

Antibody and antigen prevalence as indicators of ongoing transmission or elimination of visceral leishmaniasis: a modelling study

Luc E. Coffeng,^{a,#} Epke A. Le Rutte,^{a,b,c} Johanna Munoz,^a Emily Adams,^d Sake J. de Vlas^a

^a Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

^b Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland

^c University of Basel, Basel, Switzerland

^d Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Corresponding author: Luc E. Coffeng, Department of Public Health, Erasmus MC, University Medical Center Rotterdam, P.O. box 2040, 3000 CA Rotterdam, The Netherlands; l.coffeng@erasmusmc.nl.

Running title

Indicators of VL transmission

Summary

Risk of recrudescence of VL after scaling down control efforts can be monitored by means of population-based surveys measuring seroprevalence or antigenaemia. This study highlights the potential value of point-of-care diagnostic tools for sustainable VL control.

© The Author(s) 2021. Published by Oxford University Press for the Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Control of visceral leishmaniasis (VL) on the Indian subcontinent has been highly successful. Control efforts such as indoor residual spraying and active case detection will be scaled down or even halted over the coming years. We explored how after scale-down, potential recurrence of VL cases may be predicted based on population-based surveys of antibody or antigenaemia prevalence.

Methods: Using a stochastic age-structured transmission model of VL, we predicted trends in case incidence and biomarker prevalence over time after scaling down control efforts when the target of three successive years without VL cases has been achieved. Next, we correlated biomarker prevalence with the occurrence of new VL cases within 10 years of scale-down.

Results: Occurrence of at least one new VL case in a population of 10,000 was highly correlated with the seroprevalence and antigenaemia prevalence at the moment of scale-down, or one or two years afterwards. Receiver operating characteristic curves indicated that biomarker prevalence in adults provided the most predictive information, and seroprevalence was a more informative predictor of new VL cases than antigenaemia prevalence. Thresholds for biomarker prevalence to predict occurrence of new VL cases with high certainty were robust to variation in pre-control endemicity.

Discussion: The risk of recrudescence of VL after scaling down control efforts can be monitored and mitigated by means of population-based surveys. Our findings highlight that rapid point-of-care diagnostic tools to assess (preferably) seroprevalence or (otherwise) antigenaemia in the general population could be a key ingredient of sustainable VL control.

Key words

Visceral leishmaniasis; control; monitoring; serology; antigenaemia.

Background

Visceral leishmaniasis (VL), also known as kala-azar, is a vector-borne infection that can lead to long-lasting fevers and death if left untreated, although most infected individuals will remain asymptomatic [1,2]. In the Indian sub-continent, where the disease is considered to only occur in humans, about 5-10% of those who are treated for VL develop post-kala-azar dermal leishmaniasis (PKDL), an infectious but self-limiting skin condition [3]. In 2005, India, Nepal, and Bangladesh committed to controlling VL by signing a trilateral Memorandum of Understanding. The associated continent-wide control measures led to a steep decrease of reported VL incidence from 32,803 in 2005 to 3,128 in 2019, representing a 90% drop [4]. Control measures consist of vector control through indoor-residual spraying (IRS) of insecticide and active case detection (ACD) followed by prompt treatment, which is provided for free [5]. The goal is to achieve elimination of VL as a public health problem on the Indian subcontinent, which is defined as <1 new or recurring VL case per 10,000 individuals per year at district or sub-district level [6]. To validate achievement of the target, the World Health Organization (WHO) requires that the VL incidence in a region is below the target for three years in a row in combination with extensive testing. After validation, control measures can be scaled down [7].

Many Indian sub-districts have entered or will be entering this validation phase in the coming years. The prospect of scaling down or even halting control measures, without complete interruption of transmission, poses a risk of recrudescence of infection. This highlights the need for continued monitoring of ongoing infection. In this study we assess the predictive power of population surveys employing serological tests (e.g. the direct agglutination test or DAT for antileishmanial antibody detection) or antigen tests to predict the occurrence of new VL cases (reported or unreported) after scaling down control efforts. To this end we developed a new stochastic age-structured VL transmission model, based on an established deterministic VL transmission model [8–11].

Methods

Model structure and quantification

We use an established age-structured deterministic VL transmission model that describes the transmission of VL between humans and sand flies on the Indian subcontinent [8–11]. To realistically simulate the probability of interrupting VL transmission, we developed a new stochastic version of the model, considering a finite and discrete number of homogeneously mixing human individuals while keeping the sand fly part of the model deterministic. In the model, most infected individuals remain asymptomatic and recover without ever having symptoms; a small fraction (~1.5%) of individuals become symptomatic and will require treatment or die otherwise; a small fraction (~3%) of symptomatic cases is assumed to recover spontaneously. Transmission is driven by exposure to sand flies, where sand fly abundance is assumed to peak from July to September of each year [12]. Sand flies can pick up infection from symptomatic cases, as well as individuals with PKDL, a self-limiting but long-lasting skin condition that occurs in a fraction of individuals treated for VL (2.5% in the model [3]). In addition, we distinguish the possibility that asymptotically infected individuals

