

RESEARCH ARTICLE

The role of case proximity in transmission of visceral leishmaniasis in a highly endemic village in Bangladesh

Lloyd A. C. Chapman^{1,2,3*}, Chris P. Jewell⁴, Simon E. F. Spencer^{1,5}, Lorenzo Pellis¹, Samik Datta^{1,6}, Rajib Chowdhury^{7,8}, Caryn Bern⁹, Graham F. Medley³, T. Déirdre Hollingsworth^{1,2,10}

1 Zeeman Institute, University of Warwick, Coventry, UK, **2** School of Life Sciences, University of Warwick, Coventry, UK, **3** Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **4** Centre for Health Informatics, Computing And Statistics, Lancaster University, Lancaster, UK, **5** Department of Statistics, University of Warwick, Coventry, UK, **6** National Institute of Water and Atmospheric Research, Wellington, New Zealand, **7** National Institute of Preventive and Social Medicine (NIPSOM), Mohakhali, Dhaka, Bangladesh, **8** International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh, **9** Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA, **10** Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

* Lloyd.Chapman@lshtm.ac.uk



OPEN ACCESS

Citation: Chapman LAC, Jewell CP, Spencer SEF, Pellis L, Datta S, Chowdhury R, et al. (2018) The role of case proximity in transmission of visceral leishmaniasis in a highly endemic village in Bangladesh. *PLoS Negl Trop Dis* 12(10): e0006453. <https://doi.org/10.1371/journal.pntd.0006453>

Editor: David Poche, Genesis Laboratories, Inc., UNITED STATES

Received: January 24, 2018

Accepted: April 13, 2018

Published: October 8, 2018

Copyright: © 2018 Chapman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by the Bill and Melinda Gates Foundation in partnership with the Task Force for Global Health through the NTD Modelling Consortium [OPP1053230] (LACC, SEFS, GFM, TDH). The views, opinions, assumptions or any other information set out in this article are solely those of the authors and

Abstract

Background

Visceral leishmaniasis (VL) is characterised by a high degree of spatial clustering at all scales, and this feature remains even with successful control measures. VL is targeted for elimination as a public health problem in the Indian subcontinent by 2020, and incidence has been falling rapidly since 2011. Current control is based on early diagnosis and treatment of clinical cases, and blanket indoor residual spraying of insecticide (IRS) in endemic villages to kill the sandfly vectors. Spatially targeting active case detection and/or IRS to higher risk areas would greatly reduce costs of control, but its effectiveness as a control strategy is unknown. The effectiveness depends on two key unknowns: how quickly transmission risk decreases with distance from a VL case and how much asymptotically infected individuals contribute to transmission.

Methodology/Principal findings

To estimate these key parameters, a spatiotemporal transmission model for VL was developed and fitted to geo-located epidemiological data on 2494 individuals from a highly endemic village in Mymensingh, Bangladesh. A Bayesian inference framework that could account for the unknown infection times of the VL cases, and missing symptom onset and recovery times, was developed to perform the parameter estimation. The parameter estimates obtained suggest that, in a highly endemic setting, VL risk decreases relatively quickly with distance from a case—halving within 90m—and that VL cases contribute significantly more to transmission than asymptomatic individuals.

should not be attributed to the funders or any person connected with the funders. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusions/Significance

These results suggest that spatially-targeted interventions may be effective for limiting transmission. However, the extent to which spatial transmission patterns and the asymptomatic contribution vary with VL endemicity and over time is uncertain. In any event, interventions would need to be performed promptly and in a large radius ($\geq 300\text{m}$) around a new case to reduce transmission risk.

Author summary

Visceral leishmaniasis (VL), a fatal parasitic disease transmitted by sandflies, has been targeted for elimination as a public health problem in the Indian subcontinent by 2020. The goal has been reached in the majority of endemic regions in Bangladesh, India and Nepal, but the disease persists in several hotspots. Better understanding of spatial clustering of VL cases and the role of asymptotically infected individuals in transmission is required to improve control interventions and sustain the elimination target. To address this issue, we have fitted an individual-level spatiotemporal model of VL transmission to geolocated incidence data from Bangladesh to estimate the rate at which VL risk decreases with distance from a case and the potential contribution of asymptomatic individuals to transmission. Our results suggest that VL risk decreases quickly with distance and that symptomatic individuals are the main drivers of transmission, highlighting the potential for spatially-targeted control interventions to reduce transmission.

Introduction

Visceral leishmaniasis (VL), the world's second most lethal vector-borne parasitic disease, has been targeted for elimination as a public health problem in the Indian subcontinent (ISC) by 2020 [1]. The target is an incidence of less than 1 VL case/10,000 people/year at subdistrict level in Bangladesh and India and district level in Nepal. The reported annual number of cases in the ISC has decreased by approximately 80% since 2011 (from $\sim 37,000$ to $\sim 6,750$) [2], and as of 2017 the target had been reached in all subdistricts in Bangladesh, in approximately 85% of endemic subdistricts in India, and in all districts in Nepal for the previous 4 years [3–7]. Bangladesh has seen a more than 90% reduction in VL incidence since 2011 (from 2874 cases in 2011 to just 255 in 2016) [2], following a large epidemic wave lasting over a decade [8, 9]. The district that has consistently had the highest incidence is Mymensingh, within which Fulbaria (the source of the data analysed in this study) and Trishal have been the most endemic subdistricts and the last to reach the elimination target [9]. Despite the overall decline in incidence in the ISC, several subdistricts that have not had cases for a number of years have reported new cases in 2016 and 2017 [7, 10, 11], highlighting the need to maintain surveillance and control efforts as the target is approached and address gaps in understanding of VL transmission. Two key knowledge gaps are understanding of spatial variation in transmission and the role of asymptotically infected individuals (individuals infected with *Leishmania donovani* who do not develop clinical symptoms) in transmission [4, 12–14].

Asymptomatic individuals significantly outnumber clinical VL patients in the ISC, with estimates of the asymptomatic-to-symptomatic incidence rate ratio varying from 4:1 [15] to 17:1 [4, 16]. Parasites have been demonstrated in the blood smears of asymptomatic

individuals [17] and modelling has suggested that asymptomatic individuals could be the main drivers of transmission if they are infectious to sandflies [18–20]. Evidence of ongoing transmission, in the form of sporadic outbreaks, in regions in Nepal and Bhutan where there has historically been little symptomatic disease but asymptomatic individuals are present [21, 22] suggests that asymptomatic individuals can infect sandflies and contribute to transmission.

In recent xenodiagnosis experiments, asymptomatic individuals (defined by high anti-leishmanial antibody titres by the direct agglutination test (DAT) or rK39 enzyme-linked immunosorbent assay (ELISA) and no symptoms or prior history of VL) failed to infect sandflies fed upon them, whereas clinical VL cases and post kala-azar dermal leishmaniasis (PKDL) cases infected sandflies [23, 24]. However, only 48 asymptomatic individuals were tested and the experiments are ongoing.

Several studies have shown that proximity to past or current VL cases is a risk factor for infection and disease [15, 25–32] (see S1 Table). In addition to higher incidence of VL in households with past/current VL cases than households without VL, higher prevalence of positivity and incidence of conversion on diagnostic tests such as the rK39 ELISA, DAT, and leishmanin skin test (LST) have been observed in households with past/current VL cases, suggesting that transmission intensity is greater near VL cases. Picado et al [27] observed that incident VL was also associated with the presence of seropositive individuals in the same household at the baseline survey in their study (odds ratio (OR) for VL vs remaining DAT-negative = 2.43, 95% CI 1.55–3.79, $p < 0.001$) and the presence of other seroconvertors in the same household (OR = 2.22, 95% CI 1.38–3.58, $p = 0.001$) and within 50m of the household (OR = 8.73, 95% CI 2.59–29.41, $p < 0.001$) over the 2.5 years of the study. Incident asymptomatic DAT seroconversion was associated with presence of VL cases (OR = 1.66, 95% CI 1.16–2.36, $p = 0.005$), DAT-positive individuals (OR = 1.37, 95% CI 1.12–1.67, $p = 0.002$) and DAT seroconvertors (OR = 2.22, 95% CI 1.79–2.75, $p < 0.001$) in the same household. Incidence of asymptomatic infection and disease were more strongly associated with recent VL (i.e. VL that occurred in the relatively short time frame of the study) than VL that occurred in the 18 months before the study. These findings suggest that infection and transmission are highly spatially and temporally clustered.

Bern et al [15, 25] analysed spatial variation in risk of VL and asymptomatic infection in more detail using the dataset from Mymensingh, Bangladesh, analysed in this study. Associations between VL risk and asymptomatic infection risk (as measured by rK39 ELISA positivity or LST positivity) and distance from a VL case were assessed in terms of the odds ratios of having VL or being rK39 ELISA/LST-positive if living in the same household as, and a different household but <50m from, the closest previous VL case compared to living >50m from them. The risk of VL was substantially higher for individuals living in the same household as a previous case (OR = 6.37, 95% CI 3.30–12.28, $p < 0.0001$) and decreased relatively quickly with distance (OR = 1.85, 95% CI 0.95–3.60, $p = 0.07$, for individuals in a household without VL less than 50m from a case), while rK39 ELISA positivity and LST positivity were also strongly associated with proximity to previous VL cases. A multivariable logistic regression model comparing VL to recent asymptomatic infection (defined by rK39 ELISA positivity and LST negativity at baseline) showed that living in the same household as a previous case was associated with significantly increased risk of symptomatic infection (OR = 2.85, 95% CI 1.45–5.61, $p = 0.003$). Other studies have suggested that the risk of progression from infection to VL is higher for individuals living in households with past/current VL cases [28, 33, 34], and a study comparing transmission in households with recently treated VL cases against households with rK39-rapid-test-positive individuals or PKDL cases (or neither) found higher transmission (in terms of rK39-positivity) in VL households [35]. These studies lend support to the view that VL cases are the primary drivers of transmission, at least in highly endemic villages.

As VL incidence declines in the ISC, understanding spatial heterogeneity in transmission is particularly important for optimising control interventions. The main interventions currently employed are early case detection and treatment, and indoor residual spraying of insecticide (IRS) to try to reduce the density of the *Phlebotomus argentipes* sandfly vectors [6]. In endemic situations, the WHO recommendation is for IRS to be applied in affected villages and timed to precede maximum sandfly density, while during epidemics large-scale IRS covering all buildings, including houses and animal shelters is recommended [36]. An important question currently facing the control programme is whether more effort should be invested in improving case detection or IRS to reach and sustain the elimination target [7, 12, 20, 37]. This depends on the effectiveness of IRS, but also on the range of transmission around infected individuals and how much asymptomatic individuals contribute to transmission.

Spatially-targeted active case detection and IRS strategies are being considered as part of the national elimination programmes in India and Bangladesh due to resource constraints (the economic and practical infeasibility of performing active case detection and blanket spraying insecticide in all villages) and falling case numbers [3, 6, 9, 38, 39]. These strategies involve checking for infection and VL, and spraying insecticide, in all houses within a certain radius of a new VL case to try to limit transmission. Spatially targeted active case detection strategies have been tested in Bangladesh, India and Nepal with varying levels of success, potentially due to protocol differences (e.g. differences in the definition of ‘index’ cases, from any previous cases to only active cases, and the radii around index cases in which case detection was performed, from 50m to 200m), different levels of incidence, and implementation issues [40–42]. Currently, spatial targeting of IRS is being trialled only at village level in the Vaishali and Saran districts of Bihar state, India, with adjusted IRS village micro-plans in which all villages neighbouring (less than 500m from) a village with cases in the previous year are sprayed (based on an estimated maximum sandfly flight range of 500m [43–45]), along with all villages that have had any cases in the previous 3 years (as in the previous micro-plans) [38]. More precise estimates of the spatial and temporal range of transmission via sandfly dispersal are required to target IRS at a household level.

Although the epidemiological studies described above provide evidence of the importance of case proximity in transmission, they all use relatively simple statistical analyses with crude measures of spatial and temporal proximity to assess transmission patterns. In order to accurately account for variation in the infectious pressure on individuals in both space and time, a more fine-grained spatiotemporal model is required. Existing transmission models of VL [12, 18–20, 46, 47] fail to account for spatial heterogeneity in transmission, and as a result may overestimate the contribution of asymptomatic individuals to transmission [48]. Two key questions that need to be addressed are:

1. how does VL risk decrease with distance from a case?
2. how much do asymptomatic individuals contribute to transmission?

One approach that has been taken to estimate spatial variation in disease risk is using spatial kernel transmission models [49], in which the transmission rate between individuals is scaled by a function of the distance between them (the spatial kernel). Methods for inferring spatial kernels from geo-located incidence data were first developed for foot-and-mouth disease in livestock [50, 51], and have since been extended to handle a number of complexities, including missing data and unobserved infections [52–55], and been applied to other livestock diseases such as avian influenza and swine flu [56–58]. However, they have rarely been applied to human diseases [59] or vector-borne diseases [60] due to a number of challenges, including limited information on human movement [49]. Inference of spatial transmission of VL is

particularly challenging due to long infectious periods, the long and highly variable incubation period (reported to last anywhere between 10 days and 2 years, though generally thought to be 2–6 months [25, 61], with an estimated average duration of 5 months (95% CI 4.3–5.5 months) based on diagnostic data [62]), the potential role of asymptomatic individuals in transmission, and the sparsity of data on the flight range of *P. argentipes* sandflies.

To start to address the questions above we have developed an individual-level spatial kernel transmission model for VL and fitted it to a detailed epidemiological dataset from a highly endemic setting in Bangladesh [15, 25] using a Bayesian inference framework. Our aim in developing and parameterising the model is to eventually use it to predict the impact of spatially-targeted control interventions in the ISC, once it has been validated against data from different settings.

