doi:10.20944/preprints202208.0412.v1

Article

Capturing Household Structure and Mobility Within and Between Remote Aboriginal Communities in Northern Australia Using Longitudinal Data: A Pilot Study

Jessie Goldsmith[1], Patricia Campbell[2], Juan-Pablo Villanueva Cabezas[3], Rebecca Chisholm[3,4], Melita McKinnon[5], George Gurruwiwi[6], Roslyn Dhurrkay[7], Alfred Dockery[8], Nicholas Geard[9], Steven Tong[10], Jodie McVernon[11] and Katherine Gibney[1,8]*

[1] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; goldsmithj@student.unimelb.edu.au
[2] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia; patricia.campbell@unimelb.edu.au
[3] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; jp.villanueva@unimelb.edu.au
[4] Department of Mathematics and Statistics, La Trobe University, Bundoora, Australia; Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Australia; chisholm.r@unimelb.edu.au
[5] Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia; Melita.Mckinnon@menzies.edu.au
[6] Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia; George.Gurruwiwi@menzies.edu.au
[7] Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia; roslyn.dhurrkay@menzies.edu.au
[8] Bankwest Curtin Economics Centre, Curtin University, Perth, Australia; Mike.Dockery@cbs.curtin.edu.au
[9] School of Computing and Information Systems, Melbourne School of Engineering, University of Melbourne, Melbourne, Australia; nicholas.geard@unimelb.edu.au.
[10] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; Victorian Infectious Diseases Service, Royal Melbourne Hospital, at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia: steven.tong@mh.org.au
[11] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia; Infection Modelling, Murdoch Children's Research Institute, Melbourne, Australia; j.mcvernon@unimelb.edu.au
[12] Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; Victorian Infectious Diseases Service, Royal Melbourne Hospital, at the Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; katherine.gibney@unimelb.edu.au
* Correspondence: katherine.gibney@unimelb.edu.au

Abstract: Cultural practices and development level can influence a population's household structures and mixing patterns. Within some populations households can be organized across multiple dwellings. This likely affects the spread of infectious disease spreads through these communities, however, current demographic data collection tools do not record these data. Methods: Between June–October 2018, the Contact And Mobility Patterns in remote Aboriginal Australian communities (CAMP-remote) pilot study recruited Aboriginal mothers with infants in a remote northern Australian community to complete a monthly iPad-based contact survey. Results: Thirteen mother-infant pairs (participants) completed 69 study visits between recruitment and the end of May 2019. Participants reported they and their other children slept in 28 dwellings during the study. The median dwelling occupancy, defined as people sleeping in the same dwelling on the previous night, was nine (range: 3.5–25). Participants who completed at least three responses (n=8) slept in a median of three dwellings (range: 2–9). Each month a median of 28% (range: 0%–63%) of the participants travelled out of the community. Including these data in disease transmission models amplified estimates of infectious disease spread in the study community, compared to models parameterized using census data. Conclusions: The lack of data on mixing patterns in populations where households can be organized across dwellings may negatively impact the accuracy of infectious disease models for these communities and the efficacy of public health actions they inform.

Keywords: Aboriginal; Indigenous; contact patterns; household structure; disease transmission; household model, human mobility.

1. Introduction

The pattern of mixing between people in a population is a crucial determinant of how infectious diseases spread within it [1]. As mathematical models exploring infection dynamics gain traction as tools to guide infectious disease prevention and control activities in different populations, the adequate quantification of mixing patterns has become critical. Across the globe, the social and cultural practices of different populations result in different mixing patterns [2-7]. In Australia, the social and cultural practices of remote-living Aboriginal and Torres Strait Islander people (hereafter Aboriginal Australians) result in unique mixing patterns [2-4]. These mixing patterns may influence how infectious diseases spread through remote Aboriginal communities [8].

Community mixing patterns can be conceptualized in three parts: household mixing, intra-community mixing between people from different households, and temporary mobility into and out of the community.

First, households can be pivotal in the spread of infectious disease [9]. For example, the attack rate for SARS-CoV-2 is more than five times higher among household contacts than non-household contacts [10]. Australian demographic surveys show that on average, dwellings in remote Aboriginal communities have more occupants than dwellings in either urban Aboriginal communities or non-Aboriginal communities [11, 12]. These surveys typically assume that a household consists of a single dwelling. However, in remote Aboriginal communities, households may be organized across multiple dwellings [8]. To understand true household mixing relevant to transmission of pathogens, we may need to consider connections between dwellings. To date only one study quantified these linkages [2]. It is based on the frequency that individuals stayed in a single dwelling in a remote Aboriginal Australian community over the course of a year [2]. Using data from this study, Chisholm et al. developed an infectious diseases model that compared outbreak dynamics for populations with multi-dwelling households versus single dwelling households; the modelled outbreak scenarios with multi-dwelling households had larger and faster outbreaks [8]. Clearly, more data are needed to support this type of analysis.

Second, intra-community mixing patterns can also contribute to the spread of infectious disease [13]. Remote-living Aboriginal Australian adults are more frequently involved in cultural events, ceremonies and other community events, have more frequent face to face contact with people outside of the home and are more likely to provide support to relatives outside the household than non-remote-living Aboriginal Australians [3]. While surveys to collect intra-community mixing patterns are increasingly being used in other settings [14-17], there are no comparable studies for remote-living Aboriginal Australians [8].

Finally, temporary mobility into and out of a community provides opportunities for the introduction or exportation of pathogens. Temporary mobility can be defined as a person being absent from their usual place of residence for up to six months [18]. This is consistent with the definition of ‘usual residence’ used by the Australian census [19]. Disease transmission models that include temporary mobility typically rely on Australian Census data for parameterization [20]. Census data indicate Aboriginal Australians have higher rates of temporary mobility than Australians living in urban areas and non-Aboriginal Australians living in rural and remote areas [4]. These census data, however, do

doi:10.20944/preprints202208.0412.v1

not collect information on dynamic patterns of temporary mobility or linkages between communities.

Current demographic data collection tools do not reflect the cultural and societal norms of remote-living Aboriginal Australians. Additional data is required to characterize mixing patterns in these communities. The Contact And Mobility Patterns in remote Aboriginal Australian communities (CAMP-remote) pilot study was developed to begin addressing this knowledge gap by quantifying household mixing (including multi-dwelling households), intra-community social contact patterns and temporary mobility patterns for a remote Aboriginal Australian community. In this paper we describe the CAMP-remote pilot study, analyze the data collected, and demonstrate its application in a mathematical model that simulates the transmission of an infectious disease within the study community. We discuss the impact of using these data to enhance our understanding of mixing patterns in remote Aboriginal Australian communities.

2. Materials and Methods

The project protocol and data collection tools were developed and refined during consultation with community members and stakeholders from June 2017 to April 2018. These consultations included researchers, healthcare workers and board members from the study community, as well as staff from the University of Melbourne, Curtin University, Menzies School of Health Research, One Disease, and Miwatj Health Aboriginal Corporation. This study was approved by the Human Research Ethics Committee of Northern Territory Department of Health and Menzies School of Health Research (HREC 2017-2917), the Community Board, and the Miwatj Health Aboriginal Corporation.

2.1. Study area and study design

The study was conducted between June 2018 and May 2019 in a remote Aboriginal community in the Northern Territory, Australia. The community population is between 2000 and 2500 and it is approximately 1.5 hours flying time from Darwin, the capital of the Northern Territory. The largest employers within the community are supermarkets.

2.1.1. Study Participants

Study participants had to live in the study community and identify as female, Aboriginal and the primary carer of an infant no more than one year old. Potential participants were excluded if they were unable to give informed consent or if they usually shared a dwelling with another mother-infant pair who was already enrolled in the study.

The decision to enroll mother-infant pairs in this pilot study was made during the consultation process. This cohort was assumed to be less likely to be lost to follow-up and provided an opportunity to capture social interactions across a spectrum of ages within the community.

2.1.2. Study Procedures

Two survey instruments were used. Both were conducted in person by study staff. The baseline survey was undertaken at enrolment and gathered basic demographic information about the participant and any children for whom they were the primary caregiver. The second was the dwelling, contact and travel survey which was repeated monthly from enrolment until May 2019.

