

Social network and household exposure explain the use of malaria prevention measures in rural communities of Meghalaya, India

Elisa Bellotti¹, Andras Voros², Mattimi Passah⁴, Quinnie Doreen Nongrum⁴, Carinthia Balabet Nengnong⁴, Charishma Khongwir⁴, Annemieke van Eijk³, Anne Kessler³, Rajiv Sarkar⁴, Jane M. Carlton³, Sandra Albert⁴

¹Department of Sociology, University of Manchester, Manchester, UK

² School of Social Policy, University of Birmingham, Birmingham, UK

³Center for Genomics and Systems Biology, Department of Biology, New York University, USA

⁴Indian Institute of Public Health Shillong, Shillong, Meghalaya, India

Correspondence: Elisa Bellotti e-mail: Elisa.Bellotti@manchester.ac.uk; Sandra Albert e-mail: sandra.albert@iiphs.ac.in

Malaria remains a global concern despite substantial reduction in incidence over the past twenty years. Public health interventions to increase the uptake of preventive measures have contributed to this decline but their impact has not been uniform. To date, we know little about what determines the use of preventive measures in rural, hard-to-reach populations, which are crucial contexts for malaria eradication. We collected detailed interview data on the use of malaria preventive measures, health-related discussion networks, individual characteristics, and household composition in ten tribal, malaria-endemic villages in Meghalaya, India in 2020-2021 (n=1,530). Employing standard and network statistical models, we found that social network and household exposure were consistently positively associated with preventive measure use across villages. Network and household exposure were also the most important factors explaining behaviour, outweighing individual characteristics, opinion leaders, and network size. These results suggest that real-life data on social networks and household composition should be considered in studies of health-behaviour change.

INTRODUCTION

Following a substantial reduction in the global incidence of malaria between 2001 and 2015, the World Health Organization set a goal of global eradication by 2030 [WHO, 2015]. While large-scale interventions played a pivotal role in controlling and eliminating malaria in some countries, their impact has varied across the malarious world, resulting in little change for poor, migrant, and hard-to-reach populations [WHO, 2015]. Prevention in residual epicentres is challenging unless local populations are engaged [Dhiman, 2019]. Treating difficult pockets of transmission requires tailored and targeted approaches suited to local contexts [Gosling et al., 2020]. This is true for Meghalaya state in Northeastern India, a hilly and mountainous area of subtropical rain forests interspersed with rice-agroecosystems. The wet climate makes Meghalaya an optimal setting for malaria, with multiple *Anopheles* mosquito species transmitting *Plasmodium falciparum* and *Plasmodium vivax* malaria parasites [Kessler et al., 2018; Kessler et al., 2021]. Meghalaya is inhabited mostly by tribal peoples, with 86% of its population of approximately 3.2 million belonging to an indigenous ethnicity or Scheduled

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Tribes as referred to by the government of India [Rajagopal, 1976; Lele et al., 2009; Ministry of Tribal Affairs, 2021]. ‘Tribals’ often experience geographical marginalization, poor access to health care, and low socio-economic status [Rajagopal, 1976; Sharma et al. 2015; Kessler et al., 2018], environmental and social factors that pose a challenge to malaria prevention efforts. In Meghalaya, large-scale malaria prevention efforts are coordinated by modern, government-run public health programmes managed through local Primary Health Centres (PHC) and facilitated at the village level by Accredited Social Health Activists (ASHAs). These operate alongside largely undocumented traditional tribal treatments provided by Traditional Healers [Albert et al., 2015].

Large-scale malaria prevention measures in Meghalaya have mainly focused on reducing mosquito bites [Debboun et al., 2013] through the use of Indoor Residual Spraying (IRS) and the distribution of Long-Lasting Insecticidal Nets (LLINs), which repel and reduce the longevity of indoor-feeding and resting of *Anopheles* vector mosquitoes [Mishra et al., 2021]. Both methods have been associated with a substantial reduction in malaria incidence [Sahu et al., 2020] but are used only indoors [Lek et al., 2020; Sangoro, 2015; Rubio-Palis et al., 1992]. Villagers working or socializing outside their dwellings, especially in the evening, night, and early morning, may use other measures to prevent mosquito bites [Sangoro, 2015; Wilson et al., 2014], such as protective clothing, insect repellent spray or body cream. IRS is contested in many Meghalayan villages [Sarkar et al., 2021; Passah et al., 2022], where households could turn to other indoor measures such as air-borne repellents (coils and vaporizers) instead.

Studies that measure the adoption of – or resistance to – mosquito bite prevention methods usually evaluate how characteristics of individuals or households (gender, education, socio-economic background, number of children, living conditions, etc.) relate to knowledge and practice of malaria preventive and treatment behaviours [Nlinwe et al., 2021; Matin et al., 2020; Van Eijk et al., 2016; Shahandeh et al., 2012]. What these studies do not consider is that information and attitudes regarding both effective and ineffective preventive behaviours are likely to spread through social networks. These may facilitate support for but also resistance to new practices. Such often-overlooked network processes can help explain why some effective practices, like the use of vaporisers, may become common in certain areas [Van Eijk et al., 2016] while others, like IRS, may be opposed [Sarkar et al., 2021; Passah et al., 2022], or why practices whose effectiveness is unproven, like burning egg trays and jute bags, may thrive.

Social network research is increasingly becoming common in the study of adoption of health behaviours, including eating habits [Salvy et al., 2012], physical activity [Proestakis et al., 2018], family planning [Gayen et al., 2010; Stoebenau et al., 2003], healthy practices [Kim et al., 2015], and general diffusion of information and innovations [Rogers Everett, 1962; Coleman et al., 1957; Valente, 1996; Valente 2010; Centola, 2010; Centola, 2018]. Threshold models postulate that individuals adopt innovative behaviours depending on the proportion of the population already engaged in such behaviours [Granovetter, 1978]. They further propose that individuals comfortable with a low threshold adopt a behaviour earlier than others, while late adopters require a higher threshold. Threshold models are efficient in

predicting simple contagion, for which proximity – being close to infected subjects – is enough to trigger diffusion [Watts et al., 1998; Centola, 2010]. However, behaviours are complex phenomena that do not spread across networks in the same way an airborne disease might [Centola and Macy, 2007]. Adopting a behaviour like burning coils or jute bags is influenced both by locals who use these preventive measures as well as those who do not use them because they consider them ineffective or explicitly oppose them [Centola; 2010; Centola, 2018]. To be diffused, innovations need to be adopted and reinforced within local neighbourhoods of social networks [Coleman et al., 1957; Valente, 1996; Valente 2010; Centola; 2010; Centola, 2018].

A network exposure model of diffusion [Coleman et al., 1957; Valente, 1996; Valente 2010; Centola, 2018] does not measure the threshold level of the entire population but rather, the threshold level of the proportion of adopters within the population that individuals are directly related to. It postulates that innovative behaviour spreads relatively quickly within highly clustered and homophilous regions of the networks [Christakis and Fowler, 2007; Smith et al., 2008; Crosnoe et al., 2008; Valente et al., 2009] but the spreading slows down where such clusters are disconnected [Centola, 2011]. Central actors such as persons with numerous contacts [Freeman 1979] may not be the best located actors to promote diffusion of new practices. They are usually considered opinion leaders and champions of the status quo who are reluctant to change the way things are done within a community [Valente and Pumpang 2007]. Innovations are more likely to spin off from peripheries of social networks, where people are less constrained by collective social norms, are more likely to be connected through non-redundant weak ties, and therefore are potentially more exposed to norms of different groups [Granovetter 1978, Burt 2004].

Most studies that model diffusion of diseases or information use secondary [Weston et al., 2018] or synthetic data [Horsevad et al., 2022] to calibrate or validate models [Verelst et al., 2016; Bedson et al. 2021]. Social network information is at best inferred from physical proximity [Levy et al., 2021]. Studies that collect primary social network data, especially from remote, hard-to-reach tribal villages, are very rare [Gayen et al., 2010; Stoebenau et al., 2003; Kim et al. 2015; Holly et al., 2014; Koster, 2011]. This is so because social network data need to be gathered from the entire population of interest [Smith et al., 2008; Perkins et al., 2015] with minimal missing data, which is a huge challenge in hard-to-reach populations where individuals spend most of the day working outdoors. To guarantee coverage, the population of interest must be identified and bounded [Smith et al., 2008; Marsden, 2002], while acknowledging that boundaries are always arbitrary. In practice, an individual's social network extends beyond the village they live in, even in remote areas [Stoebenau and Valente 2003; Koster 2011].

Once populations have been identified and bounded, and data have been collected, social network analysis can be used to study the links between the structure of networks and the behaviours of its actors. This requires specialised statistical techniques [Steglich et al., 2010], since dependencies in networks and between networks and behaviour violate the independence assumption at the core of most standard (non-network) statistical methods [Bartels et al., 1979; Doreian 1980; Dow et al., 1983; Dow et al., 1984; Ord 1975; Leenders

2002]. The Stochastic Actor-oriented Model (SAOM) is one of the few methods suitable to analyse the dynamics of social networks and individual behaviour [Snijders et al., 2010]. The SAOM is increasingly popular in the social sciences and has been applied in the context of various health behaviours in recent years [adams and Schaefer, 2017; de la Haye et al., 2019; Long and Valente, 2019; Franken et al., 2023]. Despite successful applications of network models, empirical data in the area of disease prevention, and specifically malaria prevention, are still predominantly analysed with non-network methods [Nlinwe et al., 2021; Matin et al., 2020; Van Eijk et al., 2016; see Musoke et al. 2023 for a review].

