RESEARCH ARTICLE

Village-scale persistence and elimination of *gambiense* human African trypanosomiasis

Christopher N. Davis [ID]1,2, Kat S. Rock2,3, Erick Mwamba Miaka4, Matt J. Keeling [ID]2,3,5*

1 MathSys CDT, Mathematics Institute, University of Warwick, Coventry, United Kingdom, **2** Zeeman Institute (SBIDER), University of Warwick, Coventry, United Kingdom, **3** Mathematics Institute, University of Warwick, Coventry, United Kingdom, **4** Programme National de Lutte contre la Trypanosomiase Humaine Africaine (PNLTHA), Ave Coisement Liberation et Bd Triomphal No 1, Commune de Kasavubu, Kinshasa, Demecratic Republic of the Congo, **5** School of Life Sciences, University of Warwick, Coventry, United Kingdom

* m.j.keeling@warwick.ac.uk

Abstract

Gambiense human African trypanosomiasis (gHAT) is one of several neglected tropical diseases that is targeted for elimination by the World Health Organization. Recent years have seen a substantial decline in the number of globally reported cases, largely driven by an intensive process of screening and treatment. However, this infection is highly focal, continuing to persist at low prevalence even in small populations. Regional elimination, and ultimately global eradication, rests on understanding the dynamics and persistence of this infection at the local population scale. Here we develop a stochastic model of gHAT dynamics, which is underpinned by screening and reporting data from one of the highest gHAT incidence regions, Kwilu Province, in the Democratic Republic of Congo. We use this model to explore the persistence of gHAT in villages of different population sizes and subject to different patterns of screening. Our models demonstrate that infection is expected to persist for long periods even in relatively small isolated populations. We further use the model to assess the risk of recrudescence following local elimination and consider how failing to detect cases during active screening events informs the probability of elimination. These quantitative results provide insights for public health policy in the region, particularly highlighting the difficulties in achieving and measuring the 2030 elimination goal.

Author summary

Gambiense human African trypanosomiasis (gHAT) is a vector-borne infectious disease that causes sleeping sickness across many African countries. Reported gHAT cases show a continued decline, but it is unclear if this is sufficient to reach the WHO goal of stopping transmission by 2030. We develop a stochastic model necessary to address the critical question of persistence of gHAT infection at the local-scale. In contrast to other commonly studied infections, we predict long-term persistence of gHAT in small populations (< 1, 000 people) despite very low prevalence. Our local-scale predictions (together with previous larger-scale studies) suggest that, to achieve regional elimination, controls need

to be widespread and intensified in the worst affected regions, while the movement of infected people could rapidly lead to re-emergence.

Introduction

Gambiense sleeping sickness (*gambiense* human African trypanosomiasis, referred to here as gHAT) is a tsetse-borne neglected tropical disease (NTD) caused by the parasitic protozoa, *Trypanosoma brucei gambiense*. There has recently been a decline in global cases, with just 1,420 cases reported in 2017, compared to 10,466 reported in 2007 [1]. This decline is largely attributed to improvements in the active screening and treatment campaigns that have been carried out in many regions [2]. In 2012, the World Health Organization (WHO) set targets for elimination of gHAT as a public health problem; these were updated in 2017 to: reducing the area at risk of reporting more than 1 case per 10,000 people per year by 90% as compared to the baseline for 2000–2004 [2], and for fewer than 2,000 reported cases per year, by 2020 [3]. There is also a more stringent goal of stopping transmission of gHAT by 2030 [3].

Cases of gHAT primarily occur in West and Central Africa, but the distribution of infection is heterogeneous, with highly clustered incidence resulting in disease foci [4, 5]; reported prevalence often varies greatly over short distances, even between neighbouring villages [6]. This local variation suggests there is a complex spatial structure to the infection. With the observed global decline in reported cases and with many (but not all) foci likely to achieve less than 1 case per 10,000 people by 2020 [7, 8], it is increasingly important to understand both where the disease is most likely to persist and why this might be the case.

In regions where gHAT cases are no longer observed (and where local elimination has been achieved), it is possible for the disease to be re-introduced through movement of either infected humans or infected tsetse, and it may become re-established especially if active case-finding has not been maintained [9]. It is therefore clear that, while the current active screening is highly successful in many regions, an understanding of stochastic re-invasion and re-establishment in local populations is also essential to guide post-elimination policy planning.

There are two stages of gHAT, with Stage 1 following initial infection and Stage 2 defined after trypanosomes have crossed the blood–brain barrier [10]; these stages currently require very different treatments. Patients are hospitalised for the treatment duration and are advised to recover at home afterwards, on average, for a total time of 6 months [11]. Without the treatment, individuals typically progress from Stage 1 to Stage 2 after 18 months and most would likely be expected to die from meningoencephalitis after approximately 3 years [12].

Active detection of gHAT occurs through population-level screening, which is implemented in many endemic regions by mobile teams travelling to settlements and testing the available population [6]. Teams are generally able to screen substantial proportions of the local population (often over 70%) [13]; however, some socio-demographic groups (notably working adult males) frequently do not present for testing. Previous work has indicated that the individuals missed in screening are also more likely to be more highly exposed to tsetse bites [14, 15], potentially due to working in tsetse-infested forested and riverine areas. Such high risk core-group individuals pose a barrier to elimination [8, 16–18]. In parallel with active screening, passive detection occurs when individuals voluntarily attend medical facilities for testing, usually after the onset of more significant symptoms, and are thus most often in Stage 2 of the disease [11].

Mathematical models of gHAT have been beneficial in identifying the effectiveness of differing control strategies and predicting when elimination is likely to occur [19]. However,

much of the modelling work on the gHAT infection dynamics has been done in large populations using deterministic models, either for an entire regional infection focus or at a health zone level (approximately 100,000 people) [8, 15, 17, 20, 21]. Here we translate the deterministic model of Rock et al. [17] to a stochastic framework, designed to capture the infection dynamics and chance extinction at the village-scale. As such, our model is mechanistic and so captures details of the biology and epidemiology, allowing modification of model components to predict a number of different scenarios and control options.

Much of the previous work on the stochastic persistence of infection has tended to focus on measles in developed countries [22, 23]. Measles is directly transmissible, has a high reproductive ratio (12–18 compared to approximately 1–1.1 for gHAT [17]) and a high incidence before immunisation programs were introduced; yet, in contrast to gHAT, measles only persists in large populations of above approximately 300,000 and even then relies on frequent reintroductions [22, 23].

We use our model to address the dynamics of gHAT in the villages within the Yasa-Bonga and Mosango health zones of the DRC (Fig 1); a region that has an extremely high incidence of gHAT. The model has already been fitted to the available epidemiological data from the WHO HAT Atlas (Fig 1) [7, 13], with parameters inferred from regional reporting patterns and screening mimicking observed village-scale records (see Methods and S1 Appendix).

Methods

Data

The Democratic Republic of Congo (DRC) has the highest burden of gHAT cases (1,110 cases out of 1,420 reported globally in 2017 [1]), and 46% of these cases are concentrated in the former Bandundu province [9]. We focus on models and data for two high-prevalence health zones in this former province (now in Kwilu province): Yasa-Bonga and Mosango. Reported case information made available by the WHO HAT Atlas [7, 13] details the locations of settlements with estimates of population size, the years that active screenings took place, the number of people screened and the resulting newly identified gHAT cases. In Yasa-Bonga and Mosango, we consider 559 settlements, which experienced 2,701 active screenings in 2000–2012, each where a full village population was targeted to be screened, resulting in 4,875 detected gHAT cases, as well as 2,496 additional cases from detected by passive surveillance. Our model utilises these recorded screening patterns to simulate interventions, while the active and passive cases detected in each settlement are stochastically generated by the model.

