Inferring person-to-person networks of pathogen transmission: is routine surveillance data up to the task?

John H. Huber[1*], Michelle S. Hsiang[2,3,4], Nomcebo Dlamini[5], Maxwell Murphy[6], Sibonakaliso Vilakati[5], Nomcebo Nhlabathi[5], Anita Lerch[1], Rasmus Nielsen[7], Nyasatu Ntshalintshali[8], Bryan Greenhouse[6,9], T. Alex Perkins[1*]

[1]Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, United States of America

[2]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas, United States of America

[3]Malaria Elimination Initiative, Global Health Group, University of California, San Francisco, California, United States of America

[4]Department of Pediatrics, University of California, San Francisco, California, United States of America

[5]National Malaria Elimination Programme, Ministry of Health, Manzini, Eswatini

[6]Department of Medicine, University of California, San Francisco, California, United States of America

[7]Department of Integrative Biology and Statistics, University of California, Berkeley, California, United States of America

[8]Clinton Health Access Initiative, Eswatini Country Office, Mbabane, Eswatini

[9]Chan Zuckerberg Biohub, San Francisco, California, United States of America

*Corresponding Authors

Email: jhuber3@nd.edu (JHH); taperkins@nd.edu (TAP)

Short Title: Inferring person-to-person transmission networks

Abstract

Inference of person-to-person transmission networks using routinely collected surveillance data is being used increasingly to estimate spatiotemporal patterns of pathogen transmission. Several data types can be used to inform transmission network inferences, yet the sensitivity of those inferences to different data types is not routinely evaluated. We evaluated the influence of different combinations of spatial, temporal, and travel-history data on transmission network inferences for *Plasmodium falciparum*, the pathogen responsible for most human malaria. After developing a new inference framework and applying it to simulated data, we found that these data types have limited utility for inferring transmission networks and, in some combinations, tend to overestimate transmission. Only when outbreaks were highly focal in time or when travel histories were highly accurate was the inference algorithm able to accurately estimate the reproduction number under control, R_c, a key metric of transmission. Applying this approach to surveillance data from Eswatini indicated that inferences of R_c and spatiotemporal patterns therein are sensitive to the choice of data types and assumptions about the accuracy of travel-history data. Taken together, these results suggest that transmission network inferences made with routinely collected surveillance data should be interpreted with caution. As we have done here, future studies inferring transmission networks should apply their algorithm to data simulated under alternative assumptions to assess the robustness of their inferences.

Author Summary

Person-to-person networks of pathogen transmission can reveal epidemiological patterns that are not obvious based on routinely collected surveillance data alone. Several types of data can be used to inform estimates of these networks, including the timing of when cases present, the locations where those individuals reside, and whether those individuals recently traveled to an area with higher transmission. Little is known about whether transmission network inferences may depend on the types of data used to inform those inferences. To address this question, we developed an algorithm to estimate transmission networks of *Plasmodium falciparum* malaria. We validated this algorithm on simulated data and then applied it to surveillance data from Eswatini. Our results demonstrate that transmission network inferences are sensitive to which data types are used to inform them and to assumptions about how to interpret those data, such as whether data about recent travel are accurate. These results suggest that future work should carefully consider the epidemiological setting, which data types are used to inform transmission network inferences, and how those inferences are interpreted.

Introduction

Concomitant with improved epidemiological surveillance, there is growing interest to leverage the collected data to infer transmission networks for a wide range of pathogens and to use those inferences to inform public health efforts. Past studies have incorporated temporal data [1] and spatial data [2–5] to estimate pairwise probabilities of transmission between individual cases and to use those estimates to infer time-varying and spatially varying reproduction numbers, respectively. More recently, methods have been developed to incorporate this type of detailed, individual-level epidemiological data [6–8] to infer transmission networks for infectious diseases of humans, including severe acute respiratory syndrome [9] and tuberculosis [10], and of animals, such as rabies [11] and foot-and-mouth disease [12].

In addition to the diseases for which these methods have been applied to date, there is a growing need to apply similar methods to malaria in near-elimination settings. As incidence of malaria declines within a country, transmission becomes more heterogeneous in space and time [13]. Focal areas of high transmission, known as "hotspots," pose a serious risk of fueling resurgence if left untargeted, potentially reversing decades of progress towards elimination [14]. To this end, granular estimates of when and where transmission occurs are needed, as spatially aggregated estimates may obscure important heterogeneities of practical relevance to control efforts [15]. In addition to characterizing details of local transmission, measurement of progress towards malaria elimination hinges on correct classification of cases as imported or locally acquired [16,17], which is a byproduct of estimating transmission networks.

Previous work on malaria has made progress on the use of individual-level epidemiological data to infer transmission networks of *Plasmodium falciparum*, the parasite primarily responsible for human malaria in many regions of the world. Churcher *et al.* [18] used

83 temporal data to estimate the proportion of imported cases needed to confidently estimate the
84 reproduction number under control, R_c, below one and thereby provide evidence of controlled,
85 non-endemic malaria transmission. Reiner *et al.* [6] then built upon this work by incorporating
86 spatial data and inferring an individual-level transmission network of *P. falciparum* in Eswatini.
87 More recently, Routledge *et al.* [19,20] used related approaches to infer transmission networks
88 and R_c of *P. vivax* in El Salvador and China.

89 As the adoption of these methods increases, both for malaria and for other infectious
90 diseases, care should be taken to assess how the epidemiological setting and the inclusion or
91 exclusion of certain data types might affect transmission network inferences, as well as resultant
92 inferences about epidemiological quantities including R_c and spatiotemporal variation therein. To
93 date, transmission network inferences have relied on various types of epidemiological data,
94 ranging from only the use of timing of symptom onset [1] to more detailed spatiotemporal data
95 [2–6,20]. Additionally, studies that have incorporated travel-history information into
96 transmission network inferences have considered these data to be perfectly accurate, assuming
97 that all cases that reported travel were imported [6,18–20]. However, travel history may be an
98 imperfect indicator of importation owing to errors in recall [17] and the fact that travel to an area
99 of ongoing transmission alone does not guarantee that an individual was infected there [17,21].
100 Thus, how different choices about data types and different assumptions about possible errors in
101 travel-history data affect transmission network inferences remain largely unexplored [22].

102 Here, we present a Bayesian method for inferring transmission networks based on
103 temporal, spatial, and travel-history data for individual malaria cases. We use it to characterize
104 the sensitivity of transmission network inferences to the inclusion of different data types and to
105 different assumptions about the accuracy of travel histories. Our method builds upon previous

work by leveraging individual-level epidemiological data to obtain posterior estimates of transmission networks and model parameters in a way that is flexible enough to accommodate different assumptions about errors in travel histories. After establishing a proof-of-concept of our inference method on simple test cases, we applied our method to real-world surveillance data from Eswatini and additional simulated data sets to understand how the inclusion or exclusion of different data types and different assumptions about travel-history error affect our ability to infer transmission networks and estimate transmission metrics, such as R_c.

Results

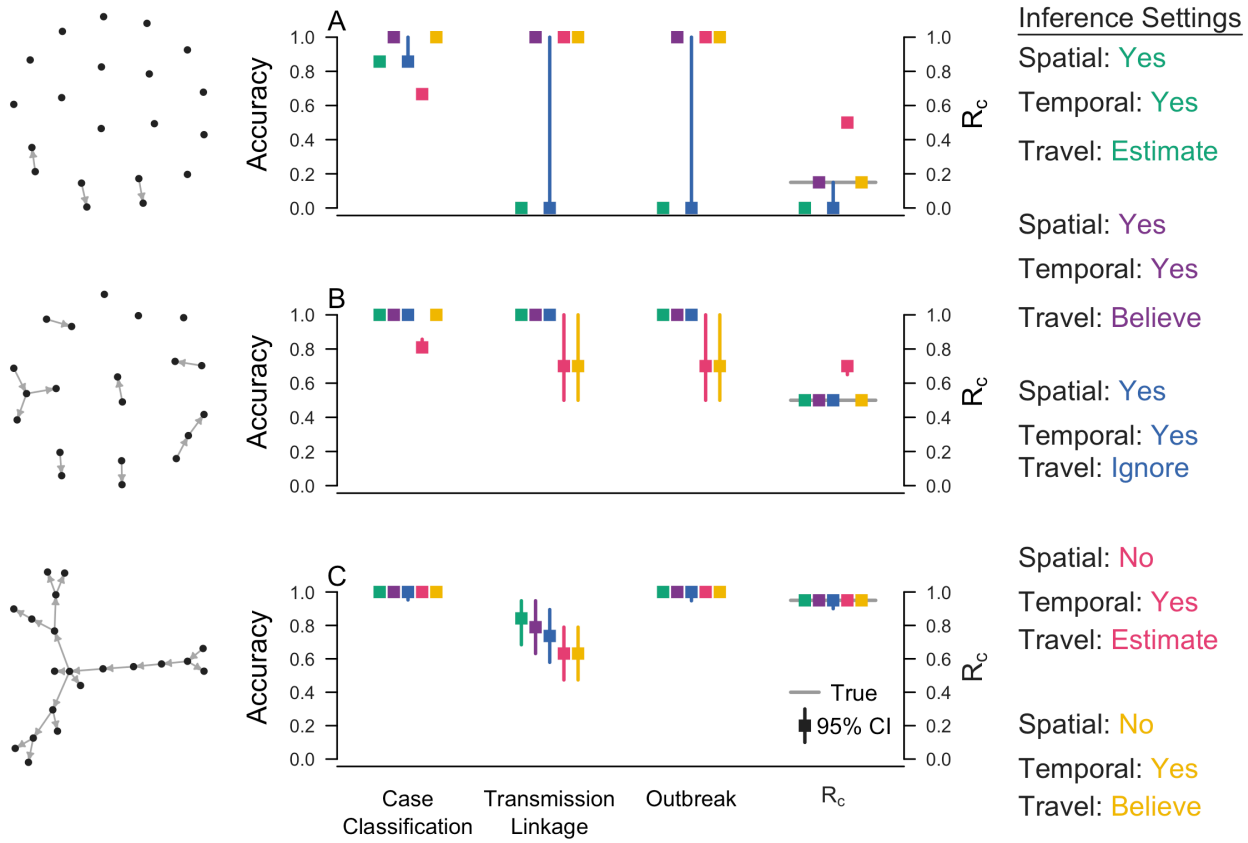
To establish proof-of-concept, we first applied our inference method on three simple test cases and evaluated how well our inferences recovered the true transmission networks. We then applied our method to surveillance data collected in Eswatini during 2013-2017 and assessed how our epidemiological conclusions changed with the inclusion or exclusion of different data types and different assumptions about travel histories. These inference settings used: (1) spatial and temporal data while estimating the accuracy of the travel history (default setting); (2) spatial and temporal data while believing the travel history; (3) spatial and temporal data alone; (4) temporal data while estimating the accuracy of the travel history; and (5) temporal data while believing the travel history. To validate the inferences based on data from Eswatini, we simulated data generated using posterior parameter estimates obtained from the data from Eswatini and evaluated the ability of our inference method to recover the true transmission networks along with the underlying parameters on those simulated data. Finally, we performed a simulation sweep across different epidemiological settings to determine the range of conditions

under which our inference method yielded reliable estimates of transmission. A full description of the analyses and additional results can be found in the S1 File.

Validation using a simple test case

We first validated our approach on three small simulated networks of twenty nodes. Although these networks varied in their proportion of imported cases, the local transmission chains were arranged optimally both temporally and spatially with perfect travel histories and complete observation of cases, providing idealized test cases to validate our inference algorithm. We measured the performance of our inference algorithm in terms of its ability to reconstruct different features of the transmission network and correctly estimate R_c. As the proportion of imported cases decreased from 90% to 5%, we found that the ability of the algorithm to correctly classify cases as imported or locally acquired improved (Fig 1). For example, when we used the default inference setting, classification accuracy improved from 85.7% (95% Uncertainty Interval: 85.7 – 85.7%) to 100% (100 – 100%). As classification accuracy improved, our estimates of R_c also improved, with all five inference settings yielding accurate estimates when imported cases comprised only 5% of total cases (Fig 1C). Similarly, the ability to identify the correct parent of each locally acquired case and assign it to the correct outbreak depended on the extent of local transmission in the network. When 90% of the cases were imported, performance was variable across inference settings (Fig 1A). Believing the travel history, regardless of whether spatial data was included, enabled us to perfectly reconstruct the transmission network, because the simulated travel-history data were accurate. However, as the proportion of imported cases decreased, the benefit of believing the travel history diminished, and using all data types resulted in the greatest accuracy. For example, in the most extreme case in which only 5% of

151 cases were imported, the accuracy of identifying the true parent ranged from 84.2% (68.4 –
152 94.7%) under the default inference settings to 63.2% (47.4 – 78.9%) using temporal data and
153 estimating the accuracy of the travel history (Fig 1C). In terms of identifying the outbreak to
154 which a case belongs, the algorithm was accurate under all inference settings, since there was
155 only one outbreak (Fig 1C).
156



157

158 **Fig 1. Inference accuracies for three simulated transmission networks with (A) 90%, (B)**
159 **50%, and (C) 5% of cases as imported.** Case classification refers to proportion of cases that are
160 correctly classified as imported vs. locally acquired. Transmission linkage denotes the proportion
161 of locally acquired cases for which the true parent is correctly identified, Outbreak is the
162 proportion of locally acquired cases for which the inferred parent belongs to the correct outbreak,
163 and R_c is the estimated reproduction number under control. Square points signify the median

posterior value, and bars are the 95% uncertainty intervals. The gray line indicates the true value of R_c.

Application to Eswatini surveillance data

We next applied our method to surveillance data collected in Eswatini during 2013-2017. Under the default inference setting, we estimated the diffusion coefficient D, which quantifies the spatial spread of transmission, to be 4.42 km^2 day^{-1} (2.92 – 6.18 km^2 day^{-1}) (Fig 2A). This corresponded to a median inferred transmission distance of 13.0 km (0.0160 – 65.9 km), a median inferred serial interval of 47 days (-33 – 150 days) (Fig 3A & 3B), and median estimates of τ_s, the probability that an imported case reported travel, of 0.61 (0.44 – 0.78) and τ_l, the probability that a locally acquired case reported travel, of 0.57 (0.53 – 0.61) (Fig 2B & 2C). That the 95% credible interval for τ_s contained 0.50 indicated that our inference algorithm found limited use of travel-history data in discriminating between imported and locally acquired cases, because that implies that imported cases have equal probabilities of reporting or not reporting travel. The algorithm estimated the proportion of imported cases to be 0.052, corresponding to R_c = 0.95. Mapping risk of importation and local transmission across Eswatini under the default inference setting, we estimated consistently low risk of importation throughout the country and transmission hotspots in the northeastern part of Eswatini, close to the border with Mozambique (Fig 4A & 4B).