In this study, we assess the importance of social networks in explaining the use of eight mosquito bite preventive measures in ten hard-to-reach tribal villages in northeast India. We compare the role of five groups of factors affecting villagers' use of the preventive measures: individual characteristics, influence from village opinion leaders, social network size (within and outside villages), network exposure, and household exposure. We first examine which factors explain the use of each preventive measure by fitting village-level logistic regression models. In the results, we present the meta-analysis of all models (Snijders et al. 2012), which shows the significant variables, across villages, in explaining the use of each preventive measure. We then apply stationary Stochastic Actor-oriented Models [Snijders and Steglich, 2015], an extension of SAOMs for cross-sectional data observed at a single time point, to jointly model individuals' use of preventive measures and their social relations with other villagers. In this step, we consider villagers' use of any measure, treating the system of social ties and measure use as a multilevel network (Lazega and Snijders, 2015:36-37). While our data is cross-sectional, the SAOM is dynamic in nature and allows inference to social processes that can explain the observed patterns of social ties, preventive behaviour, individual attributes, and household structure (see the Methods section for further details). To our knowledge, this is the first study where a social network analysis approach is used in the context of mosquito bite prevention. To date, stationary SAOMs have not been fitted to multilevel network data in the scholarly literature, while we know of only two applications of the method to a single network, both outside of the area of health-related behaviours [Snijders and Steglich, 2015; Simpson 2022]. Similar to the logistic regression models, in the results we present the meta-analysis of all SAOM models. Details of the meta-analysis procedure can be found in the Method section, while full logistic and SAOM models results are reported in the Supplementary Information, Tables 12-19 and Appendix E, Tables 41-44.

We apply the above models to data we collected in the state of Meghalaya in India between January 2020 and August 2021. Data is available from ten villages in three districts: three villages in West Khasi Hills (WK1, WK2, WK3), three villages in West Jaintia Hills (WJ1, WJ2, WJ3), and four villages in South Garo Hills (SG1, SG2, SG3, SG4). In these areas, detailed malaria epidemiological and behavioural information has only recently been generated [Kessler et al., 2018; Kessler et al., 2021]. See **Figure 1** for the geographic context of our study sites.

Map of Meghalaya

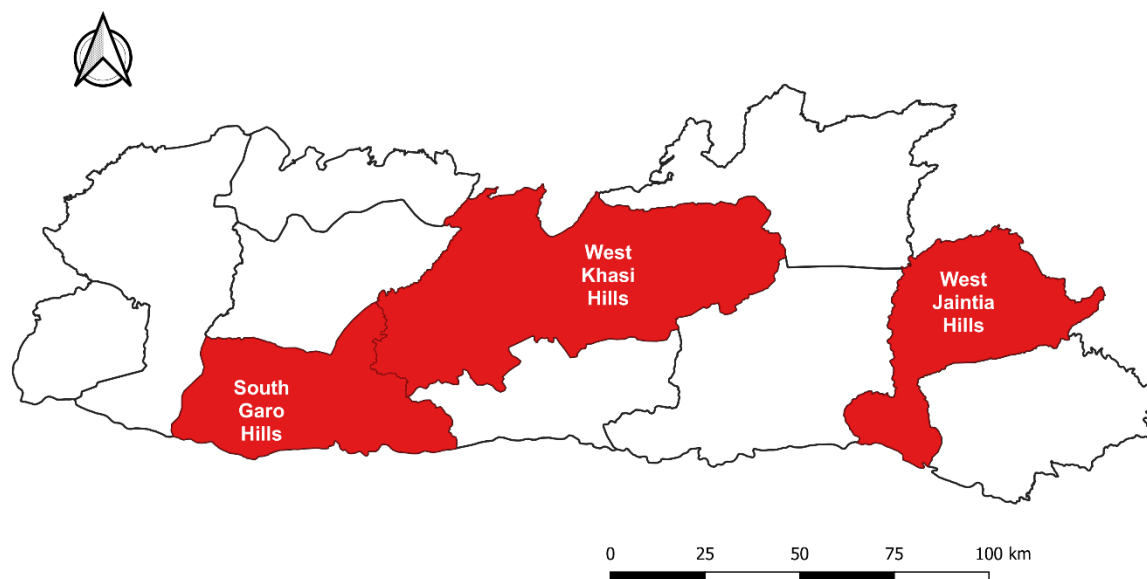


Figure 1: The map of Meghalaya state in Northeast India with the three studied districts highlighted.

A structured questionnaire was administered to every reachable adult (≥ 18 years old) in the studied villages, resulting in information from 1,530 villagers residing in 766 households. From each participant, we collected information about: (1) use of mosquito bite preventive measures (LLINs, covering clothes, boots, gloves, insecticide cream, coils, vaporizers, burning materials), (2) demographic characteristics and role(s) within household and village, (3) health-related discussion network within and outside village, and (4) household membership. The ASHA was interviewed in all 10 villages, and in the six villages where present, the Traditional Healer was also interviewed. Both the ASHA and the Traditional Healer are considered opinion leaders for their health-related expertise and their role in administering healthcare within villages. The full list of questions, details on coding and variable descriptives are reported in the Supplementary Information, Appendix B, Tables B1-B15.

By interviewing most adults in each village, we can quantify how many of the people a villager talks to about health-related matters use a given preventive measure – we refer to this as *network exposure*. Similarly, the number of household members using a preventive measure defines a villager's *household exposure* to this measure. **Figure 2** provides an illustrative example for network exposure to the use of coils, based on data from one of the villages; household exposure may be conceived similarly. We use this conceptualisation of exposure in the first part of our analysis, as it is compatible with standard logistic regression models. **Figure 3** provides a different view on exposure that allows to consider all preventive measures at once: for the same village, it presents measure use and discussion ties as a

multilevel network. In this network, ties representing usage connect villagers to any measure they report using and discussion ties connect villagers to each other. We use this conceptualisation in the second part of our analysis, as it allows to study patterns of use and discussion across measures in SAOMs.

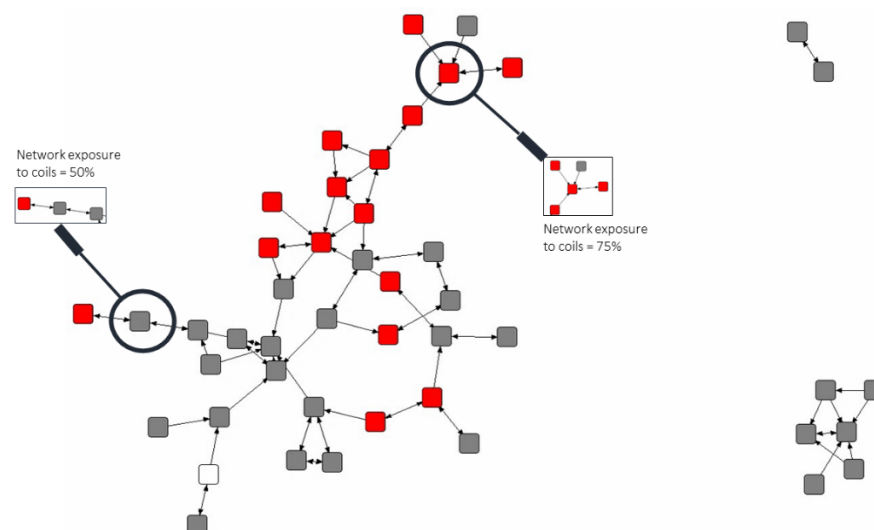


Figure 2: Example of percentage of exposure to the use of coils in WK2. Nodes represent interviewed villagers; lines represent whom the villagers talk to about health-related matters; red nodes indicate villagers who use coils; grey nodes indicate villagers who do not use coils. Network exposure to coils of the node highlighted at the top right = 75%. Network exposure to coils of the node highlighted at the left = 50%.

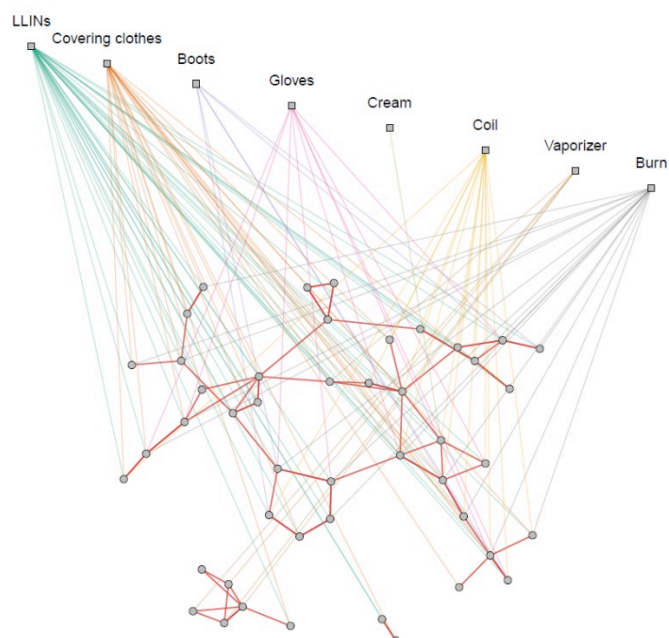


Figure 3: Multilevel network of people by people talking about health-related matters and using preventive measures in WK2. Appendix C in the Supplementary information (Figures C1-C10) reproduces multilevel network visualizations for each village.

