PLOS NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Constructing and validating a transferable epidemic risk index in data scarce environments using open data: A case study for dengue in the Philippines

Fleur Hierink[1,2]☯*, Jacopo Margutti[3]☯, Marc van den Homberg[3], Nicolas Ray[1,2]

1 Institute of Global Health, Faculty of Medicine, University of Geneva, Geneva, Switzerland, **2** Institute for Environmental Sciences, University of Geneva, Geneva, Switzerland, **3** 510, An Initiative of the Netherlands Red Cross, The Hague, The Netherlands

☯ These authors contributed equally to this work.
* fleur.hierink@unige.ch

Abstract

Epidemics are among the most costly and destructive natural hazards globally. To reduce the impacts of infectious disease outbreaks, the development of a risk index for infectious diseases can be effective, by shifting infectious disease control from emergency response to early detection and prevention.

In this study, we introduce a methodology to construct and validate an epidemic risk index using only open data, with a specific focus on scalability. The external validation of our risk index makes use of distance sampling to correct for underreporting of infections, which is often a major source of biases, based on geographical accessibility to health facilities. We apply this methodology to assess the risk of dengue in the Philippines.

The results show that the computed dengue risk correlates well with standard epidemiological metrics, i.e. dengue incidence (p = 0.002). Here, dengue risk constitutes of the two dimensions susceptibility and exposure. Susceptibility was particularly associated with dengue incidence (p = 0.048) and dengue case fatality rate (CFR) (p = 0.029). Exposure had lower correlations to dengue incidence (p = 0.193) and CFR (p = 0.162). Highest risk indices were seen in the south of the country, mainly among regions with relatively high susceptibility to dengue outbreaks.

Our findings reflect that the modelled epidemic risk index is a strong indication of subnational dengue disease patterns and has therefore proven suitability for disease risk assessments in the absence of timely epidemiological data. The presented methodology enables the construction of a practical, evidence-based tool to support public health and humanitarian decision-making processes with simple, understandable metrics. The index overcomes the main limitations of existing indices in terms of construction and actionability.

Author summary

Epidemics are among the most costly and destructive natural hazards occurring globally; currently, the response to epidemics is still focused on reaction rather than prevention or preparedness. The development of an epidemic risk index can support identifying high-risk areas and can guide prioritization of preventive action and humanitarian response. While several frameworks for epidemic risk assessment exist, they suffer from several limitations, which resulted in limited uptake by local health actors—such as governments and humanitarian relief workers—in their decision-making processes. In this study, we present a methodology to develop epidemic risk indices, which overcomes the major limitations of previous work: strict data requirements, insufficient geographical granularity, validation against epidemiological data. We take as a case study dengue in the Philippines and develop an epidemic risk index; we correct dengue incidence for underreporting based on accessibility to healthcare and show that it correlates well with the risk index (Pearson correlation coefficient 0.69, p-value 0.002). Our methodology enables the development of disease-specific epidemic risk indices at a sub-national level, even in countries with limited data availability; these indices can guide local actors in programming prevention and response activities. Our findings on the case study show that the epidemic risk index is a strong indicator of sub-national dengue disease patterns and is therefore suitable for disease risk assessments in the absence of timely and complete epidemiological data.

Introduction

Epidemic risk assessment

Epidemics are among the most costly and destructive natural hazards globally [1, 2]. Currently humanitarian action to epidemics is focused on response rather than preparedness and prevention [3–6]. Timely detection of disease cases in combination with risk assessment can support prevention measures and therefore contribute to early containment of outbreaks [4, 6]. The use of a holistic risk index for infectious diseases can reduce the impacts of epidemics on (vulnerable) communities, by shifting infectious disease control from response after emergence to early detection and prevention [1, 2, 6]. Comparable risk indices for natural hazards and humanitarian crises have proven to be effective in localizing high risk regions [7, 8] and are being used to inform preparedness programs [7, 8]. Research has shown that the risk of re-emerging infectious disease outbreaks or new spillover events (i.e. pathogen transmission from a reservoir to a new host) is increasing due to degrading ecosystems, intensification of travel and trade, climate change, population growth, and a wide variety of other factors [3, 6]. It is therefore imperative to increase our understanding of disease risk distribution at the most local level possible, so impacts can be reduced accordingly [4–6].

Epidemic risk is usually quantified by several indicators, which relate both to the probability of outbreak occurrence and to its potential impact [9–12], chosen according to the specific disease(s) under consideration. These indicators are combined and mapped to a normalized risk index, according to initial estimates (most commonly, weighted or geometric means). They can be conceptually divided into two dimensions [9]:

1. Hazard and exposure: the presence of an infectious disease and its vector (e.g. mosquitoes) and the likelihood of exposure

2. Vulnerability and coping capacity: clinical, demographic and socioeconomic data that influence health outcomes (e.g. age); the ability of a government or health system to detect, contain and respond to an outbreak (e.g. hospital capacity)

While several frameworks for epidemic risk assessment exist [9–13], they have been hardly used by health actors—such as governments and humanitarian relief workers—to prioritize intervention areas and actions, despite them being the first responders to epidemic outbreaks and thus often carrying major decision responsibilities. While dedicated research on the reasons behind this lack of adoption is missing, anecdotal evidence from humanitarian practitioners often points to one or more of the following limitations, which existing epidemic risk indices suffer from:

1. Methodology:

    (a).   rely on accurate clinical, virological and/or entomological data, which often require dedicated and in-situ data collection campaigns; these are costly, impractical and often prerogative of health authorities

    (b).   focus on a global scale, by comparing world regions or countries; this can be useful for international organisations (e.g. in long-term planning of funding by donors), but not for local ones [14]

    (c).   lack of validation against epidemiological data

2. Actionability: lack a clear connection with policy implications and practical interventions, i.e. a prescriptive aspect [15]

The methodological limitations are connected with data availability, most importantly of clinical surveillance data, whose lack of determines the difficulty, respectively, of using advanced epidemiological models [16], of modelling at a sub-national scale and, finally, of validating results. While epidemic surveillance systems that collect and aggregate this data [17] do exist, they often lack completeness and timeliness, especially in low- and middle-income countries (LMICs) [18], which carry the highest burden of infectious diseases [1]. Official numbers from such surveillance systems are often derived from clinical records of symptomatic cases [19], i.e. *passive surveillance*, and thus do not take into account asymptomatic cases, underdiagnosis and, most importantly for LMICs countries, infected individuals who do not receive treatment. This problem is often referred to as underreporting.