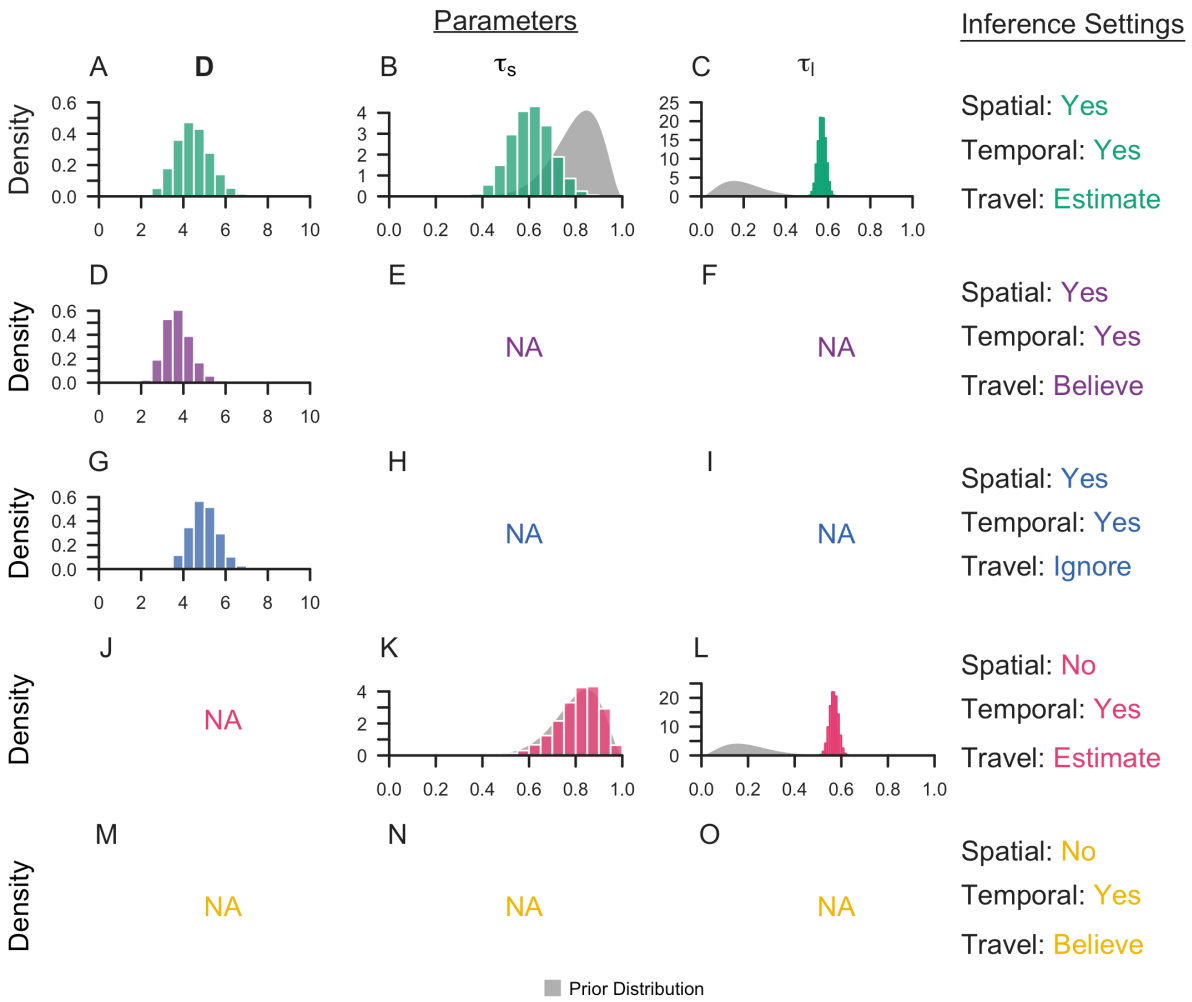


Fig 2. Marginal posterior distributions of parameters from Eswatini surveillance data. Histograms represent the marginal posterior distribution of each parameter, color-coded by the inference settings used. D is the diffusion coefficient, τ_s is the probability that an imported case reports travel, and τ_l is the probability that a locally acquired case reports travel. Gray shapes represent the prior distributions placed on each parameter. Inference settings in which a given parameter was not estimated are indicated by NA.

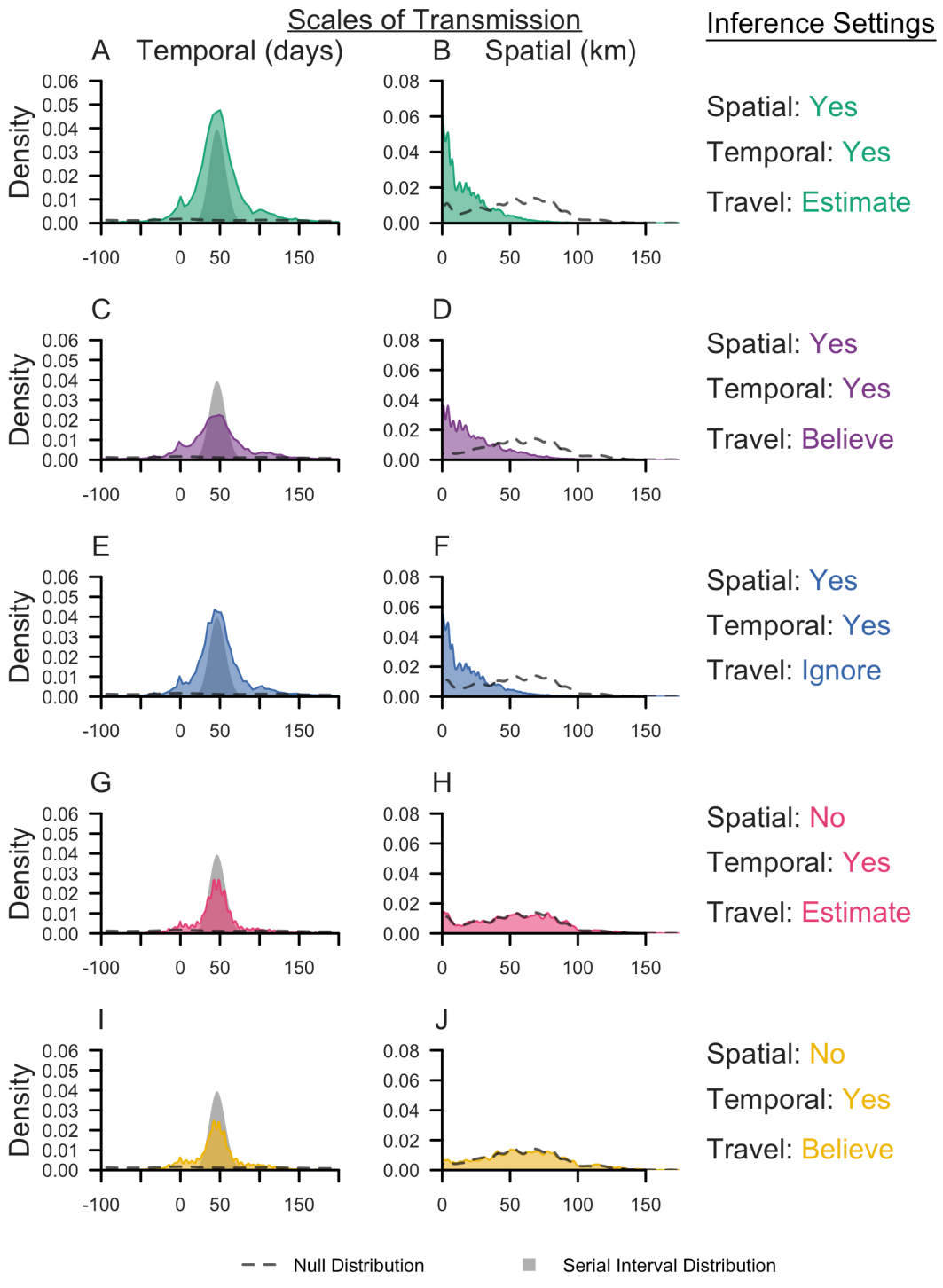


Fig 3. Spatial and temporal scales of transmission in Eswatini. Density plots of the spatial (km) and temporal (days) scales of transmission are reported and color-coded for each inference setting. Dashed lines indicate the corresponding null distribution, generated from all random

pairs of cases in the Eswatini surveillance data set. The grey shape is the serial interval distribution used in the likelihood.

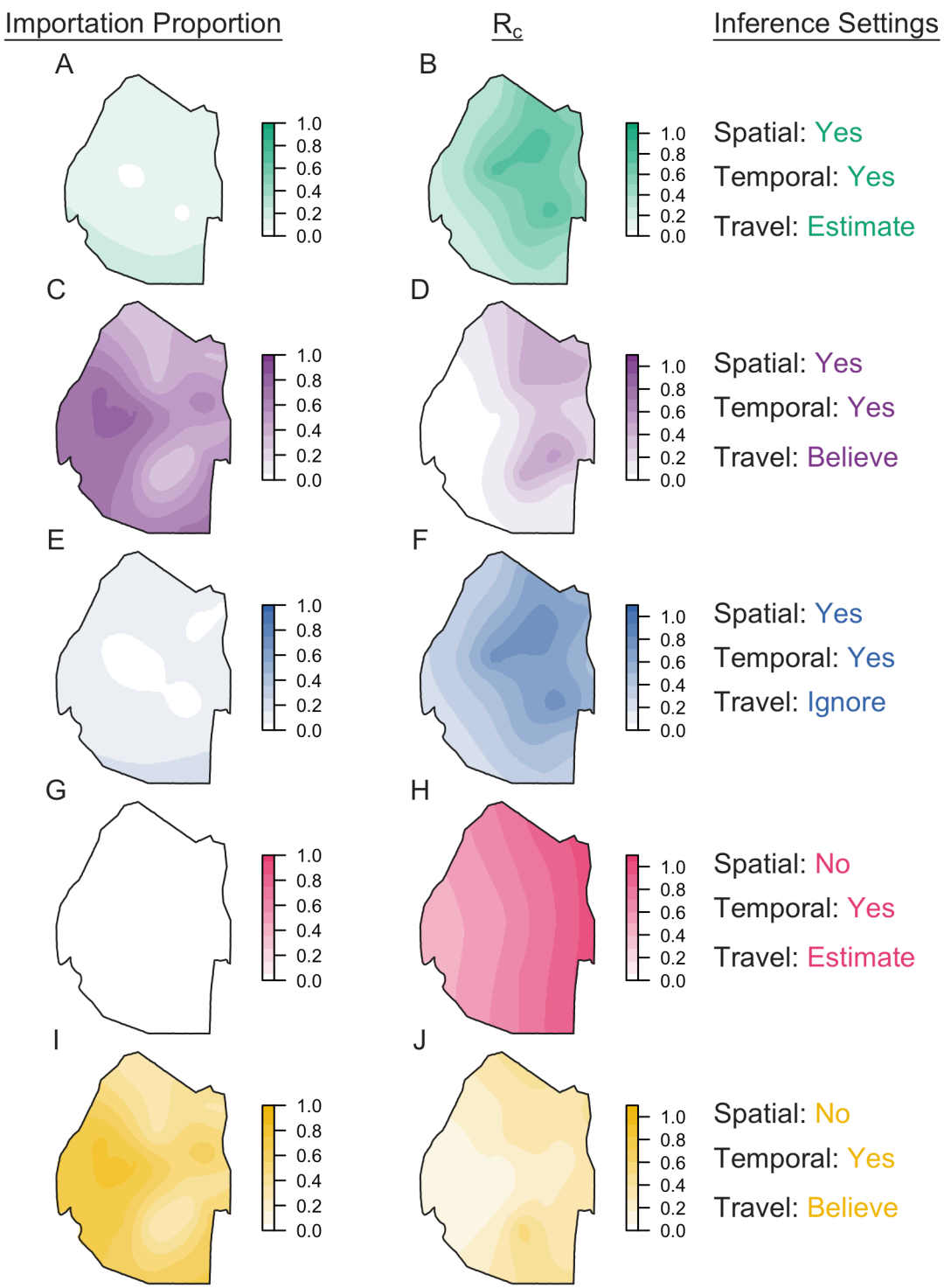


Fig 4. Spatial distribution of importation and transmission risk in Eswatini. Maps of the proportion of cases that are imported and the reproduction number under control (R_c) were

generated for each inference setting using a generalized additive model with a Gaussian process basis function setting using the mgcv package in R [23,24]. In each plot, darker colors indicate greater importation or transmission risk.

Parameter estimates and transmission network inferences differed under other inference settings. When we believed the travel history, we estimated a larger median transmission distance (Fig 3D). We attribute this increase in the spatial scale of transmission to clusters of cases with positive travel histories located near metropolitan areas. By forcing those cases to be imported, the algorithm tended to infer transmission across longer distances to explain the origins of the remainder of cases that did not report travel and were thereby inferred to be locally acquired. With respect to time, all five inference settings produced consistent serial interval estimates, though the inclusion of spatial data allowed for a wider range of transmission linkages in time (Fig 3A, 3C, & 3E). Finally, in the absence of spatial data, the model estimated higher predictive power of travel histories in identifying imported cases (τ_s: 0.83, [0.60, 0.95]), though the travel history was consistently found to be uninformative for identifying locally acquired cases (τ_l: 0.57, [0.53, 0.60]) (Fig 2K & 2L).

Classification of cases as imported or locally acquired, key information for control programs, was sensitive to the choice of inference setting. The proportion of cases classified as imported was most sensitive to different assumptions about the accuracy of the travel histories (Fig 4, left column). Believing the travel history yielded high estimates of importation in western Eswatini (Fig 4C & 4I), whereas estimating or ignoring the travel history yielded low, relatively homogeneous estimates of importation risk (Fig 4A, 4E, & 4G). For instance, using temporal data and estimating the accuracy of the travel history produced probabilities of importation that

225 ranged 0.0043 – 0.0050, suggesting that nearly all cases resulted from local transmission (Fig
226 4G). Estimates of the spatial distribution of R_c depended most on the choice of which data types
227 we included (Fig 4, right column). Notably, inclusion of spatial and temporal data produced a
228 consistent spatial distribution of relative transmission risk, with transmission hotspots in
229 northeastern Eswatini (Fig 4B, 4D, & 4F). However, believing the travel history reduced the
230 magnitude of transmission that we inferred from a median R_c of 0.95 (Fig 4B) under default
231 settings to 0.41 (Fig 4D). Omitting spatial data changed the spatial distribution of transmission.
232 Estimating the accuracy of the travel history yielded high transmission estimates (median R_c:
233 1.00) in eastern Eswatini (Fig 4H), whereas believing the travel history inferred hotspots of
234 transmission (median R_c: 0.42) in southern Eswatini (Fig 4J).
235
236 **Validation of inferences from Eswatini**
237 Reconciling the different inferences under different inference settings in Figs. 2-4 was
238 challenging because the true, underlying network and parameters were unknown. Using median
239 posterior estimates from the Eswatini data under each inference setting, we simulated data and
240 assessed the ability of the inference method to recover the underlying parameters and
241 transmission networks (Table 1). We found that the model was able to estimate τ_s and τ_l
242 reasonably well, depending on the inference setting (Fig 5). One exception was that the
243 algorithm slightly overestimated τ_s under the default inference setting. We attribute this to the
244 low proportion (0.052) of imported infections in the simulated data set and the strong prior
245 placed on this parameter. This tendency was not observed under the inference setting where
246 spatial data was excluded, because the true value of τ_s closely matched the mean of the prior
247 distribution in that case (see S2 File for further discussion).

Table 1. Characteristics of simulated data generated using the branching process model. A description of the simulated data used in the inference exercises are reported for each of the five inference settings. The total number of nodes in the network, the number of distinct outbreaks, the proportion of cases that are imported, and the underlying parameters are provided.