RESULTS

Factors associated with the use of different preventive measures

Results of meta-analyses of village-level logistic regressions explaining the use of each preventive measure are reported in **Table 1**. The table presents estimates of population parameters (μ) and standard errors (SE) of variables that were retained through backward model selection performed separately for each preventive measure. See the Methods section for details about variable selection and meta-analysis models. Heterogeneity statistics for estimates are reported in the Supplementary Information (Appendix D, Tables D19-D26).

			LLINs	Covering clothes	Boots	Gloves	Insecticide cream	Coils	Vaporizers	Burn materials
A	Intercept	μ	0.61	-0.96*	-1.51***	-2.78***	-2.93***	-1.25*	-3.06***	-1.56***
		SE	(0.6)	(0.42)	(0.35)	(0.61)	(0.64)	(0.62)	(0.51)	(0.16)
B	Female	μ			-1.15***					
		SE			(0.28)					
	Head of household	μ		-0.28					0.10	
		SE		(0.21)					(0.87)	
	Carer for a sick person	μ						0.71		
		SE						(0.47)		
	Work in fields	μ					0.04	-0.61*		
		SE					(0.34)	(0.26)		
	Age	μ					0.00			
		SE					(0.17)			
	Education	μ				0.21	0.24**	0.10		
		SE				(0.18)	(0.09)	(0.09)		
C	Talks to Asha	μ		0.38						
		SE		(0.33)						
D	Network size in village	μ		0.10		-0.22		-0.12		
		SE		(0.13)		(0.22)		(0.19)		
	Network size out village	μ		0.32***						
		SE		(0.10)						
E	Network exposure	μ	2.21***	1.04***	1.39***			1.81***	1.25	1.06*
		SE	(0.38)	(0.28)	(0.26)			(0.48)	(1.01)	(0.46)
F	Household exposure	μ	0.39	0.56**	1.04***	2.08***	1.98***	1.07**	2.38***	0.98***
		SE	0.42	(0.21)	(0.18)	(0.56)	(0.31)	(0.41)	(0.31)	(0.33)
Number of village-level models			8	10	9	6	7	9	4	8

Table 1: Results of meta-analyses of parameters from village-level logistic regression models explaining the use of each preventive measure. Variables were chosen for each measure by backward model selection as described in the Methods section; variables are grouped as **A:** intercept; **B:** Individual characteristics; **C:** Opinion leaders; **D:** Network size; **E:** Network exposure; **F:** Household exposure; μ – estimated population parameters (log odds ratios); SE – standard errors of estimates; p values: * <0.05, ** <0.01, *** <0.001.

Table 1 shows that at least two explanatory factors were retained in the model selection process for every preventive measure. The meta-analysis of parameters highlights that there were two of these that were consistently associated with measure use across the villages.

First, network exposure significantly increased the probability of using five of the eight measures, except for gloves, insecticide cream, and vaporizers (odds ratios ranging from 2.8 to 9.1). Household exposure was significantly positively associated with the use of measures in seven cases, with the only exception LLINs (odds ratios for significant estimates between 2.7 and 10.8). These effect sizes are quite substantial, considering that the average number of discussion partners and the average household size are both 2 in the entire dataset. For example, our models predict that one additional discussion partner using boots would double the odds of a villager with an average-sized network using boots (network exposure log odds ratio 1.39).

Other explanatory variables did not have a consistent impact across preventive measures. Of the considered individual characteristics, only gender (boots), working in fields (coils), and education (insecticide cream) had a significant effect on measure use and only for one measure. The number of reported discussion partners outside of one's village only had a significant positive effect on the use of covering clothes. Talking to village opinion leaders, ASHAs and Traditional Healers, did not have a significant impact on the use of any of the measures. These results do not indicate that individual variables, opinion leaders, and network size are not associated with the use of measure at all. However, the effects of these factors vary from village to village (see the Supplementary Information, Appendix D, Tables D10-D17 for village-level model results).

Relevance of network and household exposure for explaining the use of measures

Next, we examine the accuracy of village-level models to assess the importance of different factors in explaining the use of preventive measures. To do so, we fit seven nested models for each preventive measure in each village, which include different subsets of the explanatory variables, as marked by the letters A-F in **Table 1**. The Empty Model contains an intercept term only (subset A) and serves as a baseline for accuracy per measure and village. The Individual Model includes the individual variables that were used for the given measure as well (A+B). The Opinion Leader Model further includes any effects of talking to the ASHA or the Traditional Healer (A+B+C). The Network Size Model considers any included variables related to the number of discussion ties (A+B+C+D). The Network Exposure Model (A+B+C+D+E) and the Household Exposure Model (A+B+C+D+F) add the relevant exposure term to the Network Size Model. The Full Model for each preventive measure is the same as the one reported in **Table 1**.

We compare the accuracy (correct classification rate) of the above models. The results are presented in **Figure 4**. The average accuracy of the Empty Model is 77%. Adding individual (77%), opinion leader (78%), and network size variables (79%) makes only a small difference. In contrast, including either network exposure or household exposure improves average accuracy by 2% and 3% compared to the Network Size Model. The Full Model, which includes both exposure terms beyond the Network Size Model, has an average accuracy of 83%. (For village- and measure-level classification rates, see the Supplementary Information, Appendix D, Tables D29-38).

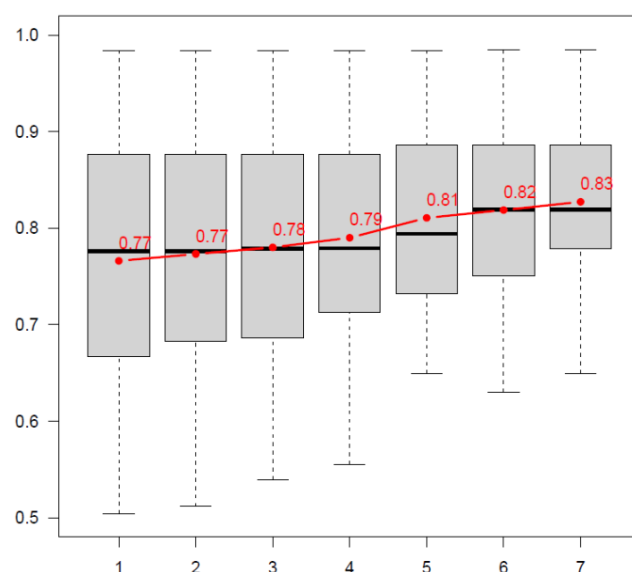


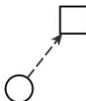
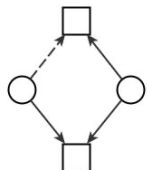
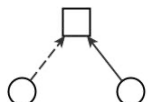

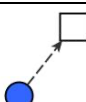
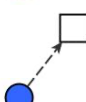
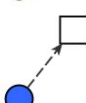
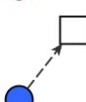
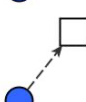
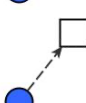
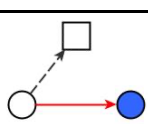
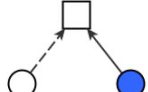
Figure 4: Distribution of the accuracy of different village-level logistic model specifications. Each bar represents models using different variable subsets of the models in **Table 1**, denoted by A-F; accordingly, each bar represents the accuracy of 61 models; the connected red points and values show the average accuracy of models; stars represent p-values for two-sided pairwise sample t-tests of the difference in the average accuracy of models adjacent in the figure: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (for full t-test statistics, see the Supplementary Information, Tables D27-D20).

Overall, the 6 percentage point improvement in accuracy between the Empty Model and the Full Model, and the 4 percentage point increase from including network and household exposure alone, may not appear high. However, these numbers should be interpreted in the context of a high baseline accuracy due to the skewness of measure use in many cases (e.g., LLINS are adopted by nearly everyone). Considering this, it is noteworthy that network and household exposure reduce the classification error of the Network Size Model by 20%, from 0.21 to 0.17. These results suggest that the two exposure factors substantially contribute to explaining the use of preventive measures.

Social mechanisms associated with preventive measure use

The results presented so far indicate that network and household exposure are key factors explaining the use of preventive measures in our villages. To explore the specific social mechanisms which may explain measure use, we fitted stationary SAOMs to the multilevel network of use and discussion ties, as defined earlier, in each village. We summarise the results of ten similarly specified SAOMs (one per village) by a meta-analysis of parameters in **Table 2**. As we are interested in the use of preventive measures, here we only discuss model effects that explain which measures are used by individuals. The models also simultaneously treat the village discussion network as an outcome that co-evolves with measure use (see the **Methods** section for further details on the model). This allows us to disentangle effect

directions between discussion and use, and thus take a step closer to identifying likely causal impacts of social networks on the use of measures. The full results of the SAOM meta-analysis, all village-level model estimates, convergence statistics, and goodness-of-fit statistics are reported in the Supplementary Information, Appendix E, Tables E1-E6.