Methods

Data

The data used in this study were collected in a longitudinal epidemiological study in a highly endemic village in Fulbaria upazila, Mymensingh district, Bangladesh between January 2002 and June 2004. Full details of the study design and data collection have been provided previously [15, 25, 62], but aspects particularly relevant to the present study are briefly described here. The data consist of detailed information on 2507 individuals living in 509 households in the 3 hamlets (or ‘paras’), out of the 9 in the village, that had the highest reported VL incidence in the years before the study. The 3 study paras are situated in an area approximately 2.7km × 2km in size. Longitudinal VL incidence data—including dates of onset of symptoms, treatment, relapse, relapse treatment and death due to VL—was collected for 183 VL cases with symptom onset between January 1999 and June 2004 (retrospectively for those with onset before 2002), and onset years were recorded for 41 individuals who had VL prior to 1999 (see S1 and S3 Data). Annual censuses to record demographic information (births and deaths) and cross-sectional diagnostic surveys using rK39 ELISA and LST were conducted on the entire study population from January–April in 2002–2004 (S1 Data). All households present in 2002 were mapped with a Global Positioning System (GPS) device accurate to ±10m, and the GPS positions of households built in the study area after 2002 were imputed as being halfway between those of the two closest households (when these were known). Thus GPS coordinates were available for 506 of the 509 households (Fig 1). All analyses were restricted to the 2494 inhabitants of these households. Individuals were assigned GPS coordinates according to the household they belonged to and pairwise distances between all individuals calculated using the haversine formula [63].

Model

We describe VL transmission using an individual-level spatiotemporal susceptible-exposed-infectious-recovered (SEIR) model. In the model, individuals can be in one of four states at any particular time:

- susceptible (i.e. uninfected and without any history of VL, and able to become infected),
- exposed or pre-symptomatic (infected with the parasite and potentially infectious but without any symptoms),
- symptomatic (i.e. with clinical symptoms of VL and infectious),
- recovered (i.e. treated for VL or recovered from asymptomatic infection);

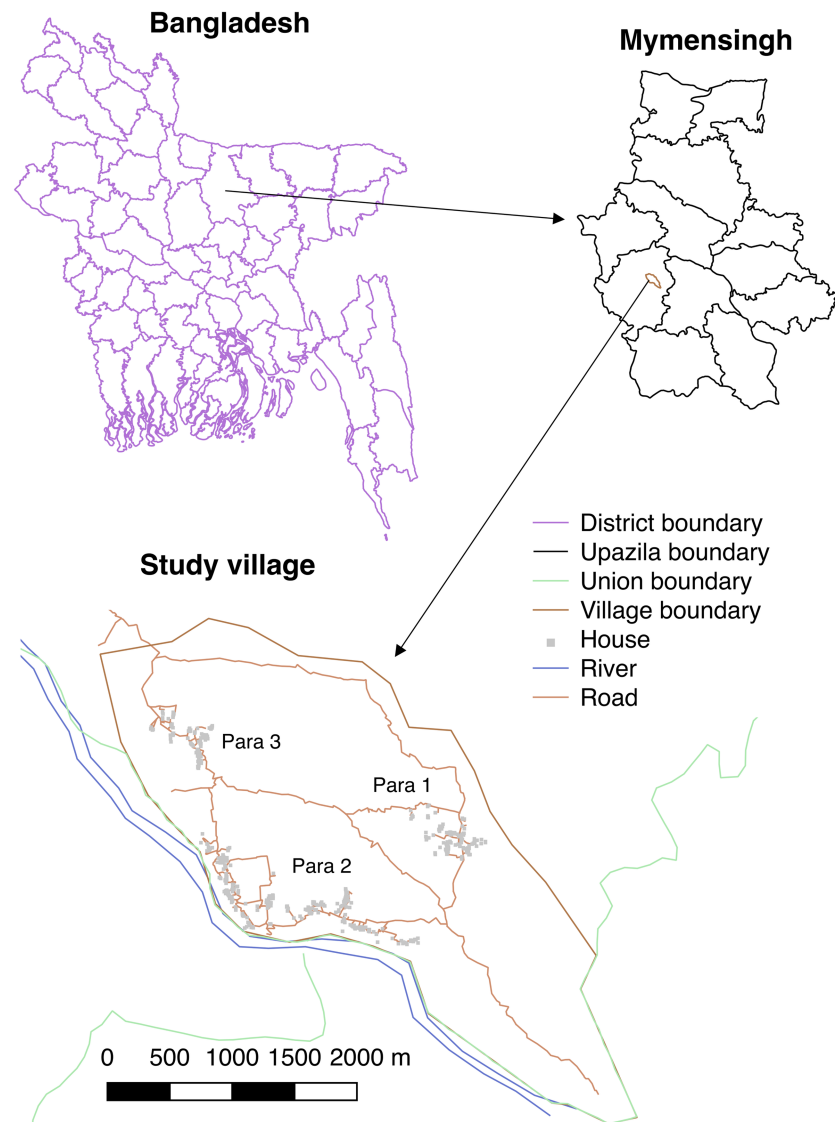


Fig 1. Location of study village and surveyed households in Fulbaria upazila, Mymensingh district, Bangladesh. District and upazila shape files from www.gadm.org/country, village details from GPS mapping during study. Maps produced using QGIS version 2.14.2 [64].

<https://doi.org/10.1371/journal.pntd.0006453.g001>

and can only progress from being susceptible to pre-symptomatically infected to symptomatic to recovered, and back to symptomatic if they suffer a relapse (Fig 2). Asymptomatically infected individuals are not explicitly modelled (see below for how transmission due to asymptomatic infection is implicitly incorporated). Note also that recovered individuals who do not relapse are assumed to possess lifelong immunity to VL. This is different to several previous transmission models of VL [18–20], which assume that individuals who have recovered from symptomatic or asymptomatic infection return to being susceptible after a number of years, and thus can be reinfected. Whilst this may be true for asymptomatic infection, multiple VL episodes are relatively rare [25] and the limited available evidence suggests that the majority of these are due to relapse rather than reinfection [65, 66]. PKDL cases are also not included,

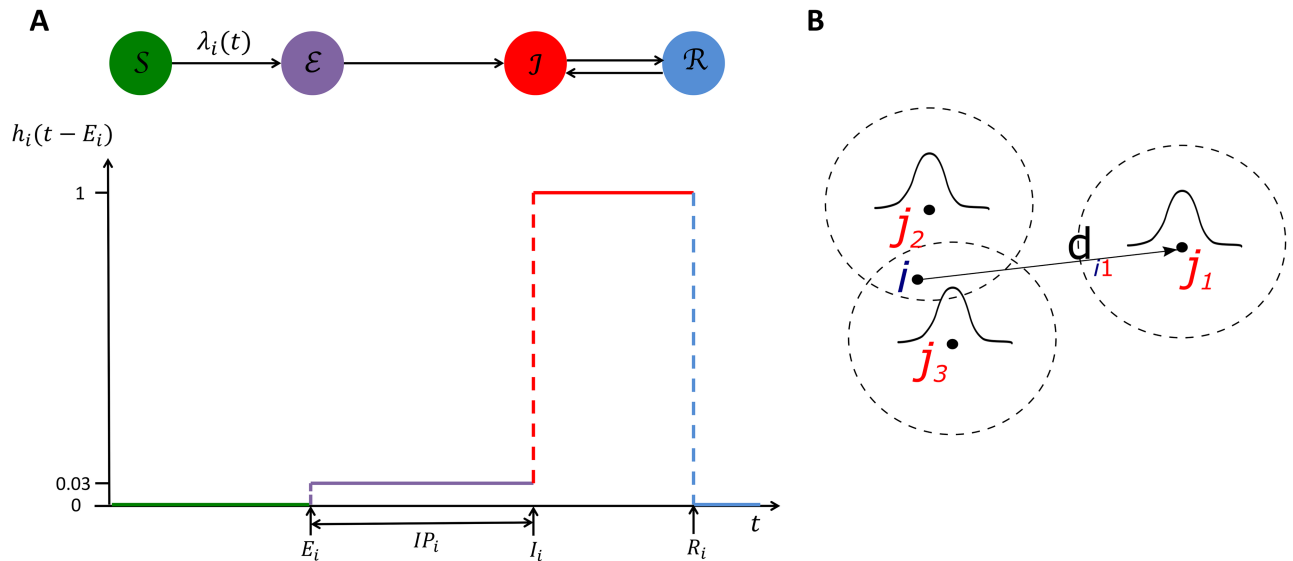


Fig 2. Schematic of the model structure. A: The different infection states in the model (susceptible \mathcal{S} , pre-symptomatically infected \mathcal{E} , symptomatically infected \mathcal{I} , treated \mathcal{R}) and their relative infectiousness $h_i(t - E_i)$ over time t for a VL case i with infection time E_i , incubation period IP_i , symptom onset time I_i and treatment time R_i . B: The infectious pressure $\lambda_i(t)$ on susceptible individual i at time t is the sum of the transmission rates from the infectious individuals around them (here j_1, j_2 and j_3), which are a function (represented by the bell-shaped curves) of the distances of the infectious individuals from i (the arrow shows the distance d_{i1} between i and j_1) and the times since their infections.

<https://doi.org/10.1371/journal.pntd.0006453.g002>

since there were only 4 confirmed cases of PKDL during the study period and no information was recorded on time of PKDL onset.

Time is measured in units of months (running from $t = 1$ (January 1998) to $t = T = 78$ (June 2004)) since this is a natural time scale for VL progression (the durations of infection and disease typically lasting for months) and is the finest time resolution to which some dates (such as birth and death) were recorded. We label individuals that developed VL during the study by $i = 1, 2, \dots, n_I$, and the remainder of the study population by $i = n_I + 1, n_I + 2, \dots, n$, where $n = 2494$ is the total study population. The vectors of birth, infection, onset, treatment, relapse, relapse treatment, and death times of all individuals are denoted by $\mathbf{B} = (B_i)_{i=1, \dots, n}$, $\mathbf{E} = (E_i)_{i=1, \dots, n}$, $\mathbf{I} = (I_i)_{i=1, \dots, n}$, $\mathbf{R} = (R_i)_{i=1, \dots, n}$, $\mathbf{I}_R = (I_i^R)_{i=1, \dots, n}$, $\mathbf{R}_R = (R_i^R)_{i=1, \dots, n}$, and $\mathbf{D} = (D_i)_{i=1, \dots, n}$. We define the sets of susceptible, pre-symptomatic, symptomatic and recovered individuals at time t by

$$\begin{aligned} \mathcal{S}(t) &:= \{i : B_i \leq t < \min(E_i, D_i)\}, \\ \mathcal{E}(t) &:= \{i : E_i \leq t < I_i\}, \\ \mathcal{I}(t) &:= \{i : I_i \leq t < \min(R_i, D_i) \text{ or } I_i^R \leq t < R_i^R\}, \quad \text{and} \\ \mathcal{R}(t) &:= \{i : R_i \leq t < \min(D_i, I_i^R) \text{ or } R_i^R \leq t < D_i\}, \quad i = 1, \dots, n, \end{aligned}$$

respectively, where we adopt the notation $E_i = I_i = R_i = I_i^R = R_i^R = \infty$ if individual i did not have VL during the study.

Transmission rates. In the model, susceptible individuals can become infected either from pre-symptomatic individuals (in \mathcal{E}) (if they are assumed to be infectious), from symptomatic individuals (in \mathcal{I}), or from ‘background’ transmission due to non-explicitly modelled factors, such as asymptomatic infection, human migration, and animal reservoirs (if any exist). It is commonly believed that transmission is anthroponotic in the ISC, i.e. that humans are the only reservoir of infection [4, 6]. Some studies have found animals positive for *L. donovani*

DNA or antibodies against the parasite [67–70], but none have demonstrated that these animals are capable of infecting sandflies or that they play any role in transmission. The time-dependent transmission rate $\lambda_{ij}(t)$ between an infected individual j and a susceptible individual i is modelled as

$$\lambda_{ij}(t) = (\beta K(d_{ij}) + \delta \mathbf{1}_{ij}) h_j(t - E_j), \quad i \in \mathcal{S}(t), j \in \mathcal{E}(t) \cup \mathcal{I}(t), \tag{1}$$

where β is the overall rate constant for spatial transmission between infected and susceptible individuals; $K(d_{ij})$ is the spatial kernel function that scales the transmission rate by the distance d_{ij} between individuals i and j (S2 Data); $\delta (\geq 0)$ is a rate constant for additional within-household transmission; $\mathbf{1}_{ij}$ is an indicator function for individuals living in the same household, i.e.

$$\mathbf{1}_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are in the same household,} \\ 0, & \text{otherwise;} \end{cases}$$

and $h_j(t - E_j)$ is the infectiousness of the infected individual at time t . Thus, transmission from person to person through sandflies and spatial variation in this transmission due to sandfly movement is not modelled explicitly, but incorporated implicitly through the spatial kernel function. We note that this description relies on the assumption that humans are generally in fixed positions relative to each other when transmission to and from sandflies occurs. We assume this to be the case as the majority of sandfly biting occurs at night, typically between the hours of 9pm and 3am [71–76], when individuals are generally asleep inside or just outside their homes.

Spatial kernel. Given the limited flight range of the sandfly and evidence from previous studies that VL risk is high close to past/active cases and lower further away from them, it is reasonable to suppose that VL risk decreases with increasing distance from a case. We test this hypothesis by fitting the following different functions for the spatial kernel K :

- a constant kernel

$$K_1(d_{ij}) = \begin{cases} 1, & \text{if } i \text{ and } j \text{ in same para,} \\ 0, & \text{if } i \text{ and } j \text{ in different paras,} \end{cases} \tag{2}$$

- a Cauchy-type kernel

$$K_2(d_{ij}) = \begin{cases} \frac{K_0}{1+(d_{ij}/\alpha)^2}, & \text{if } i \text{ and } j \text{ in same para,} \\ 0, & \text{if } i \text{ and } j \text{ in different paras,} \end{cases} \tag{3}$$

- an exponential kernel

$$K_3(d_{ij}) = \begin{cases} K_0 \exp(-d_{ij}/\alpha), & \text{if } i \text{ and } j \text{ in same para,} \\ 0, & \text{if } i \text{ and } j \text{ in different paras,} \end{cases} \tag{4}$$

where α is the parameter which determines the rate at which transmission risk decreases with distance (smaller values corresponding to a more rapid decrease); and K_0 , defined such

that

$$\sum_{i=1}^n \sum_{j=1, j \neq i}^n K(d_{ij}) = n,$$

is a normalisation constant introduced to reduce correlation between α and β that negatively affects the mixing of the Markov chain Monte Carlo (MCMC) algorithm used to estimate the parameters (see Parameter Estimation). We choose the Cauchy-type and exponential kernels as they are commonly used kernels for modelling disease spread [53, 54, 58, 59, 77], and are monotonically decreasing functions of distance parameterised by a single parameter, so do not lead to parameter identifiability issues. We assume that (in the absence of human migration) cases can only infect other individuals in the same para based on the reported maximum flight range of *P. argentipes* sandflies in the Indian subcontinent being a few hundred metres [78–81], and the minimum distance between houses in different paras being 580m. Note that potential long-range transmissions between cases in different paras that might be missed by restricting the kernel in this way are covered by the inclusion of background transmission in the model (see below).