do (model E1) or do not (model E0) contribute to transmission. The model incorporates IRS coverage through a proportional reduction in the sand fly population density, and ACD through a decrease in the average detection delay of symptomatic cases (pre-control delay of 60 days [9,13]). In the model, we assume that seropositivity based on the DAT is associated with the late asymptomatic stage of infection (i.e. after individuals have been infected for a while), as well as the symptomatic stages and the early recovered stage (after which DAT positivity, and thus the humoral immune response, is lost again) [8]. DAT positivity was defined at the 1:800 titre cut-off, instead of the standard of 1:1600, which increases test sensitivity but decreases specificity. In contrast, we assume that antigenaemia (i.e. the persistence of antigen in circulating blood) is only associated with the late asymptomatic stage and the symptomatic stages when parasite loads are presumably highest. The model assumes that no infection can be introduced from outside the population.

Model parameters were previously calibrated based on age-structured data from approximately 21,000 individuals included in the KalaNet bednet trial in India and Nepal [8,14,15]. Age patterns in DAT and PCR positivity (where PCR positivity included the stage of infection that we consider antigen positive here, along with the early stage of asymptomatic infection) were reproduced assuming that exposure to bites of sand flies increase linearly from zero at birth to a maximum value at age 20, and plateaus thereafter (crudely reflecting change of body surface area with age). The impact of IRS was previously estimated using a geographical cross-validation on decreasing case incidence in Bihar (~6,000 VL cases in 8 districts over a period of 18 months) [9,13]. A schematic representation of the model structure and an overview of model parameter values are presented in the Supplementary Information (Figure S1 and Table S1, respectively).

The model was coded in R (version 4.0.2), using the *pomp* package (version 3.1.1.7); the model code is publicly accessible at <https://gitlab.com/erasmusmc-public-health/vl-serosurveys>. We provided the Policy-Relevant Items for Reporting Models in Epidemiology of Neglected Tropical Diseases (PRIME-NTD) Summary in Table S2, which was previously established to set a standard and increase consistency among modelling studies that aim to inform policy [16].

Simulation scenarios

With each version of the stochastic transmission model (E0 and E1), we performed 10,000 repeated simulations for a stable population of 10,000 people. Transmission conditions (i.e. the sand fly to human ratio) were allowed to vary randomly between simulations such that the pre-control VL incidence ranged from 2 to 15 reported VL cases per 10,000 capita per year, as expected under the deterministic version of the model. During the pre-control phase we assumed that only passive case detection was in place, leading to an average duration between the start of symptoms and the start of treatment of 60 days [13]. To start a stochastic simulation, we extracted the expected pre-control equilibrium state of the human and fly population (i.e. distribution over model compartments) from a deterministic simulation based on the same transmission conditions. This population state was used to seed 10,000 individuals across age and disease compartments via a draw from a multinomial

distribution. Based on this seeding, we ran the stochastic model, implementing control measures as recommended by WHO until zero VL cases were observed (i.e. only reported cases) for three consecutive calendar years (i.e. a period starting on January 1st and ending in December 31st three years later). Control efforts were assumed to start with a 5-year attack phase in which we assumed that IRS covers 67% of the population and that ACD reduces the time to treatment to an average of 45 days. Subsequently, during the so-called consolidation phase, IRS coverage was reduced to 45%, but ACD efforts are further increased leading to an average duration to treatment of 30 days. Achievement of the target (three consecutive years with zero reported VL cases, with rigorous ACD) was assessed from completion of the attack phase onwards, meaning that at least eight years of control had been performed before any control programme was allowed to scale down. If a simulation did not achieve the target of three consecutive years with zero reported VL cases within 20 years of control efforts (which happens more often if asymptomatic infections do not contribute to transmission [9], i.e. in model E0), it was discarded and excluded from further analysis. For those simulations in which the target was met, control efforts were scaled down to the pre-control situation (no IRS and no ACD such that the average detection delay returned to 60 days) and the simulation was continued for another 10 years to see if any new VL cases (reported or unreported) would occur.

To see if the occurrence of at least one new VL case could be predicted based on prevalence of biomarkers in the population, we saved model-predicted trends in age-specific prevalence of DAT and antigenaemia prevalence. Receiver operating characteristic (ROC) curves as well as positive (PPV) and negative predictive values (NPV) of thresholds for biomarker prevalence were calculated for three different time points: at the moment of scaling down control efforts, and one and two years after the scale-down. For the last two time points, we excluded simulations in which one or more VL cases were already detected before the time of the survey, assuming that this observation alone would already lead to policy action. For each survey, we assumed that a random subset of 500 individuals in the population is tested, simulating 100 random surveys for each time point in each simulation. Simulated surveys sampled either only pre-school age children (ages 0-4), school-age children (ages 5-14), or adults (ages ≥ 15). Threshold values were expressed in terms of the number of biomarker-positive individuals among this sample of 500 at or above which we expect recurrence of at least one VL case in the 10 years following scaling down of control efforts.