Survey questionnaires were based on previously validated surveys (Appendix A) [14, 21]. The contact questions were based on the questionnaire used in the European POLYMOD study [14], previously adapted for use in urban Australian populations [22], while the temporary mobility questions were based on the Mobility Survey conducted in Central Australia by the Co-operative Research Centre for Remote Economic Participation [21]. Participants were asked how many people they had contact with in the 24 hours prior to the survey. Compared to previous surveys, questions on contacts were simplified to reduce the reporting burden on participants. Contacts were categorized by age-group (preschool [less than 5 years], school [5–14 years] and adult [15+ years]) and participants only reported contacts that met the following criteria:

- Dwelling contact – anyone who slept in the same dwelling as the participant the previous night.
- Social contact – anyone with whom the participant spent an hour or more during the previous 24 hours, excluding dwelling contacts.

Survey responses were collected using a custom-built iPad application. The questions were written in English, but Aboriginal community researchers were able to conduct the survey in the local language, if required. The iPad application made use of visual representation rather than text wherever possible, following feedback during the consultation process. For example, a map of the community was used to collect information about dwelling location.

2.2. Data analysis

All analyses were undertaken using the statistical software R [23]. Graphs were produced using gpplot2 and nptest was used to generate non-parametric bootstrapped confidence intervals [24, 25].

2.2.1. Clustering responses by dwelling location

The dwellings in which participants indicated they and/or their children slept over the week prior to each survey were recorded as spatial co-ordinates. A hierarchical clustering approach was used to group the spatial coordinates and identify distinct dwellings [23]. Dwellings were defined as the centroid of 50 meter radius buffers to account for variation in the coordinates resulting from a touch screen application. Spatial data were mapped using the Australian Albers project (epsg projection 3577 – gda99).

The median number of dwellings slept at by participants and their children and the proportion of surveyed nights spent at each dwelling were calculated using data from participants who responded to the monthly survey at least three times. Dwellings were classed as ‘core’, ‘frequent’ and ‘infrequent’ according to the amount of time participants spent there, where core dwellings were the dwellings that participants reported staying at most regularly. The proportion of time spent at each class of dwelling was reported with 95% confidence intervals.

2.2.2. Dwelling contact matrices

The repeated surveys collected dwelling occupancy for the night prior to each survey. The median and range of the mean occupancy was calculated per available room in each dwelling and within the room where the participant slept. In addition, the median and range of the mean occupancy per dwelling was calculated by age group. These data were presented as age-stratified matrices with 95% confidence intervals (Appendix B).

2.2.3. Intra-community mixing

The number of social contacts who spent an hour or more with the participant in the 24-hour period leading up to each study visit was captured in the repeated surveys, excluding dwelling contacts. The mean number of daily social contacts per participant was reported.

2.2.4. Temporary mobility

The repeated surveys collected data on travel by participants out of the community in the preceding month, including the primary reason for and duration of travel, and whether their baby travelled with the participant. In addition, researchers noted if a participant was out of the community during one of the monthly study visits. If a participant was absent on two successive survey visits it was assumed they had not returned to the community in the intervening period.

2.2.5. Individual based model simulation

We compared the results of simulations for an influenza-like illness and an endemic infection using an individual-based transmission model which was informed by different assumptions and data sources for community mixing. The purpose was to quantify the impact of different assumptions of household structure and community mixing on model outputs and estimate the importance of collecting multi-dwelling household structure and temporary mobility data for model accuracy in remote Aboriginal communities [8]. For each disease type, we simulated the model using three scenarios for a community with the demographic characteristics of our study community. First, we assumed that each person is equally likely to contact each other person in the community (homogenous mixing). Second, we used contact matrices developed using Census data from the Australian Bureau of Statistics on household size and age distribution, under the assumption that each household is associated with a unique dwelling [26]. Third, we used the dwelling contact matrices, multi-dwelling household structure and temporary mobility rates developed using the CAMP-remote data. Social contact data for the second and third scenarios were sourced from Kiti et al.'s study based in Kenya [15]. For the influenza-like illness, we compared the timing and magnitude of the infection peak for the three scenarios. For the endemic disease, we compared the disease prevalence between the three scenarios.

Further detail on the model structure and parameters is in Appendix C.

3. Results

3.1. Summary statistics

Thirteen mother-infant pairs were recruited: ten in June 2018, two in August 2018 and one in October 2018. Three participants withdrew over the course of the study, resulting in a 23% loss to follow-up. Sixty-nine survey responses were collected, a median of four responses per participant (IQR: 2–9). Further detail on recruitment and participation is provided in Appendix D.

At recruitment, the median age of participants was 23 years (range: 17–37 years) and infants was eight months (range: 0–22 months). Participants reported caring for a median of two children (range: 1–3).

3.2. Household structures and size

We identified 28 distinct dwellings where a participant and her infant or another of the children she cared for had slept during the week prior to a completed survey. Participants who responded to the survey at least three times (n=8) identified a median of three separate dwellings (range: 2–9, Figure 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 August 2022 doi:10.20944/preprints202208.0412.v1

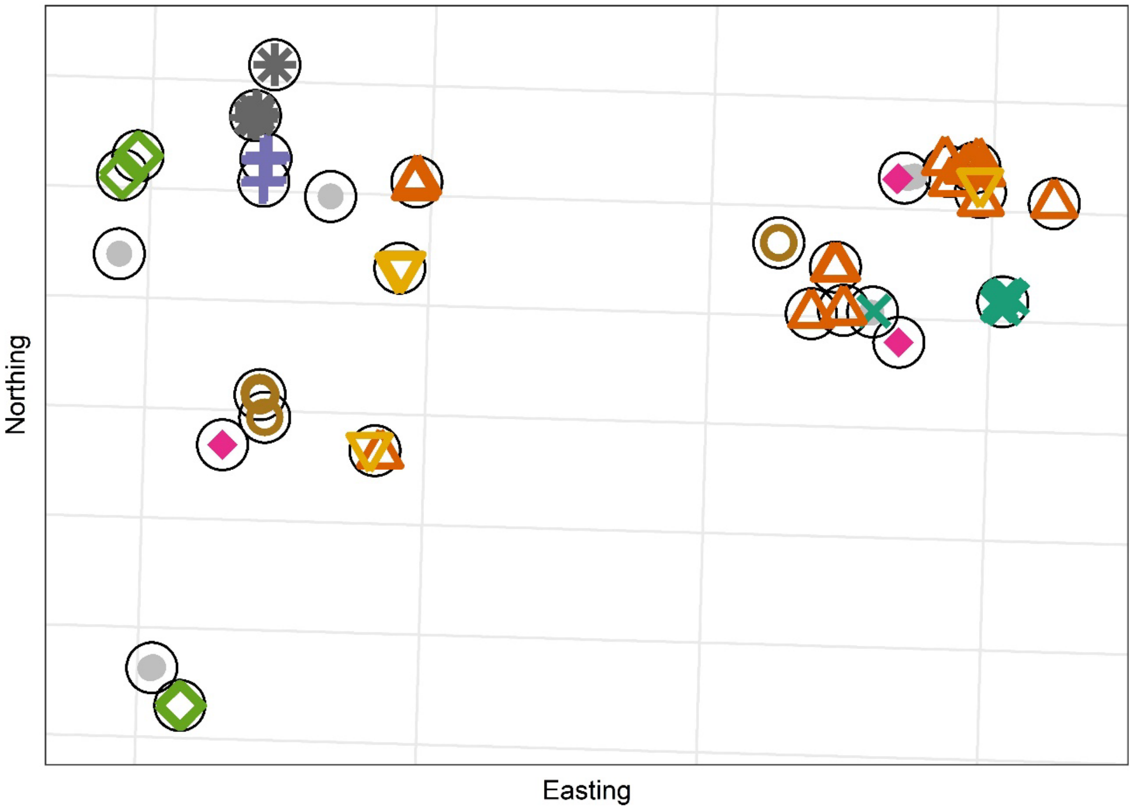


Figure 1. Clustering of spatial coordinates to identify distinct dwellings (n=28). Each symbol represents a survey response. Participants with one or two responses are depicted as grey circles, all other participants (n=8) are depicted using a unique symbol.