Population sizes for settlements are also obtained from census estimates within the WHO HAT Atlas (Fig 1B) [7, 13]; we account for an estimated yearly population growth of 2.6% [24]. Screening coverage from active surveillance is then calculated as the number of individuals screened divided by this population estimate in each year. Annual screenings larger than the estimated accessible population (yellow and red bars in Fig 1D) may indicate either multiple screenings in a year or misreporting of individuals attending from neighbouring settlements. Low annual screening coverage (at less than 20% of the population size, gray bars in Fig 1D) is assumed to represent individuals screened outside their home settlement and therefore is not considered as a complete active screening of any given village.

All relevant model data are displayed within the paper and the Supporting Information files (Table 2 in S1 Appendix). Epidemiological data for the study were provided by the WHO in the frame of the Atlas of gHAT which may be viewed at www.who.int/trypanosomiasis_african/country/risk_AFRO/en and may be requested through Jose Ramon Franco (francoj@who.int).
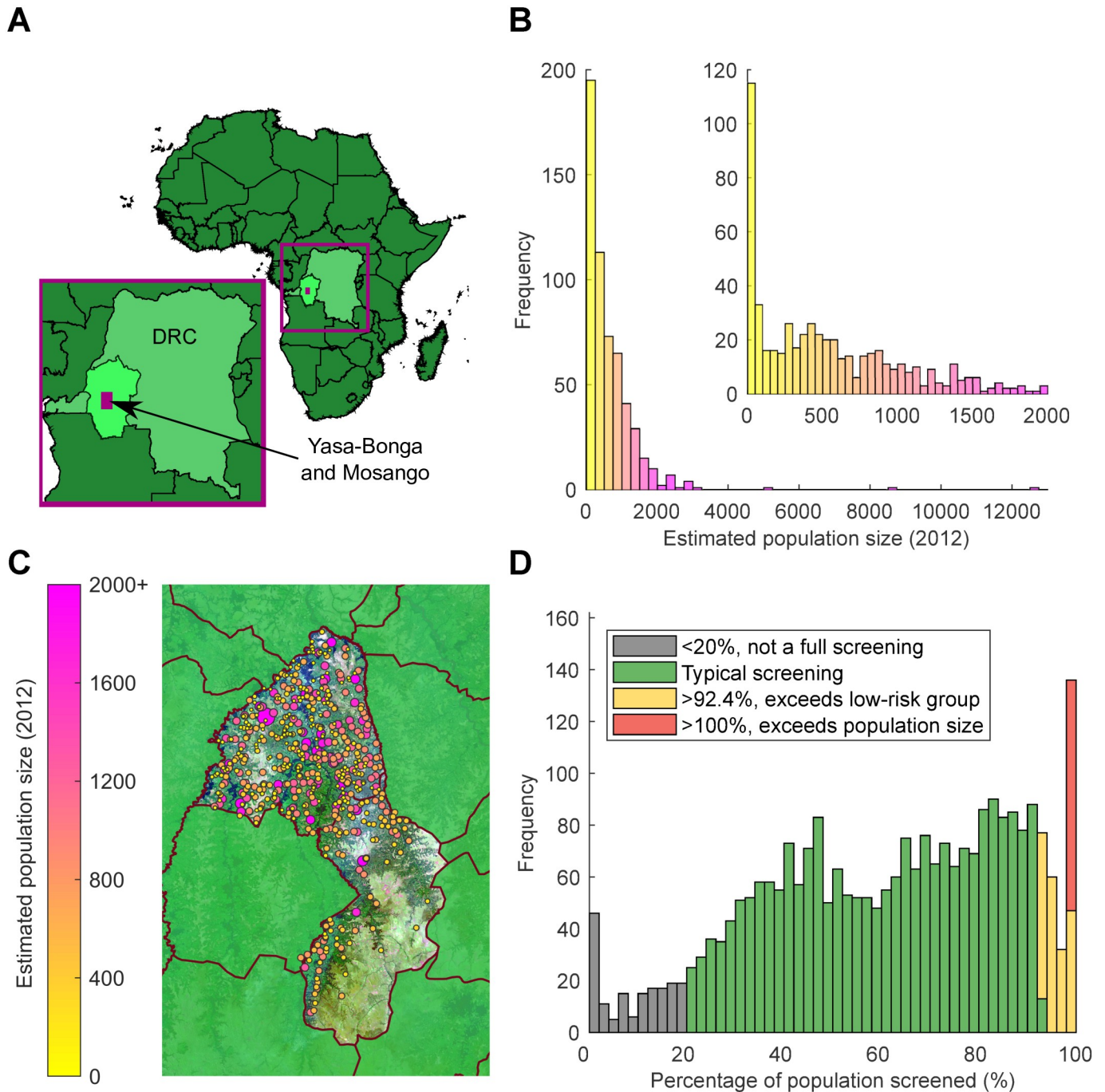
A

B

C

D

Fig 1. Visualisation of WHO HAT Atlas data for the study health zones: Yasa-Bonga and Mosango. (A) Map of Africa, showing the DRC (lighter green) with former Bandundu province highlighted in the lightest green. The area containing the health zones Yasa-Bonga and Mosango (presented in more detail in Fig 1C) is covered in a purple box. (B) Histograms of the estimated population sizes of villages in the region for 2012. The inset figure highlights the distribution of population sizes less than 2,000 individuals. Population sizes have a range of 3 to 12,645. (C) Detailed map of locations of settlements within the study region (Fig 1A, purple box), colour and radius of the circle represent population size of the individual settlements. The satellite image shown for the Yasa-Bonga and Mosango health zones is from Landsat-8 accessed through https://earthexplorer.usgs.gov/ from the U.S. Geological Survey. (D) Histogram of the coverage achieved in visits to settlements as part of the active screening programme. When annual screenings are larger than the estimated accessible population (yellow and red bars), this may indicate either multiple screenings in a given year, or misreporting, when individuals are attending from neighbouring settlements.

Modelling

The infection dynamics are described by a stochastic compartmental Ross–Macdonald-type model [25–28] extended from the previous work of Rock et al. [17] (using the formulation of Model 4 from that study) (Fig 1 in S1 Appendix). The model captures a population of humans, which is initially partitioned into those at high and low risk of being exposed to tsetse bites. Each person will transition between five different epidemiological compartments: susceptible; exposed (or latent); Stage 1 infection; Stage 2 infection; and hospitalised (and temporarily removed). We assume that there is natural mortality from all compartments, which leads to replacement of that individual as a susceptible in the population.

The risk structure is used to capture the behaviour of the small proportion of individuals that both work in the habitat of many tsetse and so have a higher biting exposure and also do not partake in active screening, thereby acting as a human reservoir of infection [16–18]. The proportion of the individuals in the high risk group is estimated through extensive model fitting to be 7.6% of the population for these health zones [17]. While the high risk group have a higher biting exposure, the probability of tsetse infection per single infective bite is the same for both risk groups. This model structure, where individuals are either high risk and non-participating in active screening or low risk and randomly participating in active screening, was selected using the deviance information criterion (DIC), which assigns low scores for models with high posterior mean log-likelihood and penalises models with more parameters.

Tsetse in the model are similarly compartmentalised into four epidemiological states: teneral (unfed); non-teneral yet uninfected; exposed (or latent); and infected. The distinction between teneral and non-teneral yet uninfected is used to capture the observation that tsetse are far more susceptible to infection at their first blood meal than at any subsequent blood meals [29]. The effect of a possible animal reservoir is not considered, since its role remains unclear [15, 17, 20, 30, 41] and its inclusion does not significantly improve the match between model outputs and currently available data in this setting [17].