To test the null hypothesis that there is no spatial or temporal variation in transmission associated with VL cases, we compare the fit of the spatial kernel models defined above against a model with background transmission only (i.e. with $\beta = 0$ and $\epsilon > 0$). The constant spatial kernel corresponds to the transmission rate in each para depending only on the number of pre-symptomatic and symptomatic individuals in the para, rather than on the positions of these individuals. Thus, testing the constant kernel against the Cauchy and exponential kernels enables an assessment of the evidence for individual-level spatial variation in transmission risk, as opposed to only para-level variation. The Cauchy kernel is a bell-shaped kernel, for which the risk is higher than the exponential kernel over short distances and also over long distances (due to its fatter tail), so we test both kernels to determine whether one provides a more accurate description of the variation in risk around a case. We also compare models with and without additional within-household transmission ($\delta > 0$ and $\delta = 0$) against each other to assess whether living in the same household as a case confers additional risk of disease, over and above that just due to distance from the case.

Infectiousness over time. The choice of the infectiousness function $h_j(t - E_j)$ warrants discussion since very little is known about how infectiousness towards sandflies changes as infection progresses or how it varies between individuals for anthroponotic VL [82]. For canine VL, infectiousness is strongly associated with skin parasite load, which increases with the duration and severity of infection, and with clinical severity, but there is substantial variation between individual dogs [83]. In humans, the picture may be similar—disease severity generally increases with the duration of symptoms, though there is again substantial variation between individuals, and preliminary xenodiagnosis studies suggest that infectiousness is correlated with disease severity (as measured by parasite loads in splenic smears) [23]. For the infectiousness of pre-symptomatic individuals, only proxy measures, such as parasite loads in peripheral blood and antibody titres, are available. Parasite loads and antibody titres appear to be higher in pre-symptomatic individuals than asymptomatic individuals, but still several orders of magnitude lower than those in clinical cases [16, 84, 85]. Although modelling studies have not generally distinguished pre-symptomatic and asymptomatic infection, they suggest that pre-symptomatic individuals are much less infectious to sandflies than symptomatic individuals [18, 20, 48]. The infectiousness of VL cases who suffer relapse is also not known, but relapse appears to be associated with a resurgence in parasitaemia to high parasite loads [86, 87]. Whether these individuals are infectious between apparent clinical cure and recurrence of

symptoms is unknown. Thus, in the absence of more detailed experimental data, we choose a simple step function for $h_j(t - E_j)$ of the form

$$h_j(t - E_j) = \begin{cases} h_0, & \text{if } E_j \leq t < I_j, \\ 1, & \text{if } I_j \leq t < R_j \text{ or } I_j^R \leq t < R_j^R, \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where h_0 is the relative infectiousness of pre-symptomatic individuals compared to symptomatic individuals, cases are assumed to become uninfected in their treatment month R_j (i.e. shortly after commencing treatment, based on observed rapid decreases in parasite loads with treatment [86, 88]), and we assume that relapse cases are as infectious upon relapse as in their first clinical episode, but uninfected between treatment and relapse. Based on modelling estimates for the relative infectiousness of pre-symptomatic/asymptomatic individuals [19, 20] we choose $h_0 = 0.03$, but also perform a sensitivity analysis for h_0 to assess how it affects the parameter estimates (see S1 Text).

Incubation periods. Since the infection times of cases were not observed and the incubation period of VL is variable, the incubation period durations of the cases, $IP_j = I_j - E_j$ ($j = 1, \dots, n_I$), are random variables. We model them as independent negative binomial random variables with fixed shape parameter $r = 3$ and ‘success’ probability parameter p (and support starting from 1 rather than 0, such that the minimum incubation period is 1 month), i.e. $IP_j \sim NB(r, p)$, where p determines the mean incubation period (since $E[NB(r, p)] = r(1 - p)/p + 1$) and is estimated along with the transmission parameters β, α, ϵ and δ (see Parameter Estimation and S1 Text for further details). We choose this form for the incubation period distribution as it implies that very short or very long incubation periods are unlikely, but there is a certain intermediate incubation period duration (the mode of the negative binomial distribution) that cases are likely to have. This is more realistic than a geometric incubation period distribution (with $r = 1$), which would imply that the most likely infection time is the month before symptom onset (since it always has a mode at 1).

Infectious pressure. The ‘infectious pressure’ on susceptible i at time t is given by the sum of the transmission rates from all infected individuals at time t (see Fig 2) plus a constant background transmission rate ϵ

$$\lambda_i(t) = \sum_{j \in \mathcal{E}(t) \cup \mathcal{I}(t)} \left((\beta K(d_{ij}) + \delta \mathbf{1}_{ij}) h_j(t - E_j) \right) + \epsilon, \quad (6)$$

where the background transmission rate is included to account for potential transmission from asymptomatic individuals (and from other factors, such as migration, that are not explicitly modelled). The transmission process is thus described by a discrete-time approximation to an inhomogeneous Poisson process with rate $\lambda_i(t) = \sum_{j \in \mathcal{S}(t)} \lambda_j(t)$, such that the probability of susceptible individual i being infected in any particular month t is:

$$p_i(t) = 1 - e^{-\lambda_i(t)}, \quad i \in \mathcal{S}(t). \quad (7)$$

Using this expression, and individuals’ imputed infection times if they had VL (see Parameter Estimation), the probability of each individual remaining susceptible or being infected in each month $t \in \{1, \dots, T\}$ can be calculated (as $q_i(t) = 1 - p_i(t)$ or $p_i(t)$ ($i = 1, \dots, n$) respectively), these monthly probabilities (and the probability of the individual’s incubation period duration if they had VL) multiplied together to give the likelihood for each individual’s infection course, and the likelihoods for all individuals multiplied together to give the overall likelihood of the data given the model (see S1 Text for full details and the likelihood expression).

Leishmanin skin test. The above model does not take into account the fact that some individuals develop immunity following asymptomatic infection. The LST is a test for previous *L. donovani* exposure and LST positivity is believed to represent durable cell-mediated immunity against the parasite, which may last for several years or for life in the absence of immunosuppression [32, 89]. The LST was applied and read at each of the annual cross-sectional surveys from 2002–2004 as described previously [62, 90]. Many individuals without past or subsequent VL were LST-positive at the first survey in 2002, and others who were seropositive and LST-negative at the first survey later became LST-positive without developing VL. It seems likely that these individuals possessed protective immunity against VL, at least for the period of the study, given that only 2 of 667 individuals without previous VL who were LST-positive at one of the surveys subsequently developed VL compared to 43 out of 1163 without previous VL who were LST-negative (relative risk of subsequent VL for LST-positives vs LST-negatives from 2002 survey = 0.05, 95% CI 0.006–0.33, $p = 0.02$).

To assess the impact on the parameter estimates of treating LST-positive individuals as immune, we incorporated the LST status data from the cross-sectional surveys into the model and re-estimated the parameters. The LST data was incorporated using a simple catalytic model for the age-prevalence distribution of asymptomatic LST-positive individuals, in which it is assumed that a certain proportion of infected individuals never become LST-positive (either because they do not develop cell-mediated immunity or because of poor sensitivity of the test); that the probability of being LST-positive otherwise only depends on age and increases monotonically with age; and that individuals who become LST-positive remain immune to VL for life. This is clearly a considerable simplification, since some studies have shown loss of LST-positivity over time [90–92] and it is uncertain whether this reflects loss of immunity or poor test sensitivity or both, and LST prevalence also varies in space and time as the force of infection changes, e.g. LST positivity is strongly associated with proximity to past VL cases [15, 32]. However, our aim is just to gauge whether immunity among asymptomatic LST-positive individuals affects estimates of the transmission rates and spatial kernel, rather than to provide an accurate model for LST conversion. Furthermore, the assumption of life-long immunity following LST conversion, as opposed to only temporary immunity, should not significantly affect the parameter estimates, as it is likely that immunity lasted much longer than the study even if it waned (see [S1 Text](#)). Further details of the catalytic model and incorporation of the LST data are provided in [S1 Text](#).

Parameter estimation

Bayesian inference framework. We developed a Bayesian inference framework for estimating the model parameters from the data based on inference techniques for partially observed stochastic epidemics originally developed by O’Neill and Roberts [93] and since extended by others [52–54, 77, 94–97]. Variations of this framework have been used to analyse the spatial transmission of diseases such as measles, foot-and-mouth disease, avian influenza and American foulbrood in honeybees [53, 54, 95, 98–100].

Although the times of symptom onset and treatment, \mathbf{I} and \mathbf{R} , are known for most of the VL cases in the study, for some cases only their year of onset is known, and the infection times \mathbf{E} of all the cases are unknown. Due to the long and variable incubation period of VL, it is necessary to impute this missing data as part of the algorithm for estimating the transmission parameters. A Bayesian approach provides a natural framework in which to do this, since model parameters are viewed as random variables and the unknown infection times and missing onset and recovery times (which we denote by \mathbf{I}' and \mathbf{R}' respectively) can be treated as extra parameters to be estimated via data augmentation [77]. With the model as

described above, we specify the likelihood $L(\theta; \mathbf{A}) := \mathbb{P}(\mathbf{A}|\theta)$ of the augmented or complete data $\mathbf{A} = (\mathbf{U}, \mathbf{Y})$, consisting of the observed data $\mathbf{Y} = (\tilde{\mathbf{I}}, \tilde{\mathbf{R}}, \mathbf{I}_R, \mathbf{R}_R)$ (with the observed onset and treatment times $\tilde{\mathbf{I}}$ and $\tilde{\mathbf{R}}$) and the unobserved data $\mathbf{U} = (\mathbf{E}, \mathbf{I}', \mathbf{R}')$, given the transmission parameters and the incubation period distribution parameter, $\theta = (\beta, \alpha, \epsilon, \delta, p)$ (see [S1 Text](#)). Using Bayes' rule, this is then combined with prior distributions for the model parameters, $\mathbb{P}(\theta)$, representing any prior information we have about their values, to give the joint posterior distribution $\mathbb{P}(\theta, \mathbf{U}|\mathbf{Y})$ of the parameters and unobserved data given the observed data:

$$\mathbb{P}(\theta, \mathbf{U}|\mathbf{Y}) \propto \mathbb{P}(\mathbf{U}, \mathbf{Y}|\theta)\mathbb{P}(\theta) = L(\theta; \mathbf{A})\mathbb{P}(\theta). \tag{8}$$

In theory we would like to perform inference for θ by calculating the posterior distribution for θ given the observed data

$$\mathbb{P}(\theta|\mathbf{Y}) \propto \mathbb{P}(\mathbf{Y}|\theta)\mathbb{P}(\theta). \tag{9}$$

However, this would require integrating the likelihood of the observed data $\mathbb{P}(\mathbf{Y}|\theta)$ over the set of all possible values of the unobserved data, Ω (which is finite),

$$\mathbb{P}(\mathbf{Y}|\theta) = \sum_{\mathbf{U} \in \Omega} \mathbb{P}(\mathbf{Y}, \mathbf{U}|\theta), \tag{10}$$

and it is not computationally feasible to calculate this sum due to the very high dimensionality of Ω . Instead, we use a MCMC data augmentation scheme to sample from the joint posterior distribution for (θ, \mathbf{U}) in [Eq \(8\)](#) by iteratively sampling from the conditional distribution of the parameters given the observed and missing data, $\mathbb{P}(\theta|\mathbf{U}, \mathbf{Y})$, and the conditional distribution of the missing data given the parameters and the observed data, $\mathbb{P}(\mathbf{U}|\mathbf{Y}, \theta)$. Full details of the MCMC algorithm are given in [S1 Text](#).

Prior distributions. To perform Bayesian inference we need to specify the prior distributions for the parameters $\theta = (\beta, \alpha, \epsilon, \delta, p)$ ([Table 1](#)). As there was little information available with which to construct priors for the transmission rate parameters β, ϵ and δ , which are non-negative, we chose weak exponential priors for all three. For the spatial kernel parameter α we chose an exponential prior with mean 50m, $\text{Exp}(1/50)$, based on Bern et al's analyses of the same dataset [[15](#), [25](#)]. Since the incubation period parameter $p \in (0, 1)$ is a probability, we chose a beta distribution as a conjugate prior for p . The parameters of the beta distribution were chosen to match the mean of the prior distribution for the incubation period ($\text{NB}(r, p)$) with the estimated mean incubation period from our previous analysis of the diagnostic data from the same study (~ 5 months) [[62](#)].

Model comparison

We compared the goodness of fit of the background-transmission-only model and the different spatial kernel models defined by [Eqs \(2\)–\(4\)](#), with and without additional within-household transmission, using the Deviance Information Criterion (DIC) [[101](#)]. DIC measures the

Table 1. Definitions and prior distributions of estimated parameters.