Underreporting in passive surveillance systems has been recognised as a major source of bias in estimates of infectious disease incidence, especially in LMICs [20, 21]. Known factors related to underreporting, other than possible asymptomatic cases [22], are sociodemographic factors that impede health-seeking behavior, such as poverty and education [23, 24], and geographical accessibility to health facilities [25–27]. Understanding and quantifying these factors is thus necessary to assess disease incidence, and thus epidemic risk, at a local level [26]. While the socioeconomic indicators which affect self-reporting are usually difficult to measure at a population-level and their relative importance is highly dependent on the local context [23], data on the location of health facilities is available virtually world-wide (but with various degrees of completeness) and their geographical accessibility can be modeled with suitable geo-spatial tools [25, 28]. Geographical accessibility models present an important and novel opportunity to bridge the data gap between reported and unreported cases, as it reflects the ability of a population to reach a health facility within a certain travel time [28]. Recently, a robust methodology has been proposed to correct for underreporting based on known covariates [26].

Dengue and the Philippines

Vector-borne diseases (VBDs), i.e. infectious diseases that are transmitted through a blood-feeding arthropod (e.g. mosquitoes, sandflies, ticks, etc.) are an important group of infectious diseases [29]. VBDs are responsible for 17% of the total burden of all communicable diseases and their prevalence disproportionately affects the poorest communities in tropical and sub-tropical regions [29, 30]. Socioeconomic, demographic and environmental indicators are known to be strongly linked to the distribution of VBD risk and an expansion of transmission patterns in the coming years is expected due to environmental changes, rapid urbanization, and globalization [30].

While the global communicable disease burden for some of the largest infectious diseases (i.e. HIV/Aids, tuberculosis, and malaria) has been tremendously reduced over the last decade, deaths due to the VBD dengue have increased by 65.5% from 2007 to 2017, with the same trend seen for dengue case fatality rates (CFR) [29, 31]. Dengue is a mosquito-borne disease with four different serotypes (i.e. DENV-1, DENV-2, DENV-3, DENV-4) and is considered as a neglected tropical disease by the World Health Organization (WHO). The disease is spread by the mosquitoes *Aedes aegypti* and *Aedes albopictus* and is responsible for an estimated 96 million cases annually, with 50% of the world's population expected to be at risk [29, 30]. Currently, the primary method for controlling dengue are vector control strategies, aimed at limiting human exposure to the transmitting mosquitoes. Targeting regions for vector control measures is of high importance to optimize and maximize the effect of the available resources [29, 30]. Understanding the distribution of dengue risk is key in tailoring and targeting intervention strategies on sub-national scales [30], but challenging due to underreporting [26]. Improved methods are needed to meet the recently updated WHO NTD roadmap target of a 0% dengue case fatality rate (CFR) by 2030 (from the 0.8% baseline in 2020) [32].

Dengue is a large scale health challenge in the Philippines [33]. The disease is endemic in the entire country with re-occurring outbreaks in all regions and the circulation of all virus strains. The country is highly vulnerable for dengue outbreaks, partly as a consequence of recurring natural hazards destructing critical infrastructures, but also because of environmental conditions favouring the life-cycle of mosquitoes [34, 35]. Dengue surveillance in the Philippines mostly represents hospitalized cases, particularly those of patients with severe dengue infections. Between 2010 and 2014 about 93% of all reported dengue cases concerned hospitalized patients of which 50% were reported from private facilities [33]. This finding highlights the fact that a large portion of the dengue cases may remain unreported, hindering a realistic understanding of dengue in the country and thus stressing the need for realistic correction methods [33].

Reliable risk estimates of dengue are needed in the Philippines to allow guided allocation of preventive measures and targeted outbreak containment [33, 36]. Research on dengue in the Philippines has focused on modeling techniques, with the goal of either describing past disease dynamics or to predict future ones [34, 37–40]. While such models could provide estimates of (future) morbidity, which is key to inform epidemic response and preparedness programs, they suffer one or more of the following limitations: relying on detailed clinical and/or entomological data, which is rarely available, and not discussing potential health outcomes, e.g. by considering the local (health) capacity. To the best of our knowledge, no research has been carried out yet on combining the different dimensions of dengue risk (hazard and exposure, vulnerability and coping capacity) into one quantitative framework, that can be applicable at a country scale. The inherent challenge, like in other data-scarce environments, is missing information on relevant risk indicators and epidemiological data.

In this study, we present a methodology to build and validate an epidemic risk index at a sub-national level, using openly available data and tools, to ensure its applicability in data-scarce settings. The methodology can be conceptually divided in three steps:

- Development of the Epidemic Risk Index: selection of indicators, normalization and aggregation

- Correction of (public) epidemiological data: estimation of relative differences in underreporting based on geographical accessibility to healthcare

- Validation of the Epidemic Risk Index against corrected epidemiological data

Materials and methods

Study region

The Philippines is an archipelago nation in the Western Pacific ocean and is subdivided into 17 administrative regions, which are further subdivided into 81 provinces, 1489 municipalities and 42,036 barangays [41, 42]. Historically, dengue cases were reported on a weekly basis by the Department of Health in surveillance reports [43]. Although we acknowledge that spatial granularity is important in risk assessment models, dengue cases have been mostly openly reported on regional level (n = 17). Therefore this study focuses on a risk index for all 17 administrative regions in the Philippines.

Epidemic risk index

The epidemic risk index was built largely following the methodology of the Water Associated Disease Index (WADI) [12], which has been developed with a specific focus on dengue and has been successfully validated against actual dengue incidence data, even at a sub-national level, in Malaysia [13] and Vietnam [44]. The risk index is defined as a weighted average of two components, exposure and susceptibility, which quantify the risk of being exposed to the pathogen and the risk of experiencing severe health outcomes, respectively. Following the methodology of [13], the weights in this average were chosen to maximize the correlation with dengue incidence.

Each component is in turn defined as the arithmetic average of one or more indicators, summarized in Table 1. The sign of each indicator was changed, if necessary, so that higher

Table 1. Indicators used to build the Epidemic Risk Index and epidemiological metrics to assess its reliability.