Additional to the epidemiological and demographic processes, we simulate the effect of active screening and passive detection of cases. Passive detection (and disease-induced mortality) is assumed to occur at a fixed (per capita) rate for all Stage 2 infected individuals [11]. Active screening takes place annually; the proportion screened is either replayed from the historic pattern for that settlement or chosen randomly from the set of all screening coverages recorded, allowing a greater range of scenarios to be explored. Since Yasa-Bonga and Mosango are high endemicity regions, we assume that the screening coverage and frequency remains constant over time but note that these quantities are somewhat affected by population size (Fig 9 in S1 Appendix).

Individuals in the low risk group are selected randomly for screening, irrespective of epidemiological status, and those that are found to be infected are moved to the hospitalised class. We assume that screening only applies to low risk individuals, such that screening coverages greater than 92.4% (the estimated proportion in the low risk group) [17] are truncated (Fig 1D). In the field, the diagnostic process is complex and multi-stage [31]; however, in the model, we collapse this into characteristics for the whole algorithm, which is assumed to be 91% sensitive [32] but 100% specific. False negatives remain undetected in the settlement, but by assuming 100% specificity, there are no false positives as we assume confirmation by microscopy will be carried out due to the low case numbers.

For the majority of this paper, we model the dynamics as a closed population, without emigration or immigration, so that once the disease has gone extinct in a population it cannot be re-introduced. This removes a critical dependency in model formulation, and greatly simplifies the presentation of results. In reality, no population is ever completely isolated; however

we show that the expected rate of infectious imports is very low and does not affect our main results or conclusions (S1 Appendix).

Parameters

Model parameters for the underlying compartmental model are taken from the previous work of Rock et al. [17]. The values of these parameters were taken from literature, where well-defined, and otherwise inferred by model fitting using a Metropolis–Hastings MCMC algorithm, which sought to match the deterministic model to observed cases and screening effort in Yasa-Bonga and Mosango. The values used in this manuscript are the median of the distributions inferred using MCMC methodology applied to the aggregate annual data from Yasa-Bonga and Mosango. Therefore, the parameter values are specific to the study region, and the model is well-matched with the incidence data from active and passive surveillance. A full list of parameters is given in Table 2 in S1 Appendix.

Results

We use our model to address the dynamics of gHAT in the villages within the Yasa-Bonga and Mosango health zones of the DRC (Fig 1), a region that has an extremely high incidence of gHAT. Using the estimated population sizes and the reported levels of active screenings (Fig 1B and 1D), we validate the stochastic model by comparing the observed and predicted screenings that did not detect any cases (which subsumes both local extinction and failure to detect). We then consider the probability of local gHAT persistence across a range of population sizes and different control scenarios, as well as the probability of re-invasion. Finally, we focus on whether not detecting cases in a series of active screens can inform on whether local elimination has been achieved, noting that WHO guidelines suggest annual active screenings until there have been three consecutive years of no new cases, followed by a further screening with no cases after three years [31].

Comparison with data

While the underlying deterministic model has been fitted to the aggregate data from this region, it is important to assess the behaviour of the stochastic model against village-scale observations. Unfortunately, local disease extinctions cannot be directly observed; failure to discover any cases does not necessarily mean that the infection is not present, simply that it has not been detected. Thus, to validate our model, we make comparisons between the simulated predictions and WHO HAT Atlas data [7, 13] for the probability of detecting no cases on an active screening (termed zero-detections for brevity), which is a combination of failure to detect and local extinction. We compare model predictions to observations by calculating the percentage of zero-detections in aggregations of 100 active screenings with similar village population sizes (Fig 2A). We find very strong agreement between model predictions and data, with a pronounced decline in zero-detections for larger populations.

For individual settlements, those where the number of zero-detections lie outside the 95th percentile of model predictions are notably spatially clustered (Fig 2B). In 2.1% of settlements (red), there are significantly fewer zero-detections than predicted and hence greater persistence; these villages are generally localised around the main river through the region. In 6.3% of settlements there are significantly more zero-detections than expected (yellow), and these are clustered far from the major rivers and in upland areas. Since tsetse are most densely distributed surrounding riverine areas [33], this spatial clustering may indicate the need for spatially heterogeneous parameters that reflect the suitability of the local environment for tsetse.
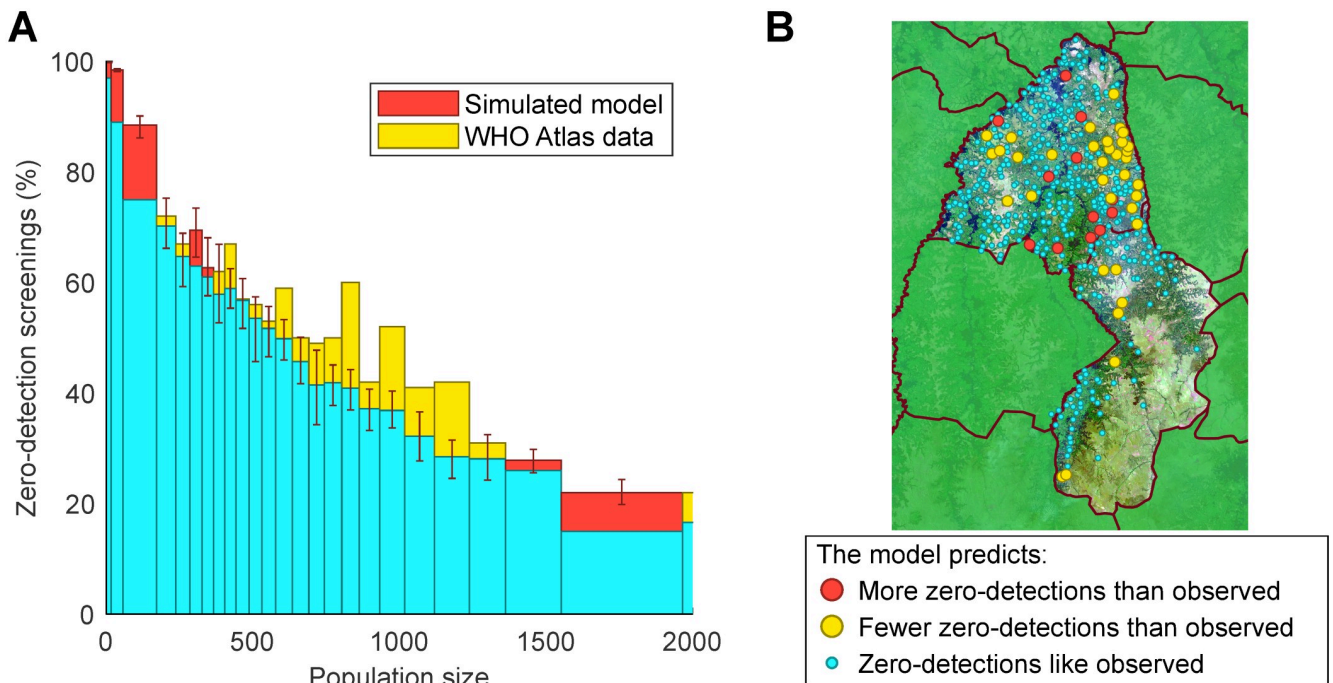
Fig 2. Comparison of model predictions and data for active screenings with no detected cases (zero-detections). (A) Histogram by population size of the percentage of active screenings that find no new gHAT cases for both the model and the HAT Atlas data. Each bar represents 100 screenings of simulated results (averaged over 10,000 replicates) from the model that uses the observed pattern of screenings and compares to the data. Values where the model predictions have more zero-detections than the data are in red, while the reverse is shown in yellow. Error bars represent the 95th percentile of model results. (B) Map of populations in Yasa-Bonga and Mosango showing the settlements with significant differences (at the 95% level) in the expected proportion of active screenings with no cases detected. Red circles are where the observed number of active screenings with zero-detections is below the 95th percentile of the model; yellow circles are where the data falls above the 95th percentile; small blue circles are for data that lie within the 95th percentile of predictions and therefore are well described by the model. The satellite image shown is from Landsat-8 accessed through https://earthexplorer.usgs.gov/ from the U.S. Geological Survey.

https://doi.org/10.1371/journal.pntd.0007838.g002

However, given that more than 90% of villages fall within our prediction intervals, we believe the homogeneous parameters capture the general stochastic behaviour of this region.

