Journal of Theoretical Biology ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Household demographic determinants of Ebola epidemic risk

Q1 Ben Adams

Q2 *Department of Mathematical Sciences, University of Bath, Bath BA27AY, UK*

HIGHLIGHTS

- Ebola transmission intense within households compared to between households.
- Models show larger households increase epidemic risk if all other conditions same.
- Control by case detection and isolation alone challenging if households large.
- Additional household quarantine can mitigate impact of large households on control.

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form
24 October 2015
Accepted 24 November 2015

Keywords:
Mathematical model
Household
Reproduction number
Quarantine
Case detection

ABSTRACT

A salient characteristic of Ebola, and some other infectious diseases such as Tuberculosis, is intense transmission among small groups of cohabitants and relatively limited indiscriminate transmission in the wider population. Here we consider a mathematical model for an Ebola epidemic in a population structured into households of equal size. We show that household size, a fundamental demographic unit, is a critical factor that determines the vulnerability of a community to epidemics, and the effort required to control them. Our analysis is based on the household reproduction number, but we also consider the basic reproduction number, intrinsic growth rate and final epidemic size. We show that, when other epidemiological parameters are kept the same, all of these quantifications of epidemic growth and size are increased by larger households and more intense within-household transmission. We go on to model epidemic control by case detection and isolation followed by household quarantine. We show that, if household quarantine is ineffective, the critical probability with which cases must be detected to halt an epidemic increases significantly with each increment in household size and may be a very challenging target for communities composed of large households. Effective quarantine may, however, mitigate the detrimental impact of large household sizes. We conclude that communities composed of large households are fundamentally more vulnerable to epidemics of infectious diseases primarily transmitted by close contact, and any assessment of control strategies for these epidemics should take into account the demographic structure of the population.

1. Introduction

The epidemiology of an infectious disease is governed by the way it is transmitted. Many respiratory infections are spread widely and fairly indiscriminately by aerosols. HIV is mainly spread through limited and well defined networks of sexual contacts. The Ebola virus is spread by direct contact with the bodily fluids of an infectious person (Aylward et al., 2014). Consequently transmission is much more intense among members of the same household than in the wider community. A study in Guinea in 2014 found that 82% of transmission occurred in the community and, of this, 81% occurred between family members (Faye et al., 2015). Therefore, the composition of the community in terms of households and the balance of transmission between and within households may be expected to have strong influences on key epidemiological characteristics such as reproduction numbers, the final epidemic size and the impact of control strategies. Here we use a mathematical model to investigate how household structure is likely to influence the epidemiological dynamics of Ebola. We show that, under otherwise equivalent epidemiological conditions, communities composed of larger households are more vulnerable to epidemics, and these epidemics are much harder to control by case detection and isolation unless the whole household is placed under effective quarantine.

Epidemiological models with household structure have been around for some time (Ball, 1999; Ball et al., 1996a,b; Becker and Dietz, 1995) and interest has steadily increased over the last decade. Household structure is appropriate to investigate any scenario

http://dx.doi.org/10.1016/j.jtbi.2015.11.025
0022-5193/

in which it is important to distinguish between random 'mass-action' transmission in the general population and transmission between members of the same family or small, well-defined group. Early household models considered stochastic epidemics. They were analysed with techniques such as branching process theory, typically to derive epidemic thresholds. More recently, deterministic household epidemic models have been introduced, using a similar framework to network epidemic models (House and Keeling, 2008). For these deterministic models the state variables of the system are the proportions of households in given epidemiological states, and ordinary differential equations describe how these variables change. Key epidemiological characteristics such as the intrinsic growth rate and distribution of the number of secondary household infections arising from an infected household can be computed using efficient methods based on the associated Markovian transition matrix (Ross et al., 2010). In recent years there has also been a concerted effort to develop theory for reproduction numbers for epidemics in populations with household structure. The household reproduction number, defined as the expected number of households infected by one infected household in a typical susceptible population, has received the most attention because it is relatively easy to define, construct and calculate. The basic reproduction number, ubiquitous throughout epidemiological theory, was elusive for household models but has recently been carefully derived to have the correct physical interpretation, although construction and calculation remain challenging (Pellis et al., 2012). Both the household and basic reproduction numbers specify epidemic thresholds – the probability of an epidemic is only non-zero if they are greater than 1. Several other reproduction numbers have also been defined for household models. These quantities offer insight into different aspects of the epidemiological dynamics and provide bounds for the basic reproduction number and the critical vaccination rate. Most of them also specify the epidemic threshold (Ball et al., 2014).

Household structured epidemic models have provided some interesting insights. House and Keeling (2008) considered an SIR epidemic model with no mortality, no demographic turnover and permanent immunity. They fixed the within-household transmission rate and, for any given household size, adjusted the between-household transmission rate to maintain a constant intrinsic growth rate. They showed that, under this constraint, household structure leads to a more sustained epidemic phase and larger final epidemic size, with the divergence from the unstructured model greatest in communities with intermediate household size. They also showed that the critical coverage of a responsive vaccination scheme depends on household size, but it is better to vaccinate randomly than target certain households. Ross et al. (2010) considered an SIRS model with no mortality, no demographic turnover and waning immunity. They showed that large households act as amplifiers of infection. So large households and high within-household transmission rates can support a positive epidemic growth rate even if between-household transmission rates are low. Black et al. (2013) considered a similar model with heterogeneous household sizes, SEEIIR infection states (where the inclusion of two exposed and infectious states allows the durations of exposure and infectiousness to be Erlang-2 distributed), and delayed antiviral treatment. They showed that epidemic prevention in communities with larger mean household sizes requires the antiviral treatment to be more effective, or be administered more quickly.