Results

Out of the 10,000 repeated stochastic simulations that were performed with each model variant, in 8,011 (model E1) and 4,705 simulations (model E0) the target of three consecutive years of zero reported VL cases was achieved within 20 years of control and control could be scaled down to passive case detection only. The probability of at least one new VL case occurring within 10 years after scaling down control was higher in model E0 ($2,936 / 4,705 = 62.4\%$) than in model E1 ($3,028 / 8,011 = 37.8\%$). If one or more new VL cases occurred, the first case typically occurred within the first year after scaling down (96.5% of simulations in model E1 and 69.3% in model E0), but with a very long right tail (up to 5 years for model E1 and 10 years for model E0) (Figure S2). The timing of

the first detected case was slightly later, although still typically within one year after scaling down (93.7% of simulations in model E1 and 60.3% in model E0).

If a new VL case occurred after scaling down control, model E1 predicted that DAT prevalences were above 0% for at least two years after scale down (100% of simulations, Table S3) and fluctuated and rose over time (Figure 1). DAT prevalences reached higher values in adults (age 15+) than in pre-school (age 0-4) and school-aged children (age 5-14), reflecting the assumed age patterns in exposure to sand fly bites. If no new VL cases occurred and DAT prevalences were above 0% at the moment of scale-down (45% of the simulations), the DAT prevalence quickly declined to values close to zero within one to two years. Trends in antigenaemia prevalence followed a similar pattern, although with overall lower prevalences compared to DAT (Figure S3), reflecting that the duration of antigen-positivity is shorter than that of seropositivity. If no new VL case occurred after scale-down, antigenaemia prevalence was rarely above 0% at the moment of scale-down (13% of simulations); if a new VL case did occur, antigenaemia prevalence was almost always above 0% (99% of simulations; Table S4). Predicted trends in biomarker prevalence and the differences in trends between simulation outcomes (absence vs. occurrence of new VL cases) were qualitatively similar between model E1 (Figure 1 and Figure S3) and model E0 (Figures S4 and S5). DAT prevalences in children < 15 were slightly lower at the turn of each year due to the fact that at that moment in the simulation, for efficiency, all individuals were assumed to simultaneously age by one year. This modelling artefact is not visible for antigenaemia prevalence due to the higher level of stochastic noise associated with overall lower prevalences.

Sensitivity and specificity of different thresholds for biomarker prevalence as predicted by model E1 for samples of 500 individuals are illustrated as ROC curves in Figure 2. Because biomarker prevalences were generally higher in adults (age 15+) and thus suffered less from stochastic noise, sensitivity and specificity of thresholds were highest when applied to prevalences in that particular age group. This pattern was qualitative similar for model E0 (Figure S6), although sensitivity and specificity were lower overall than in model E1 and the difference in ROC curves based on biomarker prevalences in adults (age 15+) and school-aged children (age 5-14) was somewhat larger. Still, the two model variants agreed that a threshold of 1 to 3 biomarker-positive cases in a sample of 500 was optimal to achieve both high sensitivity and specificity, and that DAT prevalence was a more informative predictor of recurrence of VL cases than antigenaemia prevalence. Furthermore, the model variants agreed that when biomarker prevalence was measured one or two years after scaling down control (excluding simulations in which a new VL case had already been detected before that time), sensitivity decreased strongly and specificity increased slightly.

The PPV and NPV of thresholds for biomarker prevalence, as predicted by model E1, depended on both the biomarker as well as the time of measurement (Figure 3). In general, NPVs (i.e. the probability of no more VL cases if the biomarker prevalence was under the threshold) were higher for DAT (up to 99%) than for antigenaemia prevalence (up to 95%), as the latter were limited by the fact the prevalence of antigen positivity is relatively low compared to DAT prevalence. For DAT

prevalence, a threshold value of 6 positive individuals in a sample of 500 corresponded to a $\geq 95\%$ probability that no further VL cases could occur, if DAT prevalence was measured at the moment of scale-down. For antigenaemia, this NPV of 95% could only just be achieved by allowing at most 1 positive individual. The corresponding PPVs (i.e. the probability of recurring VL cases if biomarker prevalence is above or equal to the threshold) were $\geq 94\%$ for both biomarkers. For both biomarkers, when prevalence was measured one or two years after scale-down, the PPV decreased and the NPV increased and flattened as a function of threshold value (i.e., if no new VL case had been detected by then, the a priori probability of recurrence had already strongly declined). Alternatively, when assuming that asymptomatic infections do not contribute to transmission (model E0), in general, NPVs were lower for DAT (80% to 90% for a threshold of 1 case) and particularly lower for antigenaemia (60% to 70% for a threshold of 1 case), although PPVs were still high ($\geq 90\%$) (Figure S7). In contrast to model E1, in model E0, measuring biomarker prevalence one or two years after scaling down only marginally affected the PPV, yet still increased the NPV. Last, PPVs were robust to variation in pre-control endemicity both model E1 and E0 (Figures S8 and S9), provided that the biomarker and associated threshold prevalence are chosen aiming at a PPV and NPV $\geq 90\%$. Overall, NPVs were slightly higher for settings with high pre-control endemicity, reflecting that achieving a prevalence under the threshold despite more intense transmission conditions was highly indicative of successful control.