Of the 28 dwellings identified, the composition on the night prior to the survey was collected on at least one occasion for 25 dwellings and collected five or more times for four of the dwellings. The mean occupancy by dwelling ranged from 3.5 to 25 people, with the median of the mean of ten people (IQR: 8–12, Figure 2). Most of these occupants were adults.

Figure 2. Distribution of the mean number of people per dwelling by age group. Each dot represents the mean number of individuals of a specific age across all survey responses for a given dwelling. The median of the means is represented using the thick grey line and the interquartile range of the means with the thinner grey lines.

A comparison of the dwellings where five or more nights of occupancy data were available (Figure 3) demonstrates significant variation in the number of occupants for some dwellings (e.g. Dwelling D) but not others (e.g. Dwelling A).

The median of the mean occupancy per room by dwelling was 2.7 individuals (range: 1.6–4.2), with a higher median of the mean occupancy in the participant's room of 3.3 individuals (range 2.0–5.0). Participants reported always sleeping in the same room as their baby and almost always (97% of occasions surveyed) in the same bed.

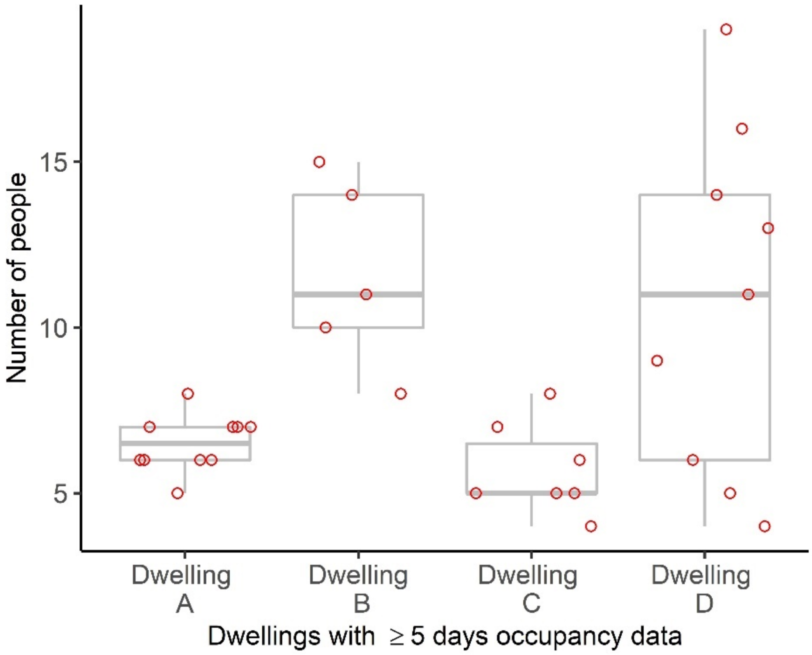


Figure 3. Distribution of the number of people per dwelling for dwellings with five or more nights of data. Each dot represents the total number of individuals in each dwelling for a specific survey response. The median is represented using the thick grey line and the interquartile range with the thinner grey lines.

3.3. Dwelling contacts

We estimate this population had 15 adult-to-adult (95% CI: 10–28) and 12 adult-to-young (95% CI: 6–16) dwelling contacts each day (Figure 4).

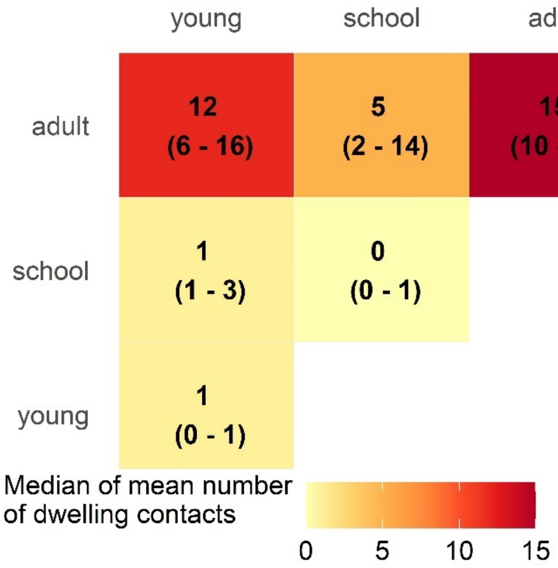


Figure 4. Median number of mean dwelling contacts between each age category per dwelling using the CAMP-remote data. 95% confidence intervals, estimated with a nonparametric bootstrap method, are indicated in brackets. Age ranges were young (less than 5 years), school-aged (5–14 years) and adult (15+ years).

3.4. Intra-community mixing

The mean number of social contacts reported per participant per day (in excess of dwelling contacts) had a median of two (range: 0–11), Figure 5.

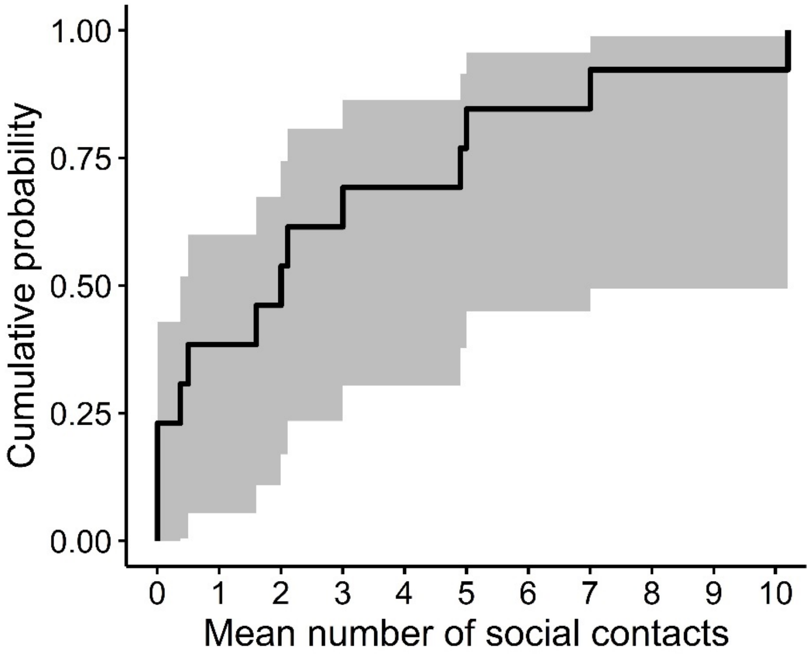


Figure 5. Cumulative distribution of the mean number of social contacts per CAMP-remote participant (n=13) per day. Shaded area represents 95% confidence interval.

3.5. Temporary mobility

A median of 28% (range: 0%–63%) of participants travelled out of the community each month. No visual evidence of seasonality was apparent between the wet (November to April) and dry (May to October) seasons (Appendix E).

From the subset of surveys in which participants reported travel (n=22), it was usually within the region where the study community was located (n=13, 59%) and lasted for a median of seven days (IQR: 5–18 days). The primary reason for travel was most frequently to attend a funeral (n=9, 41%) or visit family (n=7, 32%).