In this paper we will introduce and analyse a household-structured SEIR epidemic model for Ebola. This model incorporates within-household transmission that can be high relative to between-household transmission, and significant infection-induced mortality. Our analysis will focus on the emergent stage of the epidemic, before significant depletion of the susceptible population, although we will also briefly consider the final epidemic size. The structured framework allows us to probe, with specific reference to Ebola, how household size and the balance of transmission between and within households influence epidemic risk, understand the divergence from the epidemiological dynamics generated by conventional unstructured models (effectively households of size 1), and examine how community composition modulates the impact of control strategies based on case detection and quarantine.

2. Model description

In this section we present the household-structured epidemic model as a system of ordinary differential equations, and as a Markovian transition matrix. We discuss how these representations are used to find the intrinsic growth rate, household reproduction number and basic reproduction number, set out the parametrisation and describe the main components of our model analysis.

2.1. Household states

The model population is structured into households using the deterministic framework of House and Keeling (2008). The state of each household is defined by the number of individuals in the household with an infection status of susceptible (s), exposed but not yet infectious (e), infectious (i), and recovered with immunity (r). This state is coded as $\{s,e,i,r\}=seir$. For instance, a household composed of 2 susceptible individuals and 1 infectious individual has state 2010. If the maximum household size is n, then the set of possible household states is $\mathcal{S}=seir\,|\,0\le s,e,i,r\le n$ and $s+e+i+r\le n$. Note that we assume that initially all households are of size n, but the household size is not constant because of infection-induced mortality. The epidemiological system describes the rates of change of the proportion of households in each state, H_{seir}. The dynamics are specified by considering the rate at which individuals in each class of households experience epidemiological 'events'. Exposed individuals become infectious at rate η. Infectious individuals recover at rate γ, die at rate μ, make transmissible contact with other members of their household at rate τ, and make transmissible contact with members of other households (i.e. with the wider community) at rate α. These contact rates are frequency dependent. They are independent of the population density in the household, or in the community as a whole.

If the maximum size of all households is $n=1$ then households are equivalent to individuals and the model is equivalent to the standard SEIR model. There are four household states $S=H_{1000}, E=H_{0100}, J=H_{0010}, R=H_{0001}$ and a fifth implicit state in which the household is empty $Z=H_{0000}$. There is no within-household transmission and the system is given by Eq. (1a)–(1d).

2.1.1. Model with no within-household transmission

$$\dot{S}=-\alpha\frac{SJ}{N} \tag{1a}$$

$$\dot{E}=\alpha\frac{SJ}{N}-\eta E \tag{1b}$$

$$\dot{J}=\eta E-(\gamma+\mu)J \tag{1c}$$

$$\dot{R}=\gamma J \tag{1d}$$

where $N=S+E+J+R$ is the size of the extant population which is not constant because households/individuals may be in state Z.

Please cite this article as: Adams, B., Household demographic determinants of Ebola epidemic risk. J. Theor. Biol. (2015), http://dx.doi.org/10.1016/j.jtbi.2015.11.025

The basic reproduction number is $R_0 = \frac{\alpha}{\gamma+\mu}$, which is also the household reproduction number R_*.

For households of size n the general system is given by Eq. (2). This equation applies for all states *seir* in $\mathcal{S}$ if we adopt the convention that the proportions of households in any states generated on the right hand side of Eq. (2) that are not in $\mathcal{S}$ are set to 0. The terms in the square brackets in Eq. (2) correspond, in the order they appear, to between-household transmission, within-household transmission, progression from exposure to infectiousness, recovery and mortality. Focussing on the between-household transmission term, each of the s susceptible people in a household with state *seir* makes contact with individuals in the wider community at rate α. A proportion I of these contacts are with infectious individuals and result in transmission. Hence between-household transmission transforms state *seir* to state $(s-1)(e+1)ir$ at rate αsI. Similarly, between-household transmission transforms state $(s+1)(e-1)ir$ to state *seir* at rate $\alpha(s+1)I$. The other terms are constructed in a similar way. For within-household transmission, the proportion of a household in state *seir* that is infectious is i/m. Each of the s susceptible people in a household makes contact with other individuals in household at rate τ, a proportion $i/(m-1)$ of these contacts are with the infectious individuals. Hence the within-household transmission rate is $\tau si/(m-1)$.

2.1.2. Model with within-household transmission

$$\begin{aligned}\dot{H}_{seir} = {} & \alpha I\left[-sH_{seir}+(s+1)H_{(s+1)(e-1)ir}\right]\\ & +\tau\left[-s\frac{i}{m-1}H_{seir}+(s+1)\frac{i}{m-1}H_{(s+1)(e-1)ir}\right]\\ & +\eta\left[-eH_{seir}+(e+1)H_{s(e+1)(i-1)r}\right]\\ & +\gamma\left[-iH_{seir}+(i+1)H_{se(i+1)(r-1)}\right]\\ & +\mu\left[-iH_{seir}+(i+1)H_{se(i+1)r}\right] \qquad (2)\end{aligned}$$

where $m = s+e+i+r$ is the current household size, and I is the proportion of the total population that is infectious. So $I = \frac{\sum_{\mathcal{S}} i.H_{seir}}{\overline{n}}$ with $\overline{n} = \sum_{\mathcal{S}} m.H_{seir}$ the current average household size.

