



ELSEVIER

Contents lists available at ScienceDirect

Epidemics

journal homepage: [www.elsevier.com/locate/epidemics](http://www.elsevier.com/locate/epidemics)

# GEOFIL: A spatially-explicit agent-based modelling framework for predicting the long-term transmission dynamics of lymphatic filariasis in American Samoa

Zhijing Xu<sup>a,\*</sup>, Patricia M. Graves<sup>b</sup>, Colleen L. Lau<sup>a</sup>, Archie Clements<sup>c</sup>, Nicholas Geard<sup>d,e,f</sup>, Kathryn Glass<sup>a</sup>

<sup>a</sup> Research School of Population Health, The Australian National University, Australia

<sup>b</sup> College of Public Health, Medical and Veterinary Sciences, Division of Tropical Health and Medicine, James Cook University, Australia

<sup>c</sup> Faculty of Health Sciences, Curtin University, Australia

<sup>d</sup> School of Computing and Information Systems, The University of Melbourne, Australia

<sup>e</sup> The Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Australia

<sup>f</sup> Melbourne School of Population and Global Health, The University of Melbourne, Australia

## ARTICLE INFO

### Keywords:

Lymphatic filariasis  
Agent-based modelling  
Commuting networks  
Spatial heterogeneity  
Disease dynamics  
Vector-borne diseases

## ABSTRACT

In this study, a spatially-explicit agent-based modelling framework GEOFIL was developed to predict lymphatic filariasis (LF) transmission dynamics in American Samoa. GEOFIL included individual-level information on age, gender, disease status, household location, household members, workplace/school location and colleagues/schoolmates at each time step during the simulation. In American Samoa, annual mass drug administration from 2000 to 2006 successfully reduced LF prevalence dramatically. However, GEOFIL predicted continual increase in microfilaraemia prevalence in the absence of further intervention. Evidence from seroprevalence and transmission assessment surveys conducted from 2010 to 2016 indicated a resurgence of LF in American Samoa, corroborating GEOFIL's predictions. The microfilaraemia and antigenaemia prevalence in 6-7-yo children were much lower than in the overall population. Mosquito biting rates were found to be a critical determinant of infection risk. Transmission hotspots are likely to disappear with lower biting rates. GEOFIL highlights current knowledge gaps, such as data on mosquito abundance, biting rates and within-host parasite dynamics, which are important for improving the accuracy of model predictions.

## 1. Introduction

Lymphatic filariasis (LF) is a parasitic vector-borne disease targeted for global elimination. In American Samoa, the disease is predominantly transmitted by several mosquito species, including the highly efficient day-biting *Aedes polynesiensis* and night-biting *Aedes samoanus*. Larval worms develop over a period of up to one year into adult worms in the lymphatic system (Taylor et al., 2010). The adult worm can be reproductively active for 4–6 years (Ottesen, 2006), producing millions of microfilariae (Mf) that circulate in the blood and infect mosquitoes during biting events. The Mf then develop into L3 larvae in mosquitoes, which can be transmitted to humans in subsequent biting events. A person infected with LF may be asymptomatic but infectious to others for many years (Dreyer et al., 2000; Nutman, 2013).

Under the Pacific Programme to Eliminate LF (PacELF) initiated in

1999, seven rounds of mass drug administration (MDA) were undertaken in American Samoa from 2000 to 2006 (Lau et al., 2014; Ichimori and Graves, 2017). The WHO recommends post-MDA surveillance using transmission assessment surveys (TAS), which use critical cut-off values of numbers of antigen-positive children aged 6–7 years to determine whether transmission has been likely interrupted in defined evaluation units (World Health Organization, 2011). American Samoa passed the TAS 1 and TAS 2 in 2011 and 2015 (Won et al., 2018) but failed TAS 3 in 2016 (Sheel et al., 2018). Recent research evidence from both human and entomological surveys also suggests that transmission is still occurring (Lau et al., 2014; Schmaedick et al., 2014; Lau et al., 2016, 2017). Persistent transmission in American Samoa may be due to one or more of the following: 1) ineffective or incomplete coverage of the MDA program, either overall or in certain age or gender groups; 2) effective MDA in only parts of the islands, resulting in residual areas of ongoing transmission and subsequent spread to other areas; 3) standard TAS

\* Corresponding author at: Level 7, Tower 2, 205 Queen Street, Auckland CBD, 1010, New Zealand.

E-mail address: [xuzhijing521@gmail.com](mailto:xuzhijing521@gmail.com) (Z. Xu).

<https://doi.org/10.1016/j.epidem.2018.12.003>

Received 3 August 2018; Received in revised form 22 December 2018; Accepted 28 December 2018

Available online 29 December 2018

1755-4365/© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

target thresholds are not low enough for interruption of transmission in this context; and 4) reintroduction of parasites from external sources, e.g. from neighbouring LF-endemic countries.

Mathematical and computational models can assist researchers and program managers to better understand transmission dynamics. The three key models of LF, EPFIL (Chan et al., 1998; Norman et al., 2000), LYMFASIM (Plaisier et al., 1998) and TRANSFIL (Irvine et al., 2015), have been widely used to assess the prospects of elimination (Stolk et al., 2003; Michael et al., 2006; Stolk et al., 2013; Stone et al., 2014; Smith et al., 2017). However, none of the current models directly addresses *Aedes*-type transmission in the Pacific, and spatial aspects of transmission were not considered in their modelling frameworks (Stolk et al., 2015). This has restricted the understanding of LF transmission dynamics and spatial clustering of infections, especially at low prevalence. Furthermore, recent studies have raised concerns about the recommended target threshold ( $< 1\%$  antigenaemia, Ag) for post-MDA TAS (Lau et al., 2014; Harris and Wiegand, 2017), the utility of school-based surveys of young children versus community-based surveys for determining ongoing transmission (Sheel et al., 2018; Harris and Wiegand, 2017), and the expected size of residual hotspots of transmission (Lau et al., 2014).

In this study, a spatially-explicit agent-based modelling framework (GEOFIL) was developed using comprehensive data sources in American Samoa, where LF is diurnally subperiodic (Mf always present in the blood, but at greater densities in the day time) and transmitted mainly by both day and night biting *Aedes* mosquitoes. GEOFIL was characterized by a statistically realistic population, long-term population dynamics (including demographic changes), high-resolution geographic location information (American Samoa Department of Commerce, 2018), extensive daily commuting networks (Xu et al., 2018) and a spatially heterogeneous risk of infection with higher risk centered around locations of infectious individuals. GEOFIL was designed to predict the infection status of every individual in the synthetic population, tracking information such as age, gender, household location, household membership, workplace/school locations and colleagues/schoolmates. GEOFIL provides not only valuable insights into LF transmission dynamics, but is also a highly flexible framework to investigate potential intervention strategies, including country-wide MDA, targeted test and treat programs for high risk individuals or groups, the effectiveness of snowball surveillance (e.g. testing family members and near neighbours of infected people), and vector control programs. With moderate adjustment, the modelling framework can also be applied to a variety of infectious diseases, especially those with long incubation and infectious periods, or diseases with spatial heterogeneity in risk.