Local gHAT persistence

Local gHAT persistence, where human infection is maintained in a settlement, is affected by many factors (Fig 3), including the population size of the settlement; the vector-to-host (tsetse-to-human) ratio; the exposure to the tsetse; the screening procedure; and any movement of infected individuals between populations. We calculate the probability of persistence by stochastically simulating the epidemic for 16 years from the endemic (uncontrolled) disease equilibrium. If there is zero gHAT infection in the human population after a given number of years and no further human infection emerges in the following year from infected tsetse, we say in this simulation there is local disease elimination with no immediate threat of re-emergence (Fig 4A provides a justification for this criteria). This procedure is repeated for multiple population sizes; the proportion of simulations that retain infection after a given number of years determines the probability of persistence (Fig 3).

We focus on settlements with fewer than 2,000 inhabitants, typical of Yasa-Bonga and Mosango (Fig 1B), and use regionally specific parameters. In all scenarios investigated, we find that persistence increases with increasing population size (Fig 3). This echoes results from other infections [22, 23, 34, 35], where small populations with low incidence experience a
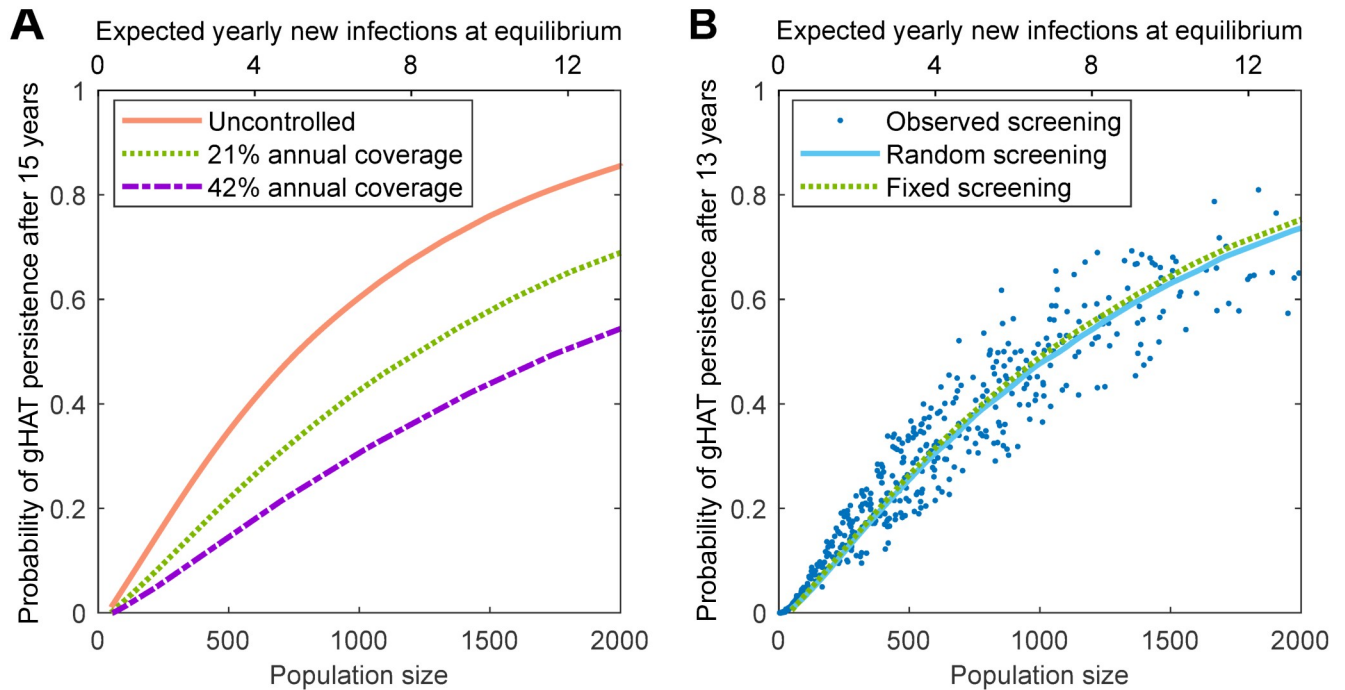
Fig 3. Predicted probability of gHAT persistence in isolated settlements. Simulations are started at the endemic (uncontrolled) equilibrium and iterated forwards (without infectious imports), while the persistence of infection is recorded. This is repeated 100,000 times for settlement population sizes between 50 and 2,000 individuals. The expected number of yearly new infections if the system were at equilibrium is proportional to the population size and is given by the top scale. (A) Impact of active screening on gHAT persistence; annual screening at a fixed coverage per year yields a drop in persistence with increased coverage. (B) Comparison of screening assumptions on the persistence of gHAT. The solid curve shows results where annual screening coverages were randomly sampled from all observed coverages; dots represent the individual settlements recorded in the WHO HAT Atlas for Yasa-Bonga and Mosango health zones [7, 13], where the reported coverage in each year is used. There were sufficient simulations such that confidence intervals are too small to be visible.

greater impact of stochasticity and chains of transmission that are more likely to be broken. In addition, given that long-term persistence relies on persistence for shorter time intervals, the probability of persistence decreases with time (Fig 10B–C in S1 Appendix). The relatively long persistence times of gHAT, compared to the frequently studied persistence of childhood diseases [22, 23], are attributable to the long time-scale of gHAT infection in the absence of active interventions [36].

The addition of active screening (leading to the treatment of detected cases) decreases the probability of persistence across all population sizes, since removing infected individuals leads to a greater chance of breaking chains of transmission (Fig 3A). Increasing levels of screening, beyond the observed 21% average, leads to further reductions for persistence. We compare three assumptions for active screening (Fig 3B): that each population is screened annually at a fixed coverage equal to the regional average (21%); that each population experiences screening coverages sampled from the regional pattern including not screening in a given year; and simply replaying the recorded pattern of active screening in each village. Despite the very different distributions of screening effort, all three of these assumptions produce comparable levels of persistence.