2.2. Markovian transition matrix

If the maximum household size is n=2, 4, 6 there are, respectively, 15, 70, 210 possible household states. For such large systems it can be expedient to write the model as a Markovian transition matrix Q. Following the methods described in Ross et al. (2010), this matrix is relatively simple to construct algorithmically. An index j is assigned to each state in $\mathcal{S}$. Element $q(j,k)$ of Q is the rate of transition from state j to state k if $j \neq k$ and $q(j,j) = -\sum_{k\neq j} q(j,k)$. It is straightforward to determine the transition rate between two states, as shown in Table 1.

2.3. Case detection and quarantine

Infectious disease epidemics can be controlled by efficient case detection, isolation and quarantine. The model described by Eq. (2) can be modified to include this process by extending the set of household states. The states in $\mathcal{S}$ are as before, but are augmented by a set $\mathcal{S}_q = \{seirq \mid 0 \le s,e,i,r \le n \text{ and } s+e+i+r \le n\}$. So each household state *seir* has a partner state *seirq* in which the household is under quarantine. Households with i infectious individuals are detected at rate ξi. Upon detection each infectious individual in the household moves to the immune state r with probability $\gamma/(\gamma+\mu)$ or dies with probability $\mu/(\gamma+\mu)$. In reality, infectious individuals would probably be removed to a treatment centre and either die or recover and return home after a delay. Approximating this process by reducing the delay to zero simplifies the model, and has minimal impact on the dynamics since the only role of immune individuals is to slightly dilute the frequency-dependent within-household transmission rate. Upon detection, infectious households are also placed under quarantine. While under quarantine, epidemiological dynamics within the household continue as before but for individuals in that household the between household contact rate is reduced to $\theta\alpha$ where $0 \le \theta \le 1$. This change affects outgoing and incoming transmission. The parameter θ is the quarantine efficiency. When $\theta = 0$ quarantine is perfect. When $\theta = 1$ quarantine has no effect, but detection still results in the removal of all infectious individuals. Households exit from quarantine at rate ζ. The transition rates involving the additional states in the quarantine model are given in Table 2. These rates can be used to construct differential equations similar to Eq. (2) in the obvious way. If the case detection rate is ξ, the probability that a case is detected (directly) before recovery or death is $\rho = \xi/(\gamma+\mu+\xi)$.

2.4. Intrinsic growth rate and reproduction numbers

Initially model epidemics (and most real epidemics) grow approximately exponentially. This exponential growth continues until depletion of the susceptible population introduces significant non-linearity into the transmission terms. In the exponential growth phase infections increase at a constant rate r, the intrinsic growth rate. The epidemiological reproduction numbers also remain constant. The intrinsic growth rate r is the largest eigenvalue of the Jacobian matrix J_0 that approximates the system by linearising about the disease-free equilibrium. Element $J_0(j,k)$ is the rate of change of state j with respect to state k. Consequently, for all transitions other than between-household infection, $J_0(j,k) = Q(k,j)$. Since we assume that all households are of size n linearisation about the disease free equilibrium means that almost all households are in state $n000$. So, between-household transmission is always a transition from state $n000$ to $(n-1)000$. Approximately, $H_{n000} = 1$, the mean household size $\overline{n} = n$ and $I = \frac{\sum_{\mathcal{S}} i.H_{seir}}{n}$. Consequently the rate of change of H_{n000} with respect to H_{seir} for any other state *seir* is $\frac{\alpha i}{n}(-n) = -\alpha i$ where i is the number of infectious individuals in a household in state *seir*. It

Table 1
Transition rates from state j to state k. The infection transition combines within and between-household transmission.

State j	State k	Transition	Rate $q(j,k)$
seir	$(s-1)(e+1)ir$	Infection	$\alpha sI + \tau s\frac{i}{s+e+i+r-1}$
seir	$s(e-1)(i+1)r$	Progression	η
seir	$se(i-1)(r+1)$	Recovery	γ
seir	$se(i-1)r$	Mortality	μ

Table 2
Transition rates for the additional states in the quarantine model. Transition rates for the *seir* states not affected by quarantine are as in Table 1 except that for all transitions associated with between-household transmission the proportion of the contactable population that is infectious is $I = \frac{\sum_{\mathcal{S}} i.H_{seir} + \theta\sum_{\mathcal{S}_q} i.H_{seirq}}{\sum_{\mathcal{S}}(s+e+i+r)H_{seir} + \theta\sum_{\mathcal{S}_q}(s+e+i+r)H_{seirq}}$.

State j	State k	Transition	Rate $q(j,k)$
seir	$se0(r+\hat{i})q$	Detection and quarantine with $0 \le \hat{i} \le i$ recoveries	$\xi i\binom{i}{\hat{i}}p^{\hat{i}}(1-p)^{i-\hat{i}}$ where $p = \gamma/(\gamma+\mu)$
seirq	*seir*	Quarantine exit	ζ
seirq	$(s-1)(e+1)irq$	Infection under quarantine	$\theta\alpha sI + \tau s\frac{i}{s+e+i+r-1}$
seirq	$s(e-1)(i+1)rq$	Progression under quarantine	η
seirq	$se(i-1)(r+1)q$	Recovery under quarantine	γ
seirq	$se(i-1)rq$	Mortality under quarantine	μ

Please cite this article as: Adams, B., Household demographic determinants of Ebola epidemic risk. J. Theor. Biol. (2015), http://dx.doi.org/10.1016/j.jtbi.2015.11.025

follows that for states $j = n100$ and $k = seir$, $J_0(j,k) = -\alpha i$, and for states $j = (n-1)100$ and $k = seir$, $J_0(j,k) = \alpha i$. It does not take long to find the eigenvalues numerically, even for large matrices but, if necessary, computationally more efficient methods using the transition matrix Q directly are detailed in Ross et al. (2010).