## 2. Material and methods

### 2.1. Population dynamics

The population generation algorithm is based on previous work to develop a synthetic population in American Samoa (Xu et al., 2017), with some improvements to make the algorithm more realistic and flexible. Instead of building households based on individuals in the synthetic population, households were generated based on family units in this study. A family unit is composed of a couple or a single (unmarried or divorced) adult, with possible underage children ( $< 15$  years) and unmarried adult children. The large household size identified in survey data (Xu et al., 2017) indicates that some households are likely to be composed of multiple family units in American Samoa. Our family unit approach allows us to optimize the family and household structure for the synthetic population. A second improvement on our prior work is related to the age structure of the synthetic population. Instead of assuming uniformly distributed individuals in each age group, the new algorithm uses kernel smoothing to produce a more realistic population structure, with smooth transition between

individual ages. Furthermore, household sizes were sampled from a zero-truncated Poisson (ZTP) distribution. The ZTP parameter was determined using maximum-likelihood estimation to fit to survey data. The application of ZTP makes the synthetic population generation algorithm more applicable to other census populations with limited household details.

The model population is characterized by major dynamic processes including population renewal (birth/death), couple formation and separation, relocation within American Samoa, and migration. Age-specific fertility rates were based on the projections in the International Database program (IDB) from the United States Census Bureau (The United States Census Bureau, 2018), while sex-specific mortality rates were based on the life table in the 2015 American Samoa Statistical Yearbook (American Samoa Department of Commerce, 2015), assuming an annual mortality improvement of 1%. Existing couples separated with a fixed small probability each year, with the female partner and any underage children assumed to move to a new household. The major change from the previous model is related to the movement of individuals between households. A large household with multiple family units can now divide into two smaller households with a probability depending on both the household size and household stability. Each year, a household experiencing an expansion whose number of members was more than the observed average household size ( $\Lambda_h = \lambda/(1 - e^{-\lambda})$  from the survey data) was deemed to be unstable. An unstable household was able to fracture into two households in accordance with a given probability. To determine the fracture probability, a second threshold ( $\Theta_h$ ) was defined as the minimum household size with which, in case of a fracture, there exist two generated households whose joint probability is larger than the probability of the original household according to the ZTP. The household fracture probability ( $q_h$ ) was defined as:  $q_h = 0.2$ ,  $\Lambda_h \leq s_h < \Theta_h$ ;  $q_h = 0.8$ ,  $s_h \geq \Theta_h$ , where  $s_h$  is the size of household  $h$ . If the size of the largest family unit ( $s_u$ ) of the household was within a reasonable range ( $1/3 \leq \max\{s_u\}/s_h < 2/3$ ), the largest family unit was assumed to move to an empty residential building to form a new household if the household fractured with probability  $q_h$ .

### 2.2. Employment and school attendance

The cultural relationship between the Samoan islands implies many similarities between American Samoa and Samoa. As the labor force participation rates (LFPRs) by age group in American Samoa were not available, the LFPRs (Samoa Bureau of Statistics (SBS) and Ministry of Commerce Industry and Labour (MCIL), 2014) of Samoa were used to estimate the employment ratio in each age group. However, American Samoa has stronger policies on ancestral land rights and fewer people work in the agricultural industry. Therefore, instead of the overall LFPRs of Samoa, the LFPRs from urban areas in Samoa (see Supplementary Fig. S1 online) were used.

For any individual  $i$  of age, let  $p_a$  be the LFPR at  $a$  and  $\phi_a$  be the employment status at  $a$  ( $\phi_a = 1$  employed, otherwise 0). The probability of individual  $i$  being employed at age  $a$  given the employment status at age  $a - 1$  ( $p(\phi_a \phi_{a-1})$ ) is given by:

- if  $p_a > p_{a-1}$ :  $p(\phi_a = 1\phi_{a-1} = 1) = 1$ ,  $p(\phi_a = 1\phi_{a-1} = 0) = (p_a - p_{a-1})/(1 - p_{a-1})$ ;
- if  $p_a < p_{a-1}$ :  $p(\phi_a = 0\phi_{a-1} = 0) = 1$ ,  $p(\phi_a = 0\phi_{a-1} = 1) = (p_{a-1} - p_a)/(1 - p_{a-1})$ .

The schooling system provides compulsory elementary (Grade 1–8) and secondary (Grade 9–12) school education in American Samoa. However, there is still a small number of children who leave secondary school and participate in the labor force. The American Samoa Community College (ASCC) offers some two-year programs to secondary school graduates. In this model, several assumptions were made about the school attendance:

- individuals aged 6–13 years would enrol at the nearest elementary school (based on Euclidean distance);
- individuals aged 14–17 years who were not employed would enrol at the nearest secondary school (based on Euclidean distance);
- individuals aged 18–19 years who were not employed would enrol at the ASCC.

### 2.3. Commuting networks

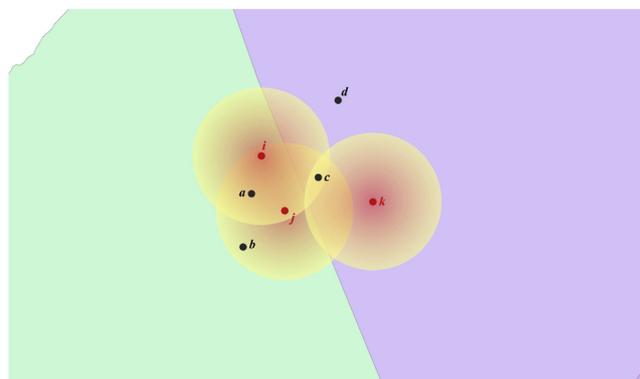
A radiation model (Simini et al., 2012) was used to predict daily commuting patterns. The average flux  $T_{ij}$  from location  $i$  to location  $j$  is given by  $\langle T_{ij} \rangle = T_i m_i n_j / ((m_i + s_{ij})(m_i + n_j + s_{ij}))$ , where  $m_i$  and  $n_j$  are the population of location  $i$  and  $j$ ;  $s_{ij}$  is the total population in the circle of radius  $r_{ij}$  centered at  $i$  (excluding the source and destination population); and  $T_i = \sum_{j \neq i} T_{ij}$ . A distance matrix  $D_{n \times n}$  (where  $n$  is the total number of locations) is required to construct radiation models. Each element  $d_{ij}$  represents the distance from location  $i$  to, where  $d_{ii} = 0$  and  $d_{ij} = d_{ji}$ . Either Euclidean or road distance can be used to formulate the distance matrix. In this study, road distance was preferred considering the mountainous landscape of the main island of American Samoa (Tutuila). An embedded problem of the radiation model is that the predicted commuting networks miss some important commuting hubs with small resident populations. For example, Atu'u is a small village where the largest non-government employer on the island (a tuna cannery) is located. In recent years, an average of 17.6% of the total employed population worked in the cannery (American Samoa Department of Commerce, 2015), greatly in excess of the employment capacity predicted by the local resident population. The number of cannery workers who resided in each village was found to be proportional to village size (Xu et al., 2018). The radiation model was therefore revised to calibrate the commuting flux to Atu'u to match the total employment in the cannery by randomly assigning workers to Atu'u proportional to the population of each village.