Discussion

We show that occurrence of new VL cases (reported or unreported) after achieving three years of zero reported VL cases and scaling down control efforts is highly correlated with the seroprevalence (based on DAT) and antigenaemia prevalence. Biomarker prevalence in adults (ages ≥ 15) provides the most predictive information on prospects of resurgence, and DAT prevalence appears to be a more informative predictor than antigenaemia prevalence. Thresholds for biomarker prevalence to predict occurrence of new VL cases with high certainty were robust to variation in pre-control endemicity, allowing thresholds to be applied without exact knowledge of transmission conditions.

For an area covering a population of 10,000 people, having no DAT-positive individuals in a sample of 500 (i.e. a threshold of <1 case) at the moment of scale-down is a conservative approach to predicting whether or not new VL cases will occur. However, the predictive power of such a criterion does depend on whether or not asymptomatic individuals do (NPV $\geq 99\%$, PPV $\geq 79\%$) or do not contribute to transmission (NPV $\geq 80\%$, PPV $\geq 92\%$). In general, the predictive value could be further boosted by sampling more individuals (and adapting the threshold criterion accordingly) or by performing repeated surveys over time (e.g. annually). Of course, as this is a modelling study and given that no monitoring programme exists as suggested here, predictions should be validated based on field data, ideally longitudinal data of biomarker prevalence from areas that have implemented the control measures as recommended by the WHO.

A major strength of this study is that the calculated threshold values for the number of biomarker-positive cases in a sample of 500 individuals can be readily used for interpretation of field data (based on the same sample size) without the need for calculating confidence bounds around prevalence estimates. This is possible because stochastic as well as statistical uncertainty are both captured in our calculations: stochastic uncertainty via many repeated simulations of the population dynamics over time, and statistical uncertainty via 100 repeated survey simulations within each transmission simulation. In addition, the methodology employed in this paper can be easily adapted to other population sizes and survey sample sizes. However, for much larger population sizes, model predictions should ideally be based on a metapopulation model that accounts for the fact that individuals live in subpopulations and do not mix homogeneously. Such a model would also address the most important limitation of this study that we assume a closed population without the possibility of introduction of infection from outside. Recrudescence of VL transmission through introduction via human mobility is a real risk, unless the strategy is successfully rolled-out at a very large scale. As such, the threshold for biomarker prevalence proposed here should be interpreted as an indicator of recurrence of VL transmission through local transmission. The metapopulation model that is needed to relax this assumption is current being developed.

Our finding that thresholds for biomarker prevalence were quite robust to variation in pre-control endemicity is extremely convenient for policy development and implementation of monitoring in field settings. This robustness is in stark contrast with, for example, thresholds for Ov16-seroprevalence to evaluate progress towards elimination of onchocerciasis, for which threshold values for a particular target predictive value strongly change with pre-control endemicity [17]. This is not the case for VL because of the precondition of three consecutive years of zero cases before scaling down control efforts. This ensures that if the target is met despite a high transmission potential, it is indicative of having interrupted transmission with high probability. Of course, this is conditional on meeting the target through high detection effort rather than low detection effort (causing fewer cases to be reported), for which we previously showed that it can lead to a “confirmatory” drop in case rates while true case numbers are rising [18]. Then again, if the monitoring strategy developed in this study would be applied to a situation where the target was met due to poor case detection, we expect that this would most likely be immediately reflected by a high biomarker prevalence.

According to our findings, measuring biomarker prevalence at one or two years after scaling down control efforts (as long as no new VL cases have been detected in the meanwhile) increases the predictive power for absence of new VL cases (i.e. NPV), which was already very high when biomarker (especially DAT) prevalence was assessed at the moment of scale-down. This finding was mostly driven by the fact that if VL transmission resurged, it did so mostly within one year after scale-down; in other words, if no new VL cases were reported in the first year after scale-down, the probability of resurgence was very low to begin with, which drove up the NPV (and lowered the PPV). It is therefore important to note that the benefit of this higher NPV comes at the cost of delayed information, giving the infection more opportunity to be transmitted further throughout the population in case of resurgence. Given that in case of VL recurrence, this typically happened within

one year of scaling down, we recommend that a biomarker survey is performed at the moment of scale-down or slightly before. However, if the biomarker survey is performed much earlier than the 3-year benchmark (say, one year before), threshold values for the number of positive cases would have to be adapted and might even depend more on pre-control endemicity.