A		Effects of structure of measure use	μ	SE	p	
	(1) Outdegree (intercept)	-2.22	0.56	0.000	***	
	(2) Agreement with villagers (4-cycles) x 100	0.02	0.20	0.919		
	(3) Popularity of measure	0.02	0.00	0.000	***	
	(4) Activity of villager	-0.11	0.06	0.086		
B		Effects of individual variables				
	(5) Female	-0.09	0.09	0.315		
	(6) Head of household	-0.09	0.09	0.368		
	(7) Carer for a sick person	-0.06	0.11	0.597		
	(8) Work in fields	0.11	0.08	0.170		
	(9) Age	0.06	0.06	0.343		
	(10) Education	0.07	0.04	0.110		
C		Effects of opinion leaders				
	(11) Talking to the ASHA	0.26	0.12	0.030	*	
	(12) ASHA's use of a specific measure	0.23	0.16	0.158		

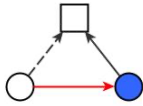
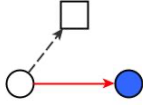
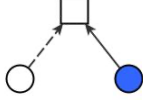
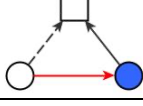
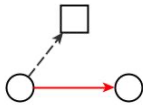
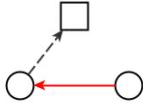
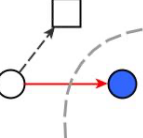
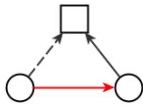
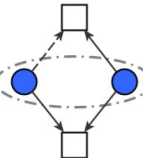
	(13) ASHA's use of a measure if talking to ASHA	0.06	0.17	0.716	
	(14) Talking to Healer	0.10	0.20	0.628	
	(15) Healer's use of a specific measures	-0.09	0.11	0.436	
	(16) Healer's use of a measure if talking to Healer	-0.01	0.33	0.995	
D <i>Effects of network size (talking to others)</i>					
	(17) Number of other villagers one talks to	-0.33	0.11	0.003	**
	(18) Number of other villagers who talk to one	0.01	0.04	0.728	
	(19) Number of non-villagers one talks to	0.10	0.03	0.000	***
E <i>Effect of network exposure</i>					
	(20) Using same measures as that one talks to	0.82	0.18	0.000	***
F <i>Effect of household exposure</i>					
	(21) Using same measures as household members	0.07	0.03	0.048	*

Table 2: Meta-analysis of estimates from multilevel stationary SAOMs in each village. Circles represent villagers; in each effect (explanatory variable), blue circles indicate villagers with specific characteristics (e.g., gender), squares represent preventive measures, red arrows represent health-related discussion ties, black solid arrows represent preventive measure use, and black dashed arrows represent the dependent tie, the probability of which is assessed by the given effect; rate parameters were fixed at 2.5 for both networks in all models; all models converged according to the criteria set out in Ripley et al. (2022) – see the Supplementary Information, Appendix E, Tables E5-E6 for further details; μ – estimated population parameters (log odds ratios); SE – standard errors of estimates; p values: * <0.05, ** <0.01, *** <0.001.

Like in any SAOM, the first set of parameters (A) control for the overall structure of the use network. The outdegree (density) effect (effect (1) in **Table 2**), which is significantly negative, is the intercept and highlights that using a measure is not very common on average: the

estimated baseline probability is 11%. The 4-cycle effect (2) expresses tendencies for similarity of use between individuals, which is not significant. The positive popularity of measures effect (3) reflects that villagers are significantly more likely to use measures which are already used by several others. The non-significant villager activity effect (4) suggests that there are no differences in villagers' tendency to use a measure based on how many other measures they are currently using.

The second set of parameters (B) models individual characteristics (5-10). None of these variables explain the use of preventive measures significantly, in line with the results of the logistic models. The third set of parameters (C) looks at the role of opinion leaders (11-16). Here, contrary to what we found for individual measures, talking to the ASHA significantly increases the odds of using any preventive measure, by 30% on average (11). However, talking to the Traditional Healer and considering the specific measures that the ASHA and the Traditional Healer uses do not seem to matter consistently across the villages. The fourth set of parameters (D) looks at effects of discussion network size (17-19). We find that talking to more people in one's own village significantly reduces (17), while talking to more people outside of the village increases the odds of using any preventive measure (19) (odds ratios 0.72 and 1.10). Finally, being talked to by many within one's village does not significantly influence the likelihood of using measures (18).

The last two sets of parameters (E-F) model network and household exposure (20-21). Both parameters are significant and positive. Network exposure (20) more than doubles the odds of use of a measure for each additional discussion partner who reports using the same measure (odds ratio 2.27). Further, if two members of a household use the same measure (21), it is more likely on average that they both use other measures as well (odds ratio 1.07). While both effects appear substantial in size, their relative importance may vary by individual, depending on the size of their discussion networks and households.

Network and household exposure explain patterns of preventive measure use and health-related discussions

The above results highlight the importance of considering villagers' discussion networks when we explain the use of preventive measures. To assess how well our village-level SAOMs capture the connections between the discussion and use networks, we compare the goodness of fit of different model specifications that include subsets of the effects in the full model in Table 2. **Figure 5** shows the distribution of SAOM goodness of fit p-values with regard to the mixed triad census of the two modelled networks (Hollway et al., 2017). We find that fit of a Baseline Model considering only endogenous structural effects in the discussion and use networks (effect group A) generally achieves a poor fit on these statistics. Sequentially adding effects of individual characteristics (A+B), opinion leaders (A+B+C), and network size (A+B+C+D) does not improve the fit. Adding network exposure (A+B+C+D+E) or household exposure (A+B+C+D+F) leads to a large improvement in model fit, while there are further gains from including both types of exposure effects (Full Model). We note that excluding network exposure from the Full Model has a statistically significant negative impact on fit (paired sample one-sided t-test: difference=-0.12, t=2.22, p=0.03, df=9), while excluding household exposure does not (difference=-0.08, t=1.63, p=0.07, df=9). These

results provide a clear indication that effects from network exposure and household exposure are crucial in explaining the links between the use preventive measures and discussion networks. Village-level goodness of fit statistics are reported in the Supplementary Information, Appendix E, Tables E5-E6.

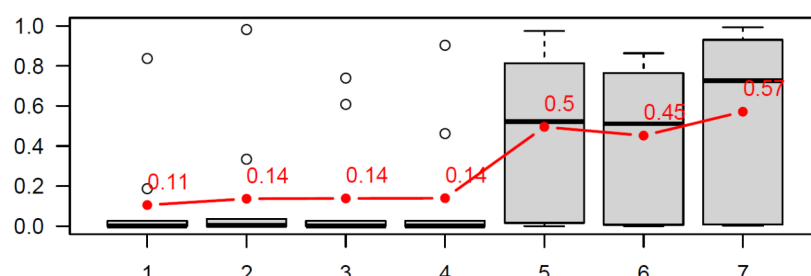


Figure 5. Comparison of village-level SAOM specifications by goodness fit (GoF) p-values regarding the combination of talk network and measure use ties. Each bar represents models using different variable subsets from **Table 2**, denoted by A-F; accordingly, each bar represents GoF results of 10 models; the considered GoF statistic is the mixed triad census of Hollway et al. (2017); the connected red points and values show the average GoF p-values of the models; stars represent p-values for one-sided pairwise sample t-tests of the difference in the average GoF p-value of models adjacent in the figure: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

DISCUSSION

Our results demonstrate that considering social networks contributes to the explanation of health-related behaviours and practices. We examined a variety of factors that may influence the uptake of eight malaria-preventive measures in ten hard-to-reach tribal communities in northeast India, using standard statistical methods and social network models. We found that exposure to the use of preventive measure in one's social network was consistently associated with own use. In both standard and network models explaining measure use, we showed that network exposure significantly improves the explanatory power of the models. Exposure in households had a comparable role in our analyses, while factors typically considered in studies of health behaviour, such as individual characteristics [Nlinwe et al., 2021; Matin et al., 2020; Van Eijk et al., 2016; Shahandeh et al, 2012] and opinion leaders [Stoebenau and Valente 2003; Valente and Pumpang 2007], had little to no effect.

The findings about social network effects are theoretically relevant as they provide empirical support for core ideas of network exposure models [Coleman et al., 1957; Valente, 1996; Valente 2010; Centola, 2018]. According to these, information and attitudes regarding both effective and ineffective preventive behaviours should be less dependent on who people are and more on whom they talk to. This suggests that research on health behaviours and social influence could benefit from further theoretical work that explores specific social mechanisms of network exposure. The effect of household members' behaviour is also

important, especially because exposure in households does not necessarily have to rely on discussions. People may use preventive measures simply because they are available in the house where they live, where exposure can be tacit and reinforced by observing use rather than talking about it. Collecting data on social networks and household composition, information on which are both not immediately available in contexts where census data are lacking, requires considerable amount of time and resources, but our analyses clearly show that these can provide novel theoretical and empirical insight into social processes explaining health behaviours.