### 2.4. Risk of infection

Due to the inefficiency of transmission of LF (Hairston and de Meillon, 1968), the likelihood of being infected after one mosquito bite is extremely small. The household, workplace and school are the three most important locations where transmission could occur. Although the time spent in the workplace/school was usually less than that in the household, the mosquito biting rates were assumed to be the approximately the same (see Table 1). This assumption is made for two reasons: firstly, the day-biting mosquitoes are highly efficient and there is potential bed-net protection (Irvine et al., 2018) during the night; secondly, there are no available data on the relative biting efficiency between the day-biting and night-biting mosquitoes in American Samoa,

**Table 1**  
Model Parameters for Estimating Risk of Infection.

Parameter	Value	Description	References
$d_{max}$	100 meters	Maximum flight range of <i>Aedes</i> mosquitoes	(Lau et al., 2016; Jachowski, 1954; Hapairai et al., 2013)
$l_m$	13 days	Length of extrinsic incubation period	(Paily et al., 2009)
$l_p$	6–12 months	Length of prepatent period	(Ottesen, 2006)
$l_i$	4–6 years	Length of infectious period	(Ottesen, 2006)
$\gamma$	$-\lambda \cdot v (0.13)/75$	Daily death rate of mated worm in prepatent period	(Hairston and de Meillon, 1968)
$s_m$	0.6	Mosquito survival probability of each feeding cycle (3 days)	(Graves et al., 1990)
$p_{L3}$	$s_m^{lm/3}$	Probability of mosquitoes surviving through the extrinsic incubation period	
$p_{if}$	0.3881	Probability that mosquitoes which survive through the extrinsic incubation period are infective	(Krishnamoorthy et al., 2004; Erickson et al., 2013)
$p_b$	0.1412	Probability of presence of mated worms due to an infective mosquito bite	(Hairston and de Meillon, 1968)
$c$	0.0–1.0	Relative mosquito exposure of individuals aged $\leq 15$ years. Kernel smoothed based on: 0.25 (0–4 yro), 0.75 (5–15 yro).	(Stone et al., 2015)
$b_1$	70	Number of effective mosquito bites per person during working hours	(Jachowski, 1954; Ramalingam, 1968; Hapairai et al., 2015)
$b_2$	70	Number of effective mosquito bites per person during off-work hours	



**Fig. 1.** Risk of infection centered on the locations of infectious individuals. The red dots ( $i$ ,  $j$  and  $k$ ) represent the household/workplace/school locations of infectious individuals. A cone risk area is shaped by each infectious individual. In this case, susceptible individuals at location  $c$  are at risk from  $i$ ,  $j$  and  $k$ ; susceptible individuals at location  $a$  are at risk from  $i$  and  $j$ ; susceptible individuals at location  $b$  only are at risk from  $j$ ; while susceptible individuals at location  $d$  have no risk of infection (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

and thus a parsimonious assumption is reasonable. Sensitivity analysis was also used to investigate the impact of varying mosquito biting rates.

For each infectious individual, mosquitoes can take in Mf during a blood meal and eventually become infective if they survive the extrinsic incubation period ( $l_m$ ). Infective mosquitoes disperse to the neighborhood area based on their flight range. Therefore, only individuals within the maximum flight distance ( $d_{max}$ ) of mosquitoes centered at the location of the infectious individual are at risk of infection (see Fig. 1). To estimate the distribution of infective mosquitoes, a linear dispersal factor ( $f_d$ ) was defined by:  $f_d = 1 - d/d_{max}$ . For LF in American Samoa, the area at risk is relatively small due to the short flight range of ~100 m for *Aedes* mosquitoes (Lau et al., 2016; Jachowski, 1954; Hapairai et al., 2013).

For any location  $j$ , the prevalence of infective mosquitoes ( $r_j$ ) is the weighted average of the prevalence of infectious individuals in the neighborhood area centered at  $j$ :

$$r_j = \frac{\sum_{k \in L_j \cup \{j\}} \left( \left( \sum_{i \in M_k} c_i I_i \right) / N_k \times f_d \right)}{\sum_{k \in L_j \cup \{j\}} f_d} \times p_{L3} \times p_{if}$$

where  $L_j$  is a location within  $d_{max}$  of  $j$ ;  $M_k$  is the set of members at  $k$ ;  $c_i$  is the relative exposure to mosquitoes of individual  $i$ ;  $I_i$  is an indicator function (1 if infectious, otherwise 0); and  $N_k$  is the number of total

individuals at  $k$ ;  $p_{L3}$  is the probability of mosquitoes surviving through the extrinsic incubation period; and  $p_{if}$  is the probability that mosquitoes that survive through the extrinsic incubation period are infected (see Table 1).

The rate of infection during working hours for an individual  $i$  at location  $j$  is given by:  $w_{ij} = b_{ij} \times c_i \times r_j \times p_b$ , where  $b_{ij}$  is the number of effective mosquito bites at  $j$  during working hours;  $c_i$  is the age-dependent relative mosquito exposure of  $i$ ;  $r_j$  is the prevalence of infective mosquitoes at  $j$ ; and  $p_b$  is the probability that an infective mosquito bite will lead to a mated pair of worms. Correspondingly, the rate of infection during off-work hours is  $w_{ij} = b_{2j} \times c_i \times r_j \times p_b$ . Due to insufficient information about the spatial heterogeneity in mosquito biting rates, we assumed the same rate throughout the island (e.g., simply  $b_1$  and  $b_2$ ). However, we retain  $j$  in the model for use in future scenarios when this information becomes available.

Hairston and de Meillon (1968) argued that the accumulation of parasites in the human body was unlikely and  $p_b$  was estimated to be 0.0325. However, our estimation indicates the rate of receiving infective mosquito bites is much higher in American Samoa (Schmaedick et al., 2014; Jachowski, 1954; Ramalingam, 1968; Hapairai et al., 2015) compared to Yangon (formerly Rangoon) Hairston and de Meillon (1968). People would most likely contract the male and female worms from multiple bites instead of one single bite, and  $p_b$  was re-estimated. Based on the data in Hairston and de Meillon (1968),  $p_b$  was estimated to be between 0.0837 (probability of receiving both sexes of worms in an infective bite) and 0.1987 (probability of receiving a single sex of worm in an infective bite). With no further knowledge,  $p_b$  was assumed to be the average of the two probabilities in our model.

The mated worms will circulate in the lymphatic and blood vessel system and the individual will finally become infectious if the mated worms survive through the prepatent period ( $l_p$ ) with probability  $p = e^{-\gamma \times l_p}$  and produce microfilariae. The length of the infectious period was assumed to be 4–6 years (Ottesen, 2006). The parameters for estimating the risk of infection are summarized in Table 1.

## 2.5. Mf and Ag prevalence

Only Mf positive individuals contribute to transmission of LF in the population. However, it is more difficult to detect Mf than to detect antigen and antibody, especially at low Mf density (Southgate, 1992). Consequently, rapid diagnostic tests for circulating filarial antigen are the primary tools used in TAS (Gass et al., 2012). However, as there are still many unknowns about how long the antigen remains detectable after the Mf positive stage, Ag prevalence was not explicitly modelled in our framework. To validate our model based on the Ag prevalence in the surveys, the ratio between the Ag prevalence and Mf prevalence was estimated. The ratio is affected by the level of Ag/Mf prevalence and whether MDA had been conducted. A recent study of LF in Papua New Guinea (Berg Soto et al., 2018) indicated that an exponential function can be used to interpret the ratio between the two prevalence values. Based on the survey data in 2007 (Coutts et al., 2017), 2014 (Lau et al., 2017) and 2016 (Sheel et al., 2018), we adopted an exponential function  $R = 5.8464 \times e^{-0.1537T} + 2.77$ , where  $R$  is the estimated ratio of Ag/Mf prevalence;  $T$  is the number of years since the last MDA to estimate the relationship between Ag and Mf prevalence (see Supplementary Fig. S2 online).