The household reproduction number R_* is straightforward to construct or calculate from the Q matrix (Ross et al., 2010). Briefly, a state is transient if the rate at which households leave that state is non-zero $q(j,j) \neq 0$. Otherwise the state is absorbing. Let $\mathcal{C}$ be the set of all transient states. The full transition matrix Q is reduced to the transition matrix of transient states $Q_{\mathcal{C}}$ by removing the rows and columns corresponding to absorbing states. A 'reward' function of the household state is defined $f(j) = i$ where i is the number of infectious individuals in a household with state j. If $X(t)$ is a continuous-time Markov process taking values in $\mathcal{C}$ then the path integral $\Gamma = \int_0^\infty f(X(t))dt$ is the 'total reward over the lifetime of the process'. The expected infectious period is $1/(\gamma+\mu)$. So, each time the Markov process $X(t)$ adds a new infectious individual, the expected 'reward' Γ increases by $1/(\gamma+\mu)$. The total expected number of infectious individuals is equal to the total expected reward multiplied by $\gamma+\mu$. The expected household epidemic size $((\gamma+\mu)\Gamma)$ in a household initially in state j, assuming there are no further infections from outside, is $H_j^\infty = (\gamma+\mu)e_j$ where the e_j are found by solving the linear system (Ross et al., 2010)

$$\sum_{k \in \mathcal{C}} q(j,k)e_k + f(j) = 0. \tag{3}$$

When the entire population is initially susceptible, and all households have size n, the state of all households is $seir = n000$ and the household reproduction number $R_* = \frac{\alpha}{\gamma+\mu} H_j^\infty = \alpha e_j$ where $\frac{\alpha}{\gamma+\mu} := \mu_G$ is the expected number of between-household infections arising from one infectious individual and state j is $(n-1)100$. For $n = 2, R_* = \alpha \frac{\gamma+\mu+2\tau}{(\gamma+\mu)(\gamma+\mu+\tau)}$. This method can also be used to find analytic expressions for R_* for larger values of n, certainly up to $n = 6$ but these expressions are so complicated that their utility is doubtful. The Supplementary Material includes the expected household epidemic sizes for all initial states when $n = 2$, and expressions for the household reproduction numbers when $n = 1, 2$ and 3.

The basic reproduction number R_0 for household models is not straightforward to either define or calculate. However, Pellis et al. (2012) have developed a method based on the next generation matrix (Diekmann et al., 2013). Briefly, the initial infected individual in a household is assigned rank 1 and generation 1. Then, as the epidemic unfolds within the household, an individual is assigned generation j if they are first infected by an individual in generation $j-1$. A susceptible individual will always be infected by the first transmissible contact event in which they are involved. However, that individual may be involved in subsequent encounters with infected individuals that would also have resulted in infection, had it not already occurred. These are also considered transmissible contact events. An individual is assigned rank j if j is the length of the shortest sequence of transmissible contact events that connects them to the initial infected individual in that household; this includes the transmissible contacts that occur after the individual was first infected. Then μ_j is defined such that $\mu_0 = 1$ is the initial infected individual in a household and μ_j is either the expected number of infections of generation j, or the expected number of infections of rank j. It follows that each infected individual for rank (or generation) j causes, on average, μ_{j+1}/μ_j new infections. Then, the basic reproduction number is $R_0 = \rho(K)$ where ρ is the spectral radius and

$$K = \begin{pmatrix} \mu_G & \mu_G & \cdots & \mu_G & \mu_G \\ \mu_1/\mu_0 & 0 & \cdots & 0 & 0 \\ 0 & \mu_2/\mu_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mu_{n-1}/\mu_{n-2} & 0 \end{pmatrix}. \tag{4}$$

Deriving expressions for the expected number of infections in each generation of the household epidemic μ_j is challenging. Pellis et al. (2012) give a method for an SIR model with no mortality. The same result holds for SEIR models with no mortality since the exposed state introduces a delay but does not affect the expected number of infections. We employ this method, as described in the Supplementary Material, to get an approximate value of R_0 for our model under the assumption of a constant household size. The result is not exact because disease-induced mortality may cause the household size to decrease which increases the contact rate between the remaining individuals. We also calculated R_0 using expected generation size, and expected rank sizes, found by stochastic simulation; for each parameter set the Gillespie method was used to simulate 10,000 household epidemics starting with one infectious individual. Since each of these epidemics is small, this method is remarkably efficient.

Other individual reproduction numbers have been suggested for household models. These are easier to construct than the basic reproduction number and, in addition to being useful in their own right, provide bounds for R_0 (Goldstein et al., 2009; Ball et al., 2014). The reproduction number R_I approximates R_0 by assuming that the initial infected individual in a household is directly responsible for all subsequent infections. If the expected size of the household epidemic, excluding the initial case, is $\mu_L = H_j^\infty - 1$ then $R_I = \frac{\mu_G + \sqrt{\mu_G^2 + 4\mu_G\mu_L}}{2}$. The reproduction number R_{HI} is the expected number of infections arising directly from a typical infected individual in a typical household. The μ_L infections in a household are caused by $\mu_L + 1$ infected individuals, including the initial infection. So, the expected number of infections per infected individual, including between-household transmission, is $R_{HI} = \mu_G + \frac{\mu_L}{1+\mu_L}$. Ball et al. (2014) show that if $R_0 > 1$ then $R_* > R_I > R_0 > R_{HI}$ and if $R_0 < 1$ then $R_* < R_I < R_0 < R_{HI}$.