Component	Indicator	Rationale for inclusion	Data source
Exposure	Fraction of the population exposed to *Aedes aegypti*	*Aedes aegypti* is the main dengue vector in the Philippines	[45, 46]
Susceptibility	Percentage of children 0–15 years of age	Children 0–15 years of age have higher susceptibility to severe dengue and CFR is higher among this group	[57]
	Female enrollment ratio to secondary school	Progression to secondary school indicates a sufficient level of education and attainment to read, interpret and act upon public health information about dengue	[58]
	Percentage of households using unimproved sanitation facilities	Unimproved sanitation indicates poorly managed water resources and poor housing quality	[59]
	Number of physicians per 1000 people	Physician density is a proxy for availability of healthcare	[60]
	Number of beds in public hospitals per 1000 people	Density of beds in public hospitals is a proxy for affordability of healthcare	[61]
Validation	Dengue incidence (number of dengue cases per person per year) and CFR; both were averaged over the years 2014–2019.	Average incidence and CFR quantify, respectively, the risk of outbreaks and their severity, in terms of health outcomes	[43]

https://doi.org/10.1371/journal.pntd.0009262.t001

values correspond to higher risk, according to the rationales listed in Table 1 (e.g. female enrollment ratio to secondary school negatively correlates with risk, so its sign was inverted, while the percentage of children 0–15 years of age was unchanged). Additionally, each indicator was transformed (unless already normalized so that it lies in the range [0, 1]) according to

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}.$$ (1)

Concerning the choice of indicators, exposure was quantified as the fraction of the population E exposed to *Aedes aegypti*. This indicator was calculated from the probability of occurrence of the main dengue vector (*Aedes aegypti*) V, modeled in a raster format [45], and the population density distribution ρ [46], according to

$$E_k = \frac{\sum_{i=1}^{N_k} V_i \, \rho_i}{\sum_{i=1}^{N_k} \rho_i},$$ (2)

where N_k is the number of raster cells within the boundaries of region k. We used high resolution population density estimates from the Facebook Connectivity Lab and Center for International Earth Science Information Network (CIESIN) [46].

Susceptibility was instead quantified combining five indicators which relate to vulnerabilities against dengue. Children 0–15 years of age have much higher chances to develop severe dengue with respect to the adult population [47] and thus constitute the most vulnerable group; their relative abundance was quantified via the fraction of the population belonging to the corresponding age group. Secondly, education has been identified among the key factors enabling health-seeking behavior [23], especially for the caregiver of the household [48], and is assumed to increase the capacity of interpreting and acting upon public health information aimed at preventing dengue. This was quantified via the female enrollment ratio to secondary school. Thirdly, the percentage of households using unimproved sanitation facilities (pit latrines without slabs or platforms or open pit, hanging latrines, bucket latrines, open defecation) was used to capture the risk of having exposed water containers in the house, which can act as breeding sites and was associated with higher risk of dengue [49]. While other factors contribute as well to the availability of breeding sites (most importantly, deficiencies in water supply and waste management [50, 51]), data on unimproved sanitation facilities is much more commonly available (including in the Philippines), as it is a standard indicator in demographic or health surveys [52], and is likely to correlate with the former. Both unimproved sanitation and low education are effectively proxies for poverty, which indirectly affects health outcomes [53]. Lastly, the density of physicians and public hospital beds was used to quantify, respectively, the availability and affordability of healthcare.

While geographical accessibility to health facilities is equally important in determining the probability of seeking and receiving adequate treatment [54], it was not included in the definition of our risk index to avoid a spurious correlation with the dengue incidence, which was corrected using accessibility data and was ultimately used to validate the index. In this study, validation refers to the correlation between the predicted risk index and dengue incidence as well as exposure and susceptibility. This was done to measure to what extent the estimated risk index reflects actual dengue incidence in the different regions.

The risk of each dengue case to develop into severe dengue and, potentially, mortality is known to be determined by a number of other factors [55], most importantly immunity and previous exposure to a different serotype of dengue, due to antibody-dependent enhancement [56]; however, to the best of our knowledge, no serotype-specific case data exists for the region and period under study and we thus had no way to quantify such effect. Ultimately, the

calculated risk index was validated against the dengue CFR and incidence [43], averaged over the years 2014–2019, by means of Pearson correlation coefficients.

Accessibility to health care.   Accessibility to health care was measured in terms of travel time (in minutes) to health facilities with dengue testing services. The applied travel scenario considered motorized travel speeds on roads and walking travel speeds on other land cover types (e.g. forest, grassland, urban landscapes) under the assumption that patients walk to the nearest road and then continue their journey with a vehicle that is readily available. Travel time rasters were computed per facility and by means of a least cost-distance algorithm in arcpy, following closely the methodology of AccessMod version 5.6.30 [28]. In order to obtain a single 110 meter resolution travel impedance surface raster, spatial data on elevation, land cover, roads, and river networks were merged in an overarching raster layer through the *merge landcover* module in AccessMod, to which the travel scenario was applied [28] (S1 Table). Each health facility coordinate (n = 4167) was then separately superimposed on the travel impedance surface to obtain a travel time raster for each individual health facility with dengue testing services.

Data preparation of all separate spatial layers was done using RStudio (R version 4.0.2). Land cover data was downloaded in tiles from Coopernicus [62] and elevation data from Shuttle Radar Topography Mission (SRTM) [63]. Both spatial raster layers were mosaiced to cover the Philippines and clipped to country borders. The two raster layers (land cover and elevation) were then re-sampled to a resolution of 110 meter, using the native resolution of the landcover as a reference, and raster cells were aligned with the elevation layer as a reference.

Vector data representing the road network and hydrography had to be separately downloaded for the Northern and Southern part of the country from Humanitarian Open Street Map [64, 65] and was enriched with data from the Open Mapping at Facebook Initiative [66]. Layers on both parts of the country were merged. Hydrographic features such as rivers and lakes were considered full barriers to movement to the population, unless a road crosses over, which was considered as a functional bridge. Road data was cleaned to only contain OpenStreetMap official road classes [67] and new integer road class values were created for each unique road type, as an essential step for the land cover merge. Health facility coordinates were downloaded from the Department of Health in the Philippines [68] and health facilities known to offer dengue testing services (i.e. "Rural Health Unit", "Hospital", "Medical Clinic"), as discussed with country representatives were filtered from the data. Coordinates falling on barriers were moved to the nearest neighbouring non-barrier cell and facilities wrongly located far outside country borders were removed from the analysis. All raster and vector datasets were projected to the Philippines' projection system (EPSG:32651, UTM51N).

Reporting probability.   All the travel time rasters (n = 4167) obtained from the accessibility model served as the input data for the multinomial calculation of the reporting probability, following a distance sampling methodology [69], modified for epidemiological studies [26]. In particular, we used the following equation to describe the reporting probability (P) as a function of travel time to health facilities with dengue testing service (t):

$$P(t) = \exp\left(a_0 + a_1 t^c\right), \tag{3}$$

where a_0, a_1 and c are free parameters. This function captures the main feature of the traditional assumption used in distance sampling methods, namely an exponential decrease. Since we did not have access to individual patient case data, we used the results of [26] to give an estimate of the free parameters in Eq 3, converting the time travel t to distance d by dividing it by the average travel time v

$$t(d) = d/v, \tag{4}$$

where v was estimated according to the aforementioned travel scenario (S1 Table).