It is important to note that for the sake of illustrating the potential predictive value of biomarkers, in our simulations, we assumed that diagnostic tests for antibodies or antigens are 100% sensitive and specific at the individual level. However, in reality, false positive and false negative results should certainly be expected to occur, adding noise to observations and lowering the PPV and NPV of biomarker prevalence for recurrence of VL cases, compared to what we estimate here. Also, it should be noted that our model predictions for DAT prevalence are quantified based on the diagnostic techniques and titre thresholds used to calibrate our model to data from the KalaNet study [14,15]. For a more sensitive monitoring strategy (i.e., a higher chance of correctly identifying settings where new VL cases will occur), it might be advantageous to use a lower titre threshold for DAT positivity, although this would come at a cost of lower specificity (i.e., a higher risk of continuing control efforts longer than strictly necessary). An added benefit would be that the use of a lower titre threshold would result in higher prevalences overall, which would allow for a higher threshold of DAT prevalence in the survey sample and thus requiring a smaller sample size to conclude that the prevalence is statistically significant under the threshold. It is possible to capture this in our model but requires further model development and calibration using more detailed data on actual DAT titres. Data from the follow-up KalaNet study [19] will be particularly valuable here, as it would also allow us to further validate model predictions over time. Antigenaemia on the other hand was not found to be as good a predictive marker as DAT. This is presumably because individuals have a high detectable antigen load for a short period of time, especially when compared to the duration of detectable antibodies, and therefore the cycle of surveillance may not capture positive cases, especially with low numbers of participants surveyed. Antigen tests could instead be used to confirm findings of these population surveys, as a highly specific test.

Our findings highlight the need for further operational research into effective monitoring to ensure sustained control of VL in the Indian subcontinent. First of all, given our finding that population-based surveys can be an extremely valuable monitoring tool when scaling down control, the question arises what geographical units such surveys should cover and how surveys should be implemented. For instance, how many communities and how many individuals should be selected per community? This would depend on the expected level of geographical clustering of VL transmission after prolonged control, which needs to be informed by field studies like the recent KalaNet follow-up study [19]. Second, how feasible is it to sample, say, 500 individuals per 10,000 population? Sampling 5% of the population may be practically too demanding or too expensive, especially when surveys have to rely on lab facilities. Therefore, there is an urgent need for rapid point-of-care diagnostic tools to assess seroprevalence or antigenaemia in the general population. Of course, to define useful threshold values for biomarker prevalence measured by such tools, their sensitivity and specificity need to be carefully quantified. Third, there is a need to define more specific policy actions for when biomarker surveys indicate ongoing VL transmission while no VL

cases have been detected yet. These actions may depend on the encountered prevalence of biomarker positivity: if biomarker positivity is highly clustered, more in-depth follow-up surveys of the locations with high biomarker prevalence may be an adequate first step; if biomarker positivity is more homogeneously distributed over a larger area, more immediate action in terms of reinstating IRS and/or ACD across the entire geographical area may be more appropriate.

We conclude that the risk of recrudescence of VL after scaling down control efforts can be monitored and mitigated by means of population-based surveys. We recommend that such surveys are based on biomarkers of current and recent infection, such as antibodies, as prevalences for such biomarkers are higher than antigen prevalences and thus provide more statistical information. Our findings highlight that rapid point-of-care diagnostic tools to assess (preferably) seroprevalence or (otherwise) antigenaemia in the general population could be a key ingredient of sustainable VL control.

Accepted Manuscript

NOTES

Authors' contributions

LEC conceived the study. JM and LEC programmed the stochastic model, ran the simulations, performed the analyses, and drafted the figures. LEC, EALR, ERA, and SJdV interpreted the results. LEC and EALR drafted the manuscript. All authors critically revised the manuscript for intellectual content, and read and approved of the final manuscript.

Funding

The authors gratefully acknowledge funding of the NTD Modelling Consortium by the Bill and Melinda Gates Foundation (OPP1184344). LEC further acknowledges funding from the Dutch Research Council (NWO, grant 016.Veni.178.023).

Competing interests

The authors declare no competing interests.