Inference Setting			Network Size	# of Outbreaks	Prop. Imported	D	τ_s	τ_l
Space	**Time**	**Travel**						
Yes	Yes	Estimate	775	43	0.052	4.42	0.61	0.57
Yes	Yes	Believe	775	492	0.59	3.70	1	0
Yes	Yes	No	775	36	0.043	4.96	NA	NA
No	Yes	Estimate	775	1	0.0013	NA	0.83	0.57
No	Yes	Believe	775	489	0.58	NA	1	0

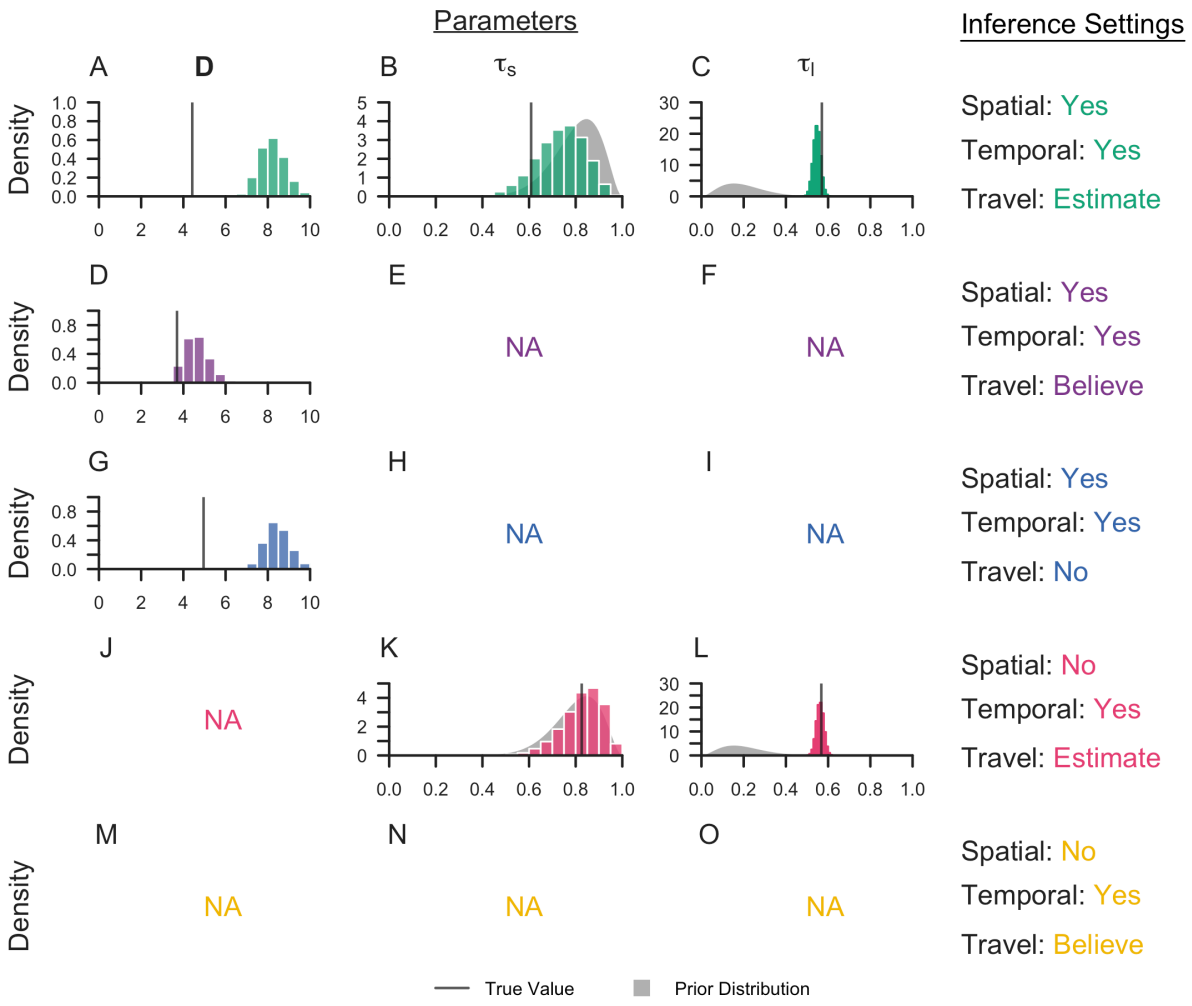


Fig 5. Marginal posterior distributions for parameters inferred from simulated data. The marginal posterior distributions are reported for each inference setting from its respective simulated data set. Each line denotes the true value of the parameter, and the grey shapes represent the prior distributions of the parameters. Inference settings in which a given parameter was not estimated are indicated by NA.

With the exception of believing the travel history, the model consistently overestimated the diffusion coefficient (Fig 5A, 5D, & 5G). We attribute the challenge of correctly estimating the diffusion coefficient to an inability to correctly estimate the underlying transmission network,

264 the extent of local transmission in the network, and a numerical insensitivity in the overall
265 likelihood to changes in *D*. When we conditioned the likelihood of *D* on the true transmission
266 network when R_c was high, the true values of *D* fell close to the range of maximum-likelihood
267 estimates, suggesting that this parameter could be estimated correctly if the true network was
268 identified (S1 Fig). The likelihood around the true value was very flat, however, making it easy
269 for *D* to be estimated incorrectly. When R_c was low, we underestimated the diffusion coefficient,
270 because the likelihood of imported cases increases as *D* decreases.

271 The overall accuracy of classifying cases as imported or locally acquired was close to one
272 (Fig 6). Though seemingly promising, these high accuracies masked a tendency to overclassify
273 cases as locally acquired, because many more cases were simulated to be locally acquired than
274 imported. For example, under the default inference setting, the accuracy of correctly classifying
275 imported cases was 0.023 (0.023 – 0.067). Similarly, the accuracies of identifying the parent of
276 each transmission linkage were poor, despite simulating under the assumptions of the model,
277 with accuracies ranging from 0.042 (0.022 – 0.065) when using temporal data and believing the
278 travel history to 0.31 (0.28 – 0.34) when incorporating spatial and temporal data and estimating
279 the accuracy of the travel history (Fig 6, circle points). This suggests that, as the number of cases
280 increases within a fixed space-time window, the information content of routinely collected
281 epidemiological data decreases and the method becomes incapable of correctly estimating the
282 transmission network. Nevertheless, under some settings, the method was able to capture higher-
283 order summaries of the network, such as case classification and R_c (Fig 6, square and diamond
284 points).
285

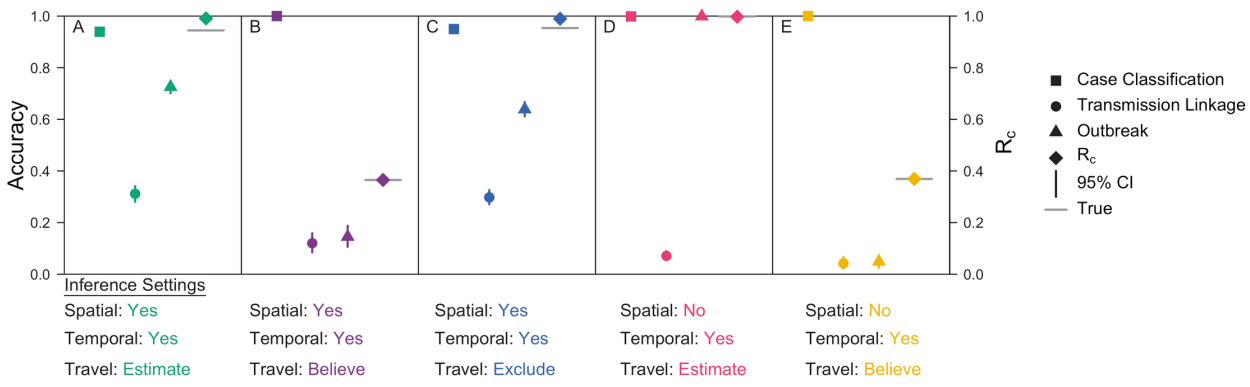


Fig 6. Inference accuracies for validation exercises. Accuracy metrics are reported for each inference setting applied to its respective simulated data set. Case Classification, represented by squares, refers to the proportion of cases that are correctly classified as imported or locally acquired. Transmission Linkage, denoted by circles, is the proportion of locally acquired cases for which the true parent is correctly identified. Outbreak, represented by triangles, is the proportion of locally acquired cases for which the inferred parent belongs to the correct outbreak. Bars denote the 95% credible intervals, and the grey line is the true R_c value of the network.

Simulation Sweep

Validation of our inference algorithm revealed that its performance varied across simulated data sets. When applied to a series of simple test cases in which the transmission networks were small and in an optimal spatiotemporal arrangement, the inference method was able to reconstruct the transmission network and correctly estimate R_c (Fig 1). When applied to larger transmission networks in which outbreaks overlapped in space and time, performance of the inference method was poor (Fig 6). This indicated that the performance of our inference algorithm depends on the epidemiological setting to which it is applied. To address this observation, we generated 2,000 simulated data sets in which we varied the proportion of imported cases, the spatiotemporal window over which imported cases were distributed, the diffusion coefficient, and the accuracies

305 of the travel history (i.e., τ_s and τ_l) (S2 Table). We then applied our inference algorithm under
306 three different inference settings and quantified the accuracy of reconstructing each transmission
307 network. The three inference settings used: (1) spatial and temporal data while estimating the
308 accuracy of the travel history (default setting); (2) spatial and temporal data while believing the
309 travel history; and (3) spatial and temporal data alone (S1 Table).

310 We observed that the accuracy of reconstructing transmission networks depended upon
311 both the inference setting used and the epidemiological features of the simulated data. When we
312 used spatial and temporal data and estimated the accuracy of the travel history or excluded it, the
313 accuracy of reconstructing transmission networks depended on the relative proportion and
314 temporal distribution of imported cases (S8 and S9 Figs). As the temporal window over which
315 imported cases are distributed increased, the accuracy of identifying the true parent and the true
316 outbreak of each locally acquired case increased. With an increasing temporal window,
317 outbreaks within the transmission network became relatively more focal in time, which made the
318 likelihoods of alternative transmission linkages more readily distinguishable. More accurate
319 estimates of R_c under these inference settings similarly depended on the temporal window over
320 which imported cases were distributed (Fig 7A, 7C). When the mean temporal interval between
321 imported infections was greater than two times the mean length of the serial interval (i.e.,
322 approximately 100 days), our estimates of R_c improved, though we generally overestimated it.
323 Furthermore, as the proportion of imported cases increased and R_c decreased, the accuracy of
324 identifying the correct outbreak of each locally acquired case decreased (S8 and S9 Figs). This
325 pattern reflected the relationship between R_c and the size of individual outbreaks. As R_c
326 decreased, the size of individual outbreaks decreased, and, consequently, the probability that the
327 inferred parent of a locally acquired case belonged to the same outbreak decreased.

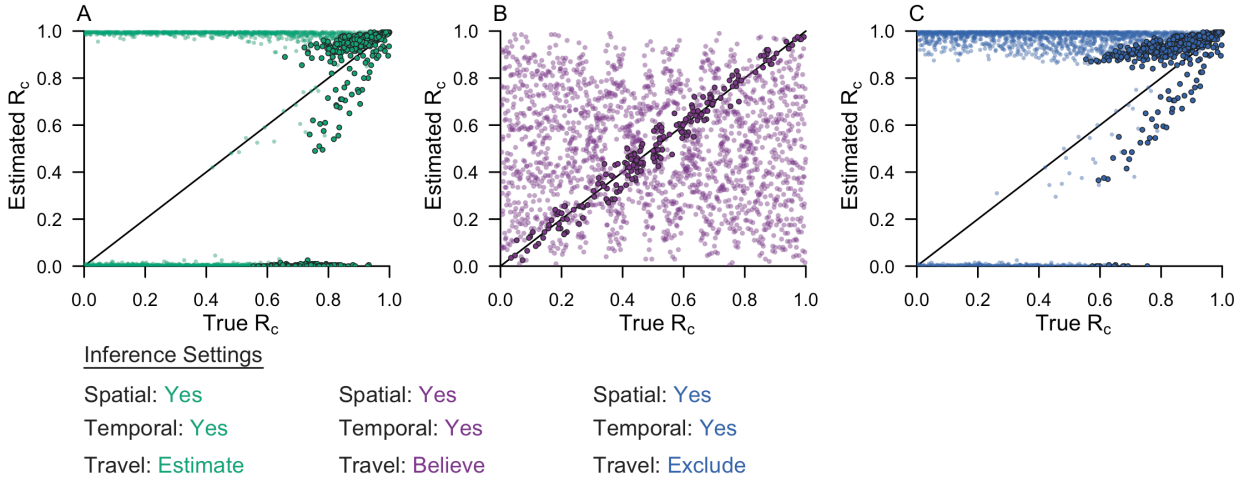


Fig 7. Comparison of R_c estimates across inference settings. The inference algorithm was applied to 2,000 simulated data sets. The estimated R_c is compared to the true R_c for each of the inference settings: (A) spatial and temporal data while estimating the accuracy of the travel history; (B) spatial and temporal data while believing the travel history; and (C) spatial and temporal data alone. Each point represents a simulated data set. The darker, accented points are simulated data sets with epidemiological features that improved performance. In (A) and (C), the darker, accented points were simulated data sets where the mean temporal interval between imported infections was greater than two times the mean serial interval. In (B), the darker, accented points were simulated data sets where the proportion of cases reporting travel was within 0.05 of the proportion of imported cases.

By contrast, when we believed the travel history, the accuracy of reconstructing transmission networks depended most strongly on the accuracies of the travel history. As the probability of reporting travel increased, the accuracy of classifying imported cases increased, and the accuracy of classifying locally acquired cases decreased (S10 Fig). Under this inference

345 setting, our estimate of R_c depended only on the proportion of cases that reported travel. When
346 the proportion of cases that reported travel matched the proportion of cases that were imported,
347 we correctly estimated R_c (Fig 7B).
348

349 ## Discussion

350 Our results show that, in most settings, routinely collected surveillance data offer limited value
351 for reconstructing individual-level transmission networks of *P. falciparum* malaria and informing
352 estimates of the reproduction number under control, R_c. Using simulated data similar to the
353 Eswatini surveillance data that we analyzed, we observed that our inference algorithm correctly
354 identified transmission linkages less than 35% of the time. We attribute this inaccuracy primarily
355 to the inherently limited information content of spatiotemporal data on *P. falciparum*. Its
356 characteristically long serial interval [25] means that an appreciable number of cases presenting
357 within a short timeframe are difficult to link to each other based on their timing, even in a
358 relatively facile test case in which the generative process assumed in the likelihood function
359 matched that used to simulate the data.

360 Under most simulated scenarios and assumptions about the accuracy of travel-history
361 data, we overestimated the number of locally acquired cases, leading to overestimates of R_c.
362 Crucially, our simulation sweep demonstrated that routinely collected surveillance data was most
363 informative of individual-level transmission networks and R_c when local outbreaks were highly
364 focal in time. Otherwise, while we were able to reconstruct the true transmission network with
365 modest accuracy, we tended to misclassify truly imported cases as locally acquired, thereby
366 overestimating R_c. Taken together, these results suggest limited use of routinely collected
367 surveillance data for informing fine-scale estimates of *P. falciparum* transmission. At broader

368 spatial scales, however, routinely collected surveillance data may still have practical value,
369 because the spatial distribution of cases can reveal epidemiological risk factors relevant for
370 targeted control interventions [26,27].

371 Although we were able to reach some general conclusions about our inference algorithm,
372 our inferences were highly sensitive to which data types we included and which assumptions we
373 made about the accuracy of travel-history data. Applying our algorithm to surveillance data from
374 Eswatini, we observed that inferred patterns of transmission depended on which data types we
375 included. With the inclusion of spatial data, we captured a spatial pattern of transmission
376 consistent with another analysis from Eswatini [28] with data from a different time period.
377 Assumptions about the travel history appeared to have a strong influence on the overall
378 magnitude of transmission that we inferred, due to the direct relationship between R_c and the
379 proportion of imported cases [16]. As a result, believing the travel history, and thereby treating it
380 as perfectly accurate as in previous approaches [6,18–20], could bias R_c estimates if there are
381 errors in travel-history data. A study comparing community travel surveys to mobile-phone data
382 in Kenya found that travel histories considerably underestimated the volume of travel, suggesting
383 high rates of false negatives in community travel surveys [29]. Believing the travel history may,
384 then, underestimate the number of imported cases and overestimate R_c.