3.6. Individual based model simulation

The household structured model parameterized with the CAMP-remote data had the highest and earliest peak of infectious individuals in the simulations of an outbreak (Figure 6A) and the highest on-going prevalence in the simulations of an endemic infection (Figure 6B). Applying the ABS data resulted in simulated outbreaks and endemic transmission that were overall similar to those for the models parameterized with the CAMP-remote data, but with a slightly lower and later peak and a slightly longer duration for the outbreak model, and slightly lower on-going prevalence for the endemic disease model. Uniform mixing assumptions resulted in simulated outbreaks with a substantially lower and later peak and the simulated endemic transmission with substantially lower on-going prevalence (Appendix F).

doi:10.20944/preprints202208.0412.v1

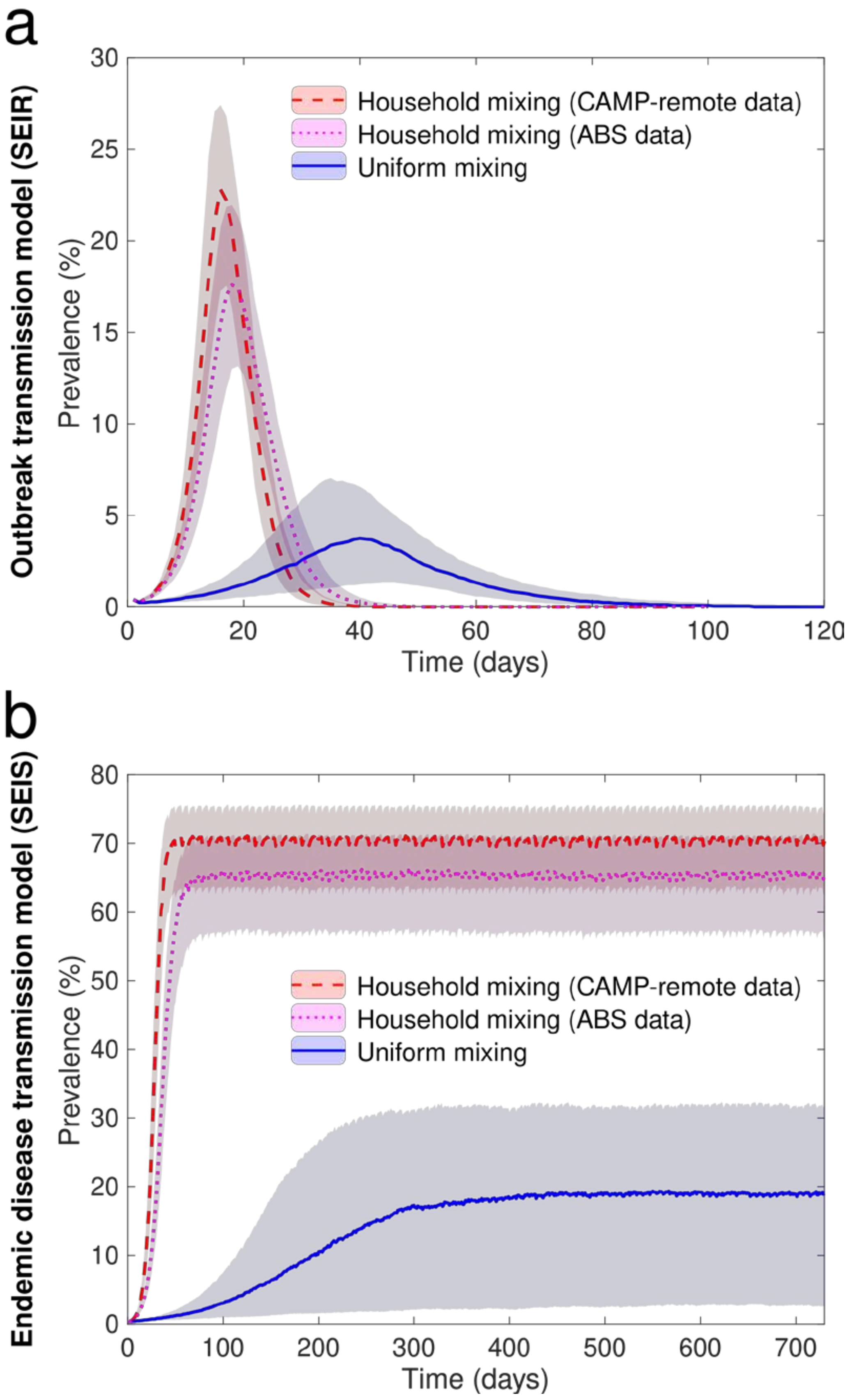


Figure 6. (a) The prevalence of infection arising from simulations of an influenza-like SEIR model under different mixing assumptions. (b)The prevalence of infection arising from simulations of an endemic disease in an SEIS model under different mixing assumptions. Both models are parameterized to reflect the study community.

4. Discussion

We documented that participants and their children slept in a median of three dwellings over the 12-month study period and more than a quarter of participants travelled out of the community each month. As we only have data on where a participant and her children slept for a small fraction of the year, and it was assumed that participants did not return to the community between surveys if they were absent for successive surveys, these estimates likely underestimate the extent of multi-dwelling households and temporary mobility in this remote Aboriginal community. Even using these lower-bound estimates, modelling has demonstrated the importance of characterizing these data for model accuracy and to subsequently inform infectious disease prevention and intervention strategies.

These socio-cultural patterns influence how infectious diseases are transmitted in a community and are an important consideration when planning and implementing a public health response. A recent systematic review found that most prior social contact surveys have been conducted in high-income countries [27]. Within Oceania, the review found that Australia was the only country that has collected data on mixing patterns [27]. Cultural norms, development level and geography influence social mixing, household size and structure [27]. For example, Vino et al. demonstrated that household contact patterns in remote-living Australian Aboriginal communities were substantially higher than those for residents of Melbourne, Australia [11].

Fluid households and substantial rates of temporary mobility are not unique to remote-living Aboriginal Australians. Similar patterns have been observed in South Africa, Tanzania, Burkina Faso and in remote Canadian Inuit communities [5-7]. In recognition of the fluidity of household membership within South Africa, the Population Intervention Platform (PIP) cohort, which was established in 2000 (African Centre Demographic Information System cohort) in rural KwaZulu Natal, uses a definition of household membership that includes resident and non-resident members and allows for individuals to belong to more than one household [28, 29]. Further studies are required that collect information on contact patterns within communities exhibiting these patterns. The CAMP-remote pilot study provides a demonstration of both a method for and the impact of collecting these data.

Our findings are consistent with those previously reported by Musharbash [2]. Musharbash’s study collected data on the number of people that stayed at a single dwelling in a remote community in Central Australia over 221 nights and, based on these data, proposed four categories of occupant: core residents (those present between 60–100% of the time); regular residents (those present between 20–34% of the time); other residents (present 4%–16% of the time); and sporadic short-term visitors [2]. While the consistency between these two studies is encouraging, further data is required to confirm the findings. The repeated survey approach of the CAMP-remote study yielded valuable data on connections between dwellings and variation in dwelling occupancy that otherwise would not have been available.

The dwelling contact matrix developed in the CAMP-remote study estimated a median of 12 adult-to-young child contacts each day per dwelling, compared with the mean of 4.14 per household for remote Australian Aboriginal communities reported in Vino et al., based on data from the Aboriginal Birth Cohort study [11]. While dwellings in the CAMP-remote study all included at least one mother and one infant (due to the recruitment strategy), dwellings in the Aboriginal Birth Cohort study all included at least one individual aged between 22 and 27 years [11]. Differences in the contact matrix calculated using the CAMP-remote data and the contact matrix in Vino et al. may be a function of the different study cohorts or differences in the age group categories, communities, study sample sizes, or study designs.

The CAMP-remote survey is the first study to collect quantitative data on intra-community mixing within a remote Aboriginal community. Direct comparisons between data collected during this study and prior studies are challenging due to differences in definitions for social contact. Kiti et al.'s study, based in Kenya, defines a social contact as anyone with whom the participant had physical contact [15]. The POLYMOD study from Europe similarly defines a social contact as anyone with whom the participant had skin-to-skin contact or an in-person two-way conversation with three or more words without skin-to-skin contact [14]. By comparison, the definition used in the CAMP-remote study was anyone with whom the participant spent an hour or more. However, when mean contacts observed in the CAMP-remote pilot study are contrasted to that of similarly aged women (20–49 years) in Kiti et al.'s study, we observe that while the scales differ, the distributions of the number of social contacts per day are similar (Appendix G). As Kiti et al. collected social contact information from all aged groups and both sexes, using Kiti et al.'s data for our simulations remains a reasonable substitute for data on intra-community mixing patterns within remote-living Aboriginal communities until more comprehensive social contact data become available for remote Aboriginal Australian communities.