## 2.6. Transmission scenarios

Two scenarios were developed based on the observed Ag prevalence by villages in a 2010 seroprevalence survey (Lau et al., 2014). For transmission scenario A, the rate ratio was estimated to determine if the Ag prevalence of a village differed from the overall Ag prevalence of all other villages at 1% significance level. In case of significant difference, the Ag prevalence of that village was used to initialize the village population; otherwise, the village population was initialized based on the

overall Ag prevalence. For transmission scenario B, all the villages were simply assumed to follow the observed Ag prevalence by villages in the 2010 survey. Due to persistent transmission prior to 2010, a considerable proportion of people may have been in the prepatent stage in 2010. The prevalence of prepatent individuals was assumed to be one-sixth of the Mf prevalence considering the relative length of the prepatent stage and infectious stage (see Table 1).

## 3. Results

### 3.1. Population dynamics

The simulated population of American Samoa decreased from 2010 to 2030 by an average annual rate of 1.27% (Supplementary Fig. S3(a) online), based on the census population (American Samoa Department of Commerce, 2015), mortality rates (American Samoa Department of Commerce, 2015) and age-specific fertility rates (The United States Census Bureau, 2018). The simulated population was found to decrease more slowly than in the estimation from the United States Census Bureau (average annual rate – 1.58%) (The United States Census Bureau, 2018). Population ageing was identified in the simulation. The simulated age distribution was consistent with the observed trend in the last two population censuses. More details on the simulated population dynamics are given in Supplementary Fig. S3 online and our previous work (American Samoa Department of Commerce, 2015).

### 3.2. Employment and school attendance

Using our model, the employment ratios of males and females aged 15–69 years were estimated to be 56.5% and 36.3%. The overall employment ratio was 46.4% in the model, close to 47.8% reported in the census (American Samoa Department of Commerce, 2015). The total number of employed persons was estimated to be 15,510, in line with census data (16,616) (American Samoa Department of Commerce, 2015). Based on school attendance assumptions, the total number of students was estimated to be 16,785, with 10,034 primary students (Year 1–8), 4876 secondary students (Year 9–12) and 1875 college students (student ratio = 5.4:2.6:1.0). Student numbers and the student ratio are consistent with the census data (total 17,885, student ratio = 4.9:2.1:1.0) (American Samoa Department of Commerce, 2015).

### 3.3. Commuting networks

The commuting network from the residence villages to the work-place villages was previously investigated and analyzed (Xu et al., 2018), indicating that workers commuted daily across the whole of the main island of Tutuila, with work hubs drawing from villages across the island. The daily commuting network for adult workers surveyed in 2010 is given in Fig. 2(a). The population covered by the network is 51,864, about 93.4% of the total population in American Samoa. The work hubs were villages with the largest populations, such as Tafuna, Pago Pago and Fagatogo, plus the village of Atu'u (where a tuna cannery, the largest non-government employer, is located). In Fig. 2(b), the major commuting routes for adult workers were predicted by a revised radiation model using road distances extracted from Google Map API (Anon, 2018). The simulated commuting network captured all of the important commuting hubs in the survey data. Although the simulated commuting network covers all villages in the island, commuting flux among pairs of villages was highly heterogeneous.

### 3.4. Transmission scenario A

The Ag prevalence in Fagali'i was found to be significantly higher than other villages ( $RR = 10.13$ ,  $p < 0.01$ ). Therefore, at the beginning of the simulation (in 2010), Mf prevalence in the adult population

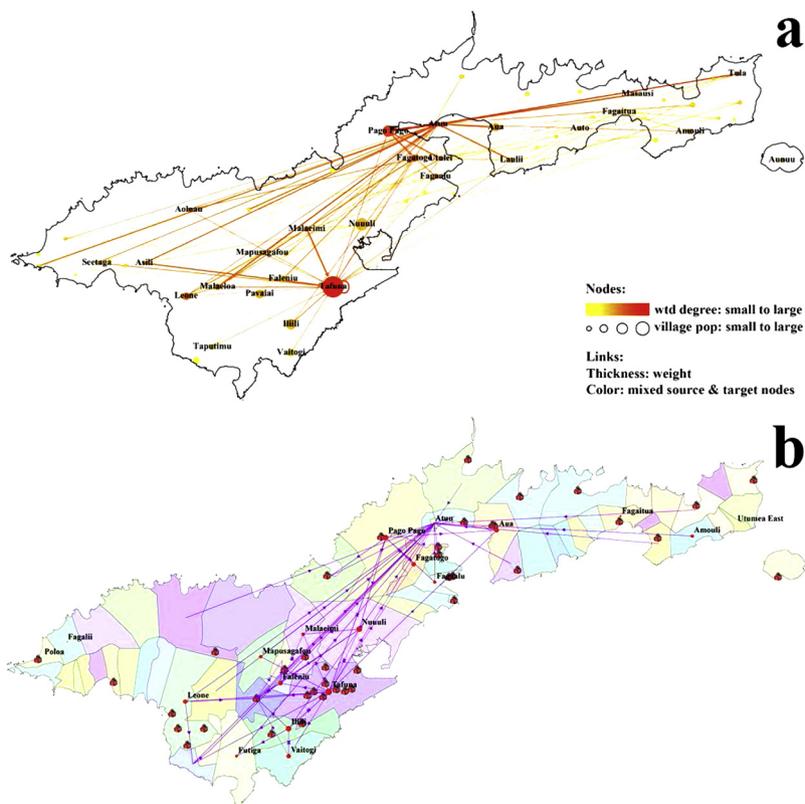


Fig. 2. Commuting network in 2010: (a) based on 2010 survey (reproduced from Ref. Xu et al. (2018)); (b) simulated with radiation model (top 60 major routes only). The directed lines represent commuting routes from residence villages to workplace villages.

(≥15 years) was assumed to be 4.76% in Fagali'i and 0.47% in other villages, based on the observed Ag prevalence in the 2010 survey (Lau et al., 2014) and the estimated ratio of Ag/Mf prevalence of 6.46. Assumptions were made for the Ag prevalence in children as there were no available data. The Mf prevalence in 8–14 year-old (yo) children was assumed to be half of the adults, according to the observed lower prevalence in children in a 2014 survey (Lau et al., 2017), and we assumed no Mf positive individuals in 0–7 yo children.

The simulated transmission dynamics over 2010–2030 were shown in Fig. 3(a). The overall Mf prevalence was 0.36% (0.31–0.41%) in 2010 and the corresponding Ag prevalence was 2.33% (2.00–2.65%). With no interventions, the overall Mf prevalence would increase to 8.58% in 2030. The Mf prevalence in 6-7-yo children was much lower than the overall prevalence. The Mf prevalence in the two suspected hotspots (Fagali'i and Ili'ili/Vaitogi/Futiga) identified in the 2010 survey (Lau et al., 2014, 2017) was shown in Fig. 3(b). The Mf prevalence in Fagali'i was found to be much higher than the overall prevalence, but the prevalence in Ili'ili/Vaitogi/Futiga did not differ from other villages. The prevalence of infective mosquitoes at specific locations was explicitly calculated in GEOFIL. The Supplementary Fig S4

showed the prevalence of infective mosquitoes in each village. The patterns agreed well with results of a previous entomological survey (Schmaedick et al., 2014).