Eq 3 was first applied to each cell of the travel time raster, to produce a reporting probability raster per health facility (n = 4167). Next, the total reporting probability raster is computed by summing the probabilities of each health facility j in each raster cell i and normalizing according to

$$P_i = \frac{\sum_{j=1}^{N_{\mathrm{hf}}} P_{i,j}}{1 + \sum_{j=1}^{N_{\mathrm{hf}}} P_{i,j}}, \tag{5}$$

where N_{hf} is the total number of health facilities.

We then computed the average reporting probability per region $\langle P \rangle$ by taking a weighted average of the reporting probability within the region boundaries and using population density ρ in each raster cell i as weight:

$$\langle P \rangle_k = \frac{\sum_{i=1}^{N_k} P_i \rho_i}{\sum_{i=1}^{N_k} \rho_i}, \tag{6}$$

where N_k is the number of cells within the boundaries of region k. We used the same population density estimates [46], resampled to 110m resolution by summing population. This technique results in the loss of population across the grid, mainly as a result of reprojecting the layer. To correct for this, the total lost population was smoothed out over the resampled population grid.

Average reporting probability was then used to correct dengue regional incidence, which was in turn estimated from official dengue case counts [43] and census data [57]. This step corrects for the major imbalance in official dengue statistics due to unequal access to healthcare. Finally, since the almost entirety of reported cases comes from hospitalized settings [70] due to the dengue case definition [71], dengue incidence was corrected for the fraction of hospital beds belonging to facilities connected to the Philippines epidemiological surveillance system [72].

Results

Validation of the epidemic risk index

The Epidemic Risk Index and its components are validated against the corrected dengue incidence and CFR in the 17 regions under study by measuring the Pearson correlation coefficient r (Table 2). The significance of each correlation is measured with p-value at the significance level of 0.05 ($p < 0.05$). Concerning incidence, a positive, significant correlation is observed between incidence and susceptibility and between incidence and risk: $r = 0.49$ ($p = 0.047$) and

Table 2. Correlation coefficients between the corrected dengue incidence, CFR, Epidemic Risk Index and its components. In bold: significant correlations ($p < 0.05$).

Variables	Pearson r	p-value
Incidence and exposure	0.33	0.193
Incidence and susceptibility	**0.49**	**0.048**
Incidence and risk	**0.69**	**0.002**
CFR and exposure	-0.35	0.162
CFR and susceptibility	**0.53**	**0.029**
CFR and risk	0.25	0.342
Exposure and risk	0.44	0.081
Susceptibility and risk	0.73	**0.001**

https://doi.org/10.1371/journal.pntd.0009262.t002

$r = 0.69$ ($p = 0.002$), respectively. Concerning CFR, a significant correlation is observed only between CFR and susceptibility, with $r = 0.53$ ($p = 0.029$).

Accessibility to healthcare.   Accessibility to dengue reporting facilities was highest in the National Capital Region (Fig 1A and 1B). Where 99.98% percent of the population (N = 12,304,651) was able to reach a health facility within 1 hour travel time. Lowest accessibility coverage was seen in Region IV-B, with 83.3% percent of the population being able to access care in 1 hour (Fig 1A and 1B).

Reporting probability.   Reporting probability (Fig 1C) was generally highest around Manila with an average reporting probability of 0.94 in the National Capital Region. However, reporting probability was lower for all other regions, with probabilities ranging from 0.61 to 0.88, implying that reported incidence was corrected with higher correction factors among all these regions (S5 Fig and S2 Table).

Geographical distribution of dengue risk.   The maps in Fig 1D–1F and S1 Fig represent the results of the calculated susceptibility, exposure, and ultimately risk index. Regions with high susceptibility, thus reflecting low coping capacity and resilience, are depicted in darker blue colors. Regions with high potential *Aedes aegypti* exposure are shown in darker green. Ultimately, regions with a relatively high risk index (Fig 1F) are shown in darker orange.

The Pearson correlation was strongest between the susceptibility dimension and dengue incidence (P = 0.048), as compared to the other covariates (Table 2). Therefore, susceptibility related variables weighted heavier on the risk index than the exposure variables. Susceptibility was highest in Region XII (0.65) and lowest in the National Capital Region (0.29), reflecting higher coping capacity of individuals and the health system around the capital.

Comparing the exposure index to the susceptibility index for instance, shows that regions with highest exposure index are mainly located in Northern regions of the country, whereas susceptibility was found to be highest in more Southern regions. The exposure index was found to be highest in the National Capital Region (0.95) and ranged from 0.44–0.95 throughout the entire country.

The modelled risk index ranged from 0.43 to 0.69 between all regions in the Philippines. All results are aggregated on regional level, firstly because dengue data was richest in terms of temporality and secondly because decision-making on resource allocation is often carried out at this level. The modelled risk index was highest for Region XII, with an index of 0.69 and the exposure and susceptibility index being 0.76 and 0.65 respectively (S1 Fig). Interestingly, CFR in this region was relatively low, being 0.43. The second highest risk index was seen in Bangsamoro, with a risk index of 0.64, and an exposure and susceptibility index of 0.67 and 0.62. In general, a cluster of higher risk indices was concentrated in Southwest Philippines (Fig 1F). When comparing this cluster of high risk indices against the susceptibility and exposure index, it becomes apparent that especially the susceptibility index is highest in these regions (Fig 1D) while higher values for the exposure index are seen among northern regions in the Philippines (Fig 1E). Highest CFRs are also concentrated in the Southwest regions of the Philippines (Fig 1G).

In general, when comparing the spatial distribution of the susceptibility and exposure index against the risk index there is no notable trend visible between the exposure index and the risk index. Yet, the susceptibility and risk indices show a more closely related trend, towards the southern regions of the country.