The definition of social contact adopted during the CAMP-remote study was developed in consultation with the community and designed to both limit reporting burden for participants and maximize the chances of on-going participation in the study. While this approach was appropriate given the longitudinal nature of the CAMP-remote study, it limited direct comparison of social contact data with previously published studies. An alternative approach would be to collect social contact data from a remote Aboriginal community using a previously validated definition (e.g. from the POLYMOD survey or study by Kiti et al.) on a single occasion to allow comparison with other, disparate, populations.

The modelling results demonstrate that including information on household characteristics and temporary mobility amplifies estimates of the spread of an infectious disease in remote Aboriginal communities. Differences were observed in the intensity of transmission between the three scenarios, with the modelling results from the simulation using CAMP-remote data having the most intense outbreaks and highest rates of endemic transmission. Both models that applied a household structure showed substantially greater intensity of transmission than the model that assumed uniform mixing, suggesting that uniform mixing assumptions are not appropriate to parametrize infection dynamics within these communities. Prior work by Chisholm et al. suggests that we would expect to observe an even greater difference in transmission intensity between the two household-structured models (parameterized by ABS versus CAMP remote data) if we were to consider a less-transmissible pathogen [8], indicating the increased importance of collecting relevant mobility, social contact, and household data for such pathogens.

Our study has several limitations. First, the small sample size means we are unable to determine the statistical significance of our results. Second, as the CAMP-remote study only recruited mother-infant pairs, data are limited to this cohort and may not be generalizable to all members of remote Aboriginal communities. To represent the household structure, intra-community mixing patterns and temporary mobility of remote-living Aboriginal Australians more accurately further, more inclusive, data collection is required. Third, the extent of loss to follow-up and missing survey responses may have introduced bias if the participants who dropped out or did not respond were different from those who did respond. Finally, to identify individual dwellings it was necessary to make assumptions about the distance between dwellings which may have resulted in dwellings being misclassified (i.e. adjacent dwelling to be classified as a single dwelling).

5. Conclusions

Our study demonstrates the gap in the current knowledge of mixing patterns within the remote Aboriginal Australian communities and the importance of collecting data that reflect the organization of households across dwellings to inform infectious disease modelling and public health interventions for these communities and others with similar patterns. Given the limited proximate health infrastructure in these settings it is vital that we develop models that are as accurate as possible to allow appropriate planning for, prevention of, and responses to an outbreak. Similarly, public health responses to endemic disease need to be based on a realistic understanding of the extent of transmission and the level of community participation required to make a meaningful difference in the health outcomes of individuals. The types of data that the CAMP-remote pilot study demonstrated could be collected contributes to this goal.

Author Contributions: Conceptualization, P.C., M.D., N.G., S.T. J.M. and K.G.; methodology, K.G.; validation, P.C. and J.P.V.C.; formal analysis, J.G.; investigation, M.M., G.G. and R.D.; resources, X.X.; data curation, J.G.; writing—original draft preparation, J.G.; writing—review and editing, P.C., J.P.V.C., N.G., S.T. and K.G.; visualization, J.G.; supervision, J.M. and K.G.; project administration, M.M. and K.G.; funding acquisition, K.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Health and Medical Research Council (NHMRC), grant number APP1098319, and the University of Melbourne.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and ethics approval for this study was provided by the Human Research Ethics Committees of the Northern Territory Department of Health and Menzies School of Health Research (protocol code 2017-2917 and Date).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available because they contain information that could compromise the privacy of research participants.

Acknowledgments: We thank the study community and especially the participants who volunteered their time and contributed to study while involved in the often-all-consuming task of caring for an infant. The App used in this project was developed by the Melbourne eResearch Group at the University of Melbourne; Duy (Daniel) Vu, Rio Susanto, Glenn Jayaputera and Richard Sinnott.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

doi:10.20944/preprints202208.0412.v1

Appendix A

All references to the study community's name have been replaced by the phrase "the study community".

Background survey (first interview)

1. What is your name?
    a. Balanda name [text]
    b. Yolngu first name [text]
    c. Yolngu last name [text]

2. What is your date of birth? [text] ________________ (enter age if DOB unknown)

3. How many days each week do you work in a paid job (on average)?
0 1 2 3 4 5 6 7

4. How many children do you care for most of the time? {we will refer to these as your children and the youngest as your baby}
1 2 3 4 5 6 more than 6 [text]

5. Please tell us about <u>each</u> of your children [start with baby and repeat for each child]
    a. Balanda name [text]
    b. Yolngu first name [text]
    c. Yolngu last name [text]
    d. Date of birth [text]___________ (enter age if unknown DOB)

    e. are you {child's name} mother?
    ○ No
    ○ Yes

    f. How many days each week does {child's name} attend (on average):
    School 0 1 2 3 4 5 6 7
    - Child care 0 1 2 3 4 5 6 7
    - Baby-hub 0 1 2 3 4 5 6 7

Household, Contact and Travel Survey (repeated monthly)

Section A: Where did you and your children sleep last night and in the last week

1. In the house you slept in last night, how many rooms do people sleep in (including the verandah)?

2. How many people slept in each room of this house last night, including you and your children? [icons on house graphic]
- *young child, school aged, adult [usual residents of this house]*
- *young child, school aged, adult [visitors to this house]*

3. How many of the 'x' visitors who slept in your house last night usually live outside *the study community*? 0 1 2……… 'x'

4. Last night, did you sleep on the same mattress or used the same bedding as:
   a. {baby}? Yes/No
   b. {child 2}? Yes/No
   c. {child 3}? Yes/No
   [more if needed]

5. Is this the house that you live in most of the time? (i.e. the house you spend most nights in)
   - ○ No
   - ○ Yes

6. How many nights in the last week did you sleep in this house?
   1 2 3 4 5 6 7

7. How many nights in the last week did you sleep in the same house as:

a. {baby}?	0	1	2	3	4	5	6	7
b. {child 2}?	0	1	2	3	4	5	6	7
c. {child 3}?	0	1	2	3	4	5	6	7

   [more if required]

8. On the map, please mark the house you slept in last night

9. On the map, please mark any other houses you slept in during the last week *{skip if Q6 = '7'}*

10. On the map, please mark any other houses your children slept in, when they were not with you *{skip if Q7 all = '7'}*
    a. {baby} [click on houses on the map] *{skip if Q7.a = '7'}*
    b. {baby} [click on houses on the map] *{skip if Q7.b = '7'}*
    c. {baby} [click on houses on the map] *{skip if Q7.c = '7'}*
    [more if required]

Section B: The people you and your baby saw yesterday

11. Was {baby} with you all day yesterday? Yes/No *{if no, skip to Q**Error! Reference source not found.**}*

12. Who was {baby} with during the day yesterday: [check all that apply, must check at least one for each part of the day]

Morning:	with me	with friends/relatives	at childcare	other [text]
Afternoon:	with me	with friends/relatives	at childcare	other [text]
Evening:	with me	with friends/relatives	at childcare	other [text]

13. During the day yesterday, did you spend at least one hour with anyone who was not in the same house as you last night? Yes/No {if no, skip Q…}

14. Please tell us about each person that you **spent at least one hour with** yesterday, who was **not** in the same house as you last night

- Record each contact person only once, even if you saw them more than once
- Don't include people who were in your household last night

Contact no.	**Age group** *Young child* *School aged* *Adult*	**Did they touch you?**		**Did they touch {baby}?**	
		Y	**N**	**Y**	**N**
1	Young child	☐	☐	☒	☐
2	School	☐	☐	☒	☐
3	Adult	☐	☐	☒	☐
4	Adult	☐	☐	☐	☒

Section C: Your travel in the last month

15. Did you travel outside *the study community* in the last month?
    - ○ No [skip to Q20]
    - ○ Yes

16. Where did you spend at least one night? [check all that apply]
    - ☐ East Arnhem (outside *the study community*)
    - ☐ Darwin
    - ☐ Elsewhere in the NT
    - ☐ Outside the NT

17. How many nights did you spend away from *the study community*?
    1 2 3 4 5 6 7 8 9 10 more [text]

18. Was {baby} with you on the nights you travelled out of *the study community*?
    - ○ Yes, all nights I was away
    - ○ Some nights I was away
    - ○ No {baby} did not travel with me

19. What was the main reason for this travel?
    - ○ Access services (e.g. shopping, health services)
    - ○ Work
    - ○ Visit family