We found a positive effect of discussion ties extending outside villages on the likelihood of using malaria preventive measures. We know from our data that these ties are mainly constituted by family and in-law ties (Of the 841 ties to people from other villages, 74% are to family members, 11% to relatives, and 15% to non-blood contacts). In Meghalaya, a matrilinear society, upon marriage sons move out of their maternal homes to stay at their spouse's home, while daughters remain close to their families of origin. The move of sons across villages creates networks of partially weak ties [Granovetter 1978]: these relationships are weak, according to Granovetter's definition [1978] in the sense that they are geographically distant and in infrequent contact. They are also potential sources of new information and attitudes about preventive techniques as they reach outside one's own, closely-knit village community. At the same time, family and relatives in different villages are emotionally close enough – a characteristic normally associated with strong ties [Granovetter 1978] to rely on each other for health-related discussions, which increases the chance that they actually influence villagers' health-related attitudes and behaviours. Therefore, intervillage ties complement the effects of network and household exposure on the use of preventive measures. The relatively small effect and explanatory importance of these ties may be due to the fact that they are informationally redundant: family and relatives likely all talk to each other, so they may end up sharing similar information or opinions. These findings point to the relevance of studying the extended family as a key social unit in shaping health-related behaviours and practices.

Interestingly, our analysis shows almost no effects from opinion leaders on using preventive measures. The literature suggests that they are rarely early adopters, but they could be instrumental in facilitating or hampering the spread of innovations [Valente and Pumpang 2007]. Our data does not allow to observe when exactly the ASHA and the Traditional Healer started using specific preventive measures, but results clearly indicate that their behaviour does not influence other villagers regardless if they talk to them or not. Talking to the ASHA does have an effect on using more measures, but this is small and it does not substantially contribute to model fit. This result may inform the planning of future interventions, as it indicates that relying solely on central actors like the ASHA may not be effective in improving disease prevention.

Similarly, we find little evidence that any of the individual characteristics we considered would consistently matter for use across villages and preventive measures. We did see indications in the logistic models that gender, type of work, and education are associated with measure use in case of a single measure. Other characteristics also appear to be

relevant, but only in some villages and in case of a few preventive measures (see the Supplementary Information, Appendix D, Tables D10-D17). Overall, our analyses do not suggest that individual factors do not matter at all, only that their effects on the use of measures may differ from community to community. By exploring these differences in a localised geographic context, large-scale public health interventions can be effectively tailored to the attitudes and practices of relevant social groups.

Our study has a few key limitations that may inform the design of future research into health-related behaviours. First, we collected self-reported data on behaviours and social ties. Self-reports are known to be sensitive to recall bias [Bernard and Killworth 1977; Small 2017] as well as socially desirable responses [Edwards 1953]. We aimed to attenuate this issue by keeping questions about prevention practices and discussion ties general. We only asked participants to report on whether they use different preventive measures or not (we queried frequency of use regarding covering clothes and insecticide cream but disregarded this detail in the analyses for comparability; see the Supplementary Information, Appendix B1 for details). We measured social networks by asking people whom they talk to about health-related matters without reference to details of these exchanges. While these questions could reduce recall problems, they did not inform us about the frequency and contexts of measure use and the content of discussions. Employing field observations [Bernard and Killworth 1977; Curtis et al., 1995] and smart sensor technologies [Cattuto et al., 2010; Elmer et al., 2019; Voros et al. 2021] may help to gather richer data. However, trade-offs between data quality and resolutions should be carefully considered, as these methods are more intrusive and may lead to lower participation rates [Voros et al. 2021] and alter behaviour [Landsberger, 1958]. Our interview-based approach helped us to gain the trust of our participants and elicit meaningful responses about general patterns of measure use and discussion ties.

Second, the validity of our findings should be assessed in the context of coverage and sample size. Our overall response rate is 68%, which is considered sufficiently high for statistical network analyses [Kossinets 2006]. However, coverage was not equal across the three regions and was only just over 60% in West Jaintia Hills. This partially compensated with the higher coverage of households (80%). Overall, our data collection efforts have resulted in reasonably high response rates that indicate the adequate internal validity of our results. In turn, our findings should not be generalised beyond the studied contexts. We only had the resources to observe a small proportion of communities in just three districts of Meghalaya state. Nonetheless, we found some common patterns in the studied communities which suggest that processes from network and household exposure may help to explain the use of preventive measures. We need larger-scale, regionally representative datasets to confirm these findings and gain a more complete picture of the social mechanisms driving health-related behaviours. Our results further suggest that future work should focus on collecting detailed data on weak ties and inter-villages relationships. Combined with a large-scale approach, this could improve our understanding of how effectively preventive measure spread across malaria endemic regions.

Third, we interpret our key results in terms of social processes, while we collected and

analysed cross-sectional data. Clearly, analysing one snapshot of networks and behaviours comes with limitations. We do not know how the use of measures changed over time. It is also possible that some measures were becoming popular, while others unpopular at the time of the data collection. We do not have information about the stability of health-related discussion ties either, although the network literature suggests that such relations should be stable in the short-term [Hogan et al., 2022]. In line with these concerns, our logistic regression results only highlight associations, without considering the effects of measure use on social networks. The Stochastic Actor-oriented Model is better suited to identify the social processes at play. Even when applied to cross-sectional data, the SAOM is dynamic: it assumes that villagers may still change their behaviour and network ties, but these changes keep the overall structure of the multilevel network in a stable state (in a “short-term dynamic equilibrium”; Snijders and Steglich 2015). This means that the model considers both the effects of network ties on measure use and vice versa (see full results the Supplementary Information, Appendix E, Tables E1-E2). By concurrently modelling the dependencies between network ties and behaviours, we may infer to the social mechanisms that are likely to maintain the network system [Lusher et al. 2012], such as network or household exposure. Any causal claims that are made on the basis of “multi-mechanistic” network models [Stadtfield and Amati 2021], such as the SAOM, are stronger than those based on results from non-network models. At the same time, collecting longitudinal data and studying the actual changes of networks and behaviour will be required to gain a clear understanding of the dynamics of measure use. Our approach can be easily adapted to serve as a basis of such a longitudinal design.

For the first time since 2015, when malaria elimination efforts were initiated in India [PIB 2020], local elimination may be within reach. However, current control efforts rely solely on mass interventions. These include the large-scale distribution of LLINS that was interrupted in India in recent years [WHO 2022], and IRS which is met with high levels of refusal in Meghalaya [Sarkar et al., 2021; Passah et al., 2022]. Our study suggests that these interventions could be complemented with targeted efforts that leverage social influence in interpersonal relations and households to increase uptake. This could help to achieve downstream elimination goals. However, further research is needed to map the patterns of measure use across different areas and social groups and to understand the social mechanisms that may facilitate or hamper the adoption of various health behaviours. This line of work may contribute to increasing the acceptance of a malaria vaccine in the future. Beyond the context of malaria, exploring the role of social networks in health behaviour could benefit the prevention of other serious mosquito-borne diseases, such as Dengue, Chikungunya and the newly emerging Zika virus [Gulland 2016], which are on the rise both globally and in India [Shepard et al. 2014].

METHODS

Data collection

From January 2020 through August 2021, we collected data from three villages in West Khasi Hills (WK1, WK2, WK3), three in West Jaintia Hills (WJ1, WJ2, WJ3) and four in South Garo Hills (SG1, SG2, SG3, SG4) in Meghalaya, India. Villages were selected based on their

manageable size (<500 eligible adults), their known willingness to participate in similar studies (we avoided villages where previous epidemiological studies encountered high rates of refusal, see Kessler et al., 2018), and their accessibility by either car or foot. Village-level coverage ranged from 53% of residents in WJ3 (80% of households) to 88% in WK3 (97% of households), with an average of 68% of individuals (n=1,530) and 80% households (a total of 766) interviewed. The topic of coverage is addressed in the Discussion section. Characteristics of non-respondents are reported in the Supplementary Information, Appendix A, Table A2.

Structured interviews were conducted in the appropriate local language (Khasi, Pnar or Garo) and translated into English by the interviewing team. Respondents were asked a series of questions regarding their individual background, household and village roles, their use of malaria preventive measures, and the people in and outside of their village they talk to about health-related issues. In total, each interview consisted of 26 questions and lasted about 30 minutes. The full list of interview questions, details about data coding and processing, and descriptive statistics for variables are reported in the Supplementary Information, Appendix B, Tables B1-B15.

The ASHA was interviewed in all ten villages, and in the six villages where present, the Traditional Healer was also interviewed. As we consider the ASHA and the Traditional Healers as opinion leaders, whose behaviour and influence may be substantially different from those of other villagers, we removed these two actors as respondents from our analyses. We still utilize information about their use of measures and about who talks to them to explain the behaviour and social networks of other villagers.

Permission to conduct the study was granted by the Headman of each village, and all respondents also signed an individual informed consent form. Identifiers of the villages and individual participants were removed from the dataset and this article to protect anonymity. Ethical approval for the study was obtained from the Institutional Review Boards (IRBs) of Martin Luther Christian University, Shillong, Meghalaya, India and New York University, New York, NY, USA.

Variables

The dependent variables in our analyses represent the use of eight malaria prevention measures by participants: Long-Lasting Insecticidal Nets (LLINs), covering clothes, boots, gloves, insecticide cream, coils, vaporisers, and burning materials. We considered how a variate of explanatory variables were associated with these outcomes in different multivariate statistical models. In total, 14 relevant explanatory variables were available in the dataset and were considered in our analyses. We reduced this set for the meta-analysis of logistic regressions by a model selection procedure described below. The fitted social network models estimated a larger number of parameters, because each variable can affect the outcome in multiple ways in these models.

