengue epidemics and therefore how to effectively target interventions, surveillance, and clinical response.

Materials and Methods

Population Data. Pakistan’s large population (182 million) is broadly divided into one capital territory, Federally Administered Tribal Area (FATA), Gilgit-Baltistan, Azad Jammu and Kashmir, and four provinces, which are further subdivided into 388 tehsils (equivalent to administrative unit 3 from 2008; Fig. 1). A 7-km kernelized period, from June 1 to December 31, 2013. The mobile operator has the largest coverage of tehsil headquarters (352 in the dataset of 388 total tehsils) across Pakistan, particularly in rural areas (two-thirds of the network). To comply with national laws and regulations of Pakistan and the privacy policy of the Telenor group, the following measures were implemented to preserve the privacy rights of Telenor Pakistan’s customers: (i) The CDR/mobility data were processed on a backup and recovery server made available to Telenor Pakistan only by Telenor employees. (ii) The detailed CDR/mobility data were taken out of Pakistan or left the premises of Telenor Pakistan. (iii) The processing of the detailed CDR/mobility data resulted in aggregated data on a tower-level granularity that was accessed only by Telenor employees. Further spatial aggregations to the level of the tehsil were made available to the remaining coauthors.

28 million subscriber SIMs were recorded as active on a given day, and of these 15.2 million subscribers generated outgoing, voice-event CDRs that encoded location information. At the time of data acquisition, the mobile phone operator had approximately a 25% market share (22% of the population) and was the second largest provider of mobile telecommunication services in Pakistan. Multi-SIM activity is common in Pakistan, but we believe that this should not create a systematic bias in mobility estimates because the geographic coverage of the operator is so extensive.

Quantifying Travel Using the Mobile Phone Data. Every caller was assigned to his or her most frequently used base station/mobile phone tower on a given day, as in previous studies (30). For a given location, defined through the location of a base station, the flux of that location is defined through the number of active callers assigned to the base station and was then aggregated to the tehsil level. On an average day, ~12% of the total population of Pakistan made a call. We measured daily travel between mobile phone towers relative to subscriber location on the previous day. Trips were aggregated to each tehsil based on the location of the origin and destination tower. We normalized trip counts by the origin tehsils’ number of active subscribers on each day (SI Text and Fig. S1). Aggregated forms of these mobility patterns can be made available upon request.

We observed two significant decreases in the recorded CDR activity, one occurring during Ramadan (July 9 to August 7, 2013), reflecting less subscriber activity (Fig. S1). The decrease in the use of mobile communication services during Ramadan has been confirmed by the mobile operator as an expected effect, and this effect is seen every year. The other decrease in activity was after October 25, due to a major system upgrade to the core mobile system infrastructure, which impacted all of the customers of the mobile operator. In the collected dataset, this led to a drop in the number of location-generating subscribers recorded per day of ~3.38 million. We adjusted the time series under the assumption that the system upgrade was a 1-d event only, based on expert opinion from the operator, and that the population in the customer base did not change its overall behavior. The flux values were adjusted after this date, assuming that the average would remain the same as in the beginning of the dataset (Fig. S1).

To analyze the relationship between mobility and dengue dynamics, we approximated travel patterns between January 1 and June 1. We simulated travel, assuming that the mean number of normalized trips (normalized by flux) remained the same, and added noise [Norm(0, σ), where σ is the variance in the number of trips between all pairs of tehsils] (SI Text and Fig. S1).

Dengue Data. Data for Punjab province were collected by the Provincial Health Department, whereas the District Health Office collected case data for Swat District (Table S1) daily. All public and private hospitals, health clinics, and laboratories reported any case of a patient presenting with dengue symptoms. The data listed each patient who was admitted in the province or district, seeking treatment for high fever, body aches, petechiae, and low platelet counts (with a cutoff value for thrombocytopenia of <50,000/mm³). These cases were then confirmed by using IgM or NS 1-Ag enzyme-linked immunoassay assay (ELISA). The data reported from Karachi were based on official public releases from the Sindh Health department. All case data were confirmed by the National Oceanic and Atmospheric Administration National Climate Data Center (Fig. S3). Temperatures peak in the middle of the year, June, July, and August; are lowest in January and December; and are variable across the country (daily averages between 9 °C and 27 °C) (for example, Figs. S2 and S3). We converted temperatures to dew-point temperature based on the Numerical Terradynamic Simulation Group proposed model at the University of Montana (31). Dew-point values and temperature were then converted to relative humidity.

Ento-Epidemiological Framework. To model dengue dynamics, we used an ordinary differential equation model based on a model by Lourenco and Recker to describe a dengue outbreak in Madeira, Portugal (10). This model captures the local dynamics of dengue between human and mosquito vectors by assigning temperature-dependent epidemiological variables as in Lourenco and Recker (SI Text and Fig. S3) (10).