- ○ Attend a funeral
- ○ Other cultural business
- ○ Holiday

Section D: Recent illness and antibiotic use

20. Have you had to go to clinic or hospital because you were unwell in the last month?
    - ○ No [skip to Q21]
    - ○ Yes

    a. What type of illness did you have?
        - ☐ Skin infection or scabies
        - ☐ Cold or ‘flu-like illness
        - ☐ Chest infection
        - ☐ Diarrhoea
        - ☐ Other illness

21. Have you taken antibiotics in the last month?
    - ○ No
    - ☐ Antibiotic injection
    - ☐ Antibiotic syrup

22. Has your baby had to go to the clinic because they were unwell in the last month?
    - ○ No [skip to Q24]
    - ○ Yes

23. What type of illness did your baby have? *[check all that apply]*
    - ☐ Skin infection or scabies
    - ☐ Cold or ‘flu-like illness
    - ☐ Chest infection
    - ☐ Diarrhoea
    - ☐ Other illness

24. Has your baby taken antibiotics in the last month? *[check all that apply]*
    - ○ No
    - ☐ Antibiotic injection
    - ☐ Antibiotic syrup

Appendix B

Age-specific contact rates were calculated using the following approach.

We have assumed each person in a dwelling encounters each other member of the dwelling in any given day. The daily number of dwelling contacts between individuals within the same age category is therefore given by $x(x-1)/2$, where x is the number of individuals in that age category. The daily number of contacts between individuals in different age categories is given by xy where x and y are the respective number of people in the two age categories. The contact matrix for an individual dwelling, which is symmetric, is therefore given by Table A.1.

Table A.1. Dwelling Contact Matrix: Number of pre-school aged children (b), school-aged children (c) and adults (a)

	Pre-school aged	School aged
Adult	ab	ac
School aged	cb	c(c−1)/2
Preschool aged	b(b−1)/2	

Appendix C

We used a modified version of the stochastic individual based model described fully in Chisholm et al [8]. The model explicitly represents each individual (agent) and each dwelling in a community. Agents are assigned a set of dynamic characteristics which includes their age, disease status (which is either susceptible, exposed, infectious, recovered) and current home location, which are updated daily according to a set of rules. They are also assigned a set of stable characteristics that do not change during simulations which includes their residency status (whether they are a permanent or temporary resident of the community), and the set of dwellings they regularly inhabit – their *core*, *frequent* (referred to as '*regular*' in Chisholm et al [8]), and *infrequent* (referred to as '*on-off*' in Chisholm et al [8]) residences.

The model simulates the introduction and spread of an infection in the community by simulating contacts between pairs of agents (according to the mixing models described below). Contacts between an infectious and susceptible agent result in transmission according to a given probability, which we refer to as the *probability of transmission*.

The model also simulates the movement of agents between dwellings in their dwelling set (which updates an agent's current residence), and into and out of the community. Migration into and out of the community is modelled using two mechanisms: *constant flow migration* (described below), and *event-based migration* whereby new agents temporarily visit the community due to events (e.g., funerals, sporting events) and leave after different mean durations of time. In the outbreak (SEIR) model we assume that all incoming temporary agents are susceptible, while in the endemic (SEIS) transmission model we assume that the prevalence of infection in incoming temporary agents is 10%.

We modified the constant flow migration mechanism (compared to the model in Chisholm et al)[1] to reflect observations of mobility in the CAMP remote study. Each day, agents are selected at random to leave the community (at a rate such that 18-36% of the population move per month). Agents stay away for a median duration of seven days. While away, these agents are replaced by temporary agents (who stay in a randomly selected dwelling for the duration of their stay) who can become infected and contribute to transmission. These agents leave once the resident they replaced returns.

From simulations of the model we report the *prevalence of infection* over time (the percentage of all agents present in the community who are infectious), and/or the *final outbreak size* (the total number of all permanent residents who were infected during an outbreak, which is related to the *attack rate = (final size / number of permanent residents) * 100%*).

Transmission scenarios (defined by mixing model and disease model)

To account for uncertainty in the model parameters, for each transmission scenario considered, we generated 1,000 samples from the model parameter space using Latin Hypercube Sampling (LHS) [30]. Sampled parameters and their distributions match those used in Chisholm et al with the exception of the infection parameters which are described in Table A.2 [8]. All other parameters were held constant.

Household/community mixing

We consider two household/community mixing models:

- The ABS model: which is parameterized by ABS (census) data only,
- The CAMP remote model: which is supplemented by CAMP remote data.

In both models, agents contact all other agents that are sharing their current residence once per day and contact all other members of the community at an age-dependent rate. Contacts between co-residents have an increased probability of transmission per contact compared to contacts between non co-residents (which is the same for both household/community mixing models).

The rate of household contacts varies between the two models due to differences in the mean core household size (which is 6.7 in the ABS model, and ten in the CAMP remote model), and the rates of between dwelling mobility (which is switched off in the ABS model).

The rate of non-household contacts is the same in both household/community mixing models, and is calculated using the rate of non-household contacts reported in a Kenyan contact study [15] and the community's age distribution, which is parameterized according to ABS data [26].

Uniform mixing between all agents in the population.

We also consider one uniform mixing model. In this model, contacts occur at an age- and household-independent rate per day for all agents and are chosen uniformly at random. The mean contact rate, M, is calculated as follows. At the end of each LHS simulation from the CAMP-remote model:

- store the distribution of the current occupancy of houses, O.
- store the distribution of the number of current residents (temporary and present, permanent) in each age group i, A_i.
- For each age group i, calculate the mean number of contacts per day per individual in age group i, C_i.
- Calculate C, the mean number of community contacts per individual for this population, $C = \text{sum}_i(C_i . A_i)/ \text{sum}_i(A_i)$.

Then, after all LHS simulations, calculate M, the mean contact rate, $M = \text{mean}(O - 1) + \text{mean}(C)$ (where the mean is taken across all LHS simulations). This works out to be 22.71 contact per day.

We also multiply the probability of transmission per contact by a factor, F, to account for some contacts occurring in households (which in both household/community mixing models have an increased probability of resulting in transmission, where the increase in risk is uniformly distributed between 3 and 5). F is calculated as:

$$F = [1/(\text{mean}(O - 1) + \text{mean}(C))] * [4 * \text{mean}(O - 1) + \text{mean}(C)] \tag{1}$$

This works out to be $F = 2.19$.