2.5. Parametrisation

For all of the numerical results presented here, parameter values were assigned as in Table 3, unless otherwise stated. The maximum household size n was fixed between 1 and 6. Our methodology readily allows higher values but at high computational cost. For reference, in 2012/2013 the average household sizes in Liberia, Sierra Leone and Guinea were 5, 5.9 and 6.3 respectively (Sierra Leone, 2013; Liberia Institute of Statistics, 2013; Institut National de

Table 3
Parameter values used for all numerical results, unless otherwise stated. The original values have rates of per day. The rescaled values are such that the expected infectious period $1/(\gamma+\mu) = 1$.

Parameter	Meaning	Original value	Rescaled value
n	Maximum household size	1–6	1–6
α	Between-household contact rate	0–0.2	0–2
τ	Within-household contact rate	0–0.5	0–5
η	Infectivity progression rate	0.11	1.1
γ	Recovery rate	0.03	0.30
μ	Infection induced mortality rate	0.07	0.70
ξ	Case detection rate	0–1	0–10
θ	Quarantine efficiency	0–1	0–1
ζ	Quarantine exit rate	0.042	0.42
r	Intrinsic growth rate	0.04	0.4

la Statistique and ICF International, 2012). Observational data from the emerging Ebola epidemic in Liberia, Sierra Leone and Guinea collected between December 2013 and September 2014 (Aylward et al., 2014) report an average duration of incubation $1/\eta = 9.4$ days. The case fatality rate was $m = 70.8\%$, the average time from onset of symptoms until death was $T_m = 7.5$ days and the average time from onset until recovery was $T_r = 16.4$ days, giving a mean infectious period of $mT_m + (1-m)T_r = 10.2$ days. So, for our model, setting the expected infectious period (time until recovery or death) $1/(\gamma+\mu) = 10.2$ days and the proportion of infections resulting in death $\mu/(\gamma+\mu) = 0.7$ gives a mortality rate $\mu = 0.07$ and recovery rate $\gamma = 0.03$. From the same observational data the basic reproduction number was estimated to be between 1.7 and 2, and epidemic doubling time between 12.8 and 17.5 days, which corresponds to an intrinsic growth rate of 0.04–0.05 per day. In August 2014, the doubling time was estimated to be between 15.7 and 30.2 days, corresponding to an intrinsic growth rate of 0.02–0.04 per day. Our numerical results were carried out using parameter values that were rescaled by dividing by $\gamma+\mu$. With these rescaled parameters, the expected duration of infection is 1 and an infectious individual is expected to make a total of α transmissible contacts between households, and τ transmissible contacts within their household (but τ infections will not occur in the household due to saturation), and the case detection probability is $\rho = \xi/(1+\xi)$. In analyses where we fix the intrinsic growth rate we use a value of $r = 0.04$ per day, which is re-scaled to $r = 0.4$.

2.6. Model analysis

The analysis of the model presented here will focus on how household size and the intensity of within-household transmission affect epidemic risk, epidemic size and epidemic management by quarantine. The epidemic risk is assessed through the reproduction numbers, primarily the household reproduction number R_*. When $R_* < 1$ significant epidemics are not expected to occur in stochastic systems; in deterministic systems the disease-free equilibrium is stable. For $R_* > 1$ higher values of R_* indicate a higher probability of a significant epidemic when an infected individual is introduced into a disease-free population; deterministically, the disease-free equilibrium is unstable. The epidemic size is calculated by solving the deterministic system (2) numerically from an initial condition consisting of a small number of infections in an otherwise susceptible population ($H_{n000} = 0.99, H_{(n-1)100} = 0.01, H_{seir} = 0$ otherwise) until the epidemic peak has passed and the infectious proportion of the population $\frac{1}{n}\sum_S i.H_{seir} < 0.01$. Two measures of the final epidemic size are examined based on the household proportions H_{seir} at the end of the epidemic. $R^\infty = 1 - \frac{1}{n}\sum_S s.H_{seir}$ is the total proportion of individuals in the initial population that were infected. $R_*^\infty = 1 - H_{n000}$ is the total proportion of households in the population in which at least one person was infected. The potential for epidemic management by case detection and quarantine is assessed by using numerical root-finding to determine the critical case detection rate ξ^* required to move the household reproduction number to the epidemic threshold $R_* = 1$ beyond which the disease-free state is stable. Although we mainly examine how household size and the balance of transmission between and within households affect R_*, R^∞ and R_*^∞, we also consider the sensitivity of R_* to all parameters using elasticity analysis. For each parameter p, the elasticity of R_* with respect to p is defined as $eR_{*p} = \frac{\partial \ln R_*}{\partial \ln p} = \frac{p}{R_*}\frac{\partial R_*}{\partial p}$ (Caswell, 2000) and quantifies the proportional response in R_* to a proportional perturbation in p.