The starting set of explanatory variables was the following:

- **Individual characteristics and roles in the household:** (1) gender, (2) being the head of the household, (3) looking after family members when they are sick (4) working in

fields, (5) age, and (6) educational background.

- **Contact with opinion leaders:** the participant or someone from their household talks to the Asha (7-8) or the Traditional Healer (9-10) of the village about health-related issues.
- **Social network size:** the number of people in and outside the village the participant talks to about health-related issues (11-12).
- **Social network exposure:** the amount of people the participant talks to who use a given preventive measure (13)
- **Household exposure:** the amount of people in the participant's household who use a given preventive measure (14)

Logistic regression model selection procedure

We considered 14 explanatory variables for each logistic model explaining the use of a given preventive measure in a specific village (see the Supplementary Information, Appendix B, for the complete list of variables). Given the size of our villages and the often skewed distribution of measure use, we found that in most cases models including all 14 explanatory factors could not be estimated. Therefore, we employed a two-stage model selection procedure aimed at reducing the set of explanatory variables while providing a single specification per measure that can be estimated in every village. First, we carried out backward model selection based on the p-values of parameter estimates for each of the 8 (measures) x 10 (villages) = 80 models. This resulted in different optimal model specifications across measures and villages. Second, considering each measure separately, we added variables back that had a significant estimate in at least two of the villages. This resulted in 8 different model specifications, one for each preventive measure, that were identical across the villages. The exact steps of our procedure are detailed in Procedure 1. Estimating the reduced models was still difficult in case of some villages and preventive measures. We omitted models from the meta-analyses reported in the text that had very large estimated parameters and/or standard errors for at least one explanatory variable. The presented results are based on 61 models. See the Supplementary Information, Appendix D, Table D18 for the excluded models.

Procedure 1. Two-stage model selection approach

Denote villages as $v_i \in V$ where $i \in \{1, \dots, 10\}$, preventive measures as $m_j \in M$ where $j \in \{1, \dots, 8\}$, and explanatory variables as $e_k \in E$ where $k \in \{1, \dots, 18\}$.

Stage I. Backward model selection

Repeat for every combination of i and j :

Estimate a logistic model in v_i with m_j as dependent and E as explanatory variables

Define $E' = E$

Repeat until $p < 0.05$ for all estimates for variables in E' :

Identify the explanatory variable e^* with the highest p-value of the estimate

$E' = E' \setminus e^*$ (exclude e^* from E')

Estimate a new logistic model with E' as the explanatory variable set

Define $E_{ij} = E'$

Return E_{ij}

Stage II. Find identical specifications for each measure across villages

Repeat for every j:

Define $d = 1$

Repeat until all parameters and standard errors can be estimated in all models:

Define E^* as set containing all elements that occur in E_{1j}, \dots, E_{10j} at least d times (for $d=1$, this is the union of the ten sets)

Estimate logistic models in every v_i with m_j as dependent and E^* as explanatory variables

If not all models can be estimated then $d = d+1$

Define $E_j^* = E^*$

Return E_j^*

571

572 Meta-analyses of village-level logistic regressions

573 To explore recurrent patterns of association between the use of different measures and our
 574 explanatory variables across villages, we meta-analysed the model results obtained from the
 575 model selection procedure described above. We perform parameter-wise meta-analyses with
 576 random-effects models estimated by restricted maximum likelihood using the metafor
 577 package (v3.8-1) in R (Viechtbauer 2010). Estimates provide information about the mean and
 578 standard error of parameters. Heterogeneity statistics are reported in the Supplementary
 579 Information, Appendix D, Tables D19-D26. The meta-analysis results are presented in Table 1.

580 Comparison of the explanatory power of nested logistic regression models

581 We compared the explanatory power of the village-level logistic models described above and
 582 of nested models that contain different subsets of the explanatory variables. We calculated
 583 the accuracy for each model as the correct classification rate: the percentage of cases for
 584 which the observed value of the dependent variable (whether a villager uses a given measure
 585 or not) equals the estimated value obtained from the model. We used the standard 0.5 cut-
 586 off for transforming probabilities into estimated binary outcomes. The results of this analysis
 587 are presented in Figure 4, while the accuracy of the village-level models for each village are
 588 reported in Supplementary Information, Appendix D, Tables D29-D38.

589 Meta-analysis of village-level stationary Stochastic Actor-oriented Models

590 For a more detailed analysis of the social mechanisms affecting the use of preventive
 591 measures, we employed stationary Stochastic Actor-oriented Models (SAOMs). Stationary
 592 SAOMs are an extension of the original model [Snijders et al. 2010; Snijders 1996] for the
 593 analysis of cross-sectional network data. In this case, it is assumed that the networks studied
 594 are in a short-term dynamic equilibrium: their structure (but not single network ties) is at
 595 least temporarily in a stable state [Snijders and Steglich 2015]. SAOMs are empirically
 596 calibrated simulation models that aim at identifying the relative strength of a set of social
 597 mechanisms, operating on network ties, individual and pairwise (dyadic) covariates, that over
 598 time could have generated an observed network. In a standard longitudinal SAOM, the model
 599 is conditional on the first observation of the network [Snijders et al. 2010]. In a stationary
 600 SAOM, the initial and final states of the network are identical, and the modelled social

processes will, at least stochastically, maintain the existing structure of the network (hence the ‘stationary’ label) [Snijders and Steglich 2015].

To apply stationary SAOMs, we defined a multilevel network that consists of two interlinked networks in each village: 1) the network of health-related discussions, which is a one-mode network where villagers A and B are connected by a binary tie if A reports talking to B; and 2) the network of preventive measure use, which is a two-mode network where villager A is connected to measure M if A reports using M. An example of this network for one of the villages is presented in Figure 3. Similar visualisations for all villages can be found in the Supplementary Information, Appendix C.

We fit stationary SAOMs using the RSiena (v1.3.0) package in R (Ripley et al. 2022). We set the rate parameter to 2.5 for both networks in all villages, as this value enabled model convergence and led to good model fit in every case. The key results are interpreted in the main text. The full list of estimates, goodness of fit tests and robustness checks for rate parameters are presented in the Supplementary Information, Appendix E. Similar to the case of logistic regressions above, we perform parameter-wise meta-analyses of village-level SAOMs with random-effects models estimated by restricted maximum likelihood using the metafor package (v3.8-1) in R (Viechtbauer 2010). Estimates provide information about the mean and standard error of parameters. Heterogeneity statistics are reported in the Supplementary Information, Appendix E, Table E4. The meta-analysis results are presented in Table 2.

Comparison of the explanatory power of nested SAOMs

We compared the explanatory power of the village-level SAOMs described above and of nested models that contain different subsets of the model effects. We calculated p-values of Mahalanobis distances for a mixed triad census fit test for different SAOM specifications in the results section. These reflect how well the distribution of simulated networks based on the fitted models represent the mixed triad census [Hollway et al., 2017] of the observed networks: a p-value close to 1 signifies very good fit and one close or equal to 0 marks poor fit [Ripley et al., 2022]. Calculations in each model are based on 5000 simulations and follow the steps described in Lospinoso and Snijders [2019]. The goodness of fit tests were carried out using the RSiena R package [Ripley et al., 2022]. The results are presented in Figure 5. Further details of the goodness of fit analyses can be found in the Supplementary Information, Appendix E, Table E5.

ACKNOWLEDGMENTS

We thank the study participants for their time, our field teams for their data collection and transcribing efforts, and the Meghalaya Malaria State Control Programme for their support and ongoing collaboration. We also thank Drs. Mark Wilson and Steven Sullivan for manuscript editing and input. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number U19AI089676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS

E.B., A.K., A.V.E., J.C., and S.A. designed the study and analyses. M.P. C.B.N, Q.D.N and C.K. conducted the questionnaires and collected the data, under the supervision of R.S. and S.A. E.B. and A.V. analysed the data and wrote the first draft of the paper. All authors finalized the manuscript. J.C. obtained the funding.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information

Correspondence and requests for materials should be addressed to Elisa Bellotti and Sandra Albert.

Data availability. TBA

Code availability. TBA

REFERENCES

- adams, j., & Schaefer, D. R. (2016). How initial prevalence moderates network-based smoking change: estimating contextual effects with stochastic actor-based models. *Journal of Health and Social Behavior*, 57(1), 22-38.
- Albert, S., Nongrum, M., Webb, E.L., Porter, J.D., Kharkongor, G.C. Medical pluralism among indigenous peoples in northeast India—implications for health policy. *Trop Med Int Health*. 20:952–60 (2015).
- Bartels, C. and Ketellapper, R. (eds) (1979) Exploratory and Explanatory Statistical Analysis of Spatial Data, The Hague: Martinus Nijhoff.
- Bedson, J., Skrip, L.A., Pedi, D. *et al.* A review and agenda for integrated disease models including social and behavioural factors. *Nat Hum Behav* 5, 834–846 (2021).
<https://doi.org/10.1038/s41562-021-01136-2>
- Burt, Ronald S. 2004. “Structural Holes and Good Ideas.” *American Journal of Sociology*, 110(2): 349–99.
- Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J-F, Vespignani A (2010) Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLoS ONE* 5(7): e11596. <https://doi.org/10.1371/journal.pone.0011596>
- Centola, D. & Macy, M. W. Complex contagions and the weakness of long ties. *Am. J. Sociol.* 113, 702–734 (2007).
- Centola, D. An Experimental Study of Homophily in the Adoption of Health Behavior. *SCIENCE* 334, 6060: 1269-1272 (2011). DOI: 10.1126/science.1207055
- Centola, D. How Behavior Spreads: The Science of Complex Contagions (Princeton University

676 Press, 2018

677 Centola, D. The spread of behavior in an online social network experiment. *Science* 329,
678 1194–1197 (2010).

679 Christakis, N.A., and Fowler, J.H. The spread of obesity in a large social network over 32 years.
680 *The New England journal of medicine* 357,4: 370-9 (2007). doi:10.1056/NEJMsa066082

681 Coleman J., Katz E., Menzel H., 1957, The Diffusion of an Innovation Among Physicians,
682 *Sociometry*, 20, 4: 253-270.