In Fig. 4(a), the simulated Ag prevalence in adults was explored. The simulated Ag prevalence in adults was 3.21% (90% range: 2.75–3.65%) in 2010 and 5.07% (3.78–6.11%) in 2016. The observed data indicated the Ag prevalence was 3.2% (95% CI: 0.6–4.7%) in 2010 (Lau et al., 2014) and 6.2% (4.5–8.6%) in 2016 (Sheel et al., 2018). In Fig. 4(b), the simulated Ag prevalence in 6-7-yo children was compared to the data from TAS. The Ag prevalence in 6–7 yo was projected to be 0.19% (0–0.47%) in 2011, 1.01% (0.40–1.88%) in 2015 and 1.05% (0.38–1.93%) in 2016. The TAS suggested the prevalence to be 0.20% (0–0.8%) in 2011 (Won et al., 2018), 0.1% (0–0.3%) in 2015 (Won et al., 2018) and 0.7% (0.3–1.8%) in 2016 (Sheel et al., 2018). In Fig. 4(c), the simulated Ag prevalence in Fagali'i was investigated. The Ag prevalence in Fagali'i was simulated to be 22.98% (10.47–39.52%) in 2010 and 20.44% (7.07–34.65%) in 2014. The observed data indicated the Ag prevalence in Fagali'i was 30.8% (9.1–61.4%) in 2010 (Lau et al., 2014) and 17.2% (8.6–29.4%) in 2014 (Lau et al., 2017). In Fig. 4(d), the simulated Ag prevalence in Ili'ili/Vaitogi/Futiga was

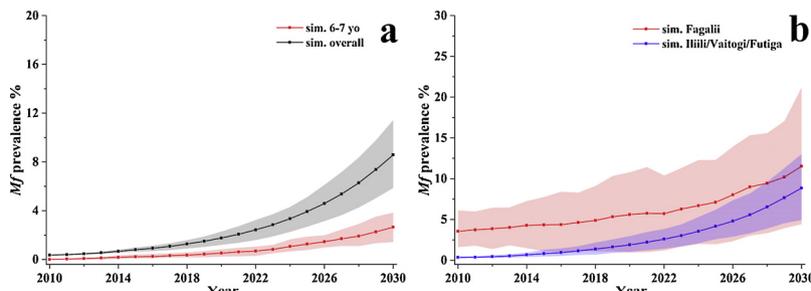
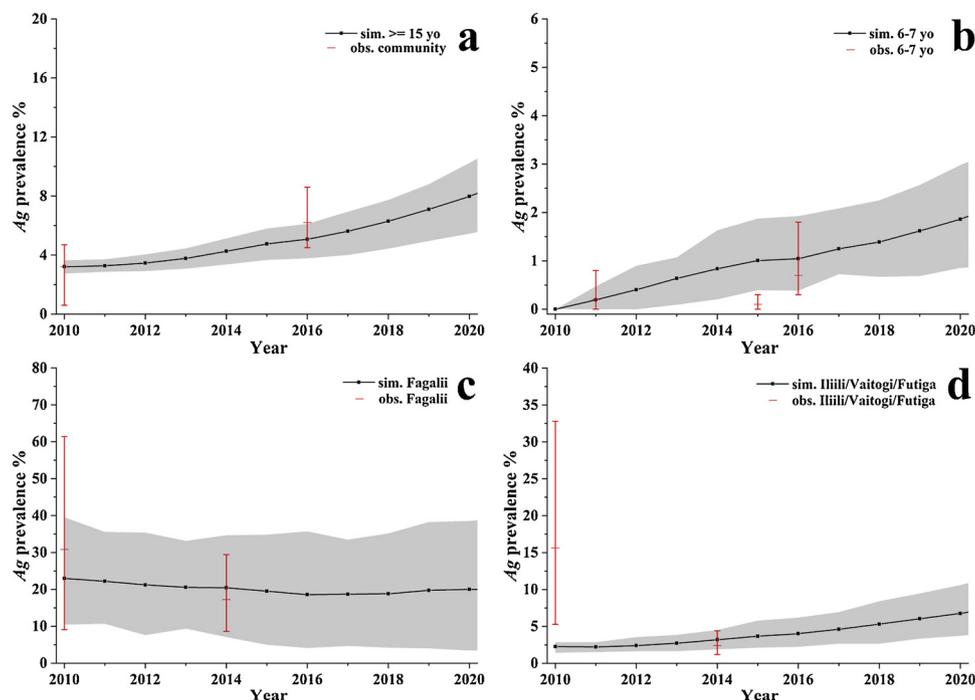


Fig. 3. Transmission scenario A: simulated (sim.) Mf prevalence in: (a) 6-7-yo children and the overall population; (b) two suspected transmission hotspots. Results were based on 50 simulations, with 90% range indicated by the bars.



**Fig. 4.** Transmission scenario A: simulated (sim.) and observed (obs.) Ag prevalence in: (a) adults ( $\geq 15$  yo); (b) 6–7 yo school children; (c) Fagali'i; (d) Ili'ili/Vaitogi/Futiga. Results were based on 50 simulations, with 90% range indicated by the bars. For the observations, the bars indicated the 95% confidence interval.

compared to the observed data. The simulation indicated the Ag prevalence was 2.27% (1.43–2.86%) in 2010 and 3.20% (1.89–4.52%) in 2014. The survey suggested the prevalence was 15.6% (5.3–32.8%) in 2010 (Lau et al., 2014) and 2.4% (1.2–4.4%) in 2014 (Lau et al., 2017).

The dynamics and evolution of the transmission risk map was shown in Fig. 5(a)–(c). The Mf prevalence was found to be highly heterogeneous among villages throughout the simulation period. To show the dynamics of LF transmission, a video of locations of households with Mf positive individuals every 30 days is available in the supplementary information online.

### 3.5. Transmission scenario B

All villages were assumed to follow the observed Ag prevalence by village in the 2010 survey (Lau et al., 2014) in transmission scenario B. Mf positive individuals were present in 17 villages (see Fig. 5(d)). The overall Mf prevalence was 0.37% (0.33–0.41%) in 2010 and the corresponding Ag prevalence was 2.39% (2.13–2.62%). The simulated transmission dynamics over 2010–2030 were shown in the Supplementary Fig. S5. In the Supplementary Fig. 5(c)–(f), the simulated Ag prevalence was compared to the observed data from surveys conducted from 2010 to 2016 (Lau et al., 2014; Won et al., 2018; Sheel et al., 2018; Lau et al., 2017). All simulated Ag prevalence agreed well with the data, except for the Ag prevalence in the suspected transmission hotspot Ili'ili/Vaitogi/Futiga, which suggests that Scenario A is more likely compared to Scenario B. In contrast to scenario A, Ili'ili/Vaitogi/Futiga was assumed to have a higher initial transmission setting in scenario B. The results suggested that Ili'ili/Vaitogi/Futiga may not have higher prevalence than the overall population in 2010, implying that the contiguous villages of Ili'ili/Vaitogi/Futiga were likely not a transmission hotspot, in agreement with data from the 2014 survey (Lau et al., 2017). Fig. 5(d)–(f) showed the dynamics and evolution of the transmission risks based on scenario B.