While accessibility in terms of travel time (Fig 1B) are highest in the Northeast of the Philippines, which might potentially reflect a poorer capacity to deal with an outbreak, the susceptibility index is generally low in this region.
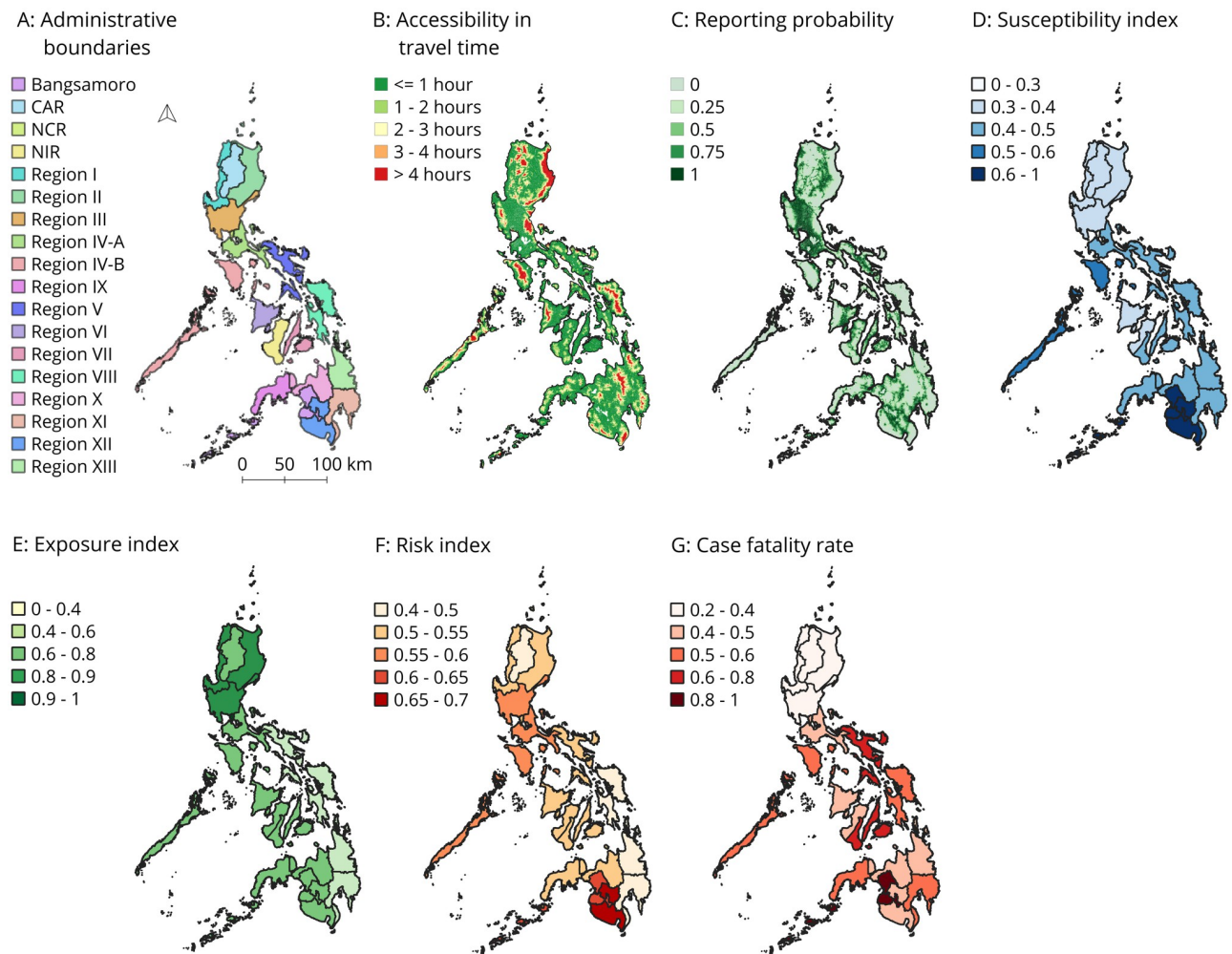
Fig 1. Overview of all results. Panel A shows the 17 administrative regions of the Philippines. Panels B-G highlight the individual results of the accessibility analysis (travel time to nearest facility for simplicity), reporting probability, the dimensions that compose the risk index, the risk index, and the case fatality rate. Enlarged versions of the figures can be found in S2–S8 Figs. Sub-national boundaries are sourced from UN-OCHA and openly available from Humanitarian Data Exchange: https://data.humdata.org/dataset/philippines-administrative-levels-0-to-3.

https://doi.org/10.1371/journal.pntd.0009262.g001

Discussion

In this study, we have constructed and validated an epidemiological risk index using openly available data, to assess exposure and vulnerability to dengue in the Philippines. The proposed methodology can be easily applied to other countries and diseases, as it does not use data which is uniquely available in the Philippines nor does it depend on specific features of dengue epidemiology. More specifically, the indicators used to construct the index are commonly captured at a sub-national level by public demographic and health surveys or, where government capacity is limited, by humanitarian programs such as USAID's DHS [52]; while different indicators might be more suitable for different diseases (e.g. elderly, not children, might be more at risk of severe health outcomes), we think that a reasonable set can be found within the aforementioned sources. The correction procedure of official epidemiological data for underreporting, which was used to validate the epidemic risk index, is also expected to be applicable in other contexts, i.e. other endemic infectious diseases and countries in which a passive surveillance system is in place.

Looking at our particular case study, we have shown how risk factors of dengue vary within the Philippines and how these correlate with epidemiological metrics. We observed, overall, that the combination of exposure and susceptibility explains, to some extent, the observed incidence and mortality rate, and it does so better than considering each of these two separately. The higher correlation between dengue incidence and risk index with regard to exposure and susceptibility alone is consistent with the hypothesis that there is an interplay between the latter two and that both need to be taken into account to correctly estimate epidemiological risk.

We also acknowledge that our study dealt with several challenges, which we discuss more in detail in the following.

Accessibility analysis

Our travel scenario may not have been representative of all populations in the Philippines. Regional specificities on modes of travel or road quality may exist, and socio-economic differences within or between regions may alter the predominant modes and speeds of travel. A finer grain study on these potential geographic disparities could improve our travel model and therefore the reporting bias estimates.

Correction for underreporting

While the models in [26] were fitted on case data of malaria in Burkina Faso, we argue that such scenario should be reasonably representative of dengue in the Philippines, at least for the purpose of this work. Dengue and malaria share indeed a high prevalence of asymptomatic cases [73, 74] which do not prompt healthcare seeking; also, they are both endemic in the Philippines and Burkina Faso, respectively. Other factors influencing health-seeking behavior, such as socio-economic ones [23], could determine a difference in reporting probability between these two countries; however, the factors that were explicitly included in [26] determined a poorer model performance with respect to including only distance, suggesting that the latter is indeed the main driver behind reporting probability. The impossibility of explicitly modelling reporting probability using data from the Philippines, which forced us to use parameters derived from another study, constitutes a limitation of the current study, which we recommend to avoid whenever individual patient case data is available. Finally, we note that our methodology aims at correcting for relative differences in reporting probability among regions in the Philippines, meaning that an absolute difference with true reporting probability might very well exist and does not influence the validity of our results.

Risk index

The risk index that we constructed is meant to be a simple metric to guide decision-making processes and resource allocation of humanitarian agencies. Simplicity comes at a price: while we show that it does correlate with both incidence and CFR, and present a methodology to test this case-by-case, it is difficult to be more specific about actual expected health outcomes in case of an outbreak, given a certain value of the risk index.