683 Crosnoe, R., Frank, K., Mueller, A.S. Gender, body size and social relations in American high
684 schools. *Soc Forces* 86 (3):1189–1216 (2008).

685 Curtis, R., et al. (1995). Street-level drug markets: network structure and HIV risk. *Social*
686 *Networks* 17: 229-49.

687 Debboun, M., Strickman, D. Insect repellents and associated personal protection for a
688 reduction in human disease. *Med Vet Entomol.* 27(1):1-9 (2013). doi: 10.1111/j.1365-
689 2915.2012.01020.x.

690 Dhiman, S. Are malaria elimination efforts on right track? An analysis of gains achieved and
691 challenges ahead. *Infect Dis Poverty* 13, 8(1):14 (2019). [https://doi.org/10.1186/s40249-019-](https://doi.org/10.1186/s40249-019-0524-x)
692 [0524-x](https://doi.org/10.1186/s40249-019-0524-x)

693 Doreian, P. (1980) 'Linear models with spatially distributed data: spatial disturbances or
694 spatial effects?', *Sociological Methods and Research* 9: 29–60.

695 Dow, M.M., Burton, M.L., White, D.R. and Reitz, K.P. (1984) 'Galton's problem as network
696 autocorrelation', *American Ethnologist* 11: 754–70.

697 Dow, M.M., White, D.R. and Burton, M.L. (1983) 'Multivariate modeling with interdependent
698 network data', *Behavior Science Research* 17: 216–45.

699 Edwards, A. L. (1953). 'The relationship between the judged desirability of a trait and the
700 probability that the trait will be endorsed', *Journal of Applied Psychology*, 37: 90-93.

701 Elmer, T., Chaitanya, K., Purwar, P. et al. The validity of RFID badges measuring face-to-face
702 interactions. *Behav Res* 51, 2120–2138 (2019). <https://doi.org/10.3758/s13428-018-1180-y>

703 Franken, R., Bekhuis, H., & Tolsma, J. (2023). Kudos make you run! How runners influence
704 each other on the online social network Strava. *Social Networks*, 72, 151-164.

705 Freeman L C (1979). 'Centrality in Social Networks: Conceptual clarification', *Social Networks*
706 1, 215-239.

707 Gayen, K. and Raeside, R., Social networks and contraception practice of women in rural
708 Bangladesh. *Social Science & Medicine* 71, 9: 1584-1592 (2010).
709 <https://doi.org/10.1016/j.socscimed.2010.08.002>.

710 Gosling, R., Chimumbwa, J., Uusiku, P. et al. District-level approach for tailoring and targeting
711 interventions: a new path for malaria control and elimination. *Malar J* 19, 125 (2020).

712 <https://doi.org/10.1186/s12936-020-03185-w>

713 Granovetter, M. Threshold models of collective behavior. *AJS* 83, 6: 1420-1443 (1978).

714 Gulland, A. WHO urges countries in dengue belt to look out for Zika. *BMJ* 352:i595, (2016).

715 de la Haye, K., Shin, H., Vega Yon, G. G., & Valente, T. W. (2019). Smoking diffusion through
716 networks of diverse, urban american adolescents over the high school period. *Journal of*
717 *Health and Social Behavior*, 60(3), 362-376.

718 Hogan, B., Janulis, P., Phillips, G., Melville, J., Mustanski, B., Contractor, N., & Birkett, M.
719 (2020). Assessing the stability of egocentric networks over time using the digital participant-
720 aided sociogram tool Network Canvas. *Network Science*, 8(2), 204-222.
721 doi:10.1017/nws.2019.27

722 Hollway, James, Alessandro Lomi, Francesca Pallotti, and Christoph Stadtfeld. "Multilevel
723 social spaces: The network dynamics of organizational fields." *Network Science* 5, no. 2
724 (2017): 187-212.

725 Holly, B.S., Christakis, N.A., and Fowler J.H. Association Between Social Network Communities
726 and Health Behavior: An Observational Sociocentric Network Study of Latrine Ownership in
727 Rural India. *American Journal of Public Health* 104, 930-937 (2014).
728 <https://doi.org/10.2105/AJPH.2013.301811>

729 Horsevad, N., Mateo, D., Kooij, R.E. et al. Transition from simple to complex contagion in
730 collective decision-making. *Nat Commun* 13, 1442 (2022). [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-022-28958-6)
731 [022-28958-6](https://doi.org/10.1038/s41467-022-28958-6)

732 Kessler A., van Eijk A. M., Limalemla J., Walton C., Carlton J. M., Albert S., Malaria in
733 Meghalaya: a systematic literature review and analysis of data from the National Vector-
734 Borne Disease Control, *Malar J* 17:411 (2018) <https://doi.org/10.1186/s12936-018-2563-3>

735 Kessler, A., Shylla, B., Singh, U.S. et al. Spatial and temporal village-level prevalence of
736 Plasmodium infection and associated risk factors in two districts of Meghalaya, India. *Malar J*
737 20, 70 (2021). <https://doi.org/10.1186/s12936-021-03600-w>

738 Kim, D.A., Hwong, A.R., Stafford, D., Hughes, D.A., O'Malley, A.J., Fowler, J.H., Christakis, N.A.
739 Social network targeting to maximise population behaviour change: a cluster randomised
740 controlled trial. *Lancet*. 386(9989):145-53 (2015). doi: 10.1016/S0140-6736(15)60095-2.

741 Kossinets, G., Effects of missing data in social networks. *Social Networks* 28, 3: 247-268
742 (2006). <https://doi.org/10.1016/j.socnet.2005.07.002>.

743 Koster, J. Interhousehold Meat Sharing among Mayangna and Miskito Horticulturalists in
744 Nicaragua. *Hum Nat* 22, 394–415 (2011). <https://doi.org/10.1007/s12110-011-9126-4>

745 Landsberger, H. A. (1958). Hawthorne Revisited. Ithaca. OCLC 61637839.

746 Lazega, E., & Snijders, T. A. (Eds.). (2015). *Multilevel network analysis for the social sciences:*
747 *Theory, methods and applications* (Vol. 12). Springer.

748 Leenders, T.T.A.J. (2002) 'Modeling social influence through network autocorrelation:
749 constructing the weight matrix', *Social Networks* 24: 21–47.

750 Lek, D., Callery, J.J., Nguon, C. et al. Tools to accelerate falciparum malaria elimination in
751 Cambodia: a meeting report. *Malar J* 19, 151 (2020). [https://doi.org/10.1186/s12936-020-](https://doi.org/10.1186/s12936-020-03197-6)
752 [03197-6](https://doi.org/10.1186/s12936-020-03197-6)

753 Lele, N., Joshi, P.K. Analyzing deforestation rates, spatial forest cover changes and identifying
754 critical areas of forest cover changes in North-East India during 1972–1999. *Environ Monit*
755 *Assess.*56:159–70 (2009).

756 Levy, D. E. et al. Social connections and the healthfulness of food choices in an employee
757 population. *Nature human behaviour* 5,10: 1349–1357 (2021). doi:10.1038/s41562-021-
758 01103-x

759 Long, E., & Valente, T. W. (2019). Perceived social acceptability and longitudinal trends in
760 adolescent cigarette smoking. *Prevention Science*, 20, 824–832.

761 Lospinoso, Josh, and Tom AB Snijders. "Goodness of fit for stochastic actor-oriented models."
762 *Methodological Innovations* 12, no. 3 (2019): 2059799119884282.

763 Marsden P. V, Egocentric and sociocentric measures of network centrality, *Social Networks*,
764 Volume 24, Issue 4, 2002, Pages 407–422, [https://doi.org/10.1016/S0378-8733\(02\)00016-3](https://doi.org/10.1016/S0378-8733(02)00016-3).

765 Matin, M.A., Sarkar, N.D.P., Phru, C.S., Ley, B., Thriemer, K., Price, R.N., Peeters Grietens, K.,
766 Ali Khan, W., Alam, M.S., Gryseels, C. Precarity at the Margins of Malaria Control in the
767 Chittagong Hill Tracts in Bangladesh: A Mixed-Methods Study. *Pathogens*. 14;9(10):840
768 (2020). doi: 10.3390/pathogens9100840.

769 Ministry of Tribal Affairs. Demographic status of scheduled tribe population of India. 2013.
770 *Government of India*. <http://www.tribal.gov.in>. Accessed 18 August 2021.