### 3.6. Sensitivity analysis

Sensitivity tests were conducted based on the transmission scenario

A. The results showed the Mf prevalence was very sensitive to the mosquito biting rates (see Fig. 6). With lower biting rates during both working and off-work hours, the overall Mf prevalence would increase at a much lower rate (Fig. 6(b)) and the hotspot Fagali'i was expected to disappear (Fig. 6(d)).

With no data on the relative mosquito biting rates during working and off-work hours, GEOFIL assumes identical biting rates during each period. Sensitivity analysis was conducted to investigate whether different mosquito biting rates during working and off-work hours would lead to distinct infection patterns in adults and children. The Supplementary Fig. S6 showed that mosquito biting rates during working and off-work hours have similar effects on the Mf prevalence in adults and children, thus we cannot infer transmission patterns based on the prevalence data.

## 4. Discussion

In this study, a fully dynamic spatially-explicit agent-based modeling framework was developed to simulate LF transmission in American Samoa, based on a synthetic population, commuting networks and spatially heterogeneous risks of infection. The predicted commuting networks, based on the radiation model (Simini et al., 2012), along with road distances from Google Map API (Anon, 2018), were able to replicate the commuting patterns identified in an empirical study (Xu et al., 2018). Based on the seroprevalence in the 2010 survey (Lau et al., 2014), and assuming no further interventions (e.g. MDA) are conducted, GEOFIL predicted a continuously increasing prevalence of Mf in the human population from 2010 to 2030. The Mf prevalence was highly sensitive to mosquito biting rates. With lower biting rates, transmission hotspots were likely to disappear.

In 2015, TAS 2 was undertaken and Ag prevalence in 6–7 yo children was 0.1% (0–0.3%) (Won et al., 2018). However, GEOFIL estimated an Ag prevalence of 1.01% (0.40–1.88%) in this age group, i.e. American Samoa should have failed the 2015 TAS based on the simulation. The discordance between the TAS 2 results and model estimates could be explained by chance, sampling issues, lower sensitivity of detection of Ag in 6–7 yo children leading to poorer accuracy in the

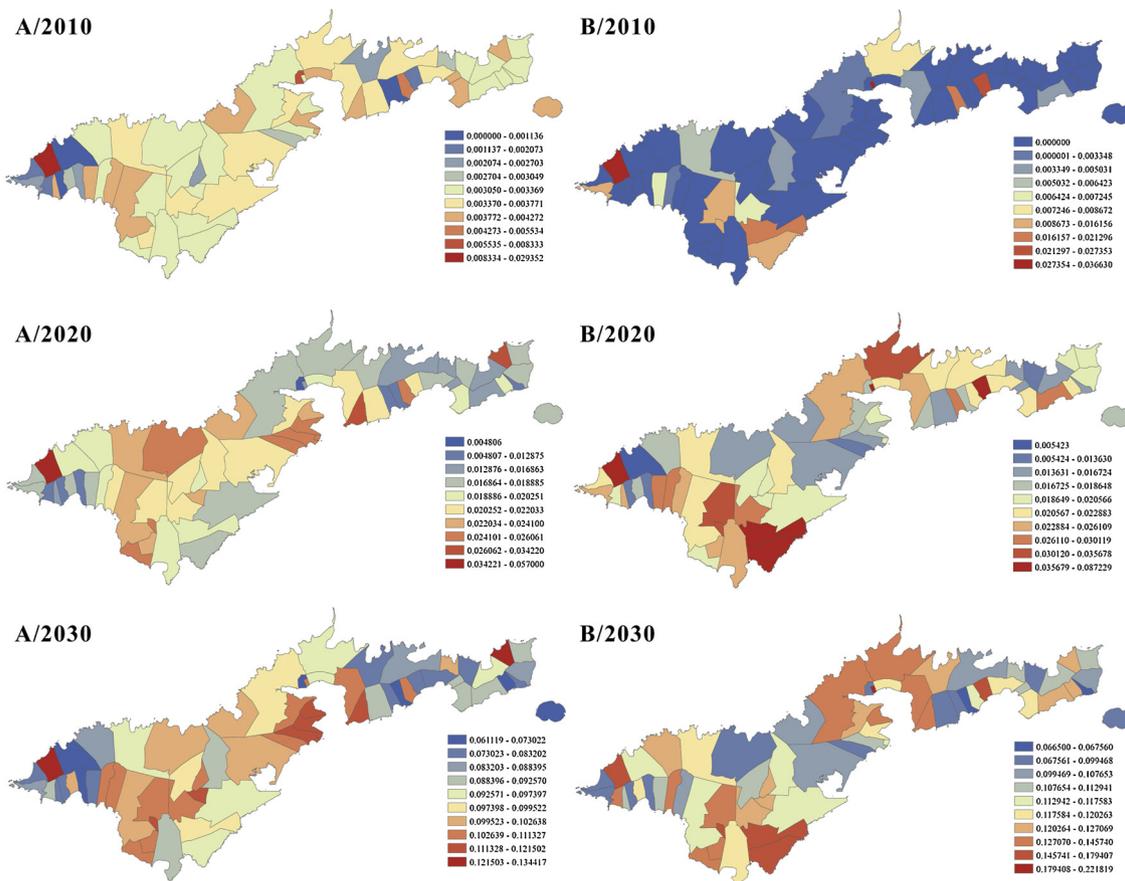


Fig. 5. Mf prevalence by villages in transmission scenario A/B in 2010, 2020 and 2030. Different color scales are used to make the spatial heterogeneity in Mf prevalence more clear. Results were based on 50 simulations.

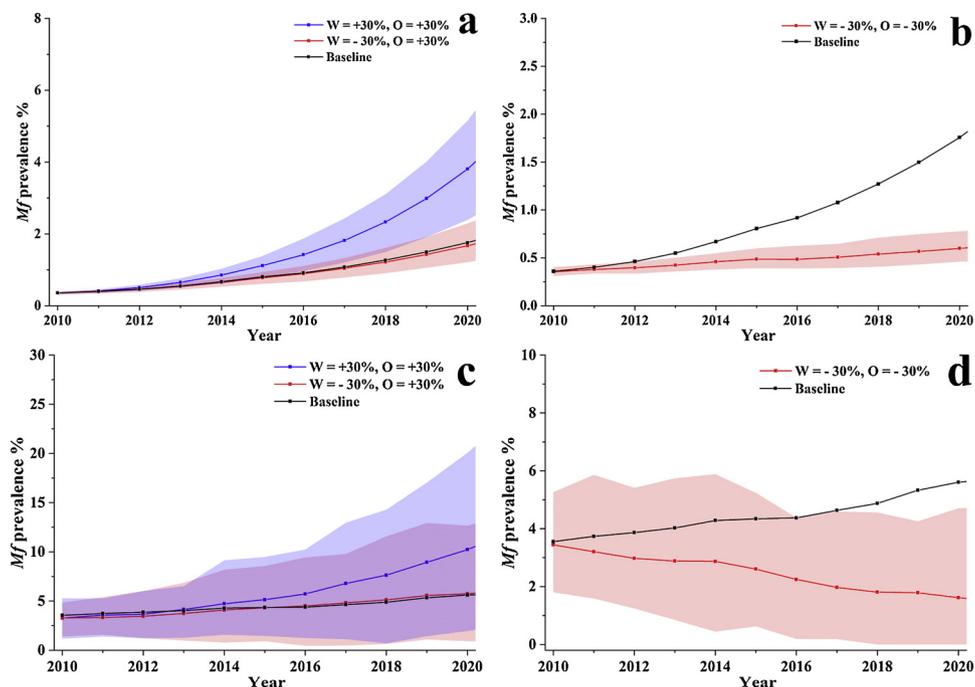


Fig. 6. Sensitivity analysis of mosquito biting rates during working (W) and off-work hours (O), simulated Mf prevalence of Scenario A in: (a)(b) the overall population; (c)(d) Fagali'i. Results were based on 50 simulations, with 90% range indicated by the bars.

prevalence estimates, or it may signal that our model needs further refinement.