385 The method that we used only considered a single spatial model to infer transmission
386 linkages and assumed complete observation of cases, both of which are factors that could have
387 affected our inferences based on the Eswatini surveillance data. The diffusion model that we
388 used to represent spatial dispersion of parasites assumed that movement is isotropic in space and
389 did not consider landscape features, such as heterogeneity in human population densities and
390 environmental factors that may affect mosquito ecology. A study analyzing self-reported

391 movement patterns in Mali, Burkina Faso, Zambia, and Tanzania found that gravity and radiation
392 models of spatial dispersion fit the data well, though the appropriateness of each model depended
393 on the type of traveler, the travel distance, and the population size of the destination considered
394 [30]. Regarding the representation of *P. falciparum* infections in our data set from Eswatini,
395 there are asymptomatic and mild infections that are unlikely to have been recorded in the
396 surveillance system yet may comprise a substantial proportion of malaria infections within
397 Eswatini [13]. Accordingly, it is possible that our assumption of complete observation of cases
398 could have biased R_c estimates, likely downward due to the fact that missing cases will tend to
399 make offspring numbers appear smaller than they actually are [31,32]. Even so, we expect that
400 our conclusions about the sensitivity of transmission network inferences to the choice of data
401 types and assumptions about travel-history data are robust to these limitations of our study. This
402 reinforces our conclusion of the need for caution in attempting to reconstruct person-to-person
403 transmission networks from routine surveillance data.
404 Given that some of the limitations of our approach may be inherent to the information
405 content of these data types in this system, one potential avenue for improving inferences of fine-
406 scale patterns of *P. falciparum* transmission could involve the integration of additional data
407 streams. For example, mobile-phone data [33] and high-resolution friction surfaces [34] could
408 more realistically characterize mobility patterns, whereas travel-history information that details
409 the dates, duration, and location of each trip that has been used in programmatic contexts [26]
410 could more accurately identify importation events. Additionally, the inclusion of pathogen
411 genetic data, which has the potential to provide a more direct signal of parasite movement, could
412 complement traditional epidemiological data [35]. There is also scope for further methodological
413 development, such as relaxing our assumption of complete observation of infections and

414 incorporating an underlying mechanistic model of transmission (as in Lau *et al*. [8]; Guzzetta *et*
415 *al*. [36]). Incorporating an underlying mechanistic model would relax our uninformative prior
416 assumption on all possible transmission networks, ruling out transmission networks that are
417 epidemiologically implausible and allowing us to account for spatial differences in transmission
418 potential due to different epidemiological factors. To this end, we envision that leveraging the
419 strengths of our method along with other, complementary methods could strengthen inferences
420 based on routinely collected epidemiological data and open up new possibilities to make use of
421 even more data types, such as serological data, prevalence surveys, and pathogen genetic data
422 [35,37,38].
423

424 **Conclusions**

425 We demonstrated the limitations of routinely collected surveillance data for the inference of
426 individual-level transmission networks of *P. falciparum*. We identified a tendency to
427 overestimate local transmission using routinely collected surveillance data, especially when
428 outbreaks overlapped in space and time. Using both real data from Eswatini and simulated data,
429 we identified strong sensitivities of our inferences to the epidemiological setting, the choice of
430 data types included, and assumptions about the accuracy of travel-history data. Our results
431 indicated that using spatial and temporal data and believing travel histories yielded the most
432 plausible estimates of transmission when applied to the Eswatini surveillance data. However, our
433 simulation sweep demonstrated that the accuracy of our inferences strongly depended on the
434 accuracy of the travel-history data when the travel-history data were assumed to be accurate.
435 These sensitivities to the choice of data types and assumptions about the accuracy of travel-
436 history data could have important programmatic implications if outputs of transmission network

437 inferences are operationalized. Although this study was specific to *P. falciparum*, the results of
438 our analyses indicate that future studies inferring transmission networks of *P. falciparum*, or any
439 pathogen, should carefully consider the epidemiological setting and the choice of data types and
440 assumptions that inform the model and should validate them using simulated data.

Methods

Bayesian framework for estimating transmission linkages

Our goal was to obtain probabilistic estimates of a transmission network N that defines transmission linkages among a set of known cases. The transmission network is defined as a directed, acyclic graph comprised of a set of directed edges represented as $N = \{N_{i,j}\}$ for all i, j. Each $N_{i,j}$ indicates that case i is hypothesized to contain parasites that are the most direct observed ancestors of the parasites contained in case j. In addition, at least one edge denoted $N_{u,j}$ must exist in the network, indicating that the parasites contained in case j have no ancestors among the parasites contained in any known local case and are instead contained in some unknown case u from some source population s, such that it is denoted u_s. To illustrate this terminology, an example transmission network is depicted in Fig 8.

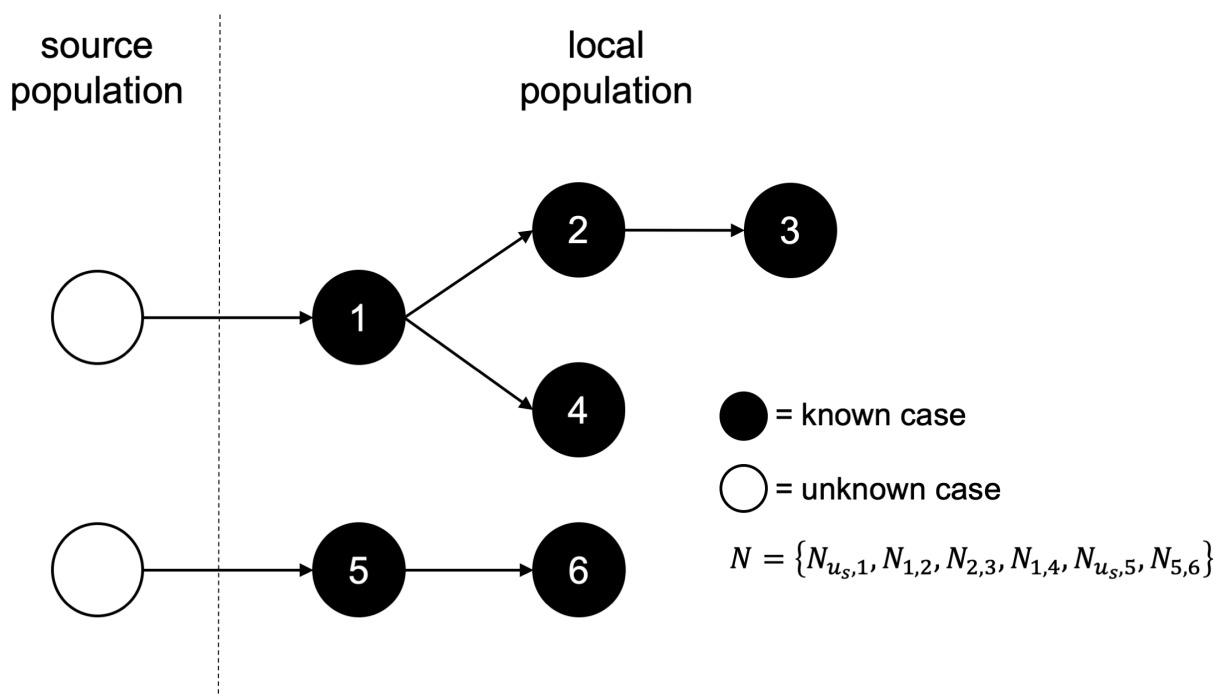


Fig 8. Schematic of a hypothetical transmission network. A hypothetical transmission network is presented along with the corresponding notation. In the schematic, white circles denote unobserved cases, and black circle denote observed cases. Arrows represent transmission between two cases.

To estimate *N*, we used spatial, temporal, and travel-history data about all cases, denoted as $\vec{X}_s$, X_t, and X_h, respectively. We did so within a Bayesian statistical framework, meaning that we sought to estimate the joint posterior probability density,

$$\Pr\left(N, \Theta \middle| X_t, \vec{X}_s, X_h\right) = \frac{\Pr\left(X_t, \vec{X}_s, X_h \middle| N, \Theta\right) \Pr\left(\mathrm{N}, \Theta\right)}{\Pr\left(X_t, \vec{X}_s, X_h\right)}, \quad (1)$$

of the transmission network defined by *N* and the model parameters Θ conditional on the data $\vec{X}_s$, X_t, and X_h. The first term in the numerator of eq. (1) is the likelihood of *N* and Θ conditional on the data. The second term in the numerator is the prior probability of *N* and Θ. The term in the denominator is the probability of the data, which is an intractable quantity to calculate directly given that it would require evaluation of an extremely high-dimensional integral over *N* and Θ. To address this, we used a Markov chain Monte Carlo algorithm to draw random samples of *N* and Θ from the posterior distribution specified in eq. (1).

The most critical piece of our inference framework is the likelihood, which we define as a function of each case *j* as

475 $$\mathcal{L}(N, \Theta | X_t, \vec{X}_s, X_h) = \prod_j \Pr\left(X_{t,j}, \vec{X}_{s,j}, X_{h,j} \middle| N_{.,j}, \Theta\right). \quad (2)$$

476

477 Below, we define the probability of the data associated with each known case *j* as a function of
478 different assumptions that are possible about how case *j* is connected to the rest of the
479 transmission network.

480

481 **Scenario 1: Local transmission between known cases *i* and *j***

482 When case *i* contains parasites that are immediate ancestors of the parasites contained in case *j*,
483 we represent its contribution to the likelihood as

484

485 $$\Pr\left(X_{t,j}, \vec{X}_{s,j}, X_{h,j} \middle| N_{i,j}, \Theta\right) = \Pr\left(X_{t,j} \middle| N_{i,j}, \Theta\right) \Pr\left(\vec{X}_{s,j} \middle| X_{t,j}, N_{i,j}, \Theta\right) \Pr\left(X_{h,j} \middle| N_{i,j}, \Theta\right), \quad (3)$$

486

487 which is the product of the probabilities of the temporal, spatial, and travel-history data given the
488 network and model parameters. This formulation assumes that those data are generated
489 independently for each individual, with the exception of a dependence of the spatial data on the
490 temporal data.

491

492 **Probability of the temporal data.** To characterize the time elapsed between two cases resulting
493 from local transmission, we used a model of the generation and serial intervals for *P. falciparum*
494 malaria by Huber *et al*. [25]. The generation interval represents the time between infection of a
495 primary and secondary case, whereas the serial interval represents the time between detection of
496 those cases. Because the timing of infection per se (i.e., an infectious mosquito inoculating a

susceptible human) is typically unknown, we focused on the serial interval as the most apropos temporal quantity relating cases.

In deriving the probability of a given length of the serial interval, Huber *et al.* [25] convolved a random variable representing variability in the generation interval (*GI*) with a random variable representing variability in the time between infection with *P. falciparum* and detection by surveillance – i.e., the infection to detection period (*IDP*). That framework yielded

$$\Pr\left(SI_{i,j} = -a + b + c\right) = \sum_{a}\sum_{b}\sum_{c} \Pr(IDP_i = a)\,\Pr\left(GI_{i,j} = b\right)\,\Pr\,(IDP_j = c), \quad (4)$$

as the probability of a serial interval of length $SI_{i,j}$. We allowed for different models of the serial interval depending upon differences in the *GI* and *IDP* for different types of primary and secondary cases. For instance, we assumed that symptomatic cases present in a clinic some number of days after infection as informed by empirical data from Zanzibar [25]. For an asymptomatic infection, we assumed that detection occurred through active surveillance at a randomly drawn day among all days where its asexual parasitemia exceeds a detection threshold [25]. The choice of *IDP* for both the primary and secondary case informs the probability of two cases separated in time by $SI_{i,j} = X_{t,j} - X_{t,i}$ days.

Probability of the spatial data. Following Reiner *et al.* [6], we assumed that a simple two-dimensional Wiener diffusion process determines the location of secondary cases relative to the location of their associated primary case. It follows that, for a given diffusion coefficient *D* with units $km^2 day^{-1}$ and generation interval $GI_{i,j}$, the two-dimensional location $\vec{X}_{s,j}$ of the secondary case *j* is described by a bivariate normal distribution with probability density

$$f\left(\vec{X}_{s,j}\,\middle|\vec{X}_{s,i}, D, GI_{i,j}, N_{i,j}, \Theta\right) = \frac{1}{2\pi\sigma^2\left(GI_{i,j}\right)} e^{-\frac{\|\vec{X}_{s,j}-\vec{X}_{s,i}\|}{2\sigma^2\left(GI_{i,j}\right)}}, \quad (5)$$

where $\sigma^2\left(GI_{i,j}\right) = DGI_{i,j}$. This formulation assumes that each spatial dimension is independent, that the variance scales linearly with the generation interval, and that movement is isotropic across a continuous landscape.

One complication to eq. (5) is that the generation interval $GI_{i,j}$ is unobserved and, therefore, cannot take on a fixed value. Instead, we must use data about the serial interval $SI_{i,j}$ to inform our generative model for $\vec{X}_{s,j}$. To do so, we take advantage of the property of normal random variables that the sum of two or more random variables is itself a normal random variable [39]. This property allows us to recast eq. (5) as a function of *SI* rather than *GI* by computing the appropriate σ^2 as

$$\sigma^2(SI) = \int \sigma^2(GI)\,\Pr(GI|SI)\,dGI, \quad (6)$$

which is effectively a weighted sum of the spatial variances associated with a given *GI* proportional to the probability of each *GI* conditional on the observed *SI*. This results in

$$f\left(\vec{X}_{s,j}\,\middle|\vec{X}_{s,i}, D, SI_{i,j}, N_{i,j}, \Theta\right) = \frac{1}{2\pi\sigma^2\left(SI_{i,j}\right)} e^{-\frac{\|\vec{X}_{s,j}-\vec{X}_{s,i}\|}{2\sigma^2\left(SI_{i,j}\right)}}, \quad (7)$$

as the probability density of the spatial data that we assume.

In the event that case i has missing spatial data, we cannot compute the spatial likelihood of eq. (7). To address this, we define a latent unobserved quantity $\tilde{X}_{s,i}$, which represents the unknown location of case i. We then integrate over the uncertainty in $\tilde{X}_{s,i}$,

$$f\left(\vec{X}_{s,j} \middle| D, SI_{i,j}, N_{i,j}, \Theta\right) = \int f\left(\vec{X}_{s,j} \middle| \tilde{X}_{s,i}, D, N_{i,j}, \Theta\right) f\left(\tilde{X}_{s,i} \middle| \vec{X}_{s,j}, D\right) d\,\tilde{X}_{s,i}, \quad (8)$$

to compute the probability density of case j with known spatial location $\vec{X}_{s,j}$ arising from case i with unknown spatial location $\tilde{X}_{s,i}$. Equation (8) is computed as the product of the probability density of the location of a known case j conditional on an unknown location $\tilde{X}_{s,i}$ and the probability density of spatial separation $\vec{X}_{s,j} - \tilde{X}_{s,i}$ conditional on the diffusion coefficient D for all $\tilde{X}_{s,i}$. Because we assume that movement is isotropic, eq. (8) is a two-dimensional Gaussian integral, simplifying to

$$f\left(\vec{X}_{s,j} \middle| D, SI_{i,j}, N_{i,j}, \Theta\right) = \frac{1}{4\pi\sigma^2(SI_{i,j})}. \quad (9)$$

In the event that case j has missing spatial data and case i has known spatial data, the latent unobserved quantity becomes $\tilde{X}_{s,j}$. We then integrate over the uncertainty in $\tilde{X}_{s,j}$ and calculate $f\left(\vec{X}_{s,j} \middle| D, SI_{i,j}, N_{i,j}, \Theta\right)$ using eq. (8-9).

Probability of the travel-history data. Although we assume in this scenario that a person's infection was locally acquired, our model must still be capable of explaining the travel-history

data $X_{h,j}$. We define a probability τ_l that case j reported travel (i.e., $X_{h,j} = 1$) even though they were not infected during that period of travel, such that

$$\Pr\left(X_{h,j} \middle| N_{i,j}, \Theta\right) = \begin{cases} \tau_l, & X_{h,j} = 1 \\ 1 - \tau_l, & X_{h,j} = 0 \end{cases}. \quad (10)$$

In the event that case j has missing travel-history data, we cannot compute the travel-history likelihood of eq. (10). To address this, we defined a latent unobserved quantity $\tilde{X}_{h,j}$, which represents the unknown travel history of case j. We then sum across the uncertainty in $\tilde{X}_{h,j}$,

$$\Pr\left(X_{h,j} = \mathrm{NA} \middle| N_{i,j}, \Theta\right) = \Pr\left(\tilde{X}_{h,j} = 1\right)\tau_l + \left(1 - \Pr\left(\tilde{X}_{h,j} = 1\right)\right)(1 - \tau_l), \quad (11)$$

to compute the probability that case j was locally acquired given an unknown travel history. In eq. (11), $\Pr\left(\tilde{X}_{h,j} = 1\right)$ was computed as the proportion of cases with a positive travel history among all cases with known travel-history data.