Accepted Manuscript

References

1. WHO SEARO. WHO | Leishmaniasis fact sheet.
2. Ostyn B, Gidwani K, Khanal B, et al. Incidence of Symptomatic and Asymptomatic *Leishmania donovani* Infections in High-Endemic Foci in India and Nepal: A Prospective Study. *PLoS Negl Trop Dis* **2011**; 5:e1284.
3. Zijlstra EE, Alves F, Rijal S, Arana B, Alvar J. Post-kala-azar dermal leishmaniasis in the Indian subcontinent: A threat to the South-East Asia Region Kala-azar Elimination Programme. *PLoS Negl Trop Dis* **2017**; 11:e0005877.
4. WHO. Eliminating visceral leishmaniasis: India takes decisive steps to overcome last-mile challenges. 2020.
5. Minister of Health & Family Welfare Government of India. National roadmap for Kala-azar elimination. Delhi: 2014.
6. WHO. Ending the neglect to attain the Sustainable Development Goals: A road map for neglected tropical diseases 2021-2030. *World Heal Organ* **2020**;
7. World Health Organization Regional Office for South-East Asia. Process of validation of elimination of kala-azar SEARO. 2016.
8. Le Rutte EA, Coffeng LE, Bontje DM, et al. Feasibility of eliminating visceral leishmaniasis from the Indian subcontinent: explorations with a set of deterministic age-structured transmission models. *Parasit Vectors* **2016**; 9:24.
9. Le Rutte EA, Chapman LAC, Coffeng LE, et al. Elimination of visceral leishmaniasis in the Indian subcontinent: a comparison of predictions from three transmission models. *Epidemics* **2017**; 18:67–80.
10. Le Rutte EA, Chapman LAC, Coffeng LE, et al. Policy Recommendations From Transmission Modeling for the Elimination of Visceral Leishmaniasis in the Indian Subcontinent. *Clin Infect Dis* **2018**; 66:S301–S308.
11. Le Rutte EA, Zijlstra EE, de Vlas SJ. Post-Kala-Azar Dermal Leishmaniasis as a Reservoir for Visceral Leishmaniasis Transmission. *Trends Parasitol* **2019**; 35:590–592.
12. Poché DM, Garlapati RB, Mukherjee S, et al. Bionomics of *Phlebotomus argentipes* in villages in Bihar, India with insights into efficacy of IRS-based control measures. *PLoS Negl Trop Dis* **2018**; 12:e0006168.
13. Jervis S, Chapman LAC, Dwivedi S, et al. Variations in visceral leishmaniasis burden, mortality and the pathway to care within Bihar, India. *Parasit Vectors* **2017**; 10:601.
14. Picado A, Das ML, Kumar V, et al. Effect of Village-wide Use of Long-Lasting Insecticidal Nets on Visceral Leishmaniasis Vectors in India and Nepal: A Cluster Randomized Trial. *PLoS Negl Trop Dis* **2010**; 4:e587.

15. Picado A, Singh SP, Rijal S, et al. Longlasting insecticidal nets for prevention of *Leishmania donovani* infection in India and Nepal: paired cluster randomised trial. *BMJ* **2010**; 341:c6760–c6760.
16. Behrend MR, Basáñez M-G, Hamley JID, et al. Modelling for policy: The five principles of the Neglected Tropical Diseases Modelling Consortium. *PLoS Negl Trop Dis* **2020**; 14:e0008033.
17. Coffeng LE, Stolk WA, Golden A, de Los Santos T, Domingo GJ, de Vlas SJ. Predictive Value of Ov16 Antibody Prevalence in Different Sub-Populations for Elimination of African Onchocerciasis. *Am J Epidemiol* **2019**; 351:843.
18. Coffeng LE, Le Rutte EA, Muñoz J, et al. Impact of Changes in Detection Effort on Control of Visceral Leishmaniasis in the Indian Subcontinent. *J Infect Dis* **2020**; 221:S546–S553.
19. Cloots K, Uranw S, Ostyn B, et al. Impact of the visceral leishmaniasis elimination initiative on *Leishmania donovani* transmission in Nepal: a 10-year repeat survey. *Lancet Glob Heal* **2020**; 8:e237–e243.

Accepted Manuscript

Figure legends

Figure 1. Model-predicted trends in age-specific DAT prevalence after scaling down control efforts against visceral leishmaniasis. Lines represent biomarker prevalence from a randomly selected subset of 500 simulations. Rows represent different age categories; columns depict simulations that resulted in occurrence (left) or absence of new VL cases (right), with the total number of simulations per outcome indicated at the top of each column (N). Predictions are based on the assumptions that both symptomatic and asymptomatic infections contribute to transmission (model E1) and that all individuals are tested. Similar predictions assuming asymptomatic infections do not contribute to transmission (model E0) can be found in Figure S4.