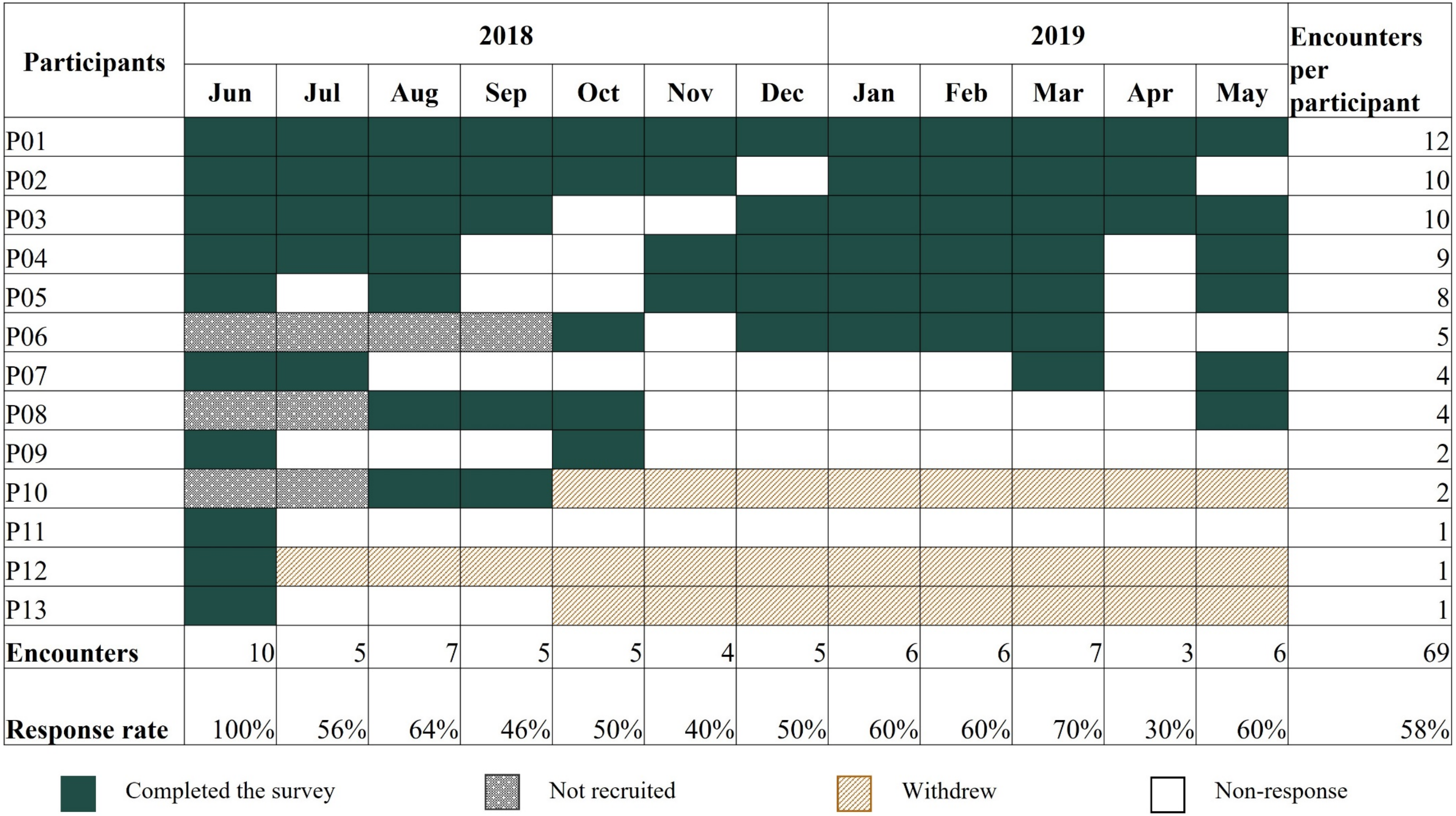
Infection parameters used in the models are summarized in Table A.2.

Table A.2. Infection parameters for the outbreak and endemic disease models

Model	**Infection parameters**	**Model inputs***
SEIR (outbreak)	Duration of latency	U (1,3) days
	Duration of infection	U (1,3) days
	Probability of transmission in the household/community mixing models (chosen so that the median attack rate in uniform model is ~66%)	U (0.01, 0.015)
SEIS (endemic):	Duration of latency	U (1,3) days
	Duration of infection	U (7,21) days
	Probability of transmission in the household/community mixing models (chosen so that the median population prevalence at endemic equilibrium in uniform model is ~20%)	U (0.0014, 0.0021)

* Where U(a,b) is a uniform distribution with minimum value a and maximum value b.

Appendix D



Participants	2018							2019					Encounters per participant
	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	
P01	■	■	■	■	■	■	■	■	■	■	■	■	12
P02	■	■	■	■	■	■	□	■	■	■	■	□	10
P03	■	■	■	■	□	□	■	■	■	■	■	■	10
P04	■	■	■	□	□	■	■	■	■	■	□	■	9
P05	■	□	■	□	□	■	■	■	■	■	□	■	8
P06	▒	▒	▒	▒	■	□	■	■	■	■	□	□	5
P07	■	■	□	□	□	□	□	□	□	■	□	■	4
P08	▒	▒	■	■	■	□	□	□	□	□	□	■	4
P09	■	□	□	□	■	□	□	□	□	□	□	□	2
P10	▒	▒	■	■	▨	▨	▨	▨	▨	▨	▨	▨	2
P11	■	□	□	□	□	□	□	□	□	□	□	□	1
P12	■	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	1
P13	■	□	□	□	▨	▨	▨	▨	▨	▨	▨	▨	1
Encounters	10	5	7	5	5	4	5	6	6	7	3	6	69
Response rate	100%	56%	64%	46%	50%	40%	50%	60%	60%	70%	30%	60%	58%

■ Completed the survey ▒ Not recruited ▨ Withdrew □ Non-response

Figure A.1. Summary of recruitment and response rates by participant and by month

Appendix E

Table A.3. Travel out of the community by month

Season	Month	Responded			Did not respond		No. of enrolled participants	Total who travelled out of the community	
		Travelled	Did not travel	%	Travelling	Not travelling		No	%
Dry	June	2	8	20%	-	-	10	2	20%
	July	1	4	20%	1	2	8	2	25%
	August	1	6	14%	2	0	10	3	30%
	September	0	5	0%	0	3	10	0	0%
	October	3	2	60%	1	1	8	4	50%
Wet	November	3	1	75%	2	1	8	5	63%
	December	3	2	60%	1	0	8	4	50%
	January	2	4	33%	0	0	8	2	25%
	February	0	6	0%	0	0	8	0	0%
	March	3	4	43%	0	0	8	3	38%
	April	1	2	33%	0	5	8	1	13%
Dry	May	3	3	50%	1	0	8	4	50%

Appendix F

Table A.4. Statistics from model scenarios of outbreaks and endemic disease for a community like the study community including total outbreak duration, attack rate and endemic prevalence.

Disease Model	Mixing model	Outbreak Duration (days)					Final size (number infected)				
		2.5p	25p	50p	75p	97.5p	2.5p	25p	50p	75p	97.5p*
SEIR	HH	30	36	41	45.5	56	2421	2472	2483	2487	2492
SEIR	HH (ABS)	36	43.5	49	55	69	2246	2386	2432	2457	2478
SEIR	U	34.5	89	107	134	220.5	45	1151	1638	1938	2211
		Endemic prevalence									
		2.5p	25p	50p	75p	97.5p					
SEIS	HH	52.0%	63.3%	70.3%	75.3%	80.7%					
SEIS	HH (ABS)	43.3%	57.1%	65.5%	70.9%	76.6%					
SEIS	U	0.2%	2.8%	19.0%	31.7%	44.9%					

*The probability of transmission under each of the three scenarios was assumed to be consistent. SEIR: (Susceptible-Exposed-Infectious-Recovered) transmission model. SEIS: (Susceptible-Exposed-Infectious-Susceptible) transmission model. HH: Household/community mixing model. U: Uniform mixing model. ABS: Australian Bureau of Statistics. p: percentile.

Appendix G

To validate the data collected during the CAMP-remote study the distribution of social contacts for adult females aged 20-49 years using the de-identified individual data provided in the supplementary material for Kiti et al was calculated for rural dwellings.[3] The median number of social contacts for each distribution is at approximately 19% of the maximum number of social contacts recorded for any individual., Figure A2.