Parameter (units)	Definition	Prior distribution
β (mnth ⁻¹)	Rate constant for spatial transmission from VL cases	Exp(1)
α (m)	Inverse distance decay rate of spatial kernel	Exp(1/50)
ϵ (mnth ⁻¹)	Background transmission rate	Exp(1)
δ (mnth ⁻¹)	Additional within-household transmission rate	Exp(1)
p	Incubation period distribution parameter	Beta(22, 28)

<https://doi.org/10.1371/journal.pntd.0006453.t001>

trade-off between model fit and model complexity, and is the Bayesian equivalent of the Akaike Information Criterion (AIC) [102]. Since some data is unobserved, we required a modified version of the standard DIC that takes into account the missing data. We used a version of DIC for missing data models proposed by Celeux et al [103] that is based on the complete data likelihood $L(\theta; \mathbf{A})$ and calculated using the modes of the posterior distributions for the parameters (see S1 Text for further details). As for AIC, lower DIC values indicate better model performance, though differences of less than 5 units between models do not suggest a substantial difference in goodness-of-fit [104]. Given that there are potential issues with using DIC for model comparison [105], we also compared the posterior distributions of the deviances of the models to assess differences in quality of fit, where the deviance is defined up to an additive constant as

$$D(\theta) = -2 \log(L(\theta; \mathbf{A})). \tag{11}$$

Results

Demographic and VL incidence data for the three paras are presented in Table 2 and Fig 3. Average household size was relatively consistent across the three paras, but there were considerable differences in VL incidence and the average number of VL cases per household with VL. All three paras had high incidence from January 1999 to June 2004, but para 1 had the highest average incidence (248/10,000/yr) and average number of cases per VL household (1.65), and para 2, the largest para, the lowest (70/10,000/yr and 1.10 respectively). Most of the difference in incidence was due to much higher incidence in para 1 between 1999 and 2002, prior to the start of the study (Fig 3A). The para-level incidences are similar to those that were observed during 1-3 year ‘micro-epidemics’ in highly endemic clusters of hamlets in the data from Muzaffarpur, Bihar State, India, analysed by Bulstra et al [106]. Forty-one households, 29 of which were in paras 1 and 3, had multiple cases between 1999 and 2004 (Fig 3B). Paras 1 and 3 also had greater maximum numbers of cases per household: all 3 households that had 4 cases were in para 1 in close proximity to each other and several households that had 3 and 2 cases, and there were 3 households that had 3 cases in para 3, but no household in para 2 had more than 2 cases. The evolution of the spatial pattern of VL cases in each para (Fig 4 and S1 and S2 Figs) from 1999-2004 suggests highly focal transmission around VL cases, with new cases in each year generally appearing very close to VL cases with onset in the previous year. Fig 5 shows the distributions of the distances between all pairs of individuals and all pairs of VL cases for each of the paras. The differences between the distributions for each para, e.g. the higher density in the case distribution at short distances (<100m) and the clumping of the case distribution, reflect the high degree of spatial clustering of cases.

Parameter estimates

Fig 6 shows the output of the MCMC algorithm for the exponential kernel model, including the log-likelihood trace plot and the posterior distributions for the transmission parameters, β ,

Table 2. Demographic and VL incidence data for the three study paras for January 1999–June 2004.

Para	No. of HH (%)	Population in 2002	No. of VL cases	No. of VL HHs (%)	No. of cases per VL HH	Mean VL incidence (cases/10,000/yr)
1	126 (24.9)	590	86	52 (40.3)	1.65	248
2	272 (53.8)	1289	53	48 (37.2)	1.10	70.4
3	108 (21.3)	546	44	29 (22.5)	1.52	135
Total	506 (100)	2425	183	129 (100)	1.42	129

HH = household, VL HH = household with ≥ 1 VL case.

<https://doi.org/10.1371/journal.pntd.0006453.t002>

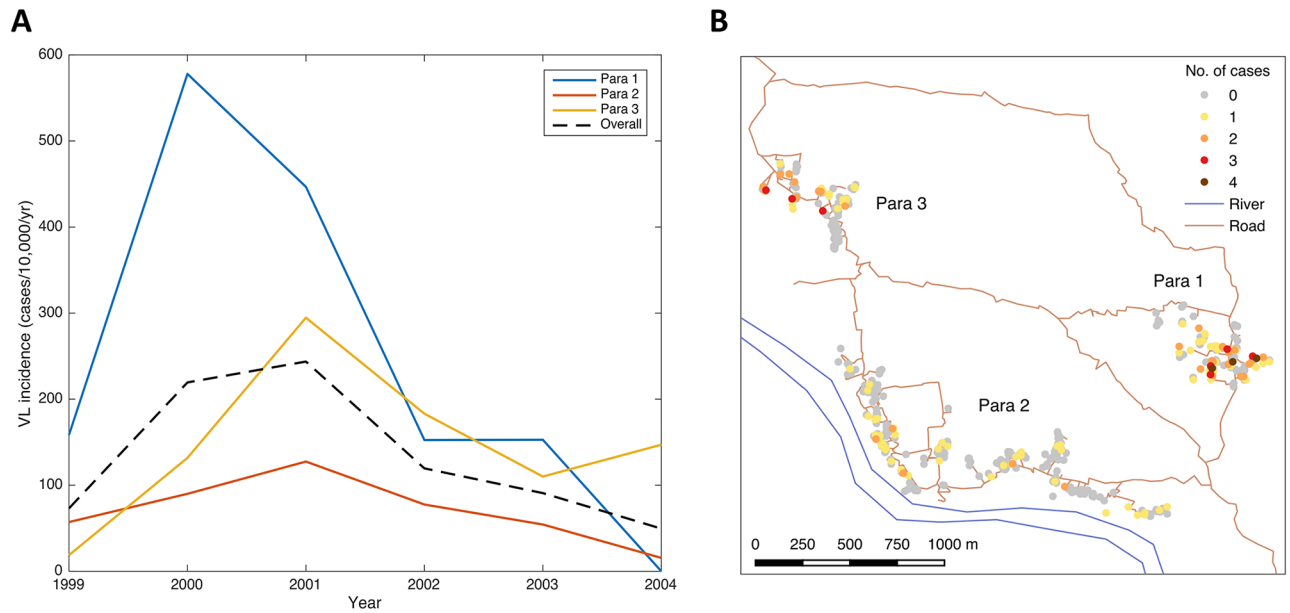


Fig 3. Para- and household-level VL incidence, January 1999–June 2004. A: Annual incidence in each para. B: Total number of VL cases in each household.

<https://doi.org/10.1371/journal.pntd.0006453.g003>

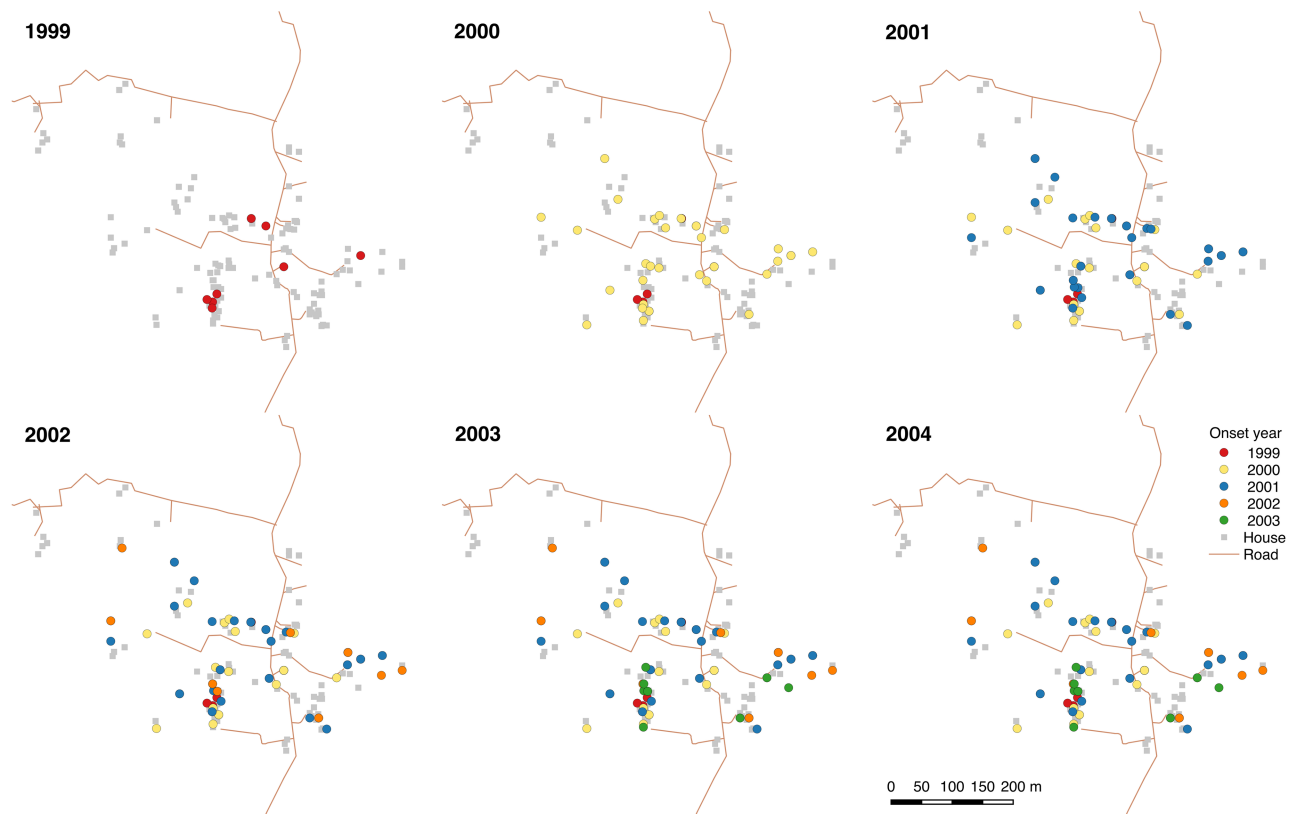


Fig 4. Locations of VL cases in para 1 by year of onset, 1999-2004. There were no VL cases with symptom onset in para 1 from January-June 2004. Maps for paras 2 and 3 are shown in S1 and S2 Figs. Locations of roads and households taken from GPS mapping during study.

<https://doi.org/10.1371/journal.pntd.0006453.g004>

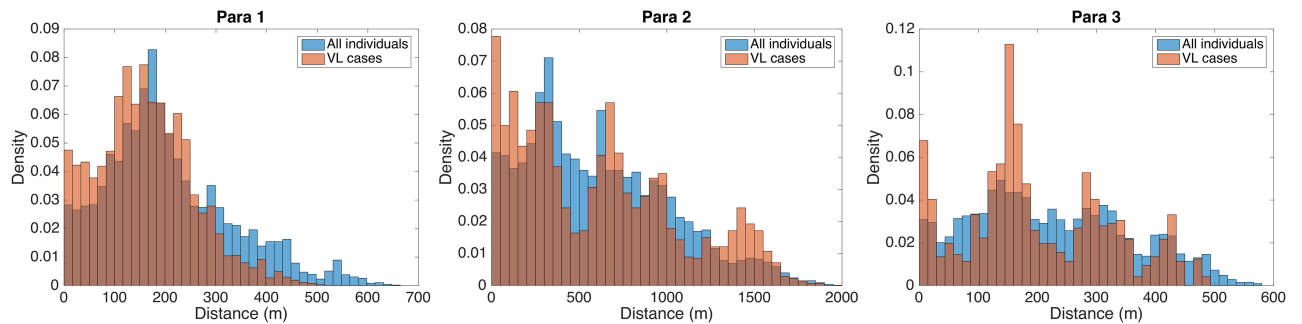


Fig 5. Distributions of pairwise distances between all individuals and between all VL cases in each para.

<https://doi.org/10.1371/journal.pntd.0006453.g005>

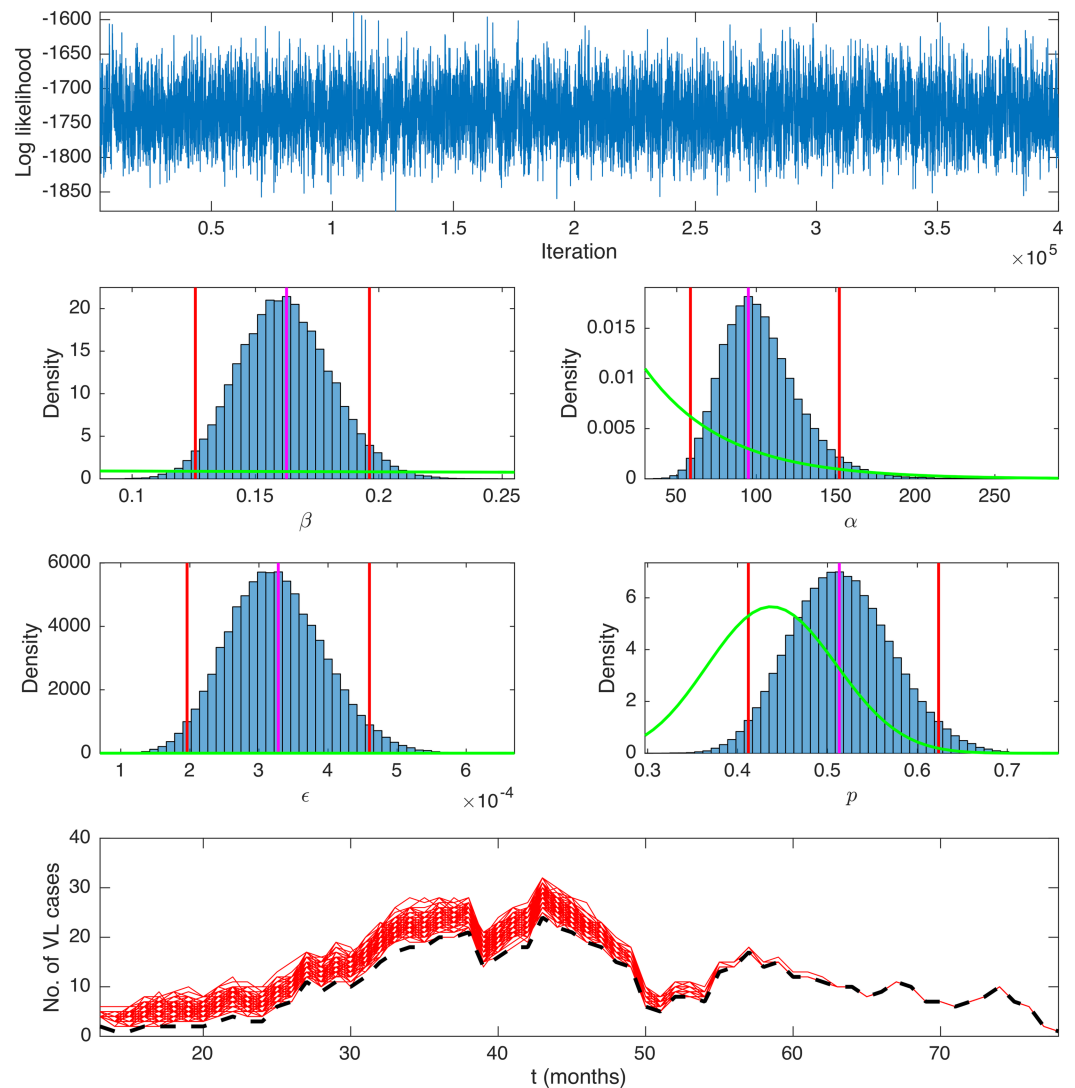


Fig 6. Output of the MCMC algorithm for the exponential kernel model. Top: Log-likelihood trace plot. 2nd-4th row: Posterior distributions for the spatial transmission rate constant β (mnh^{-1}), risk decay distance α (m), background transmission rate ϵ (mnh^{-1}), and incubation period distribution parameter p , with prior distributions (green lines), posterior modes (magenta lines) and 95% highest posterior density intervals (red lines). Bottom: Number of active VL cases over time in 100 iterations of the MCMC algorithm (red lines), with imputed missing onset and recovery times, and number of active cases excluding individuals with missing onset and/or recovery times (black dashed line).