Re-invasion probability

Following localised elimination of infection, populations remain vulnerable to re-invasion; we investigate the potential for re-establishment of sustained transmission in a settlement for
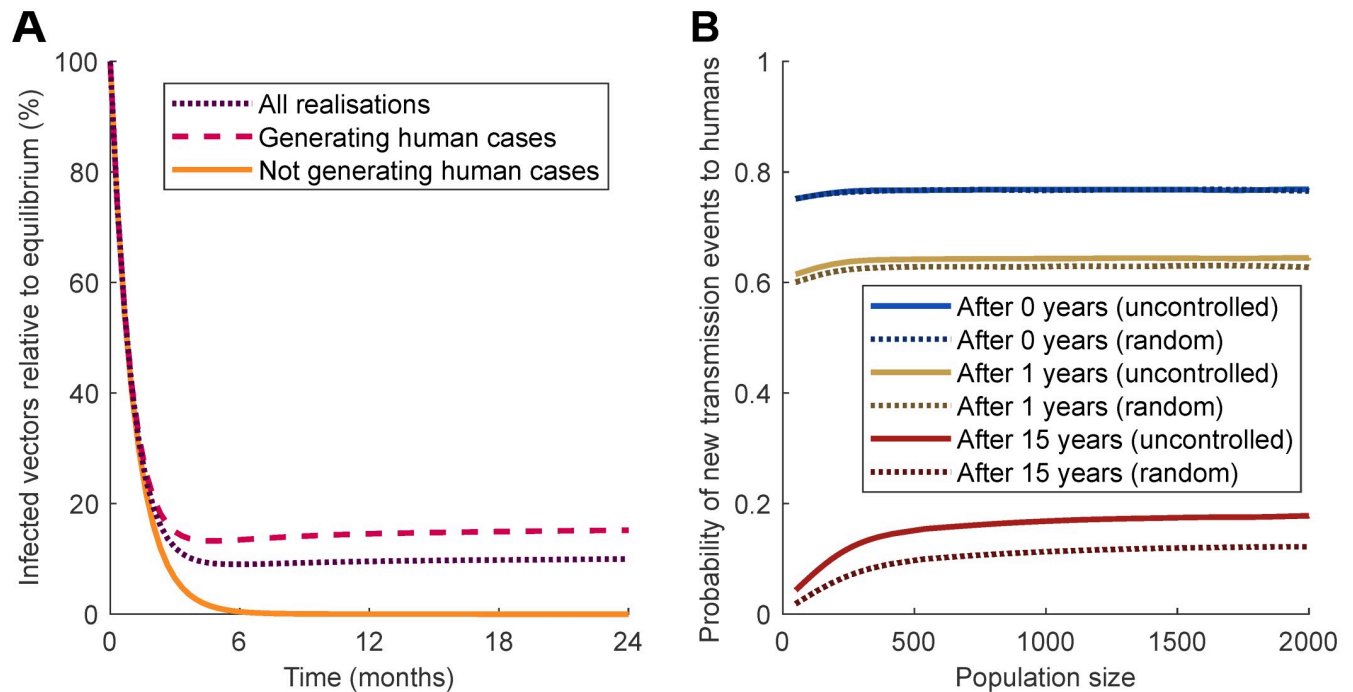
Fig 4. Dynamics of extinction and reintroduction. (A) Starting with no infected humans but the tsetse population at its endemic equilibrium and a settlement size of 1,000 individuals, the model predicts a dramatic decline in the infected tsetse population, depending on whether subsequent human cases are generated by the infected tsetse. (B) Extending this model further for a range of initial conditions, we examine the probability that at least one human case is generated and the infection persists for a given time, when starting with one infected human and no infected tsetse for both uncontrolled and random screening.

https://doi.org/10.1371/journal.pntd.0007838.g004

different invasion scenarios. For a village of 1,000 individuals, following the elimination of infection in humans, the level of infection in tsetse falls rapidly, even when starting at the endemic level in these vectors (Fig 4A). In approximately 65% of simulations, the initially infected tsetse generate human cases, and the level of infection in vectors rapidly plateaus; otherwise, infection is eliminated from the location within six months due to short vector life expectancies in comparison to the human hosts. This validates the previous simulation assumption (Fig 3) that local disease elimination can be considered achieved after a period of one year in which there are no infected humans, as the number of infected tsetse will also become negligible.

The probability that a population of tsetse, infected at the endemic equilibrium (0.02% of tsetse exposed or infected), will lead to re-establishment of infection is predicted to be a function of settlement size (Fig 11A in S1 Appendix). Small populations are unlikely to see any new human infections, and those that are generated fail to persist. However, even for large population sizes of 2,000 individuals, the chance of continued transmission beyond one year is only 55% and is less than 10% over 15 years. For lower levels of infection in the tsetse population, the risk of successful re-establishment is proportionally reduced.

In contrast, if re-invasion of an infection-free population is due to the movement of an infected person into the settlement (in the absence of infected tsetse) the probability of re-establishment over different time-frames is largely unaffected by the population size (Fig 4B). We predict a high probability (> 70%) of short term re-invasion, but a more limited chance (< 20%) that this will generate persistent infection for 15 years or more. This is typical of

stochastic dynamics of infection with low basic reproductive ratio (R_0), where, although short chains of transmission are likely, it is difficult for the infection to fully establish.

Detecting elimination

As discussed above, zero-detections may be an indication that there is no infection in a settlement; however, it may simply be that infected individuals were not screened or were false negatives. Despite this, there is a temptation to associate zero-detection with zero infection; we therefore use our simulations to tease apart this complex relationship. The model replays the observed pattern of screening in each settlement, but we perform multiple simulations to ascertain the probability that infection has been locally eliminated following one, two or three consecutive zero-detection screenings, in which at least 20% of the population are screened (Fig 5). We also insist that no passive cases were detected between the screening events. Our standard assumption, in agreement with parameter inference for this region, is that 26% of Stage 2 infections, where people either self-report or likely die of gHAT infection, are detected and reported [17], but we also investigate 0% (no passive reporting) and 100% (all passive cases and deaths are reported) for comparison.

For small settlements, given that long-term persistence is unlikely (Fig 3), even a single zero-detection screen (Fig 5, blue) is frequently associated with local elimination. For larger populations, a single zero-detection has limited predictive power and three consecutive zero-detections are needed to have any degree of confidence, in which case the reporting of passive
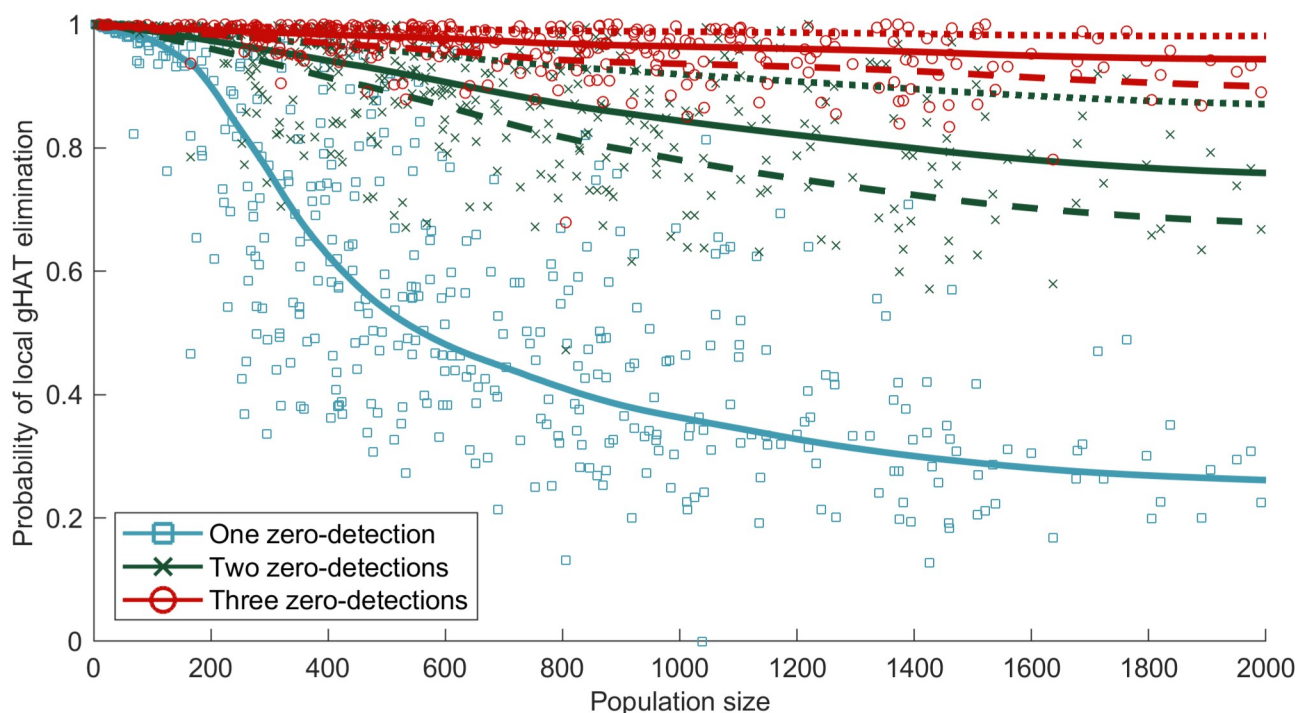