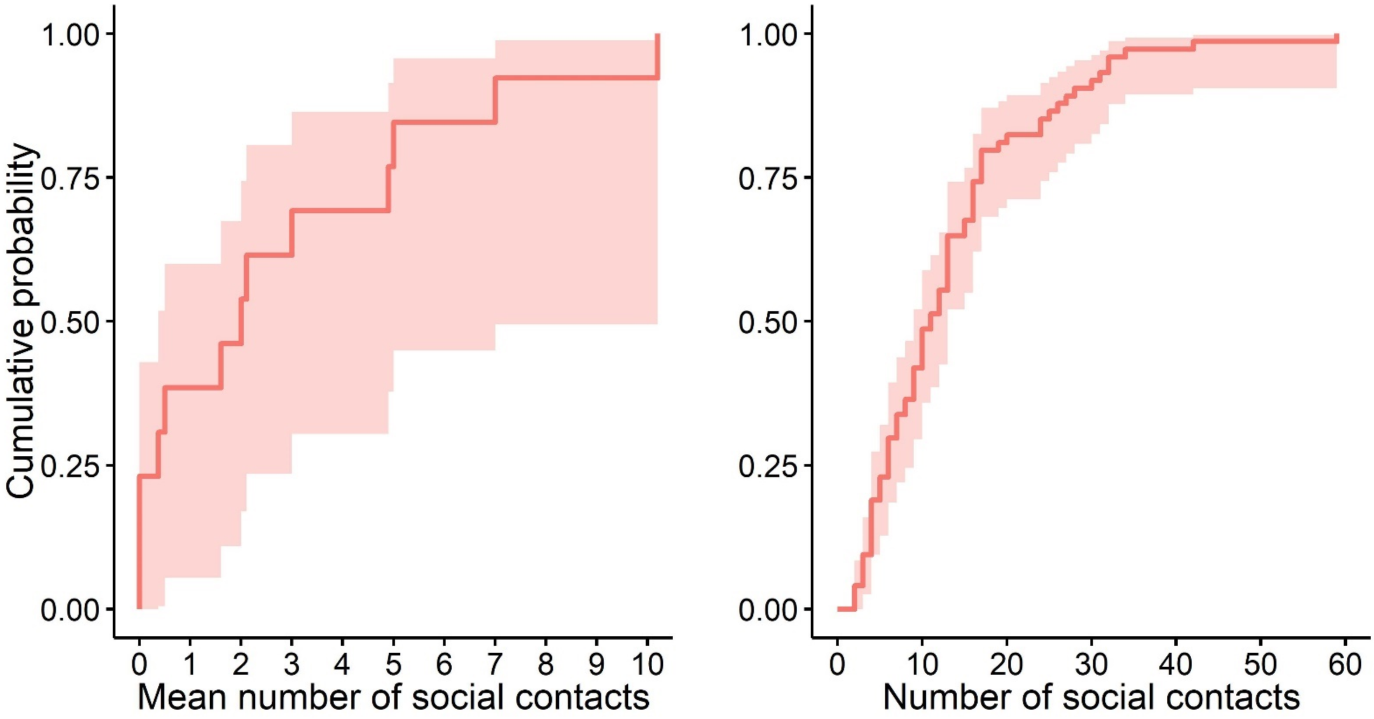


Figure A.2. (a) Mean number of social contacts per CAMP-remote participant per day (n=13). (b) Number of social contacts per adult, female participants (20–49 years) per day in a community in rural coastal Kenya collected by Kiti et al (n=74) [15]. The definition of contact for Kiti et al.'s study was anyone with whom they had physical contact [15], while the definition for the CAMP-remote study was anyone with whom they had spent an hour or more.

References

1. Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DAT. Close encounters of the infectious kind: methods to measure social mixing behaviour. *Epidemiol Infect.* **2012**; *140(12)*:2117-30. doi:10.1017/S0950268812000842.
2. Musharbash Y. *Yuendumu everyday: contemporary life in remote Aboriginal Australia*, Aboriginal Studies Press; Canberra, Australia 2008.
3. Biddle N. Measures of Indigenous social capital and their relationship with well-being. *Aust J of Rural Health.* **2012**; *20(6)*:298-304. doi: 10.1111/j.1440-1584.2012.01293
4. Biddle N, Prout S. The geography and demography of Indigenous temporary mobility: an analysis of the 2006 census snapshot. *J Popul Res (Canberra).* **2009**; *26(4)*:305-26. doi: 10.1007/s12546-010-9026-1.
5. Hosegood V, Benzler J, Solarsh GC. Population mobility and household dynamics in rural South Africa: implications for demographic and health research. *Southern Afr J Demogr.* **2005**; *10(1/2)*:43-68. http://www.jstor.org/stable/20853278.
6. Jenkins AL, Gyorkos TW, Culman KN, Ward BJ, Pekeles GS, Mills EL. An overview of factors influencing the health of Canadian Inuit infants. *Int J of Circumpolar Health.* **2003**; *62(1)*:17-39. doi: 10.3402/ijch.v62i1.17526.
7. Randall S, Coast E. Poverty in African Households: the Limits of Survey and Census Representations. *J Dev Stud.* **2015**; *51(2)*:162-77. doi: 10.1080/00220388.2014.968135
8. Chisholm RH, Crammond B, Wu Y, et al. A model of population dynamics with complex household structure and mobility: implications for transmission and control of communicable diseases. *PeerJ.* **2020**; *8*:e10203-e. doi: 10.7717/peerj.10203
9. House T, Keeling MJ. Household structure and infectious disease transmission. *Epidemiol Infect.* **2009**;*137(5)*:654-61. doi: 10.1017/S0950268808001416
10. Jing Q-L, Liu M-J, Zhang Z-B, et al. Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *Lancet Infect Dis.* **2020**; *20(10)*:1141-50. doi: 10.1101/2020.04.11.20056010.

11. Vino T, Singh GR, Davison B, et al. Indigenous Australian household structure: a simple data collection tool and implications for close contact transmission of communicable diseases. *PeerJ*. **201**; *5*:e3958-e. doi: 10.7717/peerj.3958
12. Australian Bureau of Statistics. Census of population and housing: characteristics of Aboriginal and Torres Strait Islander Australians, 2016, 'East Arnhem 2016 census story'—technical report cat no. 2076.0. Canberra, Australia 2018.
13. Jackson ML, Hart GR, McCulloch DJ, et al. Effects of weather-related social distancing on city-scale transmission of respiratory viruses: a retrospective cohort study. *BMC Infect Dis*. **2021**; *21(1)*:335. https://doi.org/10.1186/s12879-021-06028-4
14. Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*. **2008**; *5(3)*:e74. doi: 10.1371/journal.pmed.0050074
15. Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. PloS one. 2014 Aug 15;9(8):e104786. doi: 10.1371/journal.pone.0104786.
16. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proc R Soc B* **2014**; *281(1785)*:20140268. doi: 10.1098/rspb.2014.0268.
17. le Polain de Waroux O, Cohuet S, Ndazima D, et al. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. *BMC Infect Dis*. **2018**; *18(1)*:172. doi: 10.1186/s12879-018-3073-1.
18. Wesolowski A, Buckee CO, Engø-Monsen K, Metcalf CJE. Connecting mobility to infectious diseases: the promise and limits of mobile phone data. *J Infect Dis*. **2016**; *214(suppl_4)*:S414-S20. doi: 10.1093/infdis/jiw273.
19. Australian Bureau of Statistics. Census of population and housing: understanding the census and census data, Australia, 2016. Canberra, Australia 2017.
20. Zachreson C, Mitchell L, Lydeamore MJ, Rebuli N, Tomko M, Geard N. Risk mapping for COVID-19 outbreaks in Australia using mobility data. *J R Soc Interface*. **2021**; *18(174)*: 20200657. doi: 10.1098/rsif.2020.0657.
21. Dockery A, Hampton K. The dynamics of services, housing, jobs and mobility in remote Aboriginal and Torres Strait Islander communities in central Australia. CRC-REP Working Paper. 2015.
22. Bolton KJ, McCaw JM, Forbes K, et al. Influence of contact definitions in assessment of the relative importance of social settings in disease transmission risk. *PloS ONE*. **2012**; *7(2)*:e30893. doi: 10.1371/journal.pone.0030893.
23. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing. 2020.
24. Wickham H. ggplot2. Wiley Interdiscip Rev Comput Stat. 2011 Mar;3(2):180-5.
25. Helwig NE. nptest: nonparametric bootstrap and permutation tests. 1.0-3 ed2021.
26. Australian Bureau of Statistics. Life tables for Aboriginal and Torres Strait Islander Australians, 2015–2017. Canberra, Australia, 2018.
27. Hoang T, Coletti P, Melegaro A, et al. A systematic review of social contact surveys to inform transmission models of close-contact infections. *Epidemiology*. **2019**; *30(5)*:723-736. doi: 10.1097/EDE.0000000000001047
28. Gareta D, Baisley K, Mngomezulu T, et al. Cohort profile update: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *Int J of Epidemiol*. **2021**; *50(1)*:33-4. doi: 10.1093/ije/dyaa264.
29. Tanser F, Hosegood V, Bärnighausen T, et al. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *Int J of Epidemiol*. **2008**; *37(5)*:956-62. doi: 10.1093/ije/dym211.
30. Blower SM, Dowlatabadi H. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *Int Stat Rev*. **1994**;*62(2)*:229-43. doi: 10.2307/1403510.