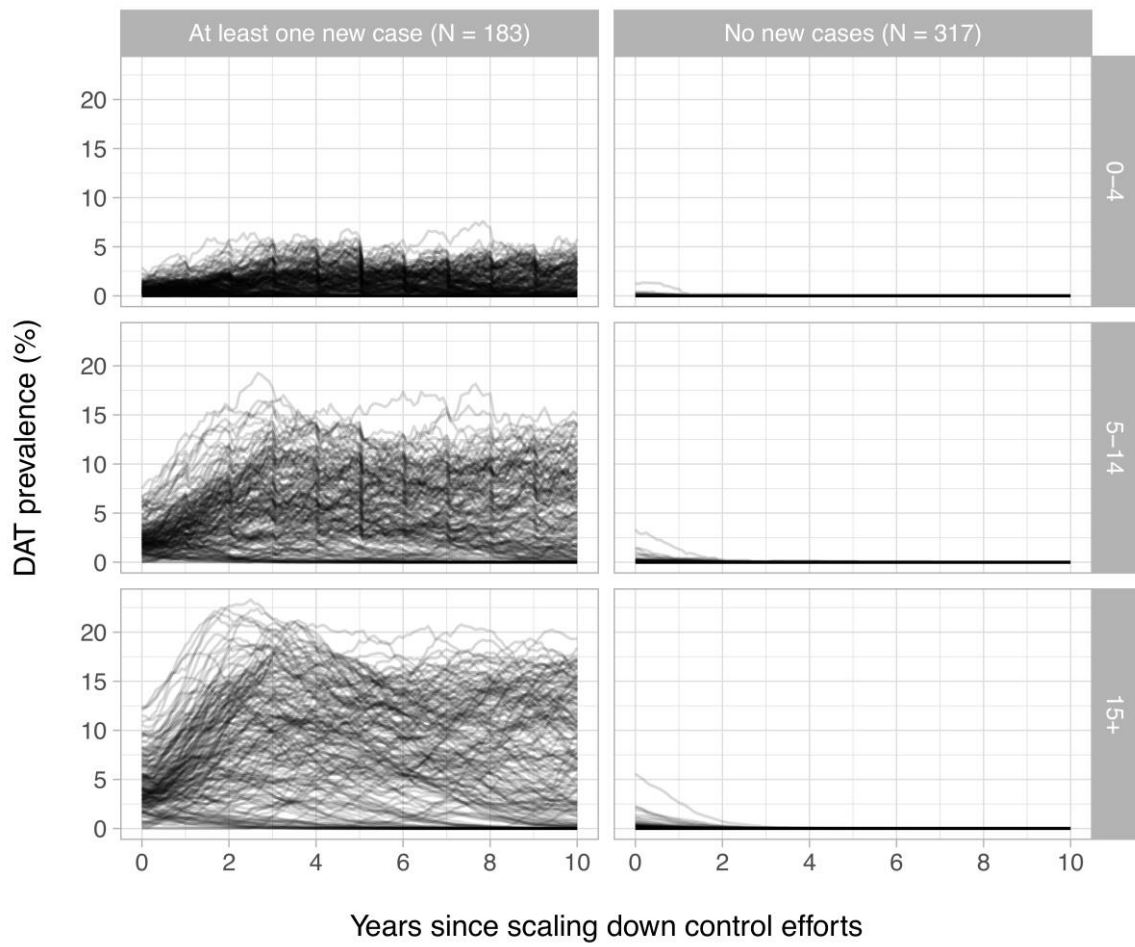
Figure 2. Receiver-operator curve for predicting occurrence of new VL cases based on age-specific prevalence of DAT or antigenaemia measured at and up to two years after scaling down control efforts. Columns depict receiver-operator curves (ROC) based on biomarkers measured on three different time points; rows depict different biomarkers. Symbols indicate thresholds for the number (N) of biomarker-positive cases at or above which the recurrence of at least one VL case was predicted. Predictions are based on the assumptions that both symptomatic and asymptomatic infections contribute to transmission (model E1) and that 500 individuals are tested for biomarker positivity. Similar predictions assuming asymptomatic infections do not contribute to transmission (model E0) can be found in Figure S6.

Figure 3. Positive (PPV) and negative predictive value (NPV) of DAT and antigenaemia prevalence in adults (age 15+) measured up to two years after scaling down control efforts, given a choice of threshold value. Columns depict curves based on biomarkers measured on three different time points; rows depict different biomarkers. Note that the predictive values based on biomarker prevalences measured one or two years after scale-down (middle and right panels) are conditional on no new VL cases having been detected since scale-down. Predictions are based on the assumptions that both symptomatic and asymptomatic infections contribute to transmission

(model E1) and that 500 individuals are tested for biomarker positivity. Similar predictions assuming asymptomatic infections do not contribute to transmission (model E0) can be found in Figure S7.