We consider the role of within-household transmission from two perspectives. Initially, we fix the between-household transmission rate α and consider the impact of varying the within-household transmission rate τ for different household sizes n. Here the intrinsic growth rate r also varies with τ. So we also consider models in which r is held constant by covariation of α and τ. The intrinsic growth rate of an epidemic can be directly observed, and this constraint produces models that are in some sense comparable. House and Keeling (2008) consider three methods to maintain a constant r: fix τ and for any given n adjust α; for any given n adjust both α and τ but fix their ratio; fix the proportion of infections in the early stages of the epidemic that occur within households. Here, for any given τ and n we adjust α to maintain $r = 0.4$.

3. Results

3.1. Within-household transmission varied, between-household transmission constant

Fig. 1a–c shows how the intrinsic growth rate r and reproduction numbers R_*, R_0 depend on the maximum household size n and the within-household transmission rate τ when the between-household transmission rate is fixed, $\alpha = 0.9$. The results are similar for other values of α. The intrinsic growth rate and reproduction numbers are higher in populations composed of larger households or with more intensive within-household transmission. The number of individuals in a household that can be infected is limited by the household size, regardless of the transmission intensity. So the impact of within-household transmission saturates, but less quickly for larger households. Elasticity analysis of R_* (Supplementary Fig. 1) shows that increasing the

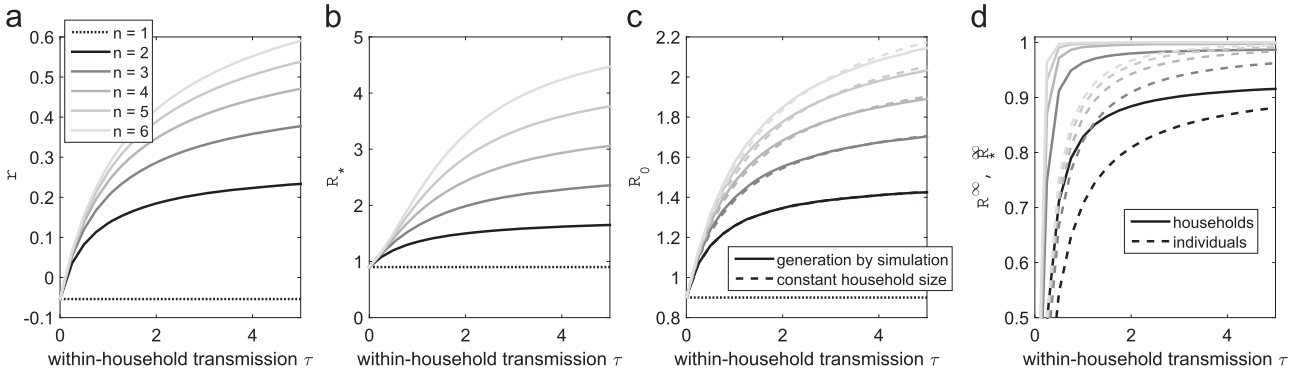

Fig. 1. Intrinsic growth rate, reproduction numbers and epidemic size when the between-household transmission rate α is fixed and the within-household transmission rate τ is varied. (a) Intrinsic growth rate r. (b) Household reproduction number R_*. (c) Basic reproduction number R_0, computed using the expected generation size from simulations, and under the approximation of a constant household population size. (d) Proportion of households R_*^∞, and proportion of individuals R^∞, infected over the course of an epidemic in an initially susceptible population. In all panels line shade indicates the household size n as specified in the legend of (a), $\alpha = 0.9$ and other parameters are as in Table 3. So when $n = 1$, $R_0 = R_* = \alpha$.

disease-induced mortality rate μ or the recovery rate γ reduce R_*, because the transmission window narrows. Increasing the within-household transmission rate τ increases R_*. The impact of all three of these parameters is strongest for large household sizes n and intermediate within-household transmission rates (around $\tau=1$). The duration of the exposed state η has no effect on R_*, and the impact of the between-household transmission rate α does not depend on the household size or the within-household transmission rate. Fig. 1c shows that the basic reproduction numbers calculated under the approximation that the household size remains constant agrees well with that calculated using the expected generation size of the household epidemic. Supplementary Fig. 1 shows that using the expected rank size rather than the expected generation size has little impact and, although the individual reproduction numbers R_{HI} and R_I do bound R_0 as expected, these bounds are quite broad, particularly when n is large. Fig. 1d shows how the final epidemic size R^∞, R_*^∞ depends on the maximum household size n and the within-household transmission rate τ. The proportion of households that experience at least one infection is higher in populations composed of larger households and with more intensive within-household transmission. For $n>2$ and $\tau>1$, almost all households experience infection. In comparison to the proportion of households, the proportion of individuals that are infected is smaller, and increases more slowly with n and τ.

Fig. 2a–c shows the critical probability with which a case must be detected before death or recovery $\rho^*=\xi^*/(1+\xi^*)$ to prevent an epidemic, or stop it during the initial exponential growth phase. When infectious individuals are isolated following detection, but the household is not quarantined ($\theta=1$), in communities of small households ($n=2$) the critical detection probability is below 0.4, even when there is intense within-household transmission ($\tau=5$). The critical detection proportion increases markedly with any increase in household size. When $\tau=5$, ρ^* increases from 0.34 when $n=2$, to 0.53 when $n=3$ to 0.82 when $n=6$. The general pattern is similar when the household is quarantined following case detection, but quarantine means that epidemic prevention is achieved at a lower case detection rate. Fully effective quarantine ($\theta=0$) reduces the critical detection probability in communities of large households considerably; when $n=6$ and $\tau=5$, ρ^* is reduced from 0.82 to 0.52. But the relative impact is smaller in communities of small households; when $n=2$ and $\tau=5$, effective quarantine only reduces the critical detection probability from 0.34 to 0.31.