<https://doi.org/10.1371/journal.pntd.0006453.g006>

Table 3. Parameter estimates for different models fitted to the data.

Model	Parameter estimates*								DIC
	β (mnth ⁻¹)	95% HPDI	α (m)	95% HPDI	ϵ ($\times 10^{-4}$ mnth ⁻¹)	95% HPDI	δ ($\times 10^{-3}$ mnth ⁻¹)	95% HPDI	
Background-only	0 [†]	-	-	-	11	(9.2–12)	0 [†]	-	3706.6
Constant kernel	0.24	(0.18–0.29)	-	-	2.4	(1.1–4.0)	0 [†]	-	3523.6
Decreasing VL risk with distance									
Cauchy kernel	0.17	(0.13–0.21)	64	(38–134)	3.0	(1.8–4.5)	0 [†]	-	3466.5
Exponential kernel	0.16	(0.13–0.20)	95	(59–152)	3.3	(2.0–4.6)	0 [†]	-	3459.5
Additional within-household transmission									
Cauchy kernel	0.16	(0.12–0.20)	80	(40–148)	3.1	(1.8–4.4)	1.6	(0.1–4.5)	3462.7
Exponential kernel	0.15	(0.12–0.19)	102	(62–160)	3.1	(1.9–4.5)	1.4	(0.1–4.4)	3459.7
LST status data included									
Cauchy kernel	0.24	(0.19–0.30)	81	(44–139)	3.5	(2.0–5.4)	0 [†]	-	- [‡]
Exponential kernel	0.23	(0.18–0.28)	100	(69–157)	3.8	(2.4–5.7)	0 [†]	-	- [‡]

* Modes of marginal posterior distributions for parameters and 95% highest posterior density intervals (HPDIs).

[†] Assumed.

[‡] No DIC calculated as additional data included in the fitting.

<https://doi.org/10.1371/journal.pntd.0006453.t003>

α , and ϵ , and incubation period distribution parameter, p . It is clear from the plots that the MCMC chain converged rapidly and mixed well, and that the parameters are well defined by the data. The chain was initialised from various regions of parameter space to assess convergence and in all cases nearly identical distributions for the parameters and log-likelihood were obtained.

The estimates (posterior modes and 95% highest posterior density intervals (HPDIs)) for the transmission parameters and DIC for each of the models tested are presented in Table 3. Comparing the DIC values for the different models, it is clear that those in which transmission risk decreases with distance from a VL case fit the data much better than the models with no variation with distance (i.e. the ‘background-only’ model and constant kernel model, for which the differences in DIC from the best-fitting model are Δ DIC = 247.1 and Δ DIC = 64.1 respectively). The best-fitting model is the exponential kernel model. However, the DIC differences between this model and the Cauchy kernel model and the models with additional within-household transmission are all small (Δ DIC < 7), and their deviance distributions are virtually completely overlapping (S3 Fig), suggesting that the data is similarly well described by a Cauchy or exponential kernel and models with additional within-household transmission.

The parameter estimates are highly consistent across the Cauchy and exponential kernel models. The estimates for the spatial transmission rate constant β , background transmission rate ϵ and additional within-household transmission rate can be interpreted as follows. In the absence of any background transmission, the estimated monthly probability of developing VL from living in the same household as a single symptomatic case was 14–16 in 10,000 (95% HPDI 6–31 in 10,000) according to the models without extra within-household transmission, and 26–28 in 10,000 (95% HPDI 7–70 in 10,000) according to models with extra within-household transmission. The monthly probability of developing VL from background transmission if there were no VL cases nearby was approximately 3 in 10,000 (95% HPDI 1.8–4.6 in 10,000). Put another way, the models suggest that at least 134–138 of the 183 VL cases in the 3 paras from 1999–2004 were infected due to being near to VL cases, while the remaining cases originated from background transmission (or at least cannot be explained by spatial proximity to active VL cases or infectious pre-symptomatic individuals).

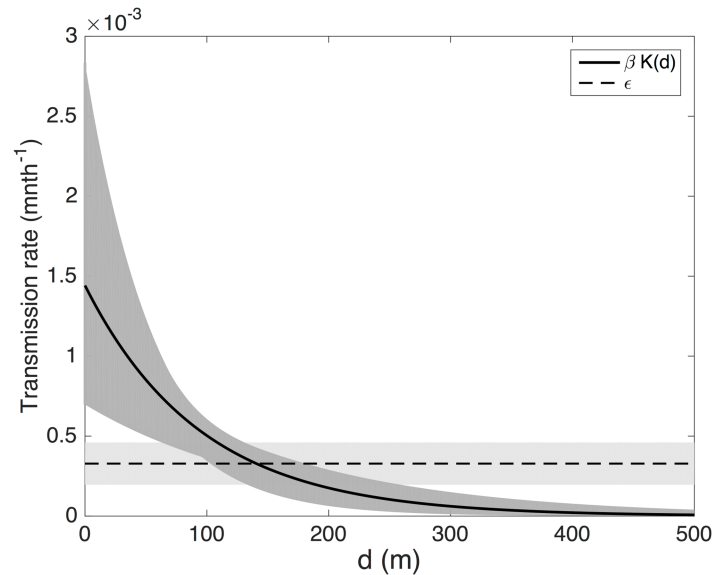


Fig 7. Estimated spatial transmission kernel for the exponential kernel model. Kernel (solid black line) shown as the transmission rate per month, $\beta K(d)$, as a function of distance d from a single VL case. Modal estimate for kernel and 95% highest posterior density interval (HPDI) (dark grey shaded region) calculated from posterior modes and HPDIs for the transmission parameters. The mode (dashed black line) and 95% HPDI (light grey shaded region) for the background transmission rate, ϵ , are also shown for comparison.

<https://doi.org/10.1371/journal.pntd.0006453.g007>

All but one of the posterior modes for the distance decay rate parameter, α , are $\leq 100\text{m}$, suggesting VL risk decreased relatively quickly with distance from a case. Although the α estimates for the exponential kernel models are greater than those for the Cauchy kernel models, they actually correspond to a very similar rate of decrease in risk, due to the different shapes of the kernels. Accounting for background transmission, the estimates for the distance at which the risk of VL halves compared to living in the same household as a case (the ‘half-risk distance’) are 79m (95% HPDI 38–197m) and 87m (95% HPDI 50–166m) for the Cauchy and exponential kernel models without extra within-household transmission, and 33m (95% HPDI 4–105m) and 0m (i.e. double the risk in the same household as a case) (95% HPDI 0–69m) for the models with additional within-household transmission. Fig 7 shows the estimated spatial kernel (in terms of the transmission rate from a single VL case $\beta K(d)$ a distance d away) and its 95% HPDI, for the exponential kernel model. The relatively rapid decay in risk with distance is clear, but there is considerable uncertainty in the estimated transmission rate up to 100m. The estimated background transmission rate and its 95% HPDI are also shown, indicating that the transmission rate from VL cases is higher than the background rate up to around 150m from a case (Box 1).

Including additional within-household transmission in the model led to an increase in the estimates for α since a flatter kernel shape was required to compensate for the extra within-household transmission. The estimate for the additional within-household transmission rate δ for the exponential kernel model suggests an approximately 2-fold increase in risk (95% HPDI 0.2–5.4) from occupying the same household as an active case compared to living right next to a case, but this result should be interpreted with caution given the wide 95% credible interval and the fact that this model did not fit significantly better than the models without additional within-household transmission (Box 1).

Including the LST data and treating individuals as immune upon becoming LST-positive did not significantly affect the estimates for α and ϵ , which both increased by only a small

Box 1. What are potential drivers for spatiotemporal clustering of VL at household level?

- Recent cases (with symptom onset within the last year) in the same or nearby households (within 150m).
- Living in the same household as a case may lead to additional VL risk over and above that due just to distance from a case, but the magnitude of any risk increase is highly uncertain.
- Poverty? No significant differences in VL risk by income, education, occupation, land ownership or asset ownership were observed in single-variable and multivariable logistic regression models of the same data [15, 25], though more frequent beef or goat consumption was found to be associated with reduced risk of progression from infection to VL [15].

amount (Table 3). However, it resulted in an increase in the estimate for β , such that the estimated monthly probability of developing VL living in the same household as a case increased (from 14 in 10,000 to 19 in 10,000 for the exponential kernel model). This is as expected, given the reduction in the susceptible population from treating LST-positive individuals as immune and the proximity of a large proportion of the LST-positive individuals to VL cases [15].

The posterior distribution for the incubation period distribution parameter, p , was very similar for all models apart from the background-only model (posterior modes 0.50–0.51, 95% HPDIs (0.40–0.62)–(0.41–0.63)), suggesting that the incubation period distribution was relatively well constrained by the data given its assumed shape. The estimated posterior distribution for p corresponds to a mean incubation period duration of approximately 4 months (95% HPDI 2.7–5.1 months).

Discussion

This study is the first individual-level fully spatiotemporal analysis of VL transmission in the Indian subcontinent and the first attempt to estimate a spatial transmission kernel for VL. It builds upon the work of several field studies that have identified proximity to a recent or current VL case as a strong risk factor for infection and disease, and given estimates of the increase in risk from living with or nearby a VL case, by providing a framework for a more refined spatiotemporal analysis of VL transmission. In particular the framework we have developed accounts for the large variation in the incubation period of the disease, the unobserved infection times of cases, the number of individuals infectious at the infection time of each new case and their distances from the new case, and potential transmission from asymptomatic individuals.

The results of our analysis suggest that the risk of being infected and developing VL decreases relatively quickly with distance from a case (halving within 90m) in a high endemicity setting. This estimate broadly agrees with the results of previous studies that used different methods of analysis. A multivariable logistic regression model of VL risk for the same dataset [15], found that the OR for VL, relative to living more than 50m from a previous case, decreased from 6.37 in the same household as a case to 1.85 in a household within 50m of a case. Another study on data from India and Nepal also found elevated risk of VL when living

within 50m of a case with VL in the previous 18 months, compared to being more than 50m away (OR = 2.14, 95% CI 1.27–3.62, $p = 0.004$) [27]. Given that the spatial kernel in our transmission model represents the spread of infection from person to person due to the sandfly movement, the estimated kernel is also in agreement with the results of recent mark-release-recapture sandfly dispersal studies performed in Bihar, India, in which most sandflies only flew short distances ($\sim 20\text{m}$), and very few more than 150m [78]. Nevertheless, the estimated distance from a case over which risk decays is relatively small, and it is possible that other aspects of the local environment, such as housing density and configuration, are important in transmission at this scale. In future work we will investigate the potential role of these fine-scale environmental factors in the spatial patterns of VL cases.

The estimated mean incubation period of 4 months is consistent with our previous estimate of ~ 5 months (95% CI 4.3–5.5 months) from a multi-state Markov model of the individual-level diagnostic data (including the serological data, which was not used in the present analysis) and case onset and treatment data from the same study [62], and falls within the 2–6 months range reported as typical in the literature [25, 61]. It also appears to agree relatively well with data from longitudinal studies on times from being measured sero-/PCR-positive to developing VL [26, 28, 33, 34, 107–109], which suggest most disease conversion occurs within 6 months.

The estimated relative contributions of transmission from nearby VL cases and background transmission to the observed number of cases suggest that VL cases contribute significantly more to transmission than asymptomatic individuals (at least three times as much) when incidence is high. Given that the background transmission rate also covers potential transmission from animals, the fact that it is much smaller than the case-to-case transmission rate suggests that the contribution of animals to transmission (if it exists) is small. Together with the relatively limited range of transmission around VL cases implied by the estimated spatial kernel, this suggests that reactive spatially-targeted interventions could be effective for reducing VL risk (Box 2). However, the relationship between reaction time to new cases and the radius around them in which interventions would need to be performed is uncertain and needs quantification [110]. It is likely that such a strategy would rely on early detection of new cases and interventions being performed soon after their detection and in a large radius around them ($\geq 300\text{m}$ according to the estimated kernel). Even short delays could render the strategy ineffective, since by the time spraying and active case detection were performed the index cases would probably have already infected the next generation of cases within their transmission range.

With regard to implications for spatially-targeted control, it is important to note that the estimated kernel is a single static estimate of how VL risk varies with distance from a case

Box 2. What do spatiotemporal patterns at household level tell us about VL transmission dynamics?

In a highly endemic setting:

- VL risk decreases relatively quickly with distance from a VL case (halving within 90m).
- VL cases appear to contribute more to transmission than asymptomatic individuals.
- Spatially-targeted interventions could be effective for reducing VL risk if performed promptly and in a sufficiently large radius around a new case.

based on incidence over 5.5 years in a highly endemic setting. This is also true of the estimated asymptomatic contribution. How spatial transmission patterns and the asymptomatic contribution vary across settings with different endemicities and longer periods of time is uncertain. The approach used here should be applied to more recent data and data from other geographical areas to assess this variation. This is particularly important given the large differences in incidence between Bangladesh, India and Nepal and the significant declines in incidence in all three countries since the study was conducted. Although existing data suggests that proximity to VL cases is a strong risk factor for infection in all three countries and across different endemicity levels [15, 26, 27, 111], it is possible that spatial patterns of transmission and the asymptomatic contribution have changed as incidence has declined, and further work is needed to assess the generalisability of the results to low-incidence and outbreak settings. Bern et al [25] observed that even over the course of the study, the pattern of VL cases spread from being highly-clustered to saturating major parts of each para (see Fig 4, S1 and S2 Figs), as a large proportion of the population became exposed and those that did not develop disease gained some level of immunity to the parasite. They suggested that this saturation could occur within 2-3 years in high transmission intensity settings, which agrees with observations of 1-3 year micro-epidemics in hamlets in Muzaffarpur, Bihar, India [106]. Thus, spatially-targeted interventions would need to be implemented as early as possible in a micro-epidemic to maximise their impact.