Fig 5. Probability of elimination in a settlement, given consecutive zero-detections with no detected passive cases. (A) Consecutive zero-detections means consecutive in the observed years of screenings, not necessarily in consecutive years, and with no passive cases detected in between. Each point represents the average from multiple simulations of individual settlements where the reported pattern of screenings is replayed. The points and solid lines assume a reporting rate of 26% [17], while the dashed and dotted lines show reporting rates of 0% and 100%, respectively. Sufficient simulations are used such that the confidence intervals are small (unobservable on the scale of this graph). Lines represent a weighted local regression fit. Active screenings where fewer than 20% of the population are assessed are excluded from our analysis due to the small sample sizes (alternative cut-offs of less than 10% and less than 50% are presented in Fig 13 in S1 Appendix).

https://doi.org/10.1371/journal.pntd.0007838.g005

cases plays a noticeable role. There is, however, significant variation between settlements, reflecting very different patterns of reported screening. Moderate population sizes of between 200 and 1,000 individuals show extreme variation in the ability to predict local elimination, while smaller villages have less variation, in part due to rarely being screened before 2009 (Fig 9C in S1 Appendix).

Importation rates

Regional persistence relies on more than independent persistence in individual settlements. It is likely that the occasional movements of infected people lead to a stochastic meta-population paradigm [40], where rare local extinctions of infection are balanced by external imports. However, the agreement between model and data (Fig 2), together with the low prevalence of infection, indicates that imports are likely to be rare.

We make this more quantitative by fitting an importation rate of infection, proportional to the size of the population. Before interventions, the presence/absence of infection in a population reflects the equilibrium balance between extinction and re-colonisation; we can therefore use the presence of infection at the first recorded active screening within a village as measure of the equilibrium state. Fitting the external importation rate to the probability of detecting infection in a village at the first active screen gives the best fit when the rate is small at just 3.4×10^{-6} per person per day (Fig 6A). Moreover, it is assumed that the importation rate declines over time as the overall prevalence in DRC reduces.

For most population sizes in the region, the expected time to importation is therefore relatively long (Fig 6B), and in many cases a single importation will not cause further transmission
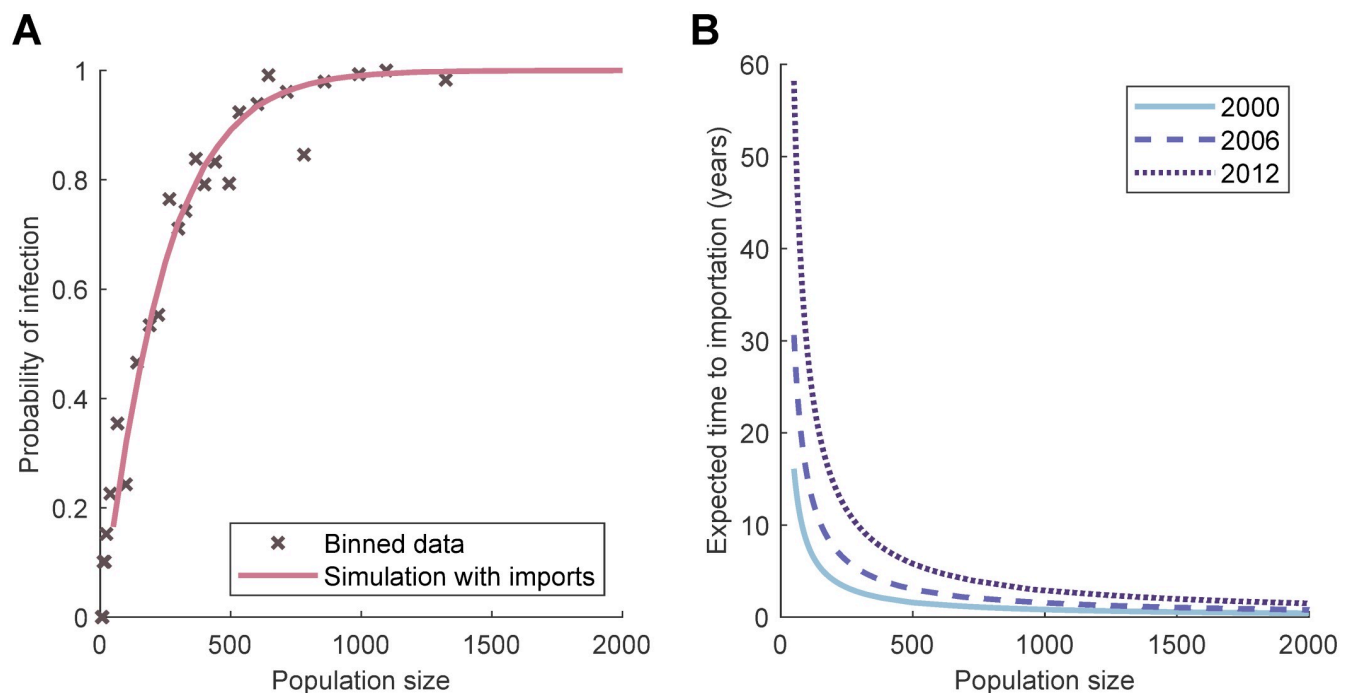

Fig 6. Simulating external importations of infection into village populations. (A) By running simulations with different values for the external infection parameter, we find the best fit—to data binned by population size from the WHO HAT Atlas on whether there are any detected cases on the first active screening—is when imports per susceptible individual are equal to 3.4×10^{-6} days^{-1} (solid line). (B) Curves of the expected time for an external importation into a village population using the fitted importation parameter. The importation parameter is assumed to decay at the same rate as total number of cases in time in the DRC (see Fig 5 in S1 Appendix).

https://doi.org/10.1371/journal.pntd.0007838.g006

events (Fig 4), leading us to conclude that in general the level of importation will not qualitatively change our results. This is made more explicit in the supporting information (Fig 6 in S1 Appendix), where it is shown that the model with the fitted level of infectious imports generates comparable results to those described in the main paper. This justifies our modelling assumptions that villages act as isolated populations and that importation of infection is unlikely to perturb the dynamics; instead, we are able to separate the processes of local elimination and re-invasion.