Taken together with the probabilities of the temporal and spatial data described above, the product of these three probabilities constitutes the entirety of the contribution of a case j infected by a known local case i to the overall likelihood of N and Θ.

Scenario 2: Importation of local case *j* from source population *s*

In the event of $N_{u_s,j}$, we represent the contribution of such a case to the overall likelihood of N and Θ as the product of the probabilities of its temporal, spatial, and travel-history data under

similar assumptions as in Scenario 1. The key difference in this scenario is that there is no information about the unknown source case that gave rise to case *j*.

Probability of the temporal data. Because the person containing parasites that are the direct ancestors of those in case *j* is unobserved and does not have an $X_{t,i}$, we are unable to compute the probability of the temporal data as described in Scenario 1. It is important though to obtain a probability comparable to that from Scenario 1 as a reference point for determining whether it is more likely that a given case arose from some other known local case or from an unknown case u_s from source population *s*. To do so, we consider the variable $\tilde{X}_{t,u_s}$, which is a latent variable describing the timing of when u_s would have been detected, had it been detected.

Because u_s is not observed, we considered it to be asymptomatic and untreated. We then calculated the probability of the timing of a known case *j* arising from an unknown case u_s as

$$\Pr\left(X_{t,j} \middle| N_{u_s,j}, \Theta\right) = \int \Pr\left(X_{t,j} \middle| \tilde{X}_{t,u_s}, N_{u_s,j}, \Theta\right) \Pr\left(\mathrm{SI} = X_{t,j} - \tilde{X}_{t,u_s}\right) d\tilde{X}_{t,u_s}, \quad (12)$$

by integrating over uncertainty in $\tilde{X}_{t,u_s}$. We represented this as the product of the probability of the timing of a known case *j* conditional on an unknown time of detection $\tilde{X}_{t,u_s}$ and the probability of the serial interval $X_{t,j} - \tilde{X}_{t,u_s}$ for all $\tilde{X}_{t,u_s}$. In equation (12), we did not distinguish between symptomatic and asymptomatic cases *j* because the calculation is identical; only the serial interval distributions differ.

Probability of the spatial data. Without an $\tilde{X}_{t,u_s}$ for the unobserved case u_s, we lacked information on the serial interval between it and case j. Consequently, we were unable to use the probability from eq. (7) in that particular form. Instead, we computed the spatial variance as a function of the diffusion coefficient alone, yielding

$$\sigma^2(D) = \int D\ GI\ \Pr(GI)\ dGI. \quad (13)$$

Equation (13) integrates across all possible generation intervals and simplifies to $D\mathbb{E}[GI]$, the product of the diffusion coefficient and the expectation of the generation interval.

We applied this spatial variance to the unobserved latent variable $\tilde{X}_{s,u_s}$, which represents the unknown location of the unobserved case u_s. We integrated over uncertainty in $\tilde{X}_{s,u_s}$ to compute the probability density,

$$f\left(X_{s,j} \middle| D, N_{u_s,j}, \Theta\right) = \int f\left(X_{s,j} \middle| \tilde{X}_{s,u_s}, D, N_{u_s,j}, \Theta\right) f\left(\tilde{X}_{s,u_s} \middle| X_{s,j}, D\right) d\,\tilde{X}_{s,u_s}, \quad (14)$$

of the location of a known case j arising from an unknown source case u_s with unknown location $\tilde{X}_{s,u_s}$. This is represented as the product of the probability density of the location of a known case j conditional on an unknown location $\tilde{X}_{s,u_s}$ and the probability density of spatial separation $X_{s,j} - \tilde{X}_{s,u_s}$ conditional on the diffusion coefficient D for all $\tilde{X}_{s,u_s}$. As in eq. (9), we treated eq. (14) as an evaluation of the Gaussian integral, evaluating to

$$f\left(X_{s,j}\middle|D, N_{u_s,j}, \Theta\right) = \frac{1}{4\pi D\mathbb{E}[GI]}. \quad (15)$$

In eq. (15), D is the diffusion coefficient and $\mathbb{E}[GI]$ is the expectation of the generation interval.

Probability of the travel-history data. We considered the travel history $X_{h,j}$ to be a binary variable with a value of 1 indicating reported travel to an area with known or assumed malaria transmission within a timeframe consistent with the person having become infected there. After defining the probability τ_s that $X_{h,j} = 1$ conditional on $N_{u_s,j}$, it follows that

$$\Pr\left(X_{h,j}\middle|N_{u_s,j}, \Theta\right) = \begin{cases} \tau_s, & X_{h,j} = 1 \\ 1 - \tau_s, & X_{h,j} = 0 \end{cases}, \quad (16)$$

which constitutes the contribution of the travel history of such a case to the overall likelihood of N and Θ. If the travel history of case j is unknown, an analogous calculation to eq. (11) is made using τ_s.

Bayesian inference

Markov Chain Monte Carlo Algorithm. To avoid evaluating the high-dimensional integral over N and Θ, we drew samples of N and Θ from their posterior distribution defined by eq. (1) using a Metropolis-Hastings Markov chain Monte Carlo (MCMC) method [40,41]. To begin the chain, N and Θ were initialized to $N^{(1)}$ and $\Theta^{(1)}$, and each subsequent step i in the chain was

denoted $N^{(i)}$ and $\Theta^{(i)}$. At each step, states N' and Θ' were proposed with $\Pr\left(\left(N^{(i)}, \Theta^{(i)}\right) \rightarrow (N', \Theta')\right)$. Proposed states were accepted with probability

$$\alpha_{\text{update}} = \min\left[1, \frac{\pi(N', \Theta') \Pr\left((N', \Theta') \rightarrow \left(N^{(i)}, \Theta^{(i)}\right)\right)}{\pi(N^{(i)}, \Theta^{(i)}) \Pr\left(\left(N^{(i)}, \Theta^{(i)}\right) \rightarrow (N', \Theta')\right)}\right], \quad (17)$$

where $\pi(N, \Theta)$ is the product of the likelihood $\Pr\,(\vec{X}_s, X_t, X_h | N, \Theta)$ of N and Θ conditional on the data and the assumed prior probability $\Pr\,(N, \Theta)$ of N and Θ. After a random draw R from a uniform distribution, the chain was updated according to

$$N^{(i+1)}, \Theta^{(i+1)} = \begin{cases} N', \Theta', & R \leq \alpha \\ N^{(i)}, \Theta^{(i)}, & R > \alpha \end{cases}. \quad (18)$$

To reduce the probability of the chain becoming stuck at a local maximum, we employed Metropolis-coupled Markov chain Monte Carlo (MC^3) [42]. Implementing MC^3 involved running multiple chains in parallel, with $\pi_c(N, \Theta)$ in chain c raised to the power β_c according to

$$\beta_c = 1 + \lambda(c - 1), \quad (19)$$

where $\lambda > 0$ is a temperature increment parameter that governs the degree to which each chain is "heated." As a result of setting $\beta_1 = 1$, $\pi_1(N, \Theta)$ is directly proportional to the joint posterior distribution and is referred to as the master or "cold" chain. This algorithm effectively flattens the likelihood in the heated chains by setting $\beta_c > 1$, allowing them to explore the parameter

space more freely and to encounter alternative high-density regions more readily than the cold chain would alone. At a pre-defined frequency, two randomly selected chains *i* and *j* were allowed to swap parameter sets according to a swap probability

$$\alpha_{\text{swap}} = \min\left[1, \frac{\pi\left(N^{(j)}, \Theta^{(j)}\right)^{\beta_i} \pi\left(N^{(i)}, \Theta^{(i)}\right)^{\beta_j}}{\pi(N^{(i)}, \Theta^{(i)})^{\beta_i} \pi(N^{(j)}, \Theta^{(j)})^{\beta_j}}\right], \quad (20)$$

where $\pi(N, \Theta)$ is the same as it was in eq. (17). A swap into the master chain only occurred if it was from one of the two randomly selected chains and $R \leq \alpha_{\text{swap}}$. We recorded a total of 100 million samples from the posterior distribution, discarding the first 50 million samples as burn-in and thinning the chain every 10,000 samples between each recorded sample.

Proposals. Proposals made by the MC^3 algorithm involved changes to the parameters (i.e., D, τ_s, and τ_l) and changes to the transmission network topology. Each proposal occurred with a fixed probability, where the sum of these proposal probabilities was equal to one.

Proposals to change parameters involved updating D, τ_s, or τ_l. To update the value of D, a new value was drawn from a normal distribution with mean set to the current value of the parameter and variance set to 2.5. Values of D proposed must be strictly nonnegative, so we rejected any proposed D that was less than zero and assigned $\alpha_{\text{update}} = 0$. Similarly, new values of τ_s and τ_l were chosen according to normal distributions with means set to their current parameter value and variance set to 0.25. Because τ_s and τ_l are probabilities, we rejected any proposed value that fell outside the range [0,1] and assigned $\alpha_{\text{update}} = 0$.

Changes proposed to the network topology involved the addition or removal of an ancestor from a randomly selected node. We assigned a uniform probability of proposing case *a* as an ancestor to a randomly selected case *i*, such that proposals to the network topology are uninformed by spatial and temporal data. Furthermore, we defined the proposal probability of removing case *a* as an ancestor to a randomly selected case *i* as

$$\Pr(\text{remove } a) = \bar{A}_i^{-1}, \quad (21)$$

where $\bar{A}_i$ represents the size of set A_i of all ancestors to case *i*. Proposed changes to the network are then accepted according to eq. (17).

Prior assumptions. We placed strong priors on τ_s and τ_l, because we assumed that travel histories were mostly, but not completely, accurate. We used a beta-distributed prior on τ_s, with parameters $\alpha_{\tau_s} = 12$ and $\beta_{\tau_s} = 3$, which resulted in a mean of 0.8 and a variance of 0.01 for this prior distribution. We also used a beta distributed prior on τ_l, with parameters $\alpha_{\tau_l} = 3$ and $\beta_{\tau_l} = 12$, which resulted in a mean of 0.2 and a variance of 0.01. We assumed a uniform prior on *D* over the interval $[10^{-3}, \infty)$ and an even prior across all possible network configurations, meaning that those prior probabilities canceled out in eqs. (17) and (20).

Assessing convergence. For *D*, τ_s, and τ_l, we assessed convergence using the Gelman-Rubin statistic [43], with values below 1.1 indicating convergence. For the transmission network *N*, we assessed convergence by calculating correlation coefficients of case-level probabilities across five chains from independent realizations of the MC^3 algorithm, for a total of 10 pairwise

comparisons across the five chains. The two case-level probabilities that we considered were the posterior probability that each case was infected by an unknown case u_s from a source population and the posterior probability that each case j was infected by each other case i. Higher values of these correlation coefficients provided stronger support for convergence.

Ethical Considerations

Ethical approval was obtained from the Eswatini Ministry of Health, the University of California, San Francisco, and the University of Notre Dame (IRB 19-06-5408). All data were analyzed anonymously.

Acknowledgments

The authors thank the National Malaria Elimination Programme as well as Nontokozo Mngadi and Deepa Pindolia from the Clinton Health Access Initiative for their support in the collection of the surveillance data used in this study. We also thank Brooke Whittemore for her support in data management.

References

1. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol. 2004;160: 509–516. doi:10.1093/aje/kwh255

2. White LF, Archer B, Pagano M. Estimating the reproductive number in the presence of spatial heterogeneity of transmission patterns. Int J Health Geogr. 2013;12: 35. doi:10.1186/1476-072X-12-35

734 3. Métras R, Baguelin M, Edmunds WJ, Thompson PN, Kemp A, Pfeiffer DU, et al.
735 Transmission Potential of Rift Valley Fever Virus over the Course of the 2010 Epidemic in
736 South Africa. Emerg Infect Dis. 2013;19: 916–924. doi:10.3201/eid1906.121641

737 4. Backer JA, Wallinga J. Spatiotemporal Analysis of the 2014 Ebola Epidemic in West
738 Africa. Lessler J, editor. PLoS Comput Biol. 2016;12: e1005210.
739 doi:10.1371/journal.pcbi.1005210

740 5. Salje H, Cummings DAT, Lessler J. Estimating infectious disease transmission distances
741 using the overall distribution of cases. Epidemics. 2016;17: 10–18.
742 doi:10.1016/j.epidem.2016.10.001

743 6. Reiner RC, Le Menach A, Kunene S, Ntshalintshali N, Hsiang MS, Perkins TA, et al.
744 Mapping residual transmission for malaria elimination. eLife. 2015;4.
745 doi:10.7554/eLife.09520

746 7. Lau MSY, Dalziel BD, Funk S, McClelland A, Tiffany A, Riley S, et al. Spatial and
747 temporal dynamics of superspreading events in the 2014-2015 West Africa Ebola epidemic.
748 Proc Natl Acad Sci USA. 2017;114: 2337–2342. doi:10.1073/pnas.1614595114

749 8. Lau MSY, Gibson GJ, Adrakey H, McClelland A, Riley S, Zelner J, et al. A mechanistic
750 spatio-temporal framework for modelling individual-to-individual transmission—With an
751 application to the 2014-2015 West Africa Ebola outbreak. Ferrari M (Matt), editor. PLoS
752 Comput Biol. 2017;13: e1005798. doi:10.1371/journal.pcbi.1005798

753 9. Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N. Bayesian
754 Reconstruction of Disease Outbreaks by Combining Epidemiologic and Genomic Data.

755 Tanaka MM, editor. PLoS Comput Biol. 2014;10: e1003457.
756 doi:10.1371/journal.pcbi.1003457

757 10. Didelot X, Gardy J, Colijn C. Bayesian Inference of Infectious Disease Transmission from
758 Whole-Genome Sequence Data. Molecular Biology and Evolution. 2014;31: 1869–1879.
759 doi:10.1093/molbev/msu121

760 11. Mollentze N, Nel LH, Townsend S, le Roux K, Hampson K, Haydon DT, et al. A Bayesian
761 approach for inferring the dynamics of partially observed endemic infectious diseases from
762 space-time-genetic data. Proc Biol Sci. 2014;281: 20133251. doi:10.1098/rspb.2013.3251

763 12. Ypma RJF, van Ballegooijen WM, Wallinga J. Relating Phylogenetic Trees to
764 Transmission Trees of Infectious Disease Outbreaks. Genetics. 2013;195: 1055–1062.
765 doi:10.1534/genetics.113.154856

766 13. Sturrock HJW, Hsiang MS, Cohen JM, Smith DL, Greenhouse B, Bousema T, et al.
767 Targeting Asymptomatic Malaria Infections: Active Surveillance in Control and
768 Elimination. PLoS Med. 2013;10: e1001467. doi:10.1371/journal.pmed.1001467

769 14. Bousema T, Griffin JT, Sauerwein RW, Smith DL, Churcher TS, Takken W, et al. Hitting
770 Hotspots: Spatial Targeting of Malaria for Control and Elimination. PLoS Med. 2012;9:
771 e1001165. doi:10.1371/journal.pmed.1001165

772 15. Bejon P, Williams TN, Nyundo C, Hay SI, Benz D, Gething PW, et al. A micro-
773 epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within
774 hotspots. eLife. 2014;3: e02130. doi:10.7554/eLife.02130

775 16. Cohen JM, Moonen B, Snow RW, Smith DL. How absolute is zero? An evaluation of
776 historical and current definitions of malaria elimination. Malar J. 2010;9: 213.
777 doi:10.1186/1475-2875-9-213

778 17. Cohen JM, Le Menach A, Pothin E, Eisele TP, Gething PW, Eckhoff PA, et al. Mapping
779 multiple components of malaria risk for improved targeting of elimination interventions.
780 Malar J. 2017;16: 459. doi:10.1186/s12936-017-2106-3

781 18. Churcher TS, Cohen JM, Novotny J, Ntshalintshali N, Kunene S, Cauchemez S. Measuring
782 the path toward malaria elimination. Science. 2014;344: 1230–1232.
783 doi:10.1126/science.1251449

784 19. Routledge I, Chevéz JER, Cucunubá ZM, Rodriguez MG, Guinovart C, Gustafson KB, et
785 al. Estimating spatiotemporally varying malaria reproduction numbers in a near elimination
786 setting. Nat Commun. 2018;9: 2476. doi:10.1038/s41467-018-04577-y

787 20. Routledge I, Lai S, Battle KE, Ghani AC, Gomez-Rodriguez M, Gustafson KB, et al.
788 Tracking progress towards malaria elimination in China: Individual-level estimates of
789 transmission and its spatiotemporal variation using a diffusion network approach. Pascual
790 M, editor. PLoS Comput Biol. 2020;16: e1007707. doi:10.1371/journal.pcbi.1007707

791 21. Marshall JM, Bennett A, Kiware SS, Sturrock HJW. The Hitchhiking Parasite: Why Human
792 Movement Matters to Malaria Transmission and What We Can Do About It. Trends in
793 Parasitology. 2016;32: 752–755. doi:10.1016/j.pt.2016.07.004

794 22. Campbell F, Cori A, Ferguson N, Jombart T. Bayesian inference of transmission chains
795 using timing of symptoms, pathogen genomes and contact data. Pitzer VE, editor. PLoS
796 Comput Biol. 2019;15: e1006930. doi:10.1371/journal.pcbi.1006930

797 23. Wood SN. Generalized additive models: an introduction with R. Boca Raton, FL: Chapman
798 & Hall/CRC; 2006.