As the first spatially explicit agent-based model of LF, we modelled

the prevalence of Mf positive individuals in the population, rather than the average worm load as in other models (Chan et al., 1998; Norman et al., 2000; Plaisier et al., 1998). To fit the model output to Ag

prevalence in the surveys, we estimated Ag prevalence based on the Ag/Mf ratios estimated in a recent study using data from Papua New Guinea (Berg Soto et al., 2018). However, the ratio of Ag/Mf prevalence (see Supplementary Fig. S2 online) may vary between settings (e.g., when predicting pre-control Ag prevalence and trends during MDA). To estimate the risk of infection at any specific location, the prevalence of infective mosquitoes was explicitly estimated based on the weighted average of the prevalence of infectious individuals in the neighboring area. This approach allowed us to validate GEOFIL using mosquito trapping data from a separate study (Schmaedick et al., 2014). When estimating commuting patterns, we found that the radiation model (Simini et al., 2012) did not capture known commuting hubs without essential adjustments, highlighting limitations of population density in capturing all travel features for small population centers.

One of the primary differences between GEOFIL and other models (EPIFIL, LYMFASIM and TRANSFIL) is that the transmission probability in other models is dependent on the average Mf load in the population. In GEOFIL, on the other hand, the rate of infection only depends on mosquito exposure, prevalence of infective mosquitoes and probability of presence of mated worms due to an infective mosquito bite. However, none of these parameters are dependent on the Mf density in the human body. Therefore, GEOFIL may overestimate the transmission capacity, especially at low worm load (e.g. post MDA). The dependence between the transmission capacity and the Mf density needs to be further investigated before applying GEOFIL for MDA strategies.

There are limitations in the ability of GEOFIL to replicate some of the observations from the field surveys. Firstly, in the 2016 survey, the Ag prevalence in males was much higher than in females (Sheel et al., 2018). Possible reasons include males spending more time outdoors (resulting in higher exposure to mosquitoes), lower rates of participation in MDA (Liang et al., 2008), hormonal reasons, and immunological reasons (Brabin, 1990). Although a relative risk of exposure could be added, there are currently limited data on risk exposures to parameterize such a difference. Secondly, mosquito biting rates during working and off-work hours were based on mosquito trapping experiments in the 1950s and 1960s (Jachowski, 1954; Ramalingam, 1968) and a recent experiment in Samoa conducted in 2011 (Hapairai et al., 2015). The biting rates were assumed to be homogeneous throughout the island, given a lack of data. As simulations are quite sensitive to mosquito biting rates, vector abundance data are likely to improve the precision of spatial model predictions and explain the emergence of transmission hotspots. Thirdly, the population was projected long into the future. The population structure relies on projections of demographics from the U.S. Census Bureau, which may not be accurate. Fourthly, there are still many unknowns concerning LF parasite dynamics in the human body, especially the accumulation and mating of male and female worms in the lymphatic system. The probability that an infective bite leads to the presence of mated worms can only be estimated. Finally, there is intensive population mobility between American Samoa and Samoa (Xu et al., 2018). The importation risk of LF may be significant. However, due to limited data on the Mf prevalence in Samoa, importation is currently not included in GEOFIL.

In conclusion, we have developed a highly flexible and extendable spatially-explicit agent-based modelling framework. The flexible population layer makes it easy to change population configurations to investigate future LF disease burden under different demographic scenarios. The high-resolution spatial locations make it possible to investigate hotspot size and the risk of infection due to the presence of infectious individuals in the neighboring area. The modeling framework further highlights some important knowledge gaps, such as mosquito abundance data and knowledge of within-host parasite dynamics, which are important for improving the accuracy of LF transmission models.

## Declaration of conflicting interest

The authors declare that there is no conflict of interest.

## Data availability

Some access restrictions apply to the data underlying the findings. American Samoa has a very small population, and high resolution geo-referenced data would potentially allow identification of individuals and households, and breach confidentiality. The datasets generated during and/or analysed during the current study may be available from the corresponding author on reasonable request providing confidentiality restrictions are met.

## Acknowledgements

ZX is funded by a NHMRC Centre of Research Excellence. CLL was supported by an Australian National Health and Medical Research Council (NHMRC) Fellowship (1109035). We thank Wayne Melrose and Luke Becker, of the James Cook University WHO Collaborating Centre for Vector Borne and NTDs, for discussion about LF infection parameters and dynamics. We thank all those in American Samoa and elsewhere who conducted or participated in the human and entomological surveys that provided data for this study.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.epidem.2018.12.003>.

## References

- American Samoa Department of Commerce, 2015. American Samoa Statistical Yearbook. American Samoa Department of Commerce, 2018. American Samoa Department of Commerce. <http://doc.as.gov/resource-management/ascmp/gis/>. <https://developers.google.com/maps/>.
- Brabin, L., 1990. Sex differentials in susceptibility to lymphatic filariasis and implications for maternal child immunity. *Epidemiol. Infect.* 105 (2), 335–353.
- Chan, M.-S., et al., 1998. Epifil: a dynamic model of infection and disease in lymphatic filariasis. *Am. J. Trop. Med. Hyg.* 59 (4), 606–614.
- Coutts, S.P., et al., 2017. Prevalence and risk factors associated with lymphatic filariasis in American Samoa after mass drug administration. *Trop. Med. Health* 45 (1), 22.
- Dreyer, G., et al., 2000. Pathogenesis of lymphatic disease in Bancroftian filariasis: a clinical perspective. *Parasitol. Today* 16 (12), 544–548.
- Erickson, S.M., et al., 2013. Mosquito-parasite interactions can shape filariasis transmission dynamics and impact elimination programs. *PLoS Negl. Trop. Dis.* 7 (9), e2433.
- Gass, K., et al., 2012. A multicenter evaluation of diagnostic tools to define endpoints for programs to eliminate Bancroftian filariasis. *PLoS Negl. Trop. Dis.* 6 (1), e1479.
- Graves, P., et al., 1990. Estimation of anopheline survival rate, vectorial capacity and mosquito infection probability from malaria vector infection rates in villages near Madang, Papua New Guinea. *J. Appl. Ecol.* 134–147.
- Hairston, N.G., de Meillon, B., 1968. On the inefficiency of transmission of *Wuchereria bancrofti* from mosquito to human host. *Bull. World Health Organ.* 38 (6), 935.
- Hapairai, L.K., et al., 2013. Population studies of the filarial vector *Aedes polynesiensis* (Diptera: Culicidae) in two island settings of French Polynesia. *J. Med. Entomol.* 50 (5), 965–976.
- Hapairai, L.K., et al., 2015. Evaluation of traps and lures for mosquito vectors and xenomonitoring of *Wuchereria bancrofti* infection in a high prevalence Samoan Village. *Parasit. Vectors* 8 (1), 287.
- Harris, J.R., Wiegand, R.E., 2017. Detecting infection hotspots: modeling the surveillance challenge for elimination of lymphatic filariasis. *PLoS Negl. Trop. Dis.* 11 (5), e0005610.
- Ichimori, K., Graves, P.M., 2017. Overview of PacELF—the Pacific programme for the elimination of lymphatic filariasis. *Trop. Med. Health* 45 (1), 34.
- Irvine, M.A., et al., 2015. Modelling strategies to break transmission of lymphatic filariasis - aggregation, adherence and vector competence greatly alter elimination. *Parasit. Vectors* 8 (1), 547.
- Irvine, M.A., et al., 2018. Understanding heterogeneities in mosquito-bite exposure and infection distributions for the elimination of lymphatic filariasis. *Proc. R. Soc. B: Biol. Sci.* 285 (1871).
- Jachowski Jr, L., 1954. Filariasis in American Samoa. Y. bionomics of the principal vector, *Aedes polynesiensis* marks. *Am. J. Hyg.* 60 (2), 186–203.
- Krishnamoorthy, K., et al., 2004. Vector survival and parasite infection: the effect of *Wuchereria bancrofti* on its vector *Culex quinquefasciatus*. *Parasitology* 129 (Pt 1), 43–50.