Concerning exposure, the probability of vector occurrence has been modeled on the basis of environmental variables, among which the degree of urbanization (urbanicity) [45]; however, such model did not explicitly take into account the abundance of breeding sites, most importantly in solid waste and plastic containers, which has recently been identified as a key ingredient of vector ecology [50, 51]. The type and coverage of solid waste management is therefore expected to be a good predictor of vector abundance, although geographically detailed information on such a topic in the study region is scarcely available. Also, we note that using climatic averages to compute vector exposure is another important limitation of [45], as

dengue incidence is known to follow seasonal patterns in the Philippines [35, 37]. However, extensive research as been conducted already on the topic [34, 40, 75] and a time-dependent exposure is easily implementable within the current framework, enabling real-time monitoring or even forecasts of the risk index throughout the year. We plan to address this in future research.

Finally the susceptibility dimension does not hold information on potential transmission dynamics of dengue to the population it represents, the social predisposition and resilience of the population in case an outbreak occurs. Therefore, it can help target regions for building prevention and preparedness strategies. While susceptibility indicators capture important aspects of health systems, they might not take into account local, specific factors that play a decisive role both in health-seeking behavior and capacity to deliver care. In the Philippines, for instance, the southwest region of Bangsamoro (previously known as Autonomous Region in Muslim Mindanao) has been plagued by years of violent conflict between tribal, political and religious group and the government [76]. Not only does this affect the local health system resilience, but it is also a major factor to consider when planning humanitarian interventions, which this risk index is meant to inform.

Conclusions

The presented methodology enables the construction of a practical, evidence-based tool to support public health and humanitarian decision-making processes with simple, understandable metrics, namely the epidemic risk index and its components. Our methodology overcomes the main limitations of existing epidemic risk indices (see Introduction): it is based on openly available data, it is localized, and results can be validated against epidemiological data. In terms of actionability, other than helping prioritizing intervention areas, we note that individual indicators contain useful information for humanitarian programs. Absolute numbers of potentially exposed and vulnerable people, for instance, can be directly extracted, together with clear indications on which interventions should be prioritized and where (e.g. vector control programs versus strengthening community-based surveillance). Investments in epidemic prevention, detection, and response are needed to advance in our capacity to deal with infectious disease outbreaks. The information captured in the epidemic risk index supports the general shift from reaction after emergence to epidemic prevention and preparedness that has been so widely advocated for, especially in light of the ongoing COVID-19 pandemic, and is transferable to other infectious diseases and settings.

Supporting information

S1 Table. Travel scenario. The speeds applied to our friction raster to calculate the accumulated travel times to dengue reporting facilities.
(XLSX)

S2 Table. Correction factors.
(CSV)

S1 Fig. Overview of gradient indices. Absolute indices for individual dimensions, risk index and case fatality ratio.
(PNG)

S2 Fig. Regional boundaries in the Philippines.
(PNG)

S3 Fig. Accessibility to dengue reporting facilities expressed as travel time.
(PNG)

S4 Fig. Reporting probability of dengue.
(PNG)

S5 Fig. Susceptibility index.
(PNG)

S6 Fig. Exposure index.
(PNG)

S7 Fig. Risk index.
(PNG)

S8 Fig. Case fatality ratio.
(PNG)

Acknowledgments

Author Contributions

Conceptualization: Fleur Hierink, Jacopo Margutti, Marc van den Homberg, Nicolas Ray.

Data curation: Fleur Hierink, Jacopo Margutti.

Formal analysis: Fleur Hierink, Jacopo Margutti.

Investigation: Fleur Hierink, Jacopo Margutti.

Methodology: Fleur Hierink, Jacopo Margutti, Marc van den Homberg, Nicolas Ray.

Supervision: Fleur Hierink, Jacopo Margutti.

Validation: Jacopo Margutti.

Visualization: Fleur Hierink.

Writing – original draft: Fleur Hierink, Jacopo Margutti, Marc van den Homberg, Nicolas Ray.

Writing – review & editing: Fleur Hierink, Jacopo Margutti, Marc van den Homberg, Nicolas Ray.

References

1.   Harvard Global Health Institute. Global Monitoring of Disease Outbreak Preparedness: Preventing the Next Pandemic. 2018.

2.   Sands P, Mundaca-Shah C, Dzau VJ. The neglected dimension of global security—a framework for countering infectious-disease crises. New England Journal of Medicine. 2016; 374(13):1281–1287. https://doi.org/10.1056/NEJMsr1600236

3.   Smith KF, Goldberg M, Rosenthal S, Carlson L, Chen J, Chen C, et al. Global rise in human infectious disease outbreaks. Journal of the Royal Society Interface. 2014; 11(101):20140950. https://doi.org/10.1098/rsif.2014.0950 PMID: 25401184

4.   Morse SS, Mazet JAK, Woolhouse M, Parrish CR, Carroll D, Karesh WB, et al. Prediction and prevention of the next pandemic zoonosis. The Lancet. 2012; 380(9857):1956–1965. https://doi.org/10.1016/S0140-6736(12)61684-5 PMID: 23200504

5.   Jonas O, Katz R, Yansen S, Geddes K, Jha A. Call for independent monitoring of disease outbreak preparedness. Bmj. 2018; 361. https://doi.org/10.1136/bmj.k2269 PMID: 29794069

6.   Allen T, Murray KA, Zambrana-Torrelio C, Morse SS, Rondinini C, Di Marco M, et al. Global hotspots and correlates of emerging zoonotic diseases. Nature communications. 2017; 8(1):1–10. https://doi.org/10.1038/s41467-017-00923-8 PMID: 29066781

7.   INFORM report 2020: Shared evidence for managing crisis and disaster. EUR 30173 EN. 2020.

8.   Birkmann J. Risk and vulnerability indicators at different scales: Applicability, usefulness and policy implications. Environmental Hazards. 2011; 7(1):20–31. https://doi.org/10.1016/j.envhaz.2007.04.002

9.   Vernaccini L, Doherty B, Marin Ferrer M. INFORM Epidemic Risk Index: Support Collaborative Risk Assessment for health threats. Publications Office of the European Union. 2017.

10. World Health Organization Regional Office for Africa. Mapping the risk and distribution of epidemics in the WHO African Region: a technical report. 2016.

11. Oppenheim B, Gallivan M, Madhav NK, Brown N, Serhiyenko V, Wolfe N, et al. Assessing global preparedness for the next pandemic: development and application of an Epidemic Preparedness Index. BMJ Global Health. 2019; 4(1). https://doi.org/10.1136/bmjgh-2018-001157 PMID: 30775006

12. Fullerton LM, Dickin SK, Schuster-Wallace CJ. Mapping global vulnerability to dengue using the water associated disease index. United Nations University. 2014; p. 16–29.