799 24. R Development Core Team. R: A Language and Environment for Statistical Computing.
800 Vienna, Austria: R Foundation for Statistical Computing; 2017.

801 25. Huber JH, Johnston GL, Greenhouse B, Smith DL, Perkins TA. Quantitative, model-based
802 estimates of variability in the generation and serial intervals of Plasmodium falciparum
803 malaria. Malaria Journal. 2016;15. doi:10.1186/s12936-016-1537-6

804 26. Cohen JM, Dlamini S, Novotny JM, Kandula D, Kunene S, Tatem AJ. Rapid case-based
805 mapping of seasonal malaria transmission risk for strategic elimination planning in
806 Swaziland. Malar J. 2013;12: 61. doi:10.1186/1475-2875-12-61

807 27. Hsiang MS, Ntuku H, Roberts KW, Dufour M-SK, Whittemore B, Tambo M, et al.
808 Effectiveness of reactive focal mass drug administration and reactive focal vector control to
809 reduce malaria transmission in the low malaria-endemic setting of Namibia: a cluster-
810 randomised controlled, open-label, two-by-two factorial design trial. The Lancet. 2020;395:
811 1361–1373. doi:10.1016/S0140-6736(20)30470-0

812 28. Sturrock HJ, Cohen JM, Keil P, Tatem AJ, Le Menach A, Ntshalintshali NE, et al. Fine-
813 scale malaria risk mapping from routine aggregated case data. Malar J. 2014;13: 421.
814 doi:10.1186/1475-2875-13-421

815 29. Wesolowski A, Stresman G, Eagle N, Stevenson J, Owaga C, Marube E, et al. Quantifying
816 travel behavior for infectious disease research: a comparison of data from surveys and
817 mobile phones. Sci Rep. 2015;4: 5678. doi:10.1038/srep05678

818 30. Marshall JM, Wu SL, Sanchez C. HM, Kiware SS, Ndhlovu M, Ouédraogo AL, et al.
819 Mathematical models of human mobility of relevance to malaria transmission in Africa. Sci
820 Rep. 2018;8: 7713. doi:10.1038/s41598-018-26023-1

821 31. Blumberg S, Lloyd-Smith JO. Inference of R0 and Transmission Heterogeneity from the
822 Size Distribution of Stuttering Chains. Ferguson N, editor. PLoS Comput Biol. 2013;9:
823 e1002993. doi:10.1371/journal.pcbi.1002993

824 32. Blumberg S, Lloyd-Smith JO. Comparing methods for estimating R0 from the size
825 distribution of subcritical transmission chains. Epidemics. 2013;5: 131–145.
826 doi:10.1016/j.epidem.2013.05.002

827 33. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the
828 Impact of Human Mobility on Malaria. Science. 2012;338: 267–270.
829 doi:10.1126/science.1223467

830 34. Weiss DJ, Nelson A, Gibson HS, Temperley W, Peedell S, Lieber A, et al. A global map of
831 travel time to cities to assess inequalities in accessibility in 2015. Nature. 2018;553: 333–
832 336. doi:10.1038/nature25181

833 35. Wesolowski A, Taylor AR, Chang H-H, Verity R, Tessema S, Bailey JA, et al. Mapping
834 malaria by combining parasite genomic and epidemiologic data. BMC Med. 2018;16: 190.
835 doi:10.1186/s12916-018-1181-9

836 36. Guzzetta G, Marques-Toledo CA, Rosà R, Teixeira M, Merler S. Quantifying the spatial
837 spread of dengue in a non-endemic Brazilian metropolis via transmission chain
838 reconstruction. Nature Communications. 2018;9. doi:10.1038/s41467-018-05230-4

839 37. Weiss DJ, Lucas TCD, Nguyen M, Nandi AK, Bisanzio D, Battle KE, et al. Mapping the
840 global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial
841 and temporal modelling study. The Lancet. 2019;394: 322–331. doi:10.1016/S0140-
842 6736(19)31097-9

843 38. Greenhouse B, Smith DL, Rodríguez-Barraquer I, Mueller I, Drakeley CJ. Taking Sharper
844 Pictures of Malaria with CAMERAs: Combined Antibodies to Measure Exposure Recency
845 Assays. The American Journal of Tropical Medicine and Hygiene. 2018;99: 1120–1127.
846 doi:10.4269/ajtmh.18-0303

847 39. Lemons DS, Langevin P. An introduction to stochastic processes in physics: containing
848 "On the theory of Brownian motion" by Paul Langevin, translated by Anthony Gythiel.
849 Baltimore: Johns Hopkins University Press; 2002.

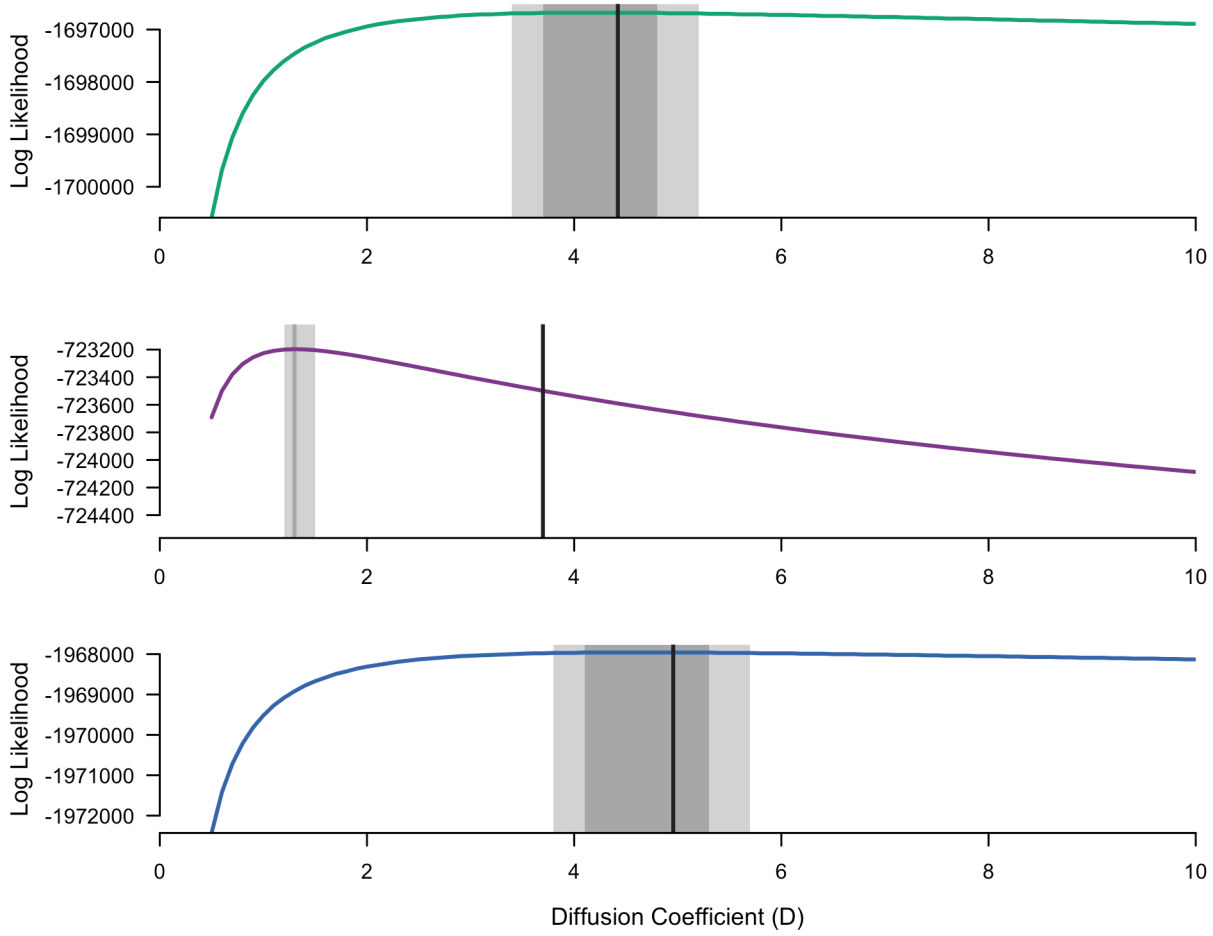
850 40. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State
851 Calculations by Fast Computing Machines. The Journal of Chemical Physics. 1953;21:
852 1087–1092. doi:10.1063/1.1699114

853 41. Hastings WK. Monte Carlo Sampling Methods Using Markov Chains and Their
854 Applications. Biometrika. 1970;57: 97. doi:10.2307/2334940

42. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics. 2004;20: 407–415. doi:10.1093/bioinformatics/btg427

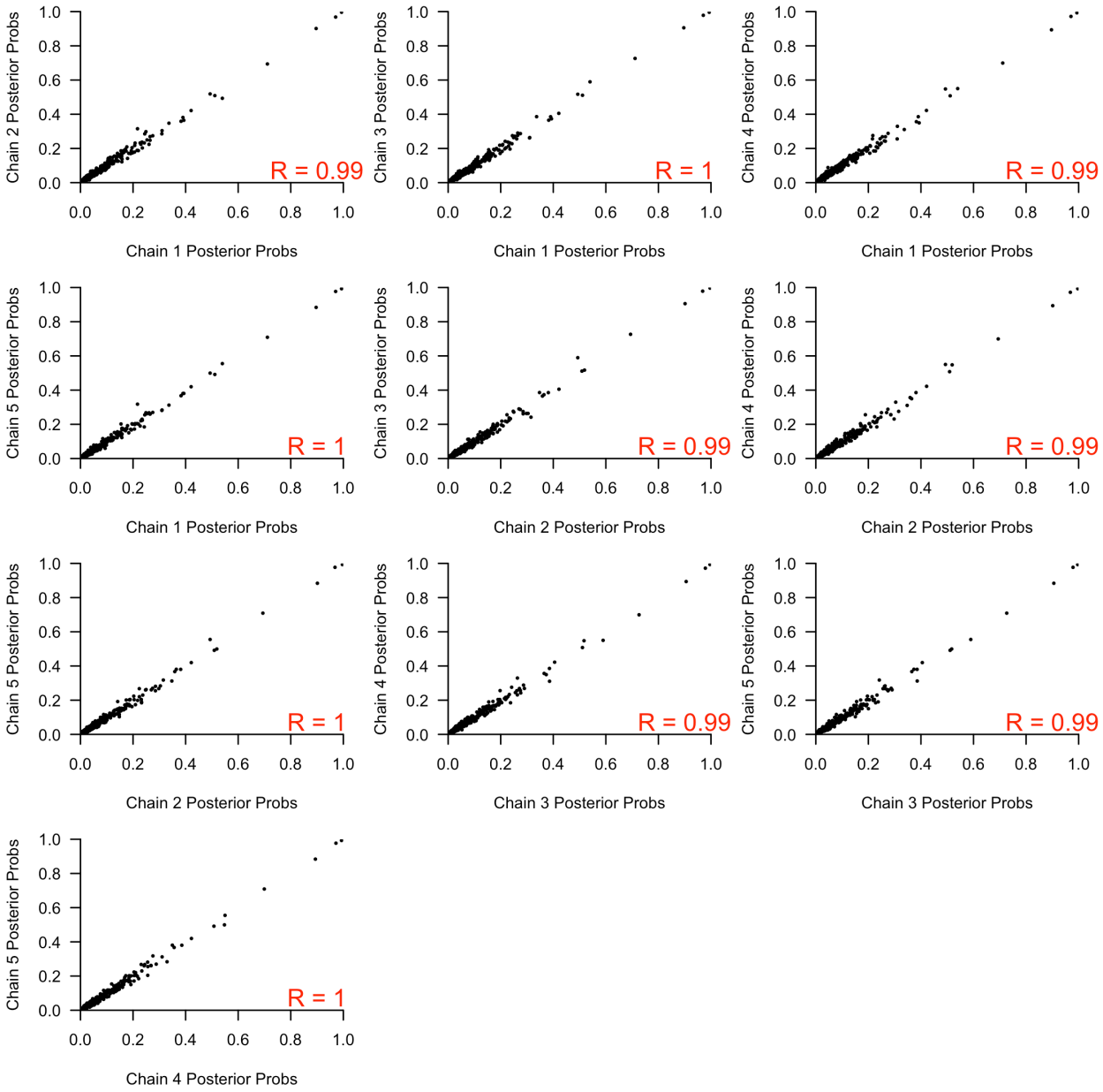
43. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992;7: 457–472. doi:10.1214/ss/1177011136

Supporting Information

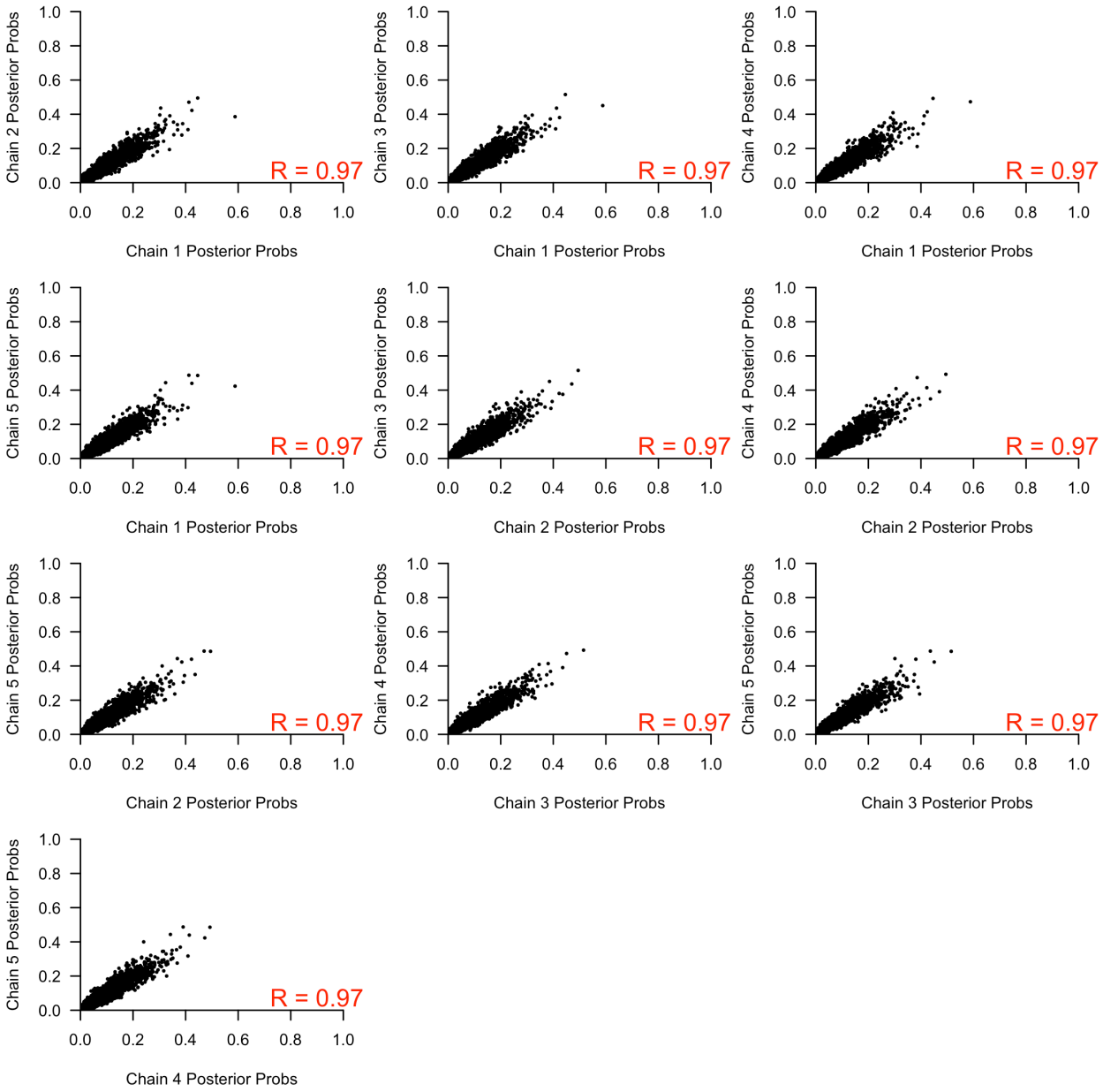


S1 Fig. Likelihood profile of the diffusion coefficient conditioning on the true network. The likelihood profile for two inference settings that incorporate spatial data are shown as a function