- Lau, C.L., et al., 2014. Seroprevalence and spatial epidemiology of lymphatic filariasis in American Samoa after successful mass drug administration. *PLoS Negl. Trop. Dis.* 8 (11), e3297.
- Lau, C.L., et al., 2016. Lymphatic filariasis elimination in American Samoa: evaluation of molecular xenomonitoring as a surveillance tool in the Endgame. *PLoS Negl. Trop. Dis.* 10 (11), e0005108.
- Lau, C.L., et al., 2017. Detecting and confirming residual hotspots of lymphatic filariasis transmission in American Samoa 8 years after stopping mass drug administration. *PLoS Negl. Trop. Dis.* 11 (9), e0005914.
- Liang, J.L., et al., 2008. Impact of five annual rounds of mass drug administration with diethylcarbamazine and albendazole on *Wuchereria bancrofti* infection in American Samoa. *Am. J. Trop. Med. Hyg.* 78 (6), 924–928.
- Michael, E., et al., 2006. Mathematical models and lymphatic filariasis control: monitoring and evaluating interventions. *Trends Parasitol.* 22 (11), 529–535.
- Norman, R.A., et al., 2000. EPIFIL: the development of an age-structured model for describing the transmission dynamics and control of lymphatic filariasis. *Epidemiol. Infect.* 124 (3), 529–541.
- Nutman, T.B., 2013. Insights into the pathogenesis of disease in human lymphatic filariasis. *Lymphat. Res. Biol.* 11 (3), 144–148.
- Ottesen, E.A., 2006. Lymphatic filariasis: treatment, control and elimination. In: Molyneux, D.H. (Ed.), *Advances in Parasitology*. Academic Press, pp. 395–441.
- Paily, K.P., Hoti, S.L., Das, P.K., 2009. A review of the complexity of biology of lymphatic filarial parasites. *J. Parasit. Dis.* 33 (1), 3–12.
- Plaisier, A.P., et al., 1998. The LYMFASIM simulation program for modeling lymphatic filariasis and its control. *Methods Inf. Med.* 37 (1), 97–108.
- Ramalingam, S., 1968. The epidemiology of filarial transmission in Samoa and Tonga. *Ann. Trop. Med. Parasitol.* 62 (3), 305–324.
- Samoa Bureau of Statistics (SBS) and Ministry of Commerce Industry and Labour (MCIL), 2014. SAMOA 2012 Labour Force Survey Report.
- Schmaedick, M.A., et al., 2014. Molecular xenomonitoring using mosquitoes to map lymphatic filariasis after mass drug administration in American Samoa. *PLoS Negl. Trop. Dis.* 8 (8), e3087.
- Sheel, M., et al., 2018. Identifying residual transmission of lymphatic filariasis after mass drug administration: comparing school-based versus community-based surveillance - American Samoa, 2016. *PLoS Negl. Trop. Dis.* 12 (7), e0006583.
- Simini, F., et al., 2012. A universal model for mobility and migration patterns. *Nature* 484, 96.
- Smith, M.E., 2017. Predicting lymphatic filariasis transmission and elimination dynamics using a multi-model ensemble framework. *Epidemics* 18 (Supplement C), 16–28.
- Berg Soto, A., et al., 2018. Combining different diagnostic studies of lymphatic filariasis for risk mapping in Papua New Guinea: a predictive model from microfilaraemia and antigenaemia prevalence surveys. *Trop. Med. Health* 46 (1), 41.
- Southgate, B.A., 1992. The significance of low density microfilaraemia in the transmission of lymphatic filarial parasites. *J. Trop. Med. Hyg.* 95 (2), 79–86.
- Stolk, W.A., et al., 2003. Prospects for elimination of bancroftian filariasis by mass drug treatment in Pondicherry, India: a simulation study. *J. Infect. Dis.* 188 (9), 1371–1381.
- Stolk, W.A., et al., 2013. Modeling the impact and costs of semiannual mass drug administration for accelerated elimination of lymphatic filariasis. *PLoS Negl. Trop. Dis.* 7 (1), e1984.
- Stolk, W.A., Stone, C., de Vlas, S.J., 2015. Modelling lymphatic filariasis transmission and control: modelling frameworks, lessons learned and future directions. *Adv. Parasitol.* 87, 249–291.
- Stone, C.M., Lindsay, S.W., Chitnis, N., 2014. How effective is integrated vector management against malaria and lymphatic filariasis where the diseases are transmitted by the same vector? *PLoS Negl. Trop. Dis.* 8 (12), e3393.
- Stone, W., et al., 2015. Assessing the infectious reservoir of falciparum malaria: past and future. *Trends Parasitol.* 31 (7), 287–296.
- Taylor, M.J., Hoerauf, A., Bockarie, M., 2010. Lymphatic filariasis and onchocerciasis. *Lancet* 376 (9747), 1175–1185.
- The United States Census Bureau, 2018. International Programs. [www.census.gov/programs-surveys/international-programs.html](http://www.census.gov/programs-surveys/international-programs.html).
- Won, K.Y., 2018. Comparison of antigen and antibody responses in repeat lymphatic filariasis transmission assessment surveys in American Samoa. *PLoS Negl. Trop. Dis.* 12 (3), e0006347.
- World Health Organization, 2011. Monitoring and Epidemiological Assessment of Mass Drug Administration in the Global Programme to Eliminate Lymphatic Filariasis: a Manual for National Elimination Programmes.
- Xu, Z., et al., 2017. A synthetic population for modelling the dynamics of infectious disease transmission in American Samoa. *Sci. Rep.* 7 (1), 16725.
- Xu, Z., et al., 2018. The extensive networks of frequent population mobility in the Samoan Islands and their implications for infectious disease transmission. *Sci. Rep.* 8 (1), 10136.