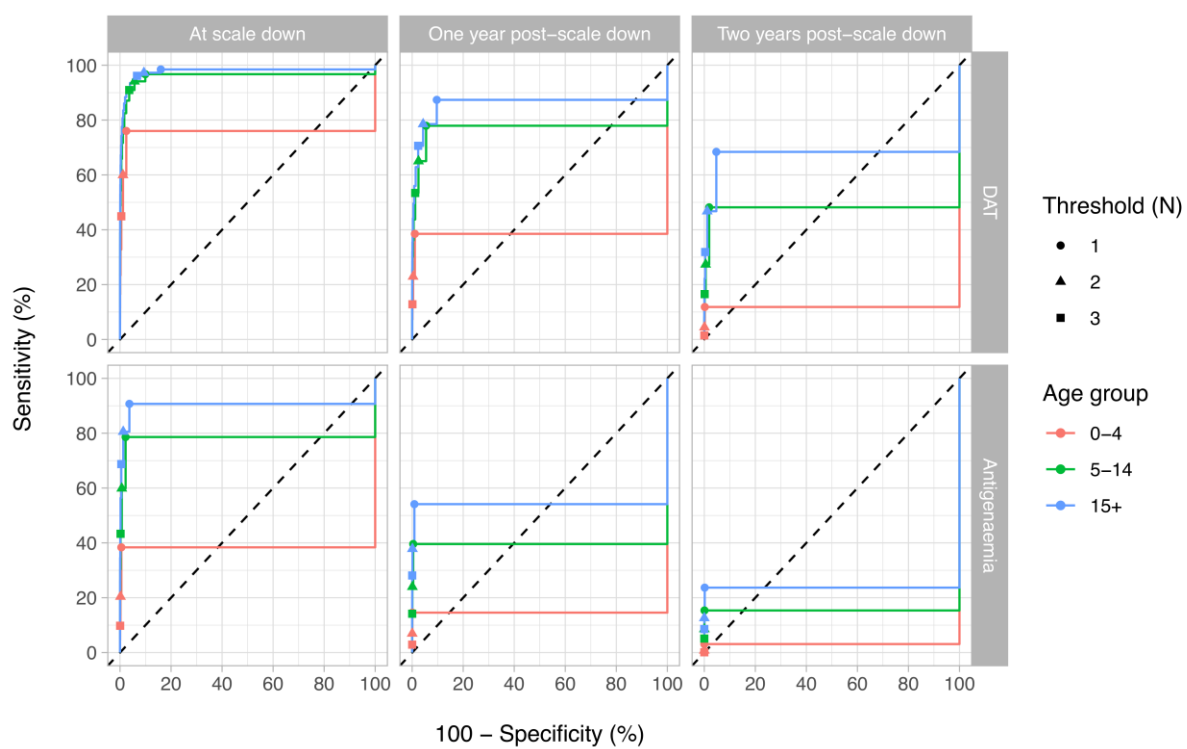
Accepted Manuscript

Figure 1 - DAT trajectory by yes-no new VL case_model E1



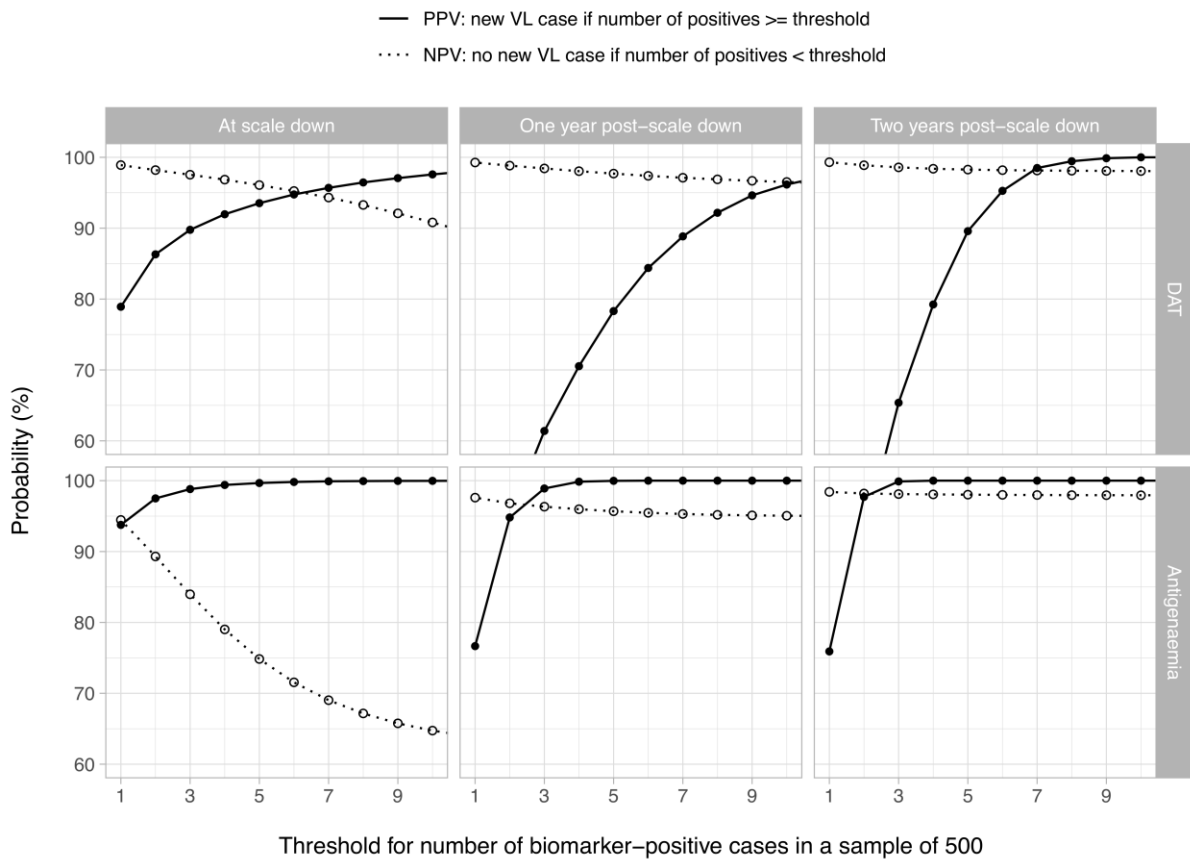
Accepted

Figure 2 - ROC by time in sample_model E1



Accepted Manuscript

Figure 3 - ppv npv in sample_model E1



Accepted