3.2. Within and between-household transmission co-varied such that intrinsic growth rate constant $r=0.4$

Fig. 3a shows that an intrinsic epidemic growth rate of $r=0.4$ is consistent with a broad range of between and within-household transmission regimes. Higher within-household transmission τ is balanced by lower between-household transmission α. But the amplification capacity of households is limited by their size so the impact of increasing τ saturates and some between-household transmission is always required. For any given intensity of within-household transmission, in communities composed of larger households the target intrinsic growth rate is achieved at lower between household transmission rates.

Fig. 3b and c shows how the reproduction numbers R_* and R_0 depend on the maximum household size n and the within-household transmission rate τ when the between-household transmission rate α is adjusted to maintain a constant intrinsic growth rate. In the absence of household structure i.e. when $\tau=0$ or $n=1$, the household reproduction number is equal to the basic reproduction number. If $\tau\neq 0$, larger households always increase the household reproduction number associated with the same intrinsic growth rate. For communities composed of large households ($n=6$) with intermediate intensity within-household transmission, the household reproduction number is up to 65% higher than the household (basic) reproduction number derived from an unstructured population composed of 'households' of size 1. For small household sizes ($n=2,3$), R_* saturates as within-household transmission increases. For larger household sizes ($n=4-6$), the between/within-household transmission trade-off is such that R_* reaches a maximum when τ is around 2 and then drops back. In contrast to the household reproduction number, household structure and within-household transmission always decreases the basic reproduction number associated with the same intrinsic growth rate. R_0 is lower when household size is larger, or within-household transmission is more intense. However, these factors have a modest impact on R_0 compared with R_*. The greatest divergence from the basic reproduction number derived without household structure occurs for communities composed of large households with high intensity within-household transmission. For $n=6$ and $\tau=5$ the community structure reduces R_0 by around 10%. Fig. 1d shows how the final epidemic size R^∞, R_*^∞ depends on the maximum household size n and the within-household transmission rate τ. The between/within-household transmission trade-off to maintain a constant intrinsic growth rate is such that the total proportion of households eventually infected is higher if it is derived from communities of larger households and, for all household sizes, lower if it is derived from communities with

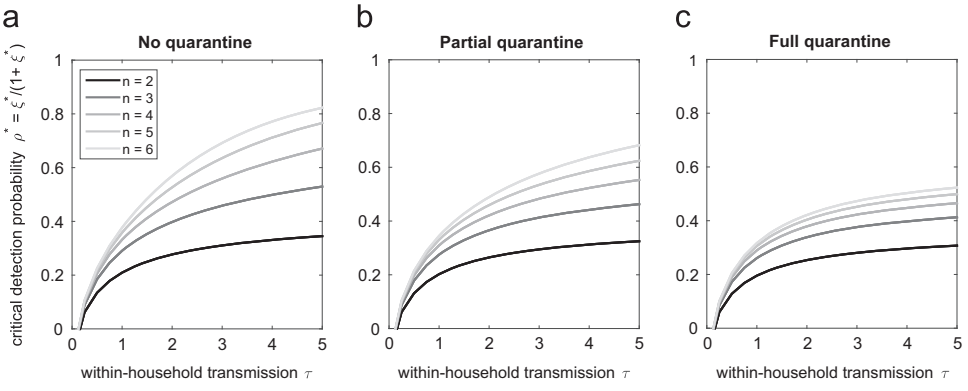

Fig. 2. Critical probability with which cases must be detected to prevent an epidemic growing ($\rho^*=\xi^*/(1+\xi^*)$ such that $R_*=1$) when the between-household transmission rate α is fixed and the within-household transmission rate τ is varied. On detection, all infected individuals are removed from the household and it may be placed under quarantine. (a) No quarantine ($\theta=1$). (b) Partially effective quarantine ($\theta=0.5$). (c) Fully effective quarantine ($\theta=0$). In all panels line shade indicates the household size n as specified in the legend of (a), $\alpha=0.9$ and other parameters are as in Table 3.

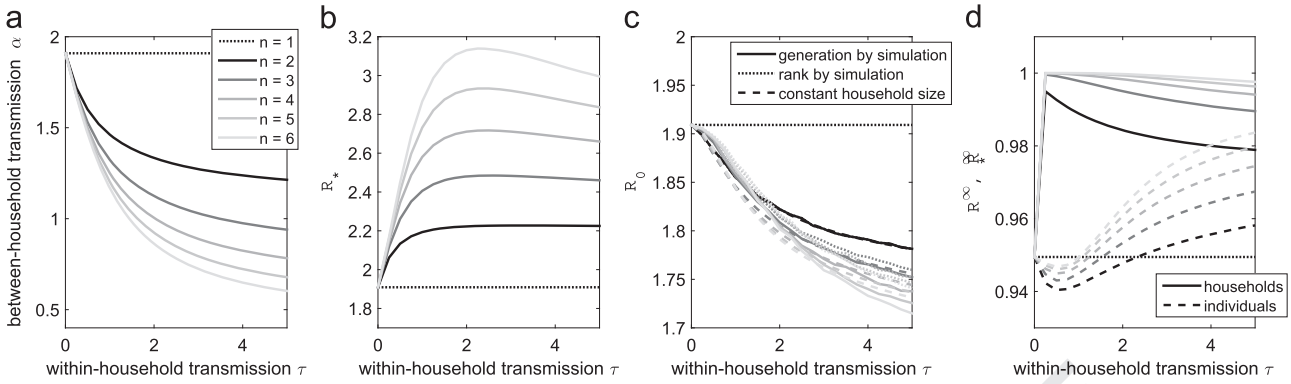

Fig. 3. Reproduction numbers and epidemic size when the between-household α and within-household τ transmission rates are co-varied such that the intrinsic growth rate remains $r=0.4$. (a) α - τ pairs that give $r=0.4$. (b) Household reproduction number R_*. (c) Basic reproduction number R_0, computed using the expected generation size and the expected rank size from simulations, and under the approximation of a constant household population size. (d) Proportion of households R_*^∞, and proportion of individuals R^∞, infected over the course of an epidemic in an initially susceptible population. In all panels line shade indicates the household size n as specified in the legend of (a). Parameters are as in Table 3.