Discussion

Despite global declines in reported cases over the last decades, *gambiense* gHAT remains a problem in many focal areas [37]. These regions, concentrated primarily in the DRC, represent a significant challenge to achieving the WHO 2020 and 2030 goals of elimination as a public health problem and zero transmission, respectively. Robust models, matched to the available data, are the only viable means of quantitatively assessing future dynamics and the long-term impact of controls [8, 17, 18, 21]. Active screening followed by treatment is one of the main control measures, but this action is deployed at the village level suggesting that village-scale models (which recognise the effects of small population size) may be needed to optimise deployment; these results can then be scaled to an infection focus or national level to measure regional elimination, which is especially important as we approach zero transmission.

We have introduced a dynamic, mechanistic, stochastic gHAT model, which is applied to 559 settlements in the Yasa-Bonga and Mosango health zones within the former Bandundu province of the DRC (Fig 1). Using parameters inferred from a deterministic model fitted to aggregated reported cases, our model reliably captures observed detection patterns at the village-scale (Fig 2). This comparison highlighted some spatial heterogeneity associated with the local environmental conditions (significantly fewer zero-detections than predicted occurred in regions close to large rivers, where the tsetse density is presumably high); however, 91.6% of settlements fell within the model (95%) prediction intervals, giving us confidence in our predictive ability. The inclusion of such local environmental factors, which modify the underlying parameters, is clearly an area for further research into refining this small-scale model and may help to practically focus localised control measures, in particular for planning tsetse control.

Throughout our simulation experiments, we consistently find that gHAT persists better in larger populations. This is as expected and agrees with theoretical work and analysis of other diseases [22, 23, 34, 35, 38, 39]; in small populations, the behaviour of the individual is more important, and hence stochastic effects are magnified. The degree of persistence predicted is, however, surprising; settlements of around 2,000 inhabitants, where yearly incidence is only 13 new infections, frequently persist for 15 years or more (Fig 3). This should be contrasted with frequently studied, highly transmissible diseases, such as measles, where local extinctions are common in population sizes of less than 300,000 [22, 23]. We attribute this pronounced difference to the much longer time scales associated with gHAT, meaning single individuals can maintain infection, and the vector-borne nature of gHAT transmission, such that the tsetse act as a short-lived reservoir. We consistently find that incorporating active screening reduces the persistence of infection (Fig 3A), although the distribution of screening across years has only a small effect (Fig 3B). Increasing the screening coverage beyond the average reported levels (of 21% per year) is predicted to lead to further reductions in persistence, but infection is still predicted to be maintained for over 15 years in many larger populations.

Throughout, we have generally ignored the impact of new infectious individuals entering the population, and indeed have shown that this rate of importation is very low. If re-invasion following local elimination is due to the movement of a single infected individual into the

settlement (Fig 4B), we predict that the probability of subsequent cases is high (70–80%) and largely independent of population size. However, only a small proportion (10–20%) of such invasions lead to long-term persistence of over 15 years. Re-establishment of infection due to the movement of a limited number of infected tsetse is even less likely. Current uncertainties about the impact of potential reservoirs—either a human reservoir of asymptomatic infections or an animal reservoir—mean there is insufficient knowledge for resurgence to be explicitly modelled by these mechanisms [41], but an animal reservoir that can maintain infection in the absence of humans is likely to represent a worst-case scenario.

A key question, as we approach the 2030 goal of zero transmission, is to ascertain when local elimination has been achieved, allowing policy-makers to scale back control if the infection is no longer present. Due to only a limited proportion of each settlement being screened and the potential for false negatives, a screening can fail to detect any cases even when there is infection in the population. We have shown that, while a single zero-detection screening provides relatively little information of the probability of local gHAT elimination, multiple consecutive zero-detection screenings are a strong indicator of elimination (Fig 5). This can be further strengthened if only large screens ($> 50\%$) are included in the analysis (Fig 13B in S1 Appendix), providing valuable public health information. This concurs with WHO guidelines for active screenings, as villages are no longer considered in planning by mobile screening teams after three consecutive years of zero-detections, followed by a further zero-detection after three years [31]; our model would predict local elimination with large probability for this level of surveillance. Consistent with the observed patterns in high endemicity regions, we have assumed that the screening coverage and frequency remains constant over these time scales but note that in other regions it may be important to consider the reduction in active screening as reported gHAT cases decline and elimination is approached.

Conclusion

The ability to capture the stochastic dynamics and persistence of *gambiense* gHAT infection at the village-scale is a major advance in public health modelling, with far-reaching consequences for informing policy decisions. This is particularly pertinent as our models operate at the same spatial scale as controls and can capture the local elimination of infection that is a prerequisite of achieving the 2030 goal of zero transmission globally.

Supporting information

S1 Appendix. Model details and supplementary results. Additional information on the formulation of the model, details of the data and supplementary results.
(PDF)

S1 Code. C++ code. The code to run the basic stochastic model.
(CPP)

Author Contributions

Conceptualization: Matt J. Keeling.

Data curation: Erick Mwamba Miaka.

Investigation: Christopher N. Davis, Kat S. Rock.

Methodology: Christopher N. Davis, Kat S. Rock, Matt J. Keeling.

Supervision: Matt J. Keeling.

Writing – original draft: Christopher N. Davis, Matt J. Keeling.

Writing – review & editing: Kat S. Rock, Erick Mwamba Miaka, Matt J. Keeling.

References

1. World Health Organization. Global health observatory data repository; 2018.

2. Franco R, Cecchi G, Priotto G, Paone M, Diarra A, Grout L, et al. Monitoring the elimination of human African trypanosomiasis: Update to 2014. PLoS Neglected Tropical Diseases. 2017; 11(5):e0005585. https://doi.org/10.1371/journal.pntd.0005585 PMID: 28531222

3. World Health Organization. Sustaining the drive to overcome the global impact of neglected tropical diseases. World Health Organization; 2013.

4. Büscher P, Cecchi G, Jamonneau V, Priotto G. Human African trypanosomiasis. The Lancet. 2017; 390:2397–2409. https://doi.org/10.1016/S0140-6736(17)31510-6

5. Smith DH, Pepin J, Stich AHR. Human African trypanosomiasis: an emerging public health crisis. British Medical Bulletin. 1998; 54(2):341–355. https://doi.org/10.1093/oxfordjournals.bmb.a011692 PMID: 9830201

6. Brun R, Blum J, Chappuis F, Burri C. Human African trypanosomiasis. The Lancet. 2010; 375 (9709):148–159. https://doi.org/10.1016/S0140-6736(09)60829-1

7. Simarro PP, Cecchi G, Franco JR, Paone M, A D, Priotto G, et al. Monitoring the progress towards the elimination of gambiense human African trypanosomiasis. PLoS Neglected Tropical Diseases. 2015; 9(6):e0003785. https://doi.org/10.1371/journal.pntd.0003785 PMID: 26056823

8. Rock KS, Pandey A, Ndeffo-Mbah ML, Atkins KE, Lumbala C, Galvani A, et al. Data-driven models to predict the elimination of sleeping sickness in former Equateur province of DRC. Epidemics. 2017; 18:101–112. https://doi.org/10.1016/j.epidem.2017.01.006 PMID: 28279451

9. Lumbala, Simarro PP, Cecchi G, Paone M, Franco JR, Mesu VKBK, et al. Human African trypanosomiasis in the Democratic Republic of the Congo: disease distribution and risk. International Journal of Health Geographics. 2015; 14(1):20. https://doi.org/10.1186/s12942-015-0013-9 PMID: 26047813

10. Kennedy PGE. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). The Lancet Neurology. 2013; 12(2):186–194. https://doi.org/10.1016/S1474-4422(12)70296-X PMID: 23260189

11. Mpanya A, Hendrickx D, Baloji S, Lutumba P. From health advice to taboo: community perspectives on the treatment of sleeping sickness in the Democratic Republic of Congo, a qualitative study. PLoS Neglected Tropical Diseases. 2015; 9(4):e0003686. https://doi.org/10.1371/journal.pntd.0003686 PMID: 25856578

12. Checchi F, Filipe J, Haydon D, Chandramohan D, Chappuis F. Estimates of the duration of the early and late stage of gambiense sleeping sickness. BMC Infectious Diseases. 2008; 8(1):16. https://doi.org/10.1186/1471-2334-8-16 PMID: 18261232

13. Simarro PP, Cecchi G, Jannin JG. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. International Journal of Health Geographics. 2010; 9:57. https://doi.org/10.1186/1476-072X-9-57 PMID: 21040555

14. Rock KS, Stone CM, Hastings IM, Keeling MJ, Torr SJ, Chitnis N. Mathematical models of Human African trypanosomiasis epidemiology. 87. Academic Press; 2015.