Another important consideration is the practical and economic feasibility of performing spatially-targeted IRS and case detection at a sub-village level, as it would require the ability to mount a rapid response to new cases and delineate the households that would receive the interventions, and focal active case detection could be time-consuming and expensive to implement. Given that VL incidence is also strongly clustered at hamlet level [106], it might make more sense to implement targeted control at this level depending on resource constraints. Based on our estimates, the policy of spraying all villages within 500m of villages with cases in the previous year that is currently being trialled in India seems sensible, as it should account for the fact that the disease may have been transmitted through more than one generation of cases by the time the spraying is performed.

Given the uncertainties in the potential impact of reactive focal IRS and active case detection and in the time window after detection of a new case in which they would be effective, and the difficulties of trialling them, in future work we will predict the impact of different spatially-targeted control strategies by simulating them with the parameterised spatiotemporal transmission model.

Limitations of this analysis

As with any spatiotemporal analysis of transmission of a disease such as VL, with a complicated natural history and transmission cycle, our analysis has a number of limitations. In particular, we have not included the serological data from the annual cross-sectional surveys in the model and have not explicitly modelled asymptomatic infection, but incorporated potential transmission from asymptomatic individuals via a background transmission rate that is constant in space and time. Although we have considered the implications of LST-convertors being immune, this was via a very simple age-dependent model for the probability of being LST-positive, which does not take account of spatial and temporal heterogeneity in the force of infection. Asymptomatic infection is known to be associated with spatial and temporal proximity to VL cases, so if asymptomatic individuals transmit it is likely that their contribution varies in space and time. Incorporating the diagnostic data into the model would enable explicit modelling of asymptomatic infection and could help to constrain the possible infection

times of cases (e.g. by assuming that individuals cannot be seropositive before they are infected). However, constructing the model and performing the inference on the full data is technically very challenging since the diagnostic surveys only cover the prospective part of the study (2002–2004), the tests (in particular the LST) have imperfect sensitivity and specificity, many individuals are missing tests, and there are a large number of observed combinations of test results. An inference approach that can account for unobserved infections, such as reversible jump MCMC [53, 54], would thus be required, along with a hidden Markov model for the true infection status underlying the observed test results [97]. Both the infection and recovery times of asymptotically infected individuals are unobserved, so it would also be necessary to impute these from the diagnostic data.

Potential edge effects in space—contributions from infected individuals in unsurveyed paras neighbouring the study paras—are not accounted for in the model. Based on evidence that *P. argentipes* flies have a limited flight range and the study paras being more than 500m apart, we assumed that there was no transmission between cases in different paras in the absence of migration. However, some of the unsurveyed paras are within the maximum flight range of the sandfly and had cases during the study period, so it is possible that there was inter-para transmission from infectious individuals in households near the edges of unsurveyed paras. A related issue is that the model does not explicitly account for migration (which is covered via the background transmission rate), and assumes individuals were infected in or very nearby their households. In reality, there is a significant amount of seasonal migration for work in rural Bangladeshi communities, which may affect transmission, and there is uncertainty about how much transmission occurs indoors vs. outdoors [14]. These issues will be addressed in future work analysing spatial transmission over a longer period in the same paras and the unsurveyed paras using geo-located incidence and migration data [112].

Another limitation is that the model treats VL risk as purely a function of the local number of infectious individuals and their spatial proximity, and does not include other potential risk factors such as age, sex, bed net use and socioeconomic status. Such risk factors can be included in the model by treating the transmission rate between individuals ($\lambda_{ij}(t)$ in Eq (1)) as a function of these variables, but were omitted here to keep the analysis as simple as possible and because case proximity was identified as the strongest risk factor in the original analyses [15, 25]. Sandfly density, which varies seasonally [113], is also likely to affect transmission risk, but sandflies are not explicitly represented in the model due to a lack of comprehensive data on sandfly abundance in the study paras, and uncertainties in the relationship between sandfly and host densities and infection prevalences [12, 14] (Box 3).

The unknown infectiousness profile of VL cases over time (before and during symptoms) and how it varies between individuals is also a potential source of uncertainty in the estimated spatial kernel and transmission rates. In households with multiple cases during the study period, there were long intervals between the onsets of the cases (longer than would be expected if the later cases were infected by the earlier cases, based on the estimated incubation period), possibly suggesting that some cases can remain infectious even after apparently successful treatment, although this is complicated by potential transmission from asymptomatic individuals and other nearby cases.

Finally, we note that the results of this analysis need to be validated. Unfortunately, there are relatively few geo-located VL incidence datasets available with which to do this. However, as part of a more formal validation of the model, in future work we will implement the equivalent stochastic spatial simulation model and test that it reproduces similar spatial patterns of incidence to those observed.

Box 3. What can we learn from this analysis for future data collection?

- Further epidemiological studies to collect geo-located VL incidence data are required to assess how spatial clustering of VL varies with endemicity and over long periods of time.
- Measurements of sandfly densities and infection prevalence should be performed alongside measurements of human infection and disease, to assess the role sandfly bio-nomics play in spatial clustering.
- More detailed studies of the dispersal and flight range of *P. argentipes* sandflies are required to determine how far they can spread the parasite and over what timescale.

Conclusion

Our analysis shows that in a high-endemicity setting VL transmission is focused around cases and cases are the main drivers of transmission. This suggests that reactive spatially-targeted control interventions could be effective at reducing transmission in high endemicity areas if implemented promptly and in a sufficiently large area around a case. Nevertheless, it is necessary to validate the model against data from different settings and to use it to predict the impact of different control strategies to assess whether spatially-targeted interventions would be effective, and under what conditions. If targeted interventions were predicted to be effective, the results could then be used to inform control policy to help achieve and maintain the 2020 elimination target for VL in the ISC.

Supporting information

S1 Table. Data on impact of case proximity on VL risk. Odds ratios for VL and asymptomatic infection risk based on proximity to infected individuals (from previous studies). (XLSX)

S1 Text. Supplementary information on model and parameter estimation. Definitions of model likelihood and deviance information criterion, and details of MCMC algorithm. (PDF)

S1 Data. Data used to fit model. Data on visceral leishmaniasis status and leishmanin skin test status of 2494 individuals in the study area, including dates of onset, diagnosis, treatment, and, where applicable, relapse and treatment for relapse for 183 VL cases. (CSV)

S2 Data. Matrix of pairwise distances between all individuals in the study. Distances are given in the same order as individuals are listed in [S1 Data](#). (ZIP)

S3 Data. Metadata for S1 and S2 Data. (CSV)

S1 Fig. Locations of VL cases in para 2 by year of onset, 1999-2004. (PDF)

S2 Fig. Locations of VL cases in para 3 by year of onset, 1999-2004.
(PDF)

S3 Fig. Deviance distributions for the different models.
(PDF)

Acknowledgments

LACC is grateful to Marinella Capriati and colleagues in the Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research at the University of Warwick, in particular Matt Keeling, Orin Courtenay, Panayiota Touloupou, Joaquin Prada, Renata Retkute, Sam Brand, Erin Dilger and Ron Crump, for very helpful discussions regarding Bayesian inference and interpretation of the data. LACC is thankful to Tim Pollington of the Centre for Complexity Science, University of Warwick, and Allen Hightower for help with producing figures and mapping.

Author Contributions

Conceptualization: Lloyd A. C. Chapman, Chris P. Jewell, Caryn Bern, Graham F. Medley, T. Déirdre Hollingsworth.

Data curation: Lloyd A. C. Chapman, Rajib Chowdhury, Caryn Bern.

Formal analysis: Lloyd A. C. Chapman, Chris P. Jewell, Simon E. F. Spencer, Lorenzo Pellis, Samik Datta, Graham F. Medley, T. Déirdre Hollingsworth.

Funding acquisition: T. Déirdre Hollingsworth.

Investigation: Lloyd A. C. Chapman, Rajib Chowdhury, Caryn Bern.

Methodology: Lloyd A. C. Chapman, Chris P. Jewell, Simon E. F. Spencer, Lorenzo Pellis, Samik Datta, Graham F. Medley, T. Déirdre Hollingsworth.

Project administration: Lloyd A. C. Chapman, Graham F. Medley, T. Déirdre Hollingsworth.

Resources: Caryn Bern, T. Déirdre Hollingsworth.

Software: Lloyd A. C. Chapman, Samik Datta.

Supervision: Chris P. Jewell, Simon E. F. Spencer, Lorenzo Pellis, Caryn Bern, Graham F. Medley, T. Déirdre Hollingsworth.

Validation: Lloyd A. C. Chapman.

Visualization: Lloyd A. C. Chapman, Graham F. Medley, T. Déirdre Hollingsworth.

Writing – original draft: Lloyd A. C. Chapman.

Writing – review & editing: Lloyd A. C. Chapman, Chris P. Jewell, Simon E. F. Spencer, Lorenzo Pellis, Caryn Bern, Graham F. Medley, T. Déirdre Hollingsworth.

References

1. World Health Organization Regional Office for South-East Asia, Government of the People's Republic Of Bangladesh, Royal Government of Bhutan, Government of India, Government of Nepal, Royal Thai Government. Memorandum of Understanding among Bangladesh, Bhutan, India, Nepal and Thailand on the Elimination of Kala-Azar from the South-East Asia Region. Dhaka; 2014.
2. World Health Organization. Global Health Observatory: Number of cases of visceral leishmaniasis reported—Data by country; 2017. Available from: <http://apps.who.int/gho/data/node.main.NTDLEISHVNUM?lang=en>.

3. National Vector Borne Disease Control Programme. Accelerated Plan for Kala-azar Elimination 2017; 2017. Available from: http://www.who.int/leishmaniasis/resources/Accelerated-Plan-Kala-azar1-Feb2017_light.pdf.
4. Hirve S, Boelaert M, Matlashewski G, Mondal D, Arana B, Kroeger A, et al. Transmission Dynamics of Visceral Leishmaniasis in the Indian Subcontinent—A Systematic Literature Review. *PLoS Neglected Tropical Diseases*. 2016; 10(8):e0004896. <https://doi.org/10.1371/journal.pntd.0004896> PMID: 27490264
5. KalaCORE. Bangladesh; 2017. Available from: <http://www.kalacore.org/where-we-work/bangladesh>.
6. World Health Organization Regional Office for South-East Asia. Process of validation of elimination of kala-azar; 2016. SEA-CD-321. Available from: http://www.who.int/leishmaniasis/resources/Process_of_validation_of_VL_elimination_SEA_CD_321.pdf.
7. Srikantiah S. Progress on VL elimination in India and challenges ahead. In: Annual Meeting of the Coalition for Operational Research on Neglected Tropical Diseases (COR-NTD 2016), Atlanta, USA, November 10-11, 2016. Atlanta; 2016.
8. Chowdhury R, Mondal D, Chowdhury V, Faria S, Alvar J. How Far Are We from Visceral Leishmaniasis Elimination in Bangladesh? An Assessment of Epidemiological Surveillance Data. *PLoS Neglected Tropical Diseases*. 2014; 8(8). <https://doi.org/10.1371/journal.pntd.0003020>
9. Ministry of Health and Family Welfare Bangladesh. National Guideline For Kala-azar Case Management; 2016. Available from: http://kalacorebd.com/wp-content/uploads/2016/04/Kala-azar-Case-Management-Guideline-2016_Latest-Draft.pdf.
10. Hirve S, Kroeger A, Matlashewski G, Mondal D, Banjara R, Das P, et al. Towards elimination of visceral leishmaniasis in the Indian subcontinent—Translating research to practice to public health. *PLoS Neglected Tropical Diseases*. 2017; 11(10):e0005889. <https://doi.org/10.1371/journal.pntd.0005889> PMID: 29023446
11. Rijal S. The diagnostic challenges and needs in the kala-azar elimination programme in South Asia. In: 6th World Congress on Leishmaniasis Abstracts Book, Toledo, Spain, May 16-20, 2017; 2017. p. 33. Available from: http://worldleish2017.org/documentos/Abstracts_Book_WL6_2017.pdf.
12. Le Rutte EA, Chapman LAC, Coffeng LE, Ruiz-Postigo JA, Olliaro PL, Adams ER, et al. Policy recommendations from transmission modelling for the elimination of visceral leishmaniasis in the Indian subcontinent. *Clinical Infectious Diseases*. 2018; 66(suppl_4):S301–S308. <http://dx.doi.org/10.1093/cid/ciy007>
13. Rock KS, Quinell RJ, Medley GF, Courtenay O. Progress in the Mathematical Modelling of Visceral Leishmaniasis. vol. 94. Elsevier Ltd; 2016. Available from: <http://dx.doi.org/10.1016/bs.apar.2016.08.001>.
14. Cameron MM, Acosta-Serrano A, Bern C, Boelaert M, Den Boer M, Burza S, et al. Understanding the transmission dynamics of *Leishmania donovani* to provide robust evidence for interventions to eliminate visceral leishmaniasis in Bihar, India. *Parasites & Vectors*. 2016; 9(1):25. <https://doi.org/10.1186/s13071-016-1309-8>
15. Bern C, Haque R, Chowdhury R, Ali M, Kurkjian KM, Vaz L, et al. The epidemiology of visceral leishmaniasis and asymptomatic leishmanial infection in a highly endemic Bangladeshi village. *American Journal of Tropical Medicine and Hygiene*. 2007; 76(5):909–914. <https://doi.org/10.4269/ajtmh.2007.76.909> PMID: 17488915
16. Hasker E, Malaviya P, Gidwani K, Picado A, Ostyn B, Kansal S, et al. Strong Association between Serological Status and Probability of Progression to Clinical Visceral Leishmaniasis in Prospective Cohort Studies in India and Nepal. *PLoS Neglected Tropical Diseases*. 2014; 8(1). <https://doi.org/10.1371/journal.pntd.0002657> PMID: 24466361
17. Sharma MC, Gupta aK, Das VNR, Verma N, Kumar N, Saran R, et al. *Leishmania donovani* in blood smears of asymptomatic persons. *Acta Tropica*. 2000; 76(2):195–196. [https://doi.org/10.1016/S0001-706X\(00\)00068-1](https://doi.org/10.1016/S0001-706X(00)00068-1) PMID: 10936579
18. Stauch A, Sarkar RR, Picado A, Ostyn B, Sundar S, Rijal S, et al. Visceral leishmaniasis in the Indian subcontinent: Modelling epidemiology and control. *PLoS Neglected Tropical Diseases*. 2011; 5(11):1–12. <https://doi.org/10.1371/journal.pntd.0001405>
19. Le Rutte EA, Coffeng LE, Bontje DM, Hasker EC, Postigo JAR, Argaw D, et al. Feasibility of eliminating visceral leishmaniasis from the Indian subcontinent: explorations with a set of deterministic age-structured transmission models. *Parasites & Vectors*. 2016; 9(1):24. <https://doi.org/10.1186/s13071-016-1292-0>.
20. Le Rutte EA, Chapman LAC, Coffeng LE, Jervis S, Hasker EC, Dwivedi S, et al. Elimination of visceral leishmaniasis in the Indian subcontinent: a comparison of predictions from three transmission models. *Epidemics*. 2017; 18:67–80. <https://doi.org/10.1016/j.epidem.2017.01.002> PMID: 28279458