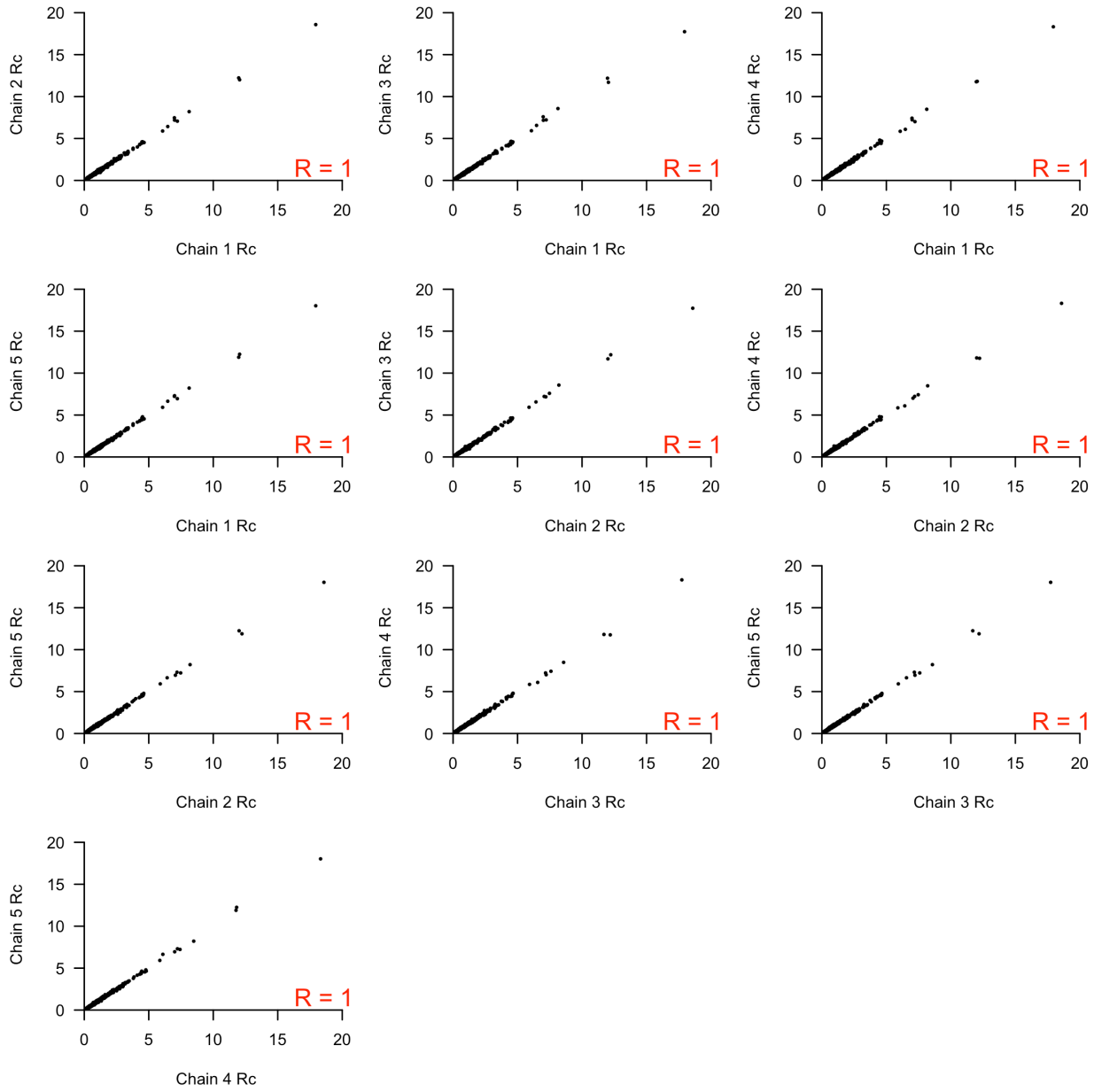
of the diffusion coefficient. The green line corresponds to the default inference setting, the purple line corresponds to the setting in which spatial and temporal data are used and the travel history is believed, and the blue line corresponds to the setting in which spatial and temporal data are used and the travel history was ignored. Black bars denote the true value of the diffusion coefficient, dark grey shapes denote the region with the maximum likelihood, and light grey shapes denote the region within 10 log likelihood units of the maximum likelihood.



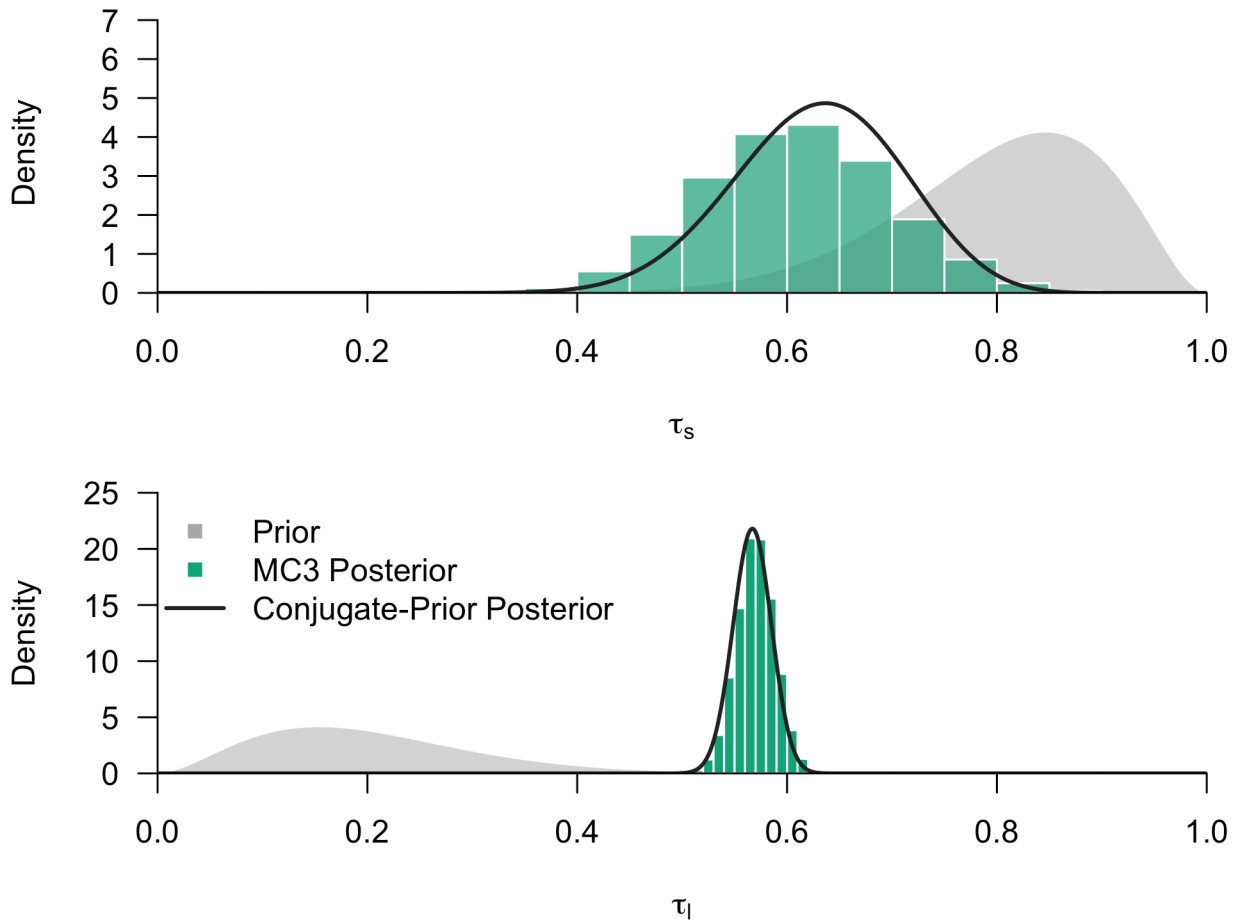
S2 Fig. Comparison of individual-level importation probabilities on Eswatini data under default settings. Pairwise scatter plots of importation probabilities are shown for each pairing of the 5 independent replicates to assess convergence. The correlation between importation probabilities for each pair of replicates is reported in red.



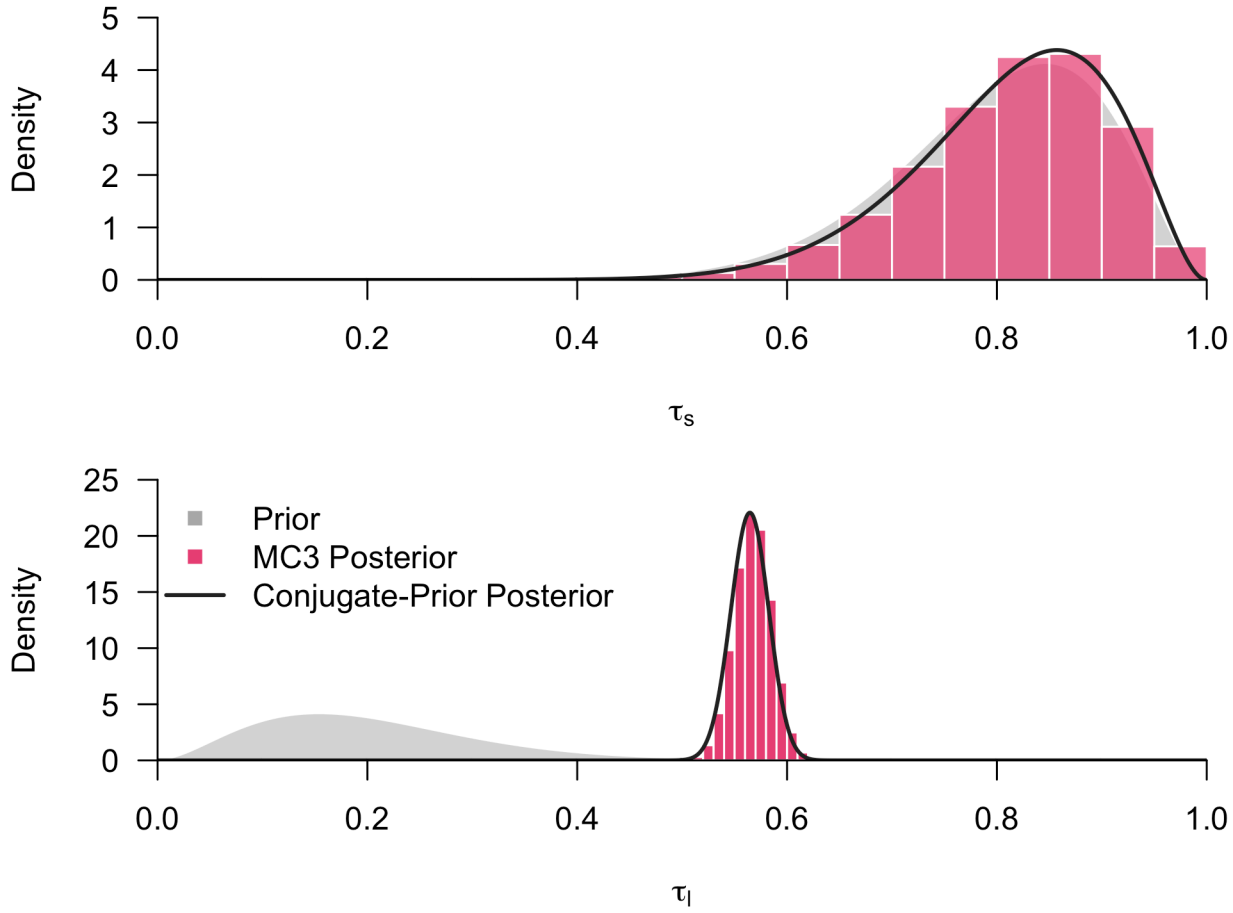
S3 Fig. Comparison of transmission linkage probabilities on Eswatini data under default settings. Pairwise scatter plots of the probability of each transmission linkage are shown for each pairing of the 5 independent replicates to assess convergence. The correlation between transmission linkage probabilities for each pair of replicates is reported in red.



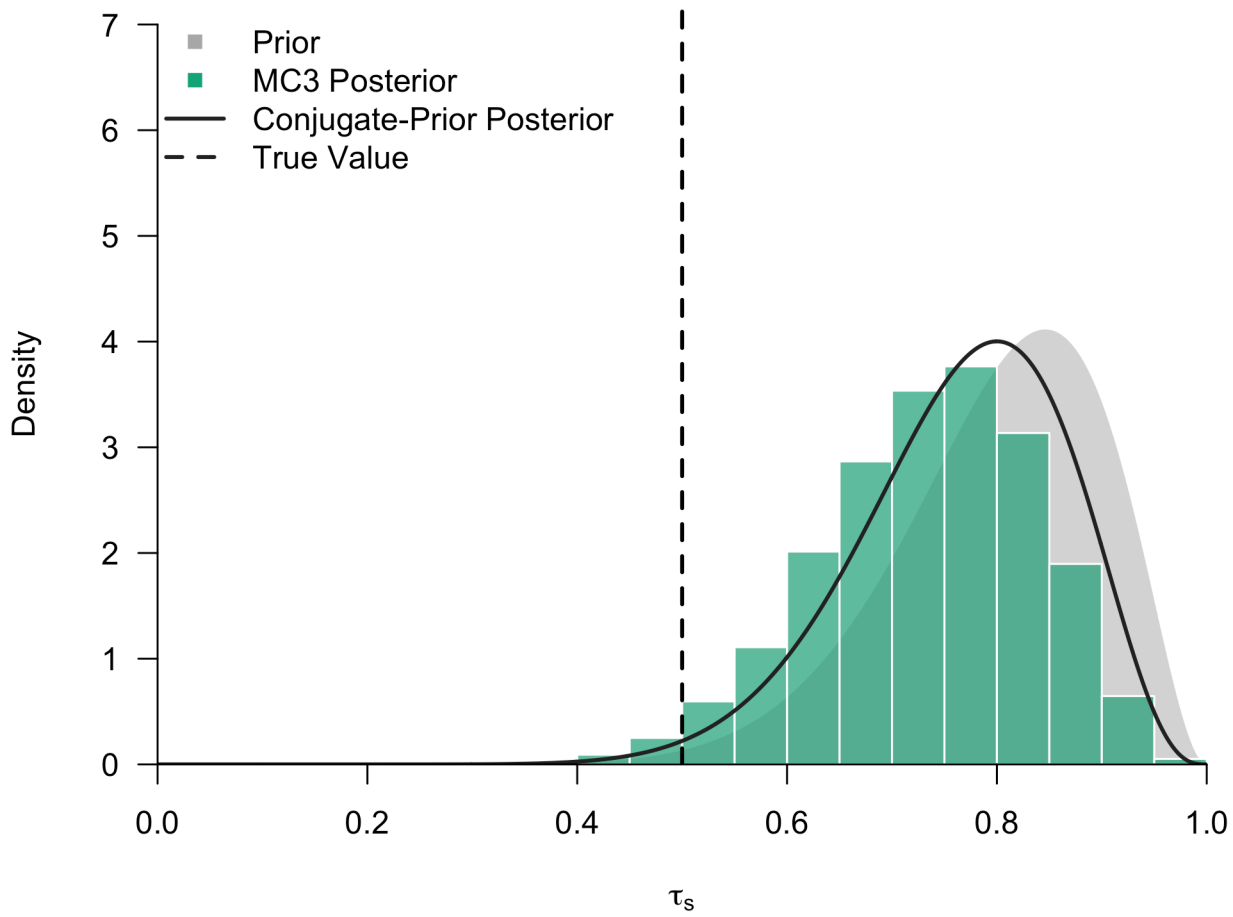
S4 Fig. Comparison of individual-level R_c estimates on Eswatini data under default settings. Pairwise scatter plots of individual-level R_c are show for each pairing of the five independent replicates to assess convergence. The correlation between individual-level R_c values is reported in red.