The relationship between relative humidity and dengue suitability has been explored in a number of environments (3, 32). We have added an additional variable to the temperature-dependent epidemiological variables, the mortality factor that is based on relative humidity (33). In contrast to ref. 10, we varied the reporting rate ρ along with the biting rate a (SI Text). We varied the reporting rate between 2% and 10% (a 2% reporting rate is shown) because low rates of dengue reporting have previously been found in southern Asia (34). We did not fit the carrying capacity (K) explicitly and performed a sensitivity analysis, changing values of K. Although various values of K changed both a and ρ, it did not change the overall shape of the epidemic (Fig. S4 and SI Text). Using the temperature data for Karachi and the population (13.4 million), we estimated that a = 0.66, ρ = 0.02 (assuming K = 566; see SI Text for the sensitivity analysis).

Importation of Infected Travelers from Endemic Areas in Southern Pakistan. To estimate the number and role of importation of cases from endemic areas in southern Pakistan (Karachi) to all other tehsils, we first modeled the dengue dynamics in Karachi, using the ento-epidemiological framework described above. We are focused only on the role of travel within Pakistan as opposed to cross-border migration (Discussion). Dengue is endemic in Karachi (16), due to year-round climatic suitability and availability of vectors, and thus we assumed the date of the first introduced was day 1 (t = 1).

We then estimated the flow of infected travelers from Karachi. We first estimated the number of infected travelers per day who have left Karachi (T(t) = m(t)ρ). Based on the mobile phone data, ~30% (in the figures 30% is shown: min, 25.6%; median, 30%; max, 34%) of subscribers have traveled outside of Karachi per day (ρ) (Fig. S5). We varied this percentage between 10% and 30% to account for uncertainty in this estimate and a possible overestimation of travel because this value was based only on mobile phone subscribers (25). The daily number of infected individuals in Karachi is based on the modeled epidemic (m). We determined the destinations of infected travelers based on the daily percentage of travelers from Karachi to all other...
tehills ($x_i$, mobile phone data; $g_{ij}$, gravity model to location $j$). We calculated these flows using either the mobile phone data or a gravity (diffusion) model.

Here, the amount of travel between two locations, $i$, $j$ estimated via a basic diffusion, naive gravity, model is $N_{diffusion_{ij}} = (pop_i \times pop_j / dist(i,j))$ and we estimated the population ($pop_i$, $pop_j$) of each tehsil, using Worldpop.org (www.worldpop.org), and calculated the distance ($dist(i,j)$) as the travel time distance between centroids of each tehsil (35). We also fitted a standard gravity model to the mobile phone data (SI Text and Fig. S2).

The raw estimates of infected travelers from Karachi to all other tehsils per day represent an upper limit. To account for variation in epidemiological and individual factors, such as the within-host viral dynamics, we assumed that a smaller subset of this upper bound of infected visitors could contribute to transmission. We sampled the actual number of effective infected travelers from a binomial distribution with a fixed probability ($Pr(T, x_{ki}/j)$ and $Pr(P_{ij} \times g_{ij}$)) for introduction of infected visitors from Karachi to each tehsil ($m_i = 0.001$ and 0.9 (0.01 is shown)), see SI Text for further details. Thus, for each day, we had an estimated number of introduced cases from Karachi to each tehsil $Pr(T, x_{ki}/j) \times m_i = Y_{t,ki}$, $Pr(P_{ij} \times g_{ij}) \times m_i = Y_{t,ij}$. The date of the first introduced case from Karachi was sensitive to both the reporting rate for Karachi and the percentage of travelers from Karachi (SI Text) (see Fig. S6 for plots of each tehsil that reported cases), although it was less sensitive to the fixed probability (Fig. S7). We also investigated the possibility of imports from Mingora into Lahore (Fig. S8) during 2013 and find that imports from Karachi to Lahore are more likely the cause of the resulting epidemic in Lahore, although these two possible causes cannot be disentangled in the case data.

Environmental Suitability and Epidemic Risk. We defined dengue suitability as a function of temperature $Z_x(T)$ that is proportional to vectorial capacity ($V$) for a given location $x$ (19) (SI Text). This measure depends on the adult vector mortality rate (determined by temperature and relative humidity) and the infectious period for adult vectors. Values of $Z_x(T)$ approximately equal to zero indicate that the environment does not permit onward transmission and can be used to determine the timing of the geographic and temporal limits of dengue transmission in Pakistan.

We created a composite measure of environmental suitability and introduction of infected travelers from Karachi, $risk_{epidemic}(x) = \sum_{i} Y_{i,ki} Y_{i,ij}$ for a location $x$, where $Z_x(T)$ is the environmental suitability for dengue on day $t$. This measure sums the environmental suitability $Z_x(T)$ times the importation of infected travelers from Karachi per day ($Y_{i,ij}$).

Acknowledgments. A.W. is funded by a James S. McDonnell Foundation postdoctoral fellowship. M.F.B. is supported by Wellcome Trust Grant 098511/Z/12/Z, C.O.B., A.W., and M.A.I. were supported by the Models of Infectious Disease Agent Study program (Cooperative Agreement U1U54GM088558).