771 Mishra, A.K., Nisar, S., Rajvanshi, H. et al. Improvement of Indoor Residual Spraying and Long-
772 Lasting Insecticidal Net services through structured monitoring and supervision as part of the
773 Malaria Elimination Demonstration Project in Mandla, Madhya Pradesh. *Malar J* 20, 101
774 (2021). <https://doi.org/10.1186/s12936-021-03639-9>

775 Musoke, D., Atusingwize, E., Namata, C. et al. Integrated malaria prevention in low- and
776 middle-income countries: a systematic review. *Malar J* 22, 79 (2023).
777 <https://doi.org/10.1186/s12936-023-04500-x>

778 Nlinwe, N.O., Singong, Y.C., Florentine, T.M.R. Evaluation of malaria preventive measures
779 among adult patients attending the Bamendjou and Foubot district hospitals of the West
780 Region of Cameroon. *Malar J*. 22;20(1):60. (2021) doi: 10.1186/s12936-021-03592-7.

781 Ord, K. (1975) Estimation methods for models of spatial interaction', *Journal of the American*
782 *Statistical Association* 70: 120–26.

783 Passah, M., Nengnong, C.B., Wilson, M.L., Carlton, J.M., Kharbamon, L., Albert, S.
784 Implementation and acceptance of government-sponsored malaria control interventions in

785 Meghalaya, India. *Malar J.* 21(1):200 (2022).

786 Perkins J. M., Subramanian S.V., Christakis N. A., Social networks and health: A systematic
787 review of sociocentric network studies in low- and middle-income countries, *Social Science &*
788 *Medicine*, Volume 125, 2015, Pages 60-78, <https://doi.org/10.1016/j.socscimed.2014.08.019>.

789 PIB 2020, <https://pib.gov.in/pressreleaseiframepage.aspx?prid=1677601> Accessed
790 22/07/2022

791 Proestakis, A., di Sorrentino, E.P., Brown, H.E. *et al.* Network interventions for changing
792 physical activity behaviour in preadolescents. *Nat Hum Behav* 2, 778–787 (2018).
793 <https://doi.org/10.1038/s41562-018-0436-y>

794 Rajagopal, R. Studies on persistent transmission of malaria in Burnihat, Meghalaya. *J Commun*
795 *Dis.* 8:235–45 (1976).

796 Ripley, R. M., Snijders, T. A. B., Boda, Z., Vörös, A. & Preciado, P. Manual for RSiena (v. 1.3.6).
797 (2022).

798 Rogers Everett, M. *Diffusion of Innovations*. New York: Free Press (1962).

799 Rubio-Palis Y., Curtis C. F.. Biting and resting behaviour of anophelines in western Venezuela
800 and implications for control of malaria transmission, *Medical and Veterinary Entomology*, 6,
801 325–334 (1992). <https://doi.org/10.1111/j.1365-2915.1992.tb00628.x>

802 Sahu SS, Thankachy S, Dash S, Nallan K, Swaminathan S, Kasinathan G, et al. Evaluation of
803 long-lasting indoor residual spraying of deltamethrin 62.5 SC-PE against malaria vectors in
804 India. *Malar J.* 2020. <https://doi.org/10.1186/s12936-020-3112-4>.

805 Salvy, S.J., de la Haye, K., Bowker, J. C., & Hermans, R. C. J. Influence of peers and friends on
806 children's and adolescents' eating and activity behaviors. *Physiology & Behavior*, 106, 369–
807 378 (2012). <https://doi.org/10.1016/j.physbeh.2012.03.022>.

808 Sangoro, P. Can repellents prevent malaria in Tanzania? PhD thesis, London School of
809 Hygiene & Tropical Medicine (2015). DOI: <https://doi.org/10.17037/PUBS.02374474>

810 Sarkar, R., Kessler, A., Mawkhlieng, B. et al. Household and individual level risk factors
811 associated with declining malaria incidence in Meghalaya, India: implications for malaria
812 elimination in low-endemic settings. *Malar J* 20, 460 (2021). [https://doi.org/10.1186/s12936-](https://doi.org/10.1186/s12936-021-03982-x)
813 [021-03982-x](https://doi.org/10.1186/s12936-021-03982-x)

814 Shahandeh, K., Basseri, H. & Shojaeizadeh, E. Using the cultural model to plan intervention
815 for malaria control in immigrants and native communities in endemic area, earmarked for
816 malaria elimination southeastern Iran. *Malar J* 11, P132 (2012).
817 <https://doi.org/10.1186/1475-2875-11-S1-P132>

818 Sharma, R.K., Thakor, H.G., Saha, K.B., Sonal, G.S., Dhariwal, A.C., Singh, N. Malaria situation
819 in India with special reference to tribal areas. *Indian J Med Res.* 141:537–45 (2015).

820 Shepard, D.S., Undurraga, E.A., Betancourt-Cravioto, M., Guzman, M.G., Halstead, S.B., Harris,
821 E., et al. Approaches to refining estimates of global burden and economics of dengue. *PLoS*

822 *Negl Trop Dis.* 8(11):e3306 (2014).

823 Simpson, C.R. Social Support and Network Formation in a Small-Scale Horticulturalist
824 Population. *Sci Data* 9, 570 (2022). <https://doi.org/10.1038/s41597-022-01516-x>

825 Smith, K.P. and Christakis, N.A. Social Networks and Health. *Annual Review of Sociology* 34:1,
826 405-429 (2008).

827 Snijders, Tom A.B. (1996) Stochastic actor-oriented models for network change. *Journal of*
828 *Mathematical Sociology*, 21, 149-172.

829 Snijders, T. A. B. & Steglich, C. E. G. Representing Micro–Macro Linkages by Actor-based
830 Dynamic Network Models. *Sociol. Methods Res.* 44, 222–271 (2015).

831 Snijders, T.A.B., Steglich, C.E.G. and Van de Bunt, G.G. (2010) ‘Introduction to actor-based
832 models for network dynamics’, *Social Networks* 32: 44–60.

833 Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and*
834 *Advanced Multilevel Modeling*. 2nd Edition. Sage, London.

835 Stadfeld, C., & Amati, V. (2021). Network mechanisms and network models. In *Research*
836 *Handbook on Analytical Sociology* (pp. 432-452). Edward Elgar Publishing.

837 Steglich, C., Snijders, T. A., & Pearson, M. (2010). 8. Dynamic networks and behavior:
838 separating selection from influence. *Sociological Methodology*, 40(1), 329-393.

839 Stoebenau, K., Valente, T.W. Using network analysis to understand community-based
840 programs: a case study from highland Madagascar. *Int Fam Plan Perspect.* 29(4):167-73
841 (2003). doi: 10.1363/ifpp.29.167.03. PMID: 14665425..

842 Valente, T. W., & Pumpuang, P. (2007). Identifying Opinion Leaders to Promote Behavior
843 Change. *Health Education & Behavior*, 34(6), 881–896.
844 <http://www.jstor.org/stable/45055988>

845 Valente, T.W. Social network thresholds in the diffusion of innovations. *Social Networks*, 18,
846 1, 69-89 (1996). [https://doi.org/10.1016/0378-8733\(95\)00256-1](https://doi.org/10.1016/0378-8733(95)00256-1).

847 Valente, Thomas W., *Social Networks and Health: Models, Methods, and Applications* (New
848 York, 2010; online edn, Oxford Academic, 1 May 2010).

849 Valente, T.W., Fujimoto, K., Chou, C.P., Spruijt-Metz, D. Adolescent Affiliations and Adiposity:
850 A Social Network Analysis of Friendships and Obesity. *J Adol Health* 45:202-204 (2009).

851 Van Eijk, A.M. et al. The use of mosquito repellents at three sites in India with declining
852 malaria transmission: surveys in the community and clinic. *Parasites & vectors* vol. 9,1 418.
853 (2016), doi:10.1186/s13071-016-1709-9

854 Verelst, F., Willem, L. & Beutels, P. Behavioural change models for infectious disease
855 transmission: a systematic review (2010-2015). *J. R. Soc. Interface* (2016).
856 <https://doi.org/10.1098/rsif.2016.0820>

857 Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of*
858 *Statistical Software*, **36**(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>

859 Vörös András, Zsófia Boda, Timon Elmer, Marion Hoffman, Kieran Mepham, Isabel J. Raabe,
860 Christoph Stadtfeld, The Swiss StudentLife Study: Investigating the emergence of an
861 undergraduate community through dynamic, multidimensional social network data, *Social*
862 *Networks*, Volume 65, 2021, Pages 71-84, ISSN 0378-8733,
863 <https://doi.org/10.1016/j.socnet.2020.11.006>.

864 Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* 393, 440–
865 442 (1998).

866 Weston, D., Hauck, K. & Amlôt, R. Infection prevention behaviour and infectious disease
867 modelling: a review of the literature and recommendations for the future. *BMC Public Health*
868 18, 336 (2018).

869 WHO. Global technical strategy for malaria 2016–2030. Geneva: World Health Organization
870 (2015).

871 WHO. World malaria report 2022. Geneva: World Health Organization; 2022. Licence: CC BY-
872 NC-SA 3.0 IGO

873 Wilson, A.L., Chen-Hussey, V., Logan, J.G., Lindsay, S.W. Are topical insect repellents effective
874 against malaria in endemic populations? A systematic review and meta-analysis. *Malar J.* 21:
875 446 (2014). doi: 10.1186/1475-2875-13-446.