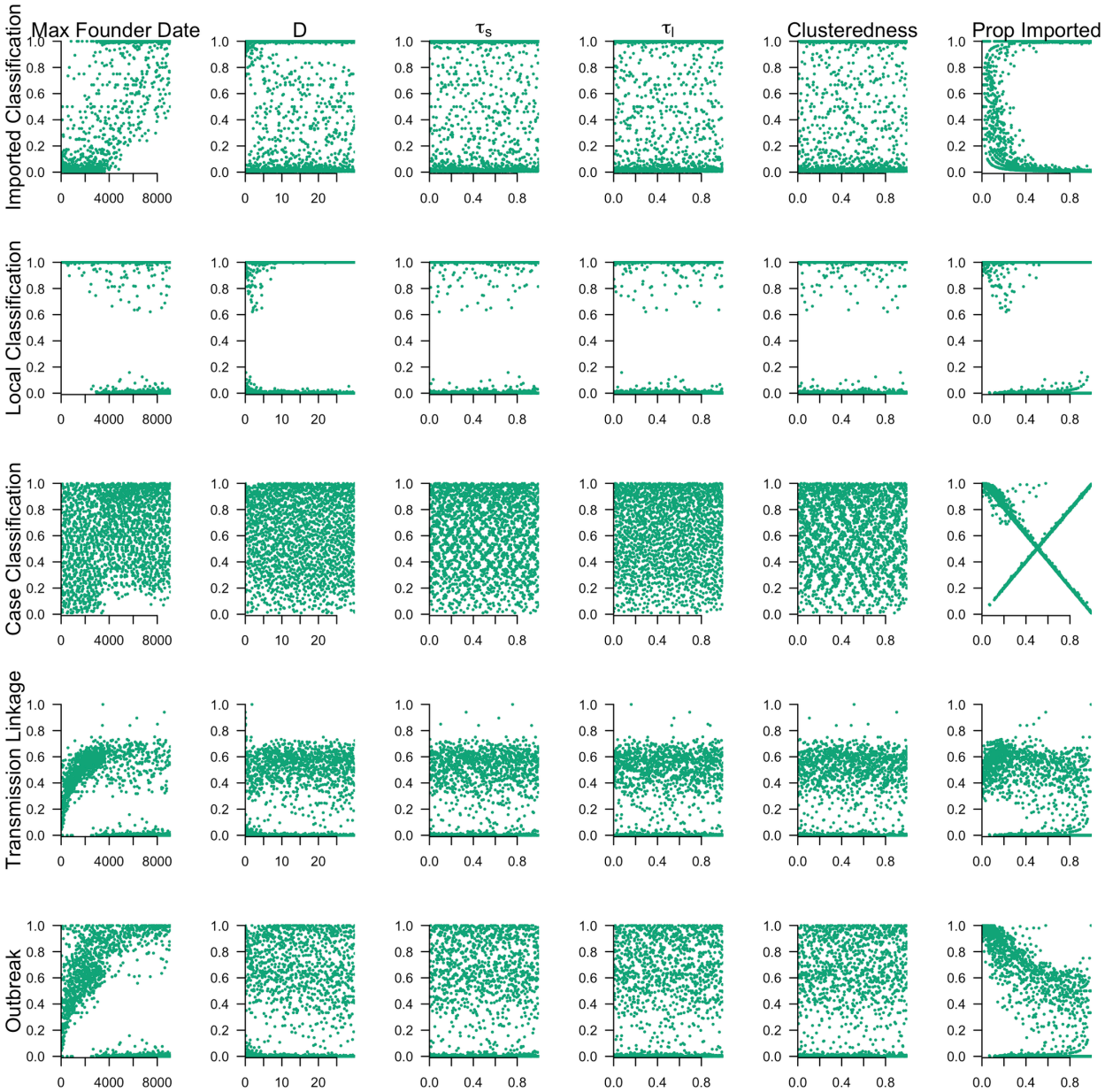
S5 Fig. Comparison of the prior and posteriors of τ_s and τ_l from the Eswatini surveillance data using spatial and temporal data and estimating the accuracy of the travel history. The prior (gray shape), the analytical posterior distribution (black line), and the numerical posterior distribution from MC3 (green histogram) are plotted for τ_s and τ_l.



S6 Fig. Comparison of the prior and posteriors of τ_s and τ_l from the Eswatini surveillance data using temporal data and estimating the accuracy of the travel history. The prior (gray shape), the analytical posterior distribution (black line), and the numerical posterior distribution from MC3 (pink histogram) are plotted for τ_s and τ_l.

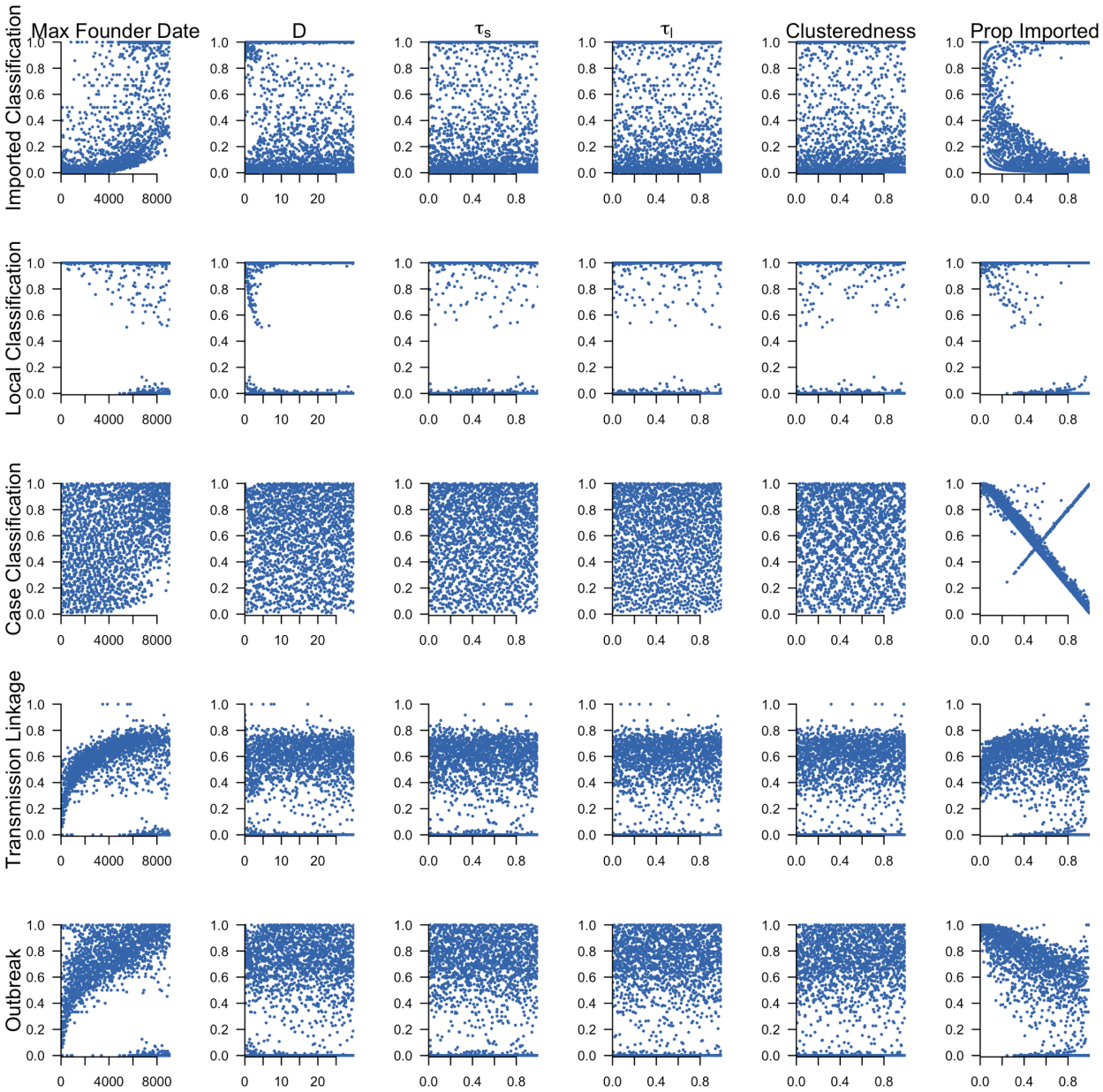


S7 Fig. Comparison of the prior and posteriors of τ_s from simulated data using spatial and temporal data and estimating the accuracy of the travel history. The prior (gray shape), the analytical posterior distribution (black line), and the numerical posterior distribution from MC3 (green histogram) are plotted for τ_s.

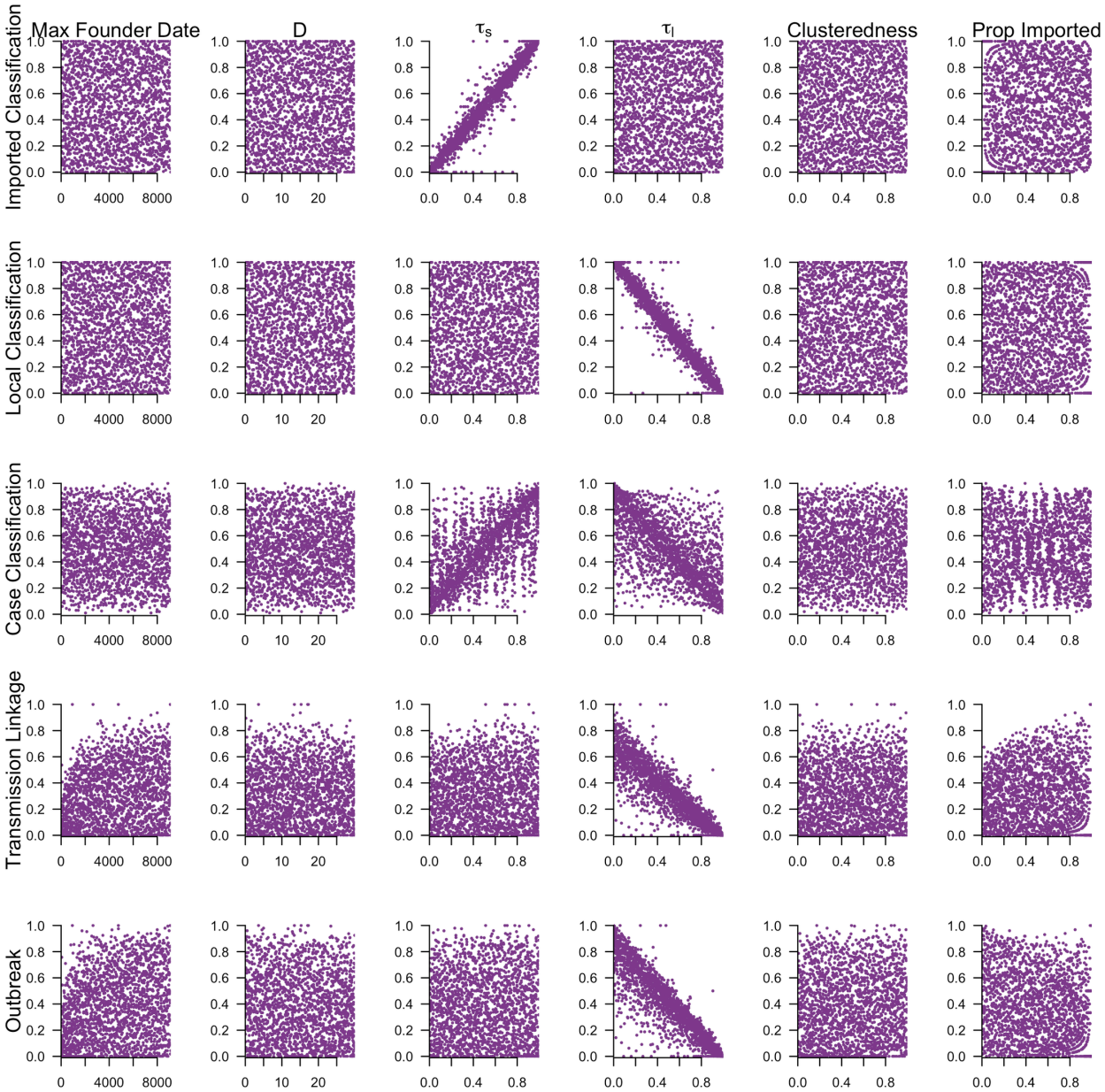


S8 Fig. Univariate relationships between accuracy metrics and simulation parameters using spatial and temporal data and estimating the accuracy of the travel history. Scatterplots of the relationship between five accuracy metrics and the simulation parameters are reported. Imported classification refers to the proportion of imported cases that are correct classified as imported, local classification refers to the proportion of locally acquired cases that are correctly classified as locally acquired, and case classification refers to the proportion of all

cases that are correctly classified as imported or locally acquired. Transmission linkage is the proportion of locally acquired cases for which the true parent is correctly identified, and outbreak is the proportion of locally acquired cases for which the inferred parent belongs to the same outbreak.



S9 Fig. Univariate relationships between accuracy metrics and simulation parameters using spatial and temporal data only. Scatterplots of the relationship between five accuracy metrics and the simulation parameters are reported. Imported classification refers to the proportion of imported cases that are correct classified as imported, local classification refers to the proportion of locally acquired cases that are correctly classified as locally acquired, and case classification refers to the proportion of all cases that are correctly classified as imported or locally acquired. Transmission linkage is the proportion of locally acquired cases for which the true parent is correctly identified, and outbreak is the proportion of locally acquired cases for which the inferred parent belongs to the same outbreak.



S10 Fig. Univariate relationships between accuracy metrics and simulation parameters using spatial and temporal data and believing the travel history. Scatterplots of the relationship between five accuracy metrics and the simulation parameters are reported. Imported classification refers to the proportion of imported cases that are correct classified as imported, local classification refers to the proportion of locally acquired cases that are correctly classified as locally acquired, and case classification refers to the proportion of all cases that are correctly

classified as imported or locally acquired. Transmission linkage is the proportion of locally acquired cases for which the true parent is correctly identified, and outbreak is the proportion of locally acquired cases for which the inferred parent belongs to the same outbreak.

S1 File. Full description of the analyses. This file contains a full description of the analyses run on both the Eswatini data sets and the simulated data sets.

S2 File. Analytical solution for the posterior distribution of τ_s and τ_l. The analytical solutions for τ_s and τ_l are derived using conjugate-prior relationships and compared to the empirical posterior distributions obtained using MC^3.