15. Mahamat MH, Peka M, Rayaisse Jb, Rock KS, Toko MA, Darnas J, et al. Adding tsetse control to medical activities contributes to decreasing transmission of sleeping sickness in the Mandoul focus (Chad). PLoS Neglected Tropical Diseases. 2017; 11(7):e0005792. https://doi.org/10.1371/journal.pntd.0005792 PMID: 28750007

16. Mpanya A, Hendrickx D, Vuna M, Kanyinda A, Lumbala C, Tshilombo V, et al. Should I get screened for sleeping sickness? A qualitative study in Kasai province. PLoS Neglected Tropical Diseases. 2012; 6(1):e1467. https://doi.org/10.1371/journal.pntd.0001467 PMID: 22272367

17. Rock KS, Torr SJ, Lumbala C, Keeling MJ. Quantitative evaluation of the strategy to eliminate human African trypanosomiasis in the DRC. Parasites & Vectors. 2015; 8(1):532. https://doi.org/10.1186/s13071-015-1131-8

18. Stone CM, Chitnis N. Implications of heterogeneous biting exposure and animal hosts on Trypanosomiasis brucei gambiense transmission and control. PLoS Computational Biology. 2015; 11(10):e1004514. https://doi.org/10.1371/journal.pcbi.1004514 PMID: 26426854

19. Hollingsworth TD, Adams ER, Anderson RM, Atkins K, Bartsch S, Basáñez Mg, et al. Quantitative analyses and modelling to support achievement of the 2020 goals for nine neglected tropical diseases. Parasites & Vectors. 2015; 8(1):630. https://doi.org/10.1186/s13071-015-1235-1

20. Pandey A, Atkins KE, Bucheton B, Camara M, Aksoy S, Galvani AP, et al. Evaluating long-term effectiveness of sleeping sickness control measures in Guinea. Parasites & Vectors. 2015; 8(1):550. https://doi.org/10.1186/s13071-015-1121-x

21. Rock KS, Torr SJ, Lumbala C, Keeling MJ. Predicting the impact of intervention strategies for sleeping sickness in two high-endemicity health zones of the Democratic Republic of Congo. PLoS Neglected Tropical Diseases. 2017; 11:e0005162. https://doi.org/10.1371/journal.pntd.0005162 PMID: 28056016

22. Bartlett MS. Measles periodicity and community size. Journal of the Royal Statistical Society Series A. 1957; 120(1):48–70. https://doi.org/10.2307/2342553

23. Keeling MJ, Grenfell BT. Disease extinction and community size: modeling the persistence of measles. Science. 1997; 275(5296):65–67. https://doi.org/10.1126/science.275.5296.65 PMID: 8974392

24. UN. World statistics pocketbook. New York: United Nations; 2016.

25. Ross R. The prevention of malaria. London: J Murray; 1911.

26. Ross R. An application of the theory of probabilities to the study of a priori pathometry. Part I. Proceedings of the Royal Society of London Series A. 1916; 92(638):204–230. https://doi.org/10.1098/rspa.1916.0007

27. Macdonald G. The analysis of infection rates in diseases in which super infection occurs. Tropical Diseases Bulletin. 1950; 47:907–915. PMID: 14798656

28. MacDonald G. The analysis of the sporozoite rate. Tropical Diseases Bulletin. 1952; 49(6):569–586. PMID: 14958825

29. Haines LR. Examining the tsetse teneral phenomenon and permissiveness to trypanosome infection. Frontiers in Cellular and Infection Microbiology. 2013; 3(November):84. https://doi.org/10.3389/fcimb.2013.00084 PMID: 24312903

30. Funk S, Nishiura H, Heesterbeek H, Edmunds WJ, Checchi F. Identifying Transmission Cycles at the Human-Animal Interface: The Role of Animal Reservoirs in Maintaining Gambiense Human African Trypanosomiasis. PLoS Computational Biology. 2013; 9(1):e1002855. https://doi.org/10.1371/journal.pcbi.1002855 PMID: 23341760

31. WHO. Control and surveillance of human African trypanosomiasis. World Health Organization; 2013.

32. Checchi F, Chappuis F, Karunakara U, Priotto G, Chandramohan D. Accuracy of five algorithms to diagnose gambiense human African trypanosomiasis. PLoS Neglected Tropical Diseases. 2011; 5(7):e1233. https://doi.org/10.1371/journal.pntd.0001233 PMID: 21750745

33. Franco JR, Simarro PP, Diarra A, Jannin JG. Epidemiology of human African trypanosomiasis. Clinical Epidemiology. 2014; 6(1):257–275. https://doi.org/10.2147/CLEP.S39728 PMID: 25125985

34. Hagenaars TJ, Donnelly CA, Ferguson NM. Spatial heterogeneity and the persistence of infectious diseases. Journal of Theoretical Biology. 2004; 229(3):349–359. https://doi.org/10.1016/j.jtbi.2004.04.002 PMID: 15234202

35. Brooks-Pollock E, Keeling M. Herd size and bovine tuberculosis persistence in cattle farms in Great Britain. Preventive Veterinary Medicine. 2009; 92(4):360–365. https://doi.org/10.1016/j.prevetmed.2009.08.022 PMID: 19762099

36. Checchi F, Filipe J, Barrett M, Chandramohan D. The natural progression of Gambiense sleeping sickness: what is the evidence? PLoS Neglected Tropical Diseases. 2008; 2(12):e303. https://doi.org/10.1371/journal.pntd.0000303 PMID: 19104656

37. Steverding D. The history of African trypanosomiasis. Parasites & Vectors. 2008; 1(1):3. https://doi.org/10.1186/1756-3305-1-3

38. Van Herwaarden OA, Grasman J. Stochastic epidemics: major outbreaks and the duration of the endemic period. Journal of Mathematical Biology. 1995; 33(6):581–601. https://doi.org/10.1007/BF00298644 PMID: 7608639

39. Grasman J. Stochastic epidemics: the expected duration of the endemic period in higher dimensional models. Mathematical Biosciences. 1998; 152(1):13–27. https://doi.org/10.1016/s0025-5564(98)10020-2 PMID: 9727295

40. Grenfell B, Harwood J. (Meta) population dynamics of infectious diseases. Trends in Ecology & Evolution. 1997; 12(10):395–399. https://doi.org/10.1016/S0169-5347(97)01174-9

41. Büscher P, Bart JM, Boelaert M, Bucheton B, Cecchi G, Chitnis N, et al. Do cryptic reservoirs threaten gambiense-sleeping sickness elimination? Trends in Parasitology. 2018; 34(3):197–207. https://doi.org/10.1016/j.pt.2017.11.008 PMID: 29396200