13. Dickin SK, Schuster-Wallace CJ, Elliott SJ. Developing a vulnerability mapping methodology: applying the water-associated disease index to dengue in Malaysia. PLoS One. 2013; 8(5):e63584. https://doi.org/10.1371/journal.pone.0063584 PMID: 23667642

14. Van Den Homberg, M., Visser, J., & Van Der Veen, M. (2017). Unpacking data preparedness from a humanitarian decision making perspective: Toward an assessment framework at subnational level. In Proc. Information Systems for Crisis Response and Management (ISCRAM) Conference.

15. Bierens S, Boersma K, van den Homberg M. The Legitimacy, Accountability, and Ownership of an Impact-Based Forecasting Model in Disaster Governance. Politics and Governance. 2020; 8(4):445–455. https://doi.org/10.17645/pag.v8i4.3161

16. Walters CE HI Meslé MMI. Modelling the global spread of diseases: A review of current practice and capability. Epidemics. 2018; 25:1–8. https://doi.org/10.1016/j.epidem.2018.05.007

17. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition). Eighth edition ed. Philadelphia: Springer; 2015. Available from: http://www.sciencedirect.com/science/article/pii/B9781455748013000326X.

18. Conte P, Natale A, Possenti F, Savini L, Danzetta L. Systems for prevention and control of epidemic emergencies. Vet Ital. 2013; 49:255–261. PMID: 24002938

19. Gibbons CL, Mangen MJJ, Plass D, Havelaar AH, Brooke RJ, Kramarz P, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. BMC Public Health. 2014; 14:147. https://doi.org/10.1186/1471-2458-14-147 PMID: 24517715

20. Runge-Ranzinger S, McCall PJ, Kroeger A, Horstick O. Dengue disease surveillance: an updated systematic literature review. Tropical Medicine & International Health. 2014; 19(9):1116–1160. https://doi.org/10.1111/tmi.12333 PMID: 24889501

21. World Health Organization. World malaria report. 2018.

22. Lindblade KA, Steinhardt L, Samuels A, Kachur SP, Slutsker L. The silent threat: asymptomatic parasitemia and malaria transmission. Expert Review of Anti-infective Therapy. 2013; 11(6):623–639. https://doi.org/10.1586/eri.13.45 PMID: 23750733

23. Lazar M, Davenport L. Barriers to Health Care Access for Low Income Families: A Review of Literature. Journal of Community Health Nursing. 2018; 35(1):28–37. https://doi.org/10.1080/07370016.2018.1404832 PMID: 29323941

24. Veugelers PJ, Yip AM. Socioeconomic disparities in health care use: Does universal coverage reduce inequalities in health? Journal of Epidemiology & Community Health. 2003; 57(6):424–428. https://doi.org/10.1136/jech.57.6.424 PMID: 12775787

25. Guagliardo MF. Spatial accessibility of primary care: concepts, methods and challenges. International Journal of Health Geographics. 2004; 3(3). https://doi.org/10.1186/1476-072X-3-3 PMID: 14987337

26. Nelli L, Guelbeogo M, Ferguson HM, Ouattara D, Tiono A, N'Fale S, et al. Distance sampling for epidemiology: an interactive tool for estimating under-reporting of cases from clinic data. International journal of health geographics. 2020; 19:1–14. https://doi.org/10.1186/s12942-020-00209-1 PMID: 32312266

27. Hierink F, Okiro EA, Flahault A, Ray N. The winding road to health: A systematic scoping review on the effect of geographical accessibility to health care on infectious diseases in low-and middle-income countries. Plos one. 2021; 16(1):e0244921. https://doi.org/10.1371/journal.pone.0244921 PMID: 33395431

28. Ray N, Ebener S. AccessMod 3.0: computing geographic coverage and accessibility to health care services using anisotropic movement of patients. International journal of health geographics. 2008; 7 (1):63. https://doi.org/10.1186/1476-072X-7-63 PMID: 19087277

29. Wilson AL, Courtenay O, Kelly-Hope LA, Scott TW, Takken W, Torr SJ, et al. The importance of vector control for the control and elimination of vector-borne diseases. PLoS neglected tropical diseases. 2020; 14(1):e0007831. https://doi.org/10.1371/journal.pntd.0007831 PMID: 31945061

30. World Health Organization. Global vector control response 2017-2030. 2017.

31. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018; 392(10159):1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7

32. World Health Organization. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. World Health Organization; 2020.

33. Undurraga EA, Edillo FE, Erasmo JNV, Alera MTP, Yoon IK, Largo FM, et al. Disease burden of dengue in the Philippines: adjusting for underreporting by comparing active and passive dengue surveillance in Punta Princesa, Cebu City. The American journal of tropical medicine and hygiene. 2017; 96(4):887–898. https://doi.org/10.4269/ajtmh.16-0488 PMID: 28093542

34. Buczak AL, Baugher B, Babin SM, Ramac-Thomas LC, Guven E, Elbert Y, et al. Prediction of high incidence of dengue in the Philippines. PLoS Negl Trop Dis. 2014; 8(4):e2771. https://doi.org/10.1371/journal.pntd.0002771 PMID: 24722434

35. Bravo L, Roque VG, Brett J, Dizon R, L'Azou M. Epidemiology of dengue disease in the Philippines (2000–2011): a systematic literature review. PLoS Negl Trop Dis. 2014; 8(11):e3027. https://doi.org/10.1371/journal.pntd.0003027 PMID: 25375119

36. Shepard DS, Undurraga EA, Halasa YA, Stanaway JD. The global economic burden of dengue: a systematic analysis. The Lancet infectious diseases. 2016; 16(8):935–941. https://doi.org/10.1016/S1473-3099(16)00146-8 PMID: 27091092

37. Agrupis KA, Ylade M, Aldaba J, Lopez AL, Deen J. Trends in dengue research in the Philippines: A systematic review. PLoS neglected tropical diseases. 2019; 13(4):e0007280. https://doi.org/10.1371/journal.pntd.0007280 PMID: 31022175

38. Iguchi JA, Seposo XT, Honda Y. Meteorological factors affecting dengue incidence in Davao, Philippines. BMC public health. 2018; 18(1):1–10. https://doi.org/10.1186/s12889-018-5532-4 PMID: 29764403

39. de los Reyes AA, Escaner JML IV. Dengue in the Philippines: model and analysis of parameters affecting transmission. Journal of biological dynamics. 2018; 12(1):894–912. https://doi.org/10.1080/17513758.2018.1535096

40. Sumi A, Telan EFO, Chagan-Yasutan H, Piolo MB, Hattori T, Kobayashi N. Effect of temperature, relative humidity and rainfall on dengue fever and leptospirosis infections in Manila, the Philippines. Epidemiology & Infection. 2017; 145(1):78–86. https://doi.org/10.1017/S095026881600203X PMID: 27608858

41. Philippines statistics authority. Provincial summary: number of provinces, cities, municipalities and barangays by region, as of 30 June 2017. Accessed: 26 January 2021.