21. Yangzom T, Cruz I, Bern C, Argaw D, den Boer M, Vélez ID, et al. Endemic transmission of visceral leishmaniasis in Bhutan. *American Journal of Tropical Medicine and Hygiene*. 2012; 87(6):1028–37. <https://doi.org/10.4269/ajtmh.2012.12-0211> PMID: 23091191
22. Ostyn B, Uranw S, Bhattarai NR, Das ML, Rai K, Tersago K, et al. Transmission of *Leishmania donovani* in the Hills of Eastern Nepal, an Outbreak Investigation in Okhaldhunga and Bhojpur Districts. *PLOS Neglected Tropical Diseases*. 2015; 9(8):e0003966. <https://doi.org/10.1371/journal.pntd.0003966> PMID: 26252494
23. Sundar S, Singh OP, Tiwary P, Singh S, Kushwaha AK, Kumar D, et al. Infectiousness of visceral leishmaniasis, post kala-azar dermal leishmaniasis and asymptomatic subjects with their capacity to infect sand fly vectors by xenodiagnosis.; 2017. *American Society of Tropical Medicine and Hygiene*: 66th Annual Meeting, Nov 5-9, 2017. Available from: <http://www.abstractsonline.com/pp8/#!/4395/presentation/5342> [cited 15/12/2017].
24. Molina R, Ghosh D, Carrillo E, Monnerat S, Bern C, Mondal D, et al. Infectivity of Post-Kala-azar Dermal Leishmaniasis Patients to Sand Flies: Revisiting a Proof of Concept in the Context of the Kala-azar Elimination Program in the Indian Subcontinent. *Clinical Infectious Diseases*. 2017; 65(1):150–153. <https://doi.org/10.1093/cid/cix245> PMID: 28520851
25. Bern C, Hightower AW, Chowdhury R, Ali M, Amann J, Wagatsuma Y, et al. Risk factors for kala-azar in Bangladesh. *Emerging Infectious Diseases*. 2005; 11(5):655–662. <https://doi.org/10.3201/eid1105.040718> PMID: 15890115
26. Gidwani K, Kumar R, Rai M, Sundar S. Longitudinal Seroepidemiologic Study of Visceral Leishmaniasis in Hyperendemic Regions of Bihar, India. *American Journal of Tropical Medicine and Hygiene*. 2009; 80(3):345–346. <https://doi.org/10.4269/ajtmh.2009.80.345> PMID: 19270279
27. Picado A, Ostyn B, Singh SP, Uranw S, Hasker E, Rijal S, et al. Risk factors for visceral leishmaniasis and asymptomatic *Leishmania donovani* infection in India and Nepal. *PLoS ONE*. 2014; 9(1):1–8. <https://doi.org/10.1371/journal.pone.0087641>
28. Bimal S, Das VNR, Sinha PK, Gupta AK, Verma N, Ranjan A, et al. Usefulness of the direct agglutination test in the early detection of subclinical *Leishmania donovani* infection: a community-based study. *Annals of Tropical Medicine & Parasitology*. 2005; 99(8):743–749. <https://doi.org/10.1179/136485905X65107>
29. Bhattarai NR, Das ML, Rijal S, van der Auwera G, Picado A, Khanal B, et al. Natural infection of *Phlebotomus argentipes* with *Leishmania* and other trypanosomatids in a visceral leishmaniasis endemic region of Nepal. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2009; 103(11):1087–92. <https://doi.org/10.1016/j.trstmh.2009.03.008> PMID: 19345387
30. Khanal B, Rijal S, Ostyn B, Picado A, Gidwani K, Menten J, et al. Serological markers for *Leishmania donovani* infection in Nepal: Agreement between direct agglutination test and rK39 ELISA. *Tropical Medicine and International Health*. 2010; 15(11):1390–1394. <https://doi.org/10.1111/j.1365-3156.2010.02631.x> PMID: 21998875
31. Rijal S, Uranw S, Chappuis F, Picado A, Khanal B, Paudel IS, et al. Epidemiology of *Leishmania donovani* infection in high-transmission foci in Nepal. *Tropical Medicine and International Health*. 2010; 15(Suppl. 2):21–28. <https://doi.org/10.1111/j.1365-3156.2010.02518.x> PMID: 20487421
32. Nandy A, Neogy AB, Chowdhury AB. Leishmanin test survey in an endemic village of Indian kala-azar near Calcutta. *Annals of Tropical Medicine & Parasitology*. 1987; 81(6):693–699. <https://doi.org/10.1080/00034983.1987.11812172>
33. Topno RK, Das VNR, Ranjan A, Pandey K, Singh D, Kumar N, et al. Asymptomatic infection with visceral leishmaniasis in a disease-endemic area in Bihar, India. *American Journal of Tropical Medicine and Hygiene*. 2010; 83(3):502–506. <https://doi.org/10.4269/ajtmh.2010.09-0345> PMID: 20810810
34. Das VNR, Siddiqui Na, Verma RB, Topno RK, Singh D, Das S, et al. Asymptomatic infection of visceral leishmaniasis in hyperendemic areas of Vaishali district, Bihar, India: A challenge to kala-azar elimination programmes. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2011; 105(11):661–666. <https://doi.org/10.1016/j.trstmh.2011.08.005> PMID: 21945327
35. Das A, Karthick M, Dwivedi S, Banerjee I, Mahapatra T. Epidemiologic Correlates of Mortality among Symptomatic Visceral Leishmaniasis Cases: Findings from Situation Assessment in High Endemic Foci in India. *PLoS Neglected Tropical Diseases*. 2016; 10(11): e0005150. <https://doi.org/10.1371/journal.pntd.0005150>
36. World Health Organization. Control of the leishmaniasis. World Health Organization Technical Report Series; 2010. 949. Available from: http://apps.who.int/iris/bitstream/handle/10665/44412/WHO_TRS_949_eng.pdf.
37. Poche DM, Garlapati RB, Mukherjee S, Torres-Poche Z, Hasker E, Rahman T, et al. Bionomics of *Phlebotomus argentipes* in villages in Bihar, India with insights into efficacy of IRS-based control

- measures. *PLoS Neglected Tropical Diseases*. 2018; 12(1):e0006168. <https://doi.org/10.1371/journal.pntd.0006168>
38. Kumar V, Mandal R, Kesari S, Das S, Das P. Reaching the elimination target in the district of Vaishali, Bihar, India. In: 6th World Congress on Leishmaniasis Abstracts Book, Toledo, Spain, May 16–20, 2017; 2017. p. 23. Available from: http://worldleish2017.org/documentos/Abstracts_Book_WL6_2017.pdf.
 39. Ahmed BN, Nabi SG, Rahman M, Selim S, Bashar A. Kala-azar (Visceral Leishmaniasis) Elimination in Bangladesh: Successes and Challenges. *Current Tropical Medicine Reports*. 2014; p. 163–169. <https://doi.org/10.1007/s40475-014-0027-6>
 40. Singh SP, Hirve S, Huda MM, Banjara MR, Kumar N, Mondal D, et al. Options for active case detection of visceral leishmaniasis in endemic districts of India, Nepal and Bangladesh, comparing yield, feasibility and costs. *PLoS Neglected Tropical Diseases*. 2011; 5(2). <https://doi.org/10.1371/journal.pntd.0000960>
 41. Das A, Harries A, Hinderaker S, Zachariah R, Ahmed B, Shah G, et al. Active and passive case detection strategies for the control of leishmaniasis in Bangladesh. *Public Health Action*. 2014; 4(1):15–21. <https://doi.org/10.5588/pha.13.0084> PMID: 26423756
 42. Huda MM, Hirve S, Siddiqui NA, Malaviya P, Banjara MR, Das P, et al. Active case detection in national visceral leishmaniasis elimination programs in Bangladesh, India, and Nepal: feasibility, performance and costs. *BMC Public Health*. 2012; 12(1):1001. <https://doi.org/10.1186/1471-2458-12-1001> PMID: 23164318
 43. Dhiman R, Sen A. Epidemiology of kala-azar in rural Bihar (India) using village as a component unit of study. *Indian Journal of Medical Research*. 1991; 93:155–160. PMID: 1937592
 44. Maroli M, Feliciangeli MD, Bichaud L, Charrel RN, Gradoni L. Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern. *Medical and Veterinary Entomology*. 2013; 27(2):123–147. <https://doi.org/10.1111/j.1365-2915.2012.01034.x> PMID: 22924419
 45. Sharma U, Singh S. Insect vectors of Leishmania: Distribution, physiology and their control. *Journal of Vector Borne Diseases*. 2008; 45(4):255–272. PMID: 19248652
 46. Stauch A, Duerr HP, Dujardin JC, Vanaerschot M, Sundar S, Eichner M. Treatment of visceral leishmaniasis: model-based analyses on the spread of antimony-resistant *L. donovani* in Bihar, India. *PLoS Neglected Tropical Diseases*. 2012; 6(12):e1973. <https://doi.org/10.1371/journal.pntd.0001973> PMID: 23285309
 47. Stauch A, Duerr HP, Picado A, Ostyn B, Sundar S, Rijal S, et al. Model-based investigations of different vector-related intervention strategies to eliminate visceral leishmaniasis on the Indian subcontinent. *PLoS Neglected Tropical Diseases*. 2014; 8(4):e2810. <https://doi.org/10.1371/journal.pntd.0002810> PMID: 24762676
 48. Medley GF, Hollingsworth TD, Olliaro PL, Adams ER. Health-seeking behaviour, diagnostics and transmission dynamics in the control of visceral leishmaniasis in the Indian subcontinent. *Nature*. 2015; 528(7580):S102–S108. <https://doi.org/10.1038/nature16042> PMID: 26633763
 49. Keeling MJ, Rohani P. *Modeling infectious diseases in humans and animals*. Princeton University Press; 2011.
 50. Keeling MJ, Woolhouse MEJ, Shaw DJ, Matthews L, Chase-topping M, Haydon DT, et al. Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science*. 2001; 294(October):813–817. <https://doi.org/10.1126/science.1065973> PMID: 11679661
 51. Ferguson NM, Donnelly CA, Anderson RM. Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature*. 2001; 413(October):542–548. <https://doi.org/10.1038/35097116> PMID: 11586365
 52. Boys RJ, Giles PR. Bayesian inference for stochastic epidemic models with time-inhomogeneous removal rates. *Journal of Mathematical Biology*. 2007; 55(2):223–247. <https://doi.org/10.1007/s00285-007-0081-y> PMID: 17361423
 53. Jewell CP, Keeling MJ, Roberts GO. Predicting undetected infections during the 2007 foot-and-mouth disease outbreak. *Journal of the Royal Society Interface*. 2009; 6:1145–1151. <https://doi.org/10.1098/rsif.2008.0433>
 54. Jewell CP, Kyraios T, Christley RM, Roberts GO. A novel approach to real-time risk prediction for emerging infectious diseases: A case study in Avian Influenza H5N1. *Preventive Veterinary Medicine*. 2009; 91:19–28. <https://doi.org/10.1016/j.prevetmed.2009.05.019> PMID: 19535161
 55. Cauchemez S, Ferguson NM. Methods to infer transmission risk factors in complex outbreak data. *Journal of the Royal Society Interface*. 2012; 9(68):456–469. <https://doi.org/10.1098/rsif.2011.0379>