42. Philippine Statistics Authority and ICF. Philippines National Demographic and Health Survey 2017. 2018.

43. Republic of the Philippines Department of Health—Statistics.

44. Pham N, Nguyen C, Vu D, Nakamura K. Mapping of dengue vulnerability in the Mekong Delta region of Viet Nam using a water-associated disease index and remote sensing approach. APN Science Bulletin. 2018; 8(1). https://doi.org/10.30852/sb.2018.480

45. Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. elife. 2015; 4:e08347. https://doi.org/10.7554/eLife.08347 PMID: 26126267

46. Facebook Connectivity Lab and Center for International Earth Science Information Network—CIESIN—Columbia University. Philippines: High Resolution Population Density Maps + Demographic Estimates.

47. Guzmán M, Kouri G, Bravo J, Valdes L, Vázquez S, Halstead S. Effect of age on outcome of secondary dengue 2 infections. International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases. 2002; 6:118–24. https://doi.org/10.1016/S1201-9712(02)90072-X PMID: 12121599

48. Porterfield SL, McBride TD. The Effect of Poverty and Caregiver Education on Perceived Need and Access to Health Services Among Children With Special Health Care Needs. American Journal of Public Health. 2007; 97(2):323–329. https://doi.org/10.2105/AJPH.2004.055921 PMID: 17194872

49. Schmidt W-P, Suzuki M, Dinh Thiem V, White RG, Tsuzuki A, Yoshida L-M, et al. Population Density, Water Supply, and the Risk of Dengue Fever in Vietnam: Cohort Study and Spatial Analysis. PLOS Medicine. 2011; 8:1–10. https://doi.org/10.1371/journal.pmed.1001082 PMID: 21918642

50. Krystosik A, Njoroge G, Odhiambo L, Forsyth JE, Mutuku F, LaBeaud AD. Solid Wastes Provide Breeding Sites, Burrows, and Food for Biological Disease Vectors, and Urban Zoonotic Reservoirs: A Call to Action for Solutions-Based Research. Front Public Health. 2020; 7(405):564–81.

51. Banerjee S, Aditya G, Saha GK. Household Wastes as Larval Habitats of Dengue Vectors: Comparison between Urban and Rural Areas of Kolkata, India. PLoS One. 2015; 10(10):564–81. https://doi.org/10.1371/journal.pone.0138082

52. The Demographic and Health Surveys (DHS) Program.

53. Peters DH, Garg A, Bloom G, Walker DG, Brieger WR, Rahman MH. Poverty and access to health care in developing countries. Ann N Y Acad Sci. 2008; 1136:161–171. https://doi.org/10.1196/annals.1425.011 PMID: 17954679

54. Kelly C, Hulme C, Farragher T, Clarke G. Are differences in travel time or distance to healthcare for adults in global north countries associated with an impact on health outcomes? A systematic review. BMJ Open. 2016; 6(11). https://doi.org/10.1136/bmjopen-2016-013059

55. Martina BE, Koraka P, Osterhaus AD. Dengue virus pathogenesis: an integrated view. Clinical microbiology reviews. 2009; 22(4):564–81. https://doi.org/10.1128/CMR.00035-09 PMID: 19822889

56. Halstead SB, Nimmannitya S, Cohen SN. Observations related to pathogenesis of dengue hemorrhagic fever. IV. Relation of disease severity to antibody response and virus recovered. Yale J Biol Med. 1970; 42(5):311–28. PMID: 5419206

57. Philippines Statistics Authority. 2015 Census of Population. 2015.

58. Philippines Statistics Authority. OpenSTAT. 2015.

59. UN OCHA. https://data.humdata.org/dataset/philippines-pre-disaster-indicators. 2019.

60. UN OCHA. Philippines: Health workers by profession and geographical location. 2018.

61. Philippines Center for Investigative Journalism. Dengue, doctors, hospital beds: Ne'er the twain shall meet? 2019.

62. Buchhorn M, Smets B, Bertels L, Lesiv M, Tsendbazar NE, Herold M, et al. Copernicus Global Land Service: Land Cover 100m: Epoch 2015: Globe. Version V2 02. 2019.

63. Jarvis A. Hole-field seamless SRTM data, International Centre for Tropical Agriculture (CIAT). http://srtm.csi.cgiar.org. 2008.

64. Humanitarian OpenStreetMap Team (HOT). Philippines (north) Roads, accessed: 19 July 2020.

65. Humanitarian OpenStreetMap Team (HOT). Philippines (south) Roads, accessed: 19 July 2020.

66. Open Mapping at Facebook. AI-Assisted Road Tracing: Philippines, accessed: 19 July 2020.

67. OpenStreetMap Wiki. Key: Highway, accessed: 19 July 2020.

68. Geoportal Philippines. National Health Facility Registry (NHFR) of the Department of Health, accessed: 30 April 2020.

69. Buckland ST, Anderson DR, Burnham KP, Laake JLee, Borchers DL, Thomas L, et al. Introduction to distance sampling: estimating abundance of biological populations. 2001.

70. Edillo FE, Halasa YA, Largo FM, Erasmo JNV, Amoin NB, Alera MTP, et al. Economic cost and burden of dengue in the Philippines. Am J Trop Med Hyg. 360–366; 92(2). https://doi.org/10.4269/ajtmh.14-0139 PMID: 25510723

71. Republic of the Philippines Department of Health—Dengue.

72.   National Epidemiology Center—Philippines Department of Health. Manual of Procedures for the Philippine Integrated Disease Surveillance and Response. 2014.

73.   van Eijk AM, Sutton PL, Ramanathapuram L, Sullivan SA, Kanagaraj D, Priya SRL, et al. The burden of submicroscopic and asymptomatic malaria in India revealed from epidemiology studies at three varied transmission sites in India. Sci Rep. 2019; 9(17095). https://doi.org/10.1038/s41598-019-53386-w PMID: 31745160

74.   ten Bosch QA, Clapham HE, Lambrechts L, Duong V, Buchy P, Benjamin MA, et al. Contributions from the silent majority dominate dengue virus transmission. PLoS Pathogens. 2018. https://doi.org/10.1371/journal.ppat.1006965 PMID: 29723307

75.   Van Panhuis WG, Choisy M, Xiong X, Chok NS, Akarasewi P, Iamsirithaworn S, et al. Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. Proceedings of the National Academy of Sciences. 2015; 112(42):13069–13074. https://doi.org/10.1073/pnas.1501375112 PMID: 26438851

76.   Strachan AL. Conflict analysis of Muslim Mindanao. GSDRC, University of Birmingham. 2015.