56. Boender GJ, Hagenaars TJ, Bouma A, Nodelijk G, Elbers ARW, De Jong MCM, et al. Risk maps for the spread of highly pathogenic avian influenza in poultry. *PLoS Computational Biology*. 2007; 3(4):704–712. <https://doi.org/10.1371/journal.pcbi.0030071>
57. Hill EM, House T, Dhingra MS, Kalpravidh W, Morzaria S, Osmani MG, et al. Modelling H5N1 in Bangladesh across spatial scales: Model complexity and zoonotic transmission risk. *Epidemics*. 2017; 20:37–55. <https://doi.org/10.1016/j.epidem.2017.02.007> PMID: 28325494
58. Gamado K, Marion G, Porphyre T. Data-Driven Risk Assessment from Small Scale Epidemics: Estimation and Model Choice for Spatio-Temporal Data with Application to a Classical Swine Fever Outbreak. *Frontiers in Veterinary Science*. 2017; 4(February):1–14.
59. Salje H, Lessler J, Paul KK, Azman AS, Rahman MW, Rahman M, et al. How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study. *Proceedings of the National Academy of Sciences*. 2016; 113(47):13420–13425. <https://doi.org/10.1073/pnas.1611391113>
60. Jewell CP, Brown RG. Bayesian data assimilation provides rapid decision support for vector-borne diseases. *Journal of the Royal Society Interface*. 2015; 12(108):20150367. <https://doi.org/10.1098/rsif.2015.0367>
61. Bora D. Epidemiology of visceral leishmaniasis in India. *The National Medical Journal of India*. 1999; 12(2):62–68. PMID: 10416321
62. Chapman LAC, Dyson L, Courtenay O, Chowdhury R, Bern C, Medley GF, et al. Quantification of the natural history of visceral leishmaniasis and consequences for control. *Parasites & Vectors*. 2015; 8(1):521. <https://doi.org/10.1186/s13071-015-1136-3>
63. Sinnott RW. *Virtues of the Haversine*. Sky and Telescope. 1984; 68:159.
64. QGIS Development Team. QGIS Geographic Information System Version 2.14.2-Essen; 2016. Available from: <http://qgis.osgeo.org>.
65. Rijal S, Ostyn B, Uranw S, Rai K, Bhattarai NR, Dorlo TPC, et al. Increasing Failure of Miltefosine in the Treatment of Kala-azar in Nepal and the Potential Role of Parasite Drug Resistance, Reinfection, or Noncompliance. *Clinical Infectious Diseases*. 2013; 56:1530–1538. <https://doi.org/10.1093/cid/cit102> PMID: 23425958
66. Rai K, Bhattarai NR, Vanaerschot M, Imamura H, Gebru G, Khanal B, et al. Single locus genotyping to track *Leishmania donovani* in the Indian subcontinent: Application in Nepal. *PLoS Neglected Tropical Diseases*. 2017; 11(3): e0005420. <https://doi.org/10.1371/journal.pntd.0005420>
67. Sanyal RK. Leishmaniasis in the Indian sub-continent. In: Chang KP, Bray RS, editors. *Leishmaniasis*. Amsterdam: Elsevier; 1985. p. 443–468.
68. Bhattarai NR, Auwera GVD, Rijal S, Picado A, Speybroeck N, Khanal B, et al. Domestic Animals and Epidemiology of Visceral Leishmaniasis, Nepal. *Emerging Infectious Diseases*. 2010; 16(2). <https://doi.org/10.3201/eid1602.090623>
69. Singh N, Mishra J, Singh R, Singh S. Animal reservoirs of visceral leishmaniasis in India. *Journal of Parasitology*. 2013; 99(1):64–67. <https://doi.org/10.1645/GE-3085.1> PMID: 22765517
70. Alam MS, Ghosh D, Khan GM, Islam MF. Survey of domestic cattle for anti-*Leishmania* antibodies and *Leishmania* DNA in a visceral leishmaniasis endemic area of Bangladesh. *BMC Veterinary Research*. 2011; 7(1):27. <https://doi.org/10.1186/1746-6148-7-27> PMID: 21651757
71. Dinesh D, Ranjan A, Palit A, Kishore K, Kar SK. Seasonal and nocturnal landing/biting behaviour of *Phlebotomus argentipes* (Diptera: Psychodidae). *Annals of Tropical Medicine & Parasitology*. 2001; 95(2):197–202. <https://doi.org/10.1080/00034983.2001.11813629>
72. Ghosh K, Das S, Hati A. Studies on seasonal man sandfly (*Phlebotomus argentipes*) contact at night. *Journal of the Indian Association for Communicable Diseases*. 1982; 5:14–18.
73. Srinivasan R, Panicker K. Biting rhythm & biting activity of phlebotomid sandflies. *Indian Journal of Medical Research*. 1992; 95:301–304. PMID: 1291465
74. Shrestha S, Pant S. Seasonal distribution of Phlebotomine sand flies-vector of Visceral Leishmaniasis. *Journal of Nepal Medical Association*. 1994; 32(112):237–246.
75. Das ML, Deb M, Karki BMS, Sarif M, Khanal B, Bhattacharya SK, et al. Use of RK39 for diagnosis of post kala-azar dermal leishmaniasis in Nepal. *Southeast Asian Journal of Tropical Medicine and Public Health*. 2007; 38(4):619–625. PMID: 17882997
76. Palit A, Chakraborty S, Bhattacharya S, Chowdhury D, Hati A. Bait preference of *Phlebotomus argentipes* (Ann. & Brunn.). *Indian Journal of Public Health*. 1988; 32(1):41. PMID: 3243629
77. Kypraios T. *Efficient Bayesian Inference for Partially Observed Stochastic Epidemics and A New Class of Semi-Parametric Time Series Models*. Lancaster University; 2007. Available from: <http://eprints.lancs.ac.uk/26392/1/kypraios.pdf>.

78. Poche D, Garlapati R, Torres Z, Grant W, Wang HH, Mukherjee S, et al. The importance of systems ecology: implications of vector-based field studies and quantitative analysis of vector control programs. In: 6th World Congress on Leishmaniasis Abstracts Book, Toledo, Spain, May 16-20, 2017; 2017. p. 374. Available from: http://worldleish2017.org/documentos/Abstracts_Book_WL6_2017.pdf.
79. Mandal R, Das P, Kumar V, Kesari S. Spatial Distribution of Phlebotomus argentipes (Diptera: Psychodidae) in Eastern India, a Case Study Evaluating Multispatial Resolution Remotely Sensed Environmental Evidence and Microclimatic Data. *Journal of Medical Entomology*. 2017; p. 1–10.
80. Singh NS, Phillips-Singh D. Relative abundance of Phlebotominae sandflies with emphasis on vectors of kala-azar. *Asian Pacific Journal of Tropical Medicine*. 2010; 3(4):270–271. [https://doi.org/10.1016/S1995-7645\(10\)60065-X](https://doi.org/10.1016/S1995-7645(10)60065-X)
81. National Vector Borne Disease Control Programme. Guidelines on Vector Control in Kala-Azar Elimination;. Available from: <http://nvbdcp.gov.in/Doc/Guidelines-Vector-Control-KA-Elimination-FINAL.pdf> [cited 12/01/18].
82. Courtenay O, Peters NC, Rogers ME, Bern C. Combining epidemiology with basic biology of sand flies, parasites, and hosts to inform leishmaniasis transmission dynamics and control. *PLoS Pathogens*. 2017; 13(10):e1006571. <https://doi.org/10.1371/journal.ppat.1006571> PMID: 29049371
83. Courtenay O, Carson C, Calvo-Bado L, Garcez LM, Quinell RJ. Heterogeneities in Leishmania infantum Infection: Using Skin Parasite Burdens to Identify Highly Infectious Dogs. *PLoS Neglected Tropical Diseases*. 2014; 8(1):26. <https://doi.org/10.1371/journal.pntd.0002583>
84. Sudarshan M, Singh T, Singh AK, Chourasia A, Singh B, Wilson ME, et al. Quantitative PCR in Epidemiology for Early Detection of Visceral Leishmaniasis Cases in India. *PLoS Neglected Tropical Diseases*. 2014; 8(12):e3366. <https://doi.org/10.1371/journal.pntd.0003366> PMID: 25503103
85. Sudarshan M, Sundar S. Parasite load estimation by qPCR differentiates between asymptomatic and symptomatic infection in Indian visceral leishmaniasis. *Diagnostic Microbiology and Infectious Disease*. 2014; 80(1):40–42. <https://doi.org/10.1016/j.diagmicrobio.2014.01.031> PMID: 25023070
86. Sudarshan M, Weirather JL, Wilson ME, Sundar S. Study of parasite kinetics with antileishmanial drugs using real-time quantitative PCR in Indian visceral leishmaniasis. *Journal of Antimicrobial Chemotherapy*. 2011; 66(May):1751–1755. <https://doi.org/10.1093/jac/dkr185> PMID: 21609983
87. Rai K, Cuypers B, Bhattarai NR, Uranw S, Berg M, Ostyn B, et al. Relapse after treatment with Miltefosine for visceral Leishmaniasis is associated with increased infectivity of the infecting Leishmania donovani strain. *mBio*. 2013; 4(5):7–10. <https://doi.org/10.1128/mBio.00611-13>
88. Verma S, Kumar R, Katara GK, Singh LC, Negi NS, Ramesh V, et al. Quantification of parasite load in clinical samples of leishmaniasis patients: Il-10 level correlates with parasite load in visceral leishmaniasis. *PLoS ONE*. 2010; 5(4). <https://doi.org/10.1371/journal.pone.0010107>
89. Pampiglione S, Manson-Bahr PEC, La Placa M, Borgatti MA, Musumeci S. Studies in mediterranean leishmaniasis. 3. The leishmanin skin test in kala-azar. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1975; 69(1):60–68. PMID: 1145719
90. Bern C, Amann J, Haque R, Chowdhury R, Ali M, Kurkjian KM, et al. Loss of leishmanin skin test antigen sensitivity and potency in a longitudinal study of visceral leishmaniasis in Bangladesh. *American Journal of Tropical Medicine and Hygiene*. 2006; 75(4):744–8. <https://doi.org/10.4269/ajtmh.2006.75.744> PMID: 17038705
91. Ali A, Ashford RW. Visceral leishmaniasis in Ethiopia. II. Annual leishmanin transformation in a population. Is positive leishmanin reaction a life-long phenomenon? *Annals of Tropical Medicine & Parasitology*. 1993; 87(2):163–167.
92. Hailu A, Gramiccia M, Kager PA. Visceral leishmaniasis in Aba-Roba, south-western Ethiopia: prevalence and incidence of active and subclinical infections. *Annals of Tropical Medicine & Parasitology*. 2009; 103(8):659–670. <https://doi.org/10.1179/000349809X12554106963555>
93. O'Neill PD, Roberts GO. Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society A*. 1999; 162:121–129. <https://doi.org/10.1111/1467-985X.00125>
94. Britton T, O'Neill PD. Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure. *Scandinavian Journal of Statistics*. 2002; 29(1998):375–390. <https://doi.org/10.1111/1467-9469.00296>
95. Neal PJ, Roberts GO. Statistical inference and model selection for the 1861 Hagelloch measles epidemic. *Biostatistics*. 2004; 5(2):249–261. <https://doi.org/10.1093/biostatistics/5.2.249> PMID: 15054029
96. Jewell CP, Kypraios T, Neal P, Roberts GO. Bayesian Analysis for Emerging Infectious Diseases. *Bayesian Analysis*. 2009; 4(4):465–496. <https://doi.org/10.1214/09-BA417>

97. Spencer SEF, Besser TE, Cobbold RN, French NP, Spencer SEF. 'Super' or just 'above average'? Supershedders and the transmission of *Escherichia coli* O157:H7 among feedlot cattle. *Journal of the Royal Society Interface*. 2015; 12:20150446. <https://doi.org/10.1098/rsif.2015.0446>
98. Chis Ster I, Singh BK, Ferguson NM. Epidemiological inference for partially observed epidemics: The example of the 2001 foot and mouth epidemic in Great Britain. *Epidemics*. 2009; 1(1):21–34. <https://doi.org/10.1016/j.epidem.2008.09.001> PMID: 21352749
99. Walker PGT, Cauchemez S, Métras R, Dung DH, Pfeiffer D, Ghani AC. A Bayesian Approach to Quantifying the Effects of Mass Poultry Vaccination upon the Spatial and Temporal Dynamics of H5N1 in Northern Vietnam. *PLoS Computational Biology*. 2010; 6(2):e1000683. <https://doi.org/10.1371/journal.pcbi.1000683> PMID: 20174604
100. Datta S, Bull JC, Budge GE, Keeling MJ. Modelling the spread of American foulbrood in honeybees. *Journal of the Royal Society Interface*. 2013; p. 20130650. <https://doi.org/10.1098/rsif.2013.0650>
101. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*. 2002; 64(4):583–639. <https://doi.org/10.1111/1467-9868.00353>
102. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*. CRC Press; 2014.
103. Celeux G, Forbes F. Deviance information criteria for missing data models. *Bayesian Analysis*. 2006; 1(4):651–673. <https://doi.org/10.1214/06-BA122>
104. MRC Biostatistics Unit, University of Cambridge. The BUGS Project: DIC; 2006. Available from: <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/#q9> [cited 15/01/2018].
105. Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD. The deviance information criterion: 12 years on (with discussion). *Journal of the Royal Statistical Society: Series B*. 2014; 64:485–493.
106. Bulstra CA, Malaviya P, Le Rutte EA, Hasker EC, Picado A, Singh OP, et al. Visceral leishmaniasis: spatiotemporal heterogeneity and drivers underlying the hotspots in Muzaffarpur, Bihar, India. *PLoS Neglected Tropical Diseases*. 2018; Forthcoming
107. Ostyn B, Gidwani K, Khanal B, Picado A, Chappuis F, Singh SP, et al. Incidence of symptomatic and asymptomatic *Leishmania donovani* infections in High-Endemic foci in India and Nepal: A prospective study. *PLoS Neglected Tropical Diseases*. 2011; 5(10):1–7. <https://doi.org/10.1371/journal.pntd.0001284>
108. Singh S, Kumari V, Singh N. Predicting Kala-Azar Disease Manifestations in Asymptomatic Patients with Latent *Leishmania donovani* Infection by Detection of Antibody against Recombinant K39 Antigen. *Clinical and Diagnostic Laboratory Immunology*. 2002; 9(3):568–572. <https://doi.org/10.1128/CDLI.9.3.568-572.2002> PMID: 11986261
109. Saha P, Ganguly S, Chatterjee M, Das SB, Kundu K, Guha SK, et al. Asymptomatic leishmaniasis in kala-azar endemic areas of Malda district, West Bengal, India. *PLoS Neglected Tropical Diseases*. 2017; 11(2): e0005391. <https://doi.org/10.1371/journal.pntd.0005391>
110. Schwab SR, Stone CM, Fonseca DM, Fefferman NH. The importance of being urgent: the impact of surveillance target and scale on mosquito-borne disease control. *Epidemics*. 2017;.
111. Barnett PG, Singh SP, Bern C, Hightower AW, Sundar S. Virgin soil: the spread of visceral leishmaniasis into Uttar Pradesh, India. *American Journal of Tropical Medicine and Hygiene*. 2005; 73(4):720–725. <https://doi.org/10.4269/ajtmh.2005.73.720> PMID: 16222016
112. Islam S, Kenah E, Bhuiyan MAA, Rahman KM, Goodhew B, Ghalib CM, et al. Clinical and immunological aspects of post-kala-azar dermal leishmaniasis in Bangladesh. *American Journal of Tropical Medicine and Hygiene*. 2013; 89(2):345–353. <https://doi.org/10.4269/ajtmh.12-0711> PMID: 23817330
113. Chowdhury R, Dotson E, Blackstock AJ, McClintock S, Maheswary NP, Faria S, et al. Comparison of insecticide-treated nets and indoor residual spraying to control the vector of visceral leishmaniasis in Mymensingh District, Bangladesh. *American Journal of Tropical Medicine and Hygiene*. 2011; 84(5):662–667. <https://doi.org/10.4269/ajtmh.2011.10-0682> PMID: 21540372