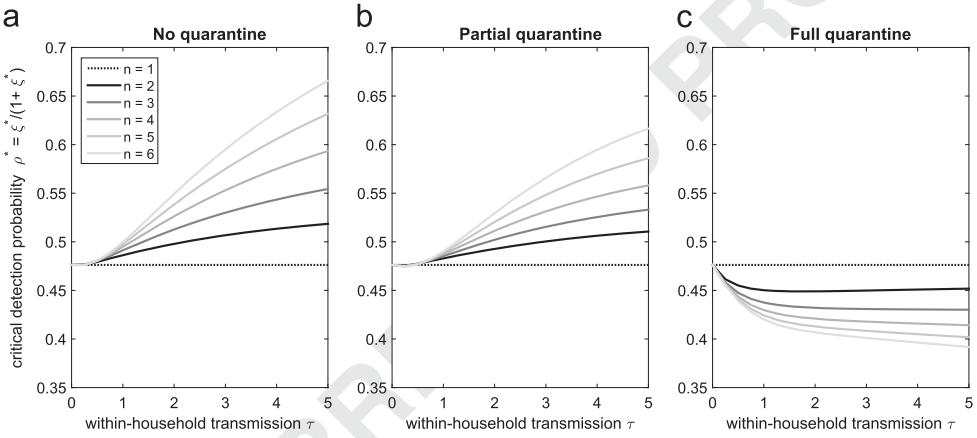

Fig. 4. Critical detection probability ($\rho^* = \xi^*/(1+\xi^*)$ such that $R_* = 1$) when the between-household α and within-household τ transmission rates are co-varied such that the intrinsic growth rate remains $r=0.4$. On detection, all infected individuals are removed from the household and it may be placed under quarantine. (a) No quarantine ($\theta = 1$). (b) Partially effective quarantine ($\theta = 0.5$). (c) Fully effective quarantine ($\theta = 0$). In all panels line shade indicates the household size n as specified in the legend of (a). Other parameters are as in Table 3.

more intense within-household transmission. Conversely, the total proportion of individuals eventually infected is lowest when derived from communities with weak, but non-zero, within-household transmission and generally increases when within-household transmission is more intense. Larger household sizes amplify these effects.

Fig. 4a–c shows the critical probability with which cases must be detected before death or recovery $\rho^* = \xi^*/(1+\xi^*)$ to prevent an epidemic when the within and between-household transmission rates are co-varied to maintain a constant intrinsic growth rate. When infectious individuals are isolated following detection but the household is not quarantined ($\theta = 1$) the critical detection probabilities derived from communities with household structure are always higher than ρ^* derived from an unstructured community. The divergence is greater when households are larger and within-household transmission is more intense. When $n=1$, $\rho^* = 0.38$ but this increases to $\rho^* = 0.61$ when $n=6$ and $\tau = 5$. When quarantine is partially effective ($\theta = 0.5$) the critical detection probabilities are generally lower. Values of ρ^* derived from communities with household structure are higher than those derived from an unstructured community unless within-household transmission is weak ($0 < \tau < 0.5$) in which case they are marginally lower. When quarantine is fully effective ($\theta = 0$) the critical detection probabilities derived from communities with household structure are always lower than ρ^* derived from an unstructured community, although the divergence is relatively small.

4. Discussion

We have used an epidemic model in which the population is structured into households to examine how demography, in terms of the size of the households that make up a community, affects the vulnerability to epidemics of Ebola and other infectious diseases where a significant proportion of transmission occurs between cohabitants. Much of our analysis has been based on the household reproduction number. This quantity specifies an epidemic threshold. Large outbreaks are not expected to occur if $R_* < 1$. The assumptions behind the household reproduction number mean that it is meaningful as long as the susceptible population is sufficiently large that almost all of the between-household transmissible contacts made by an infectious individual are with susceptible individuals. This may be a reasonable approximation, and so the insights from our model will be applicable, for quite an extended period in the emergent phase of an epidemic. So R_* can be used to assess control strategies that may be implemented at any time during that period, as well as to quantify the vulnerability of a disease-free population to an epidemic. Here we have considered control case identification and

Please cite this article as: Adams, B., Household demographic determinants of Ebola epidemic risk. J. Theor. Biol. (2015), http://dx.doi.org/10.1016/j.jtbi.2015.11.025

quarantine, but the model framework could also be used to examine other measures such as 'lock-down'.

We have shown that, all else being equal, in the early stages of an epidemic, the intrinsic growth rate, the basic reproduction number and the household reproduction number are all higher if households are larger. This divergence is accentuated by more intense within-household transmission. At the end of the epidemic, almost all households will have had at least one infection unless the community is composed of small households or within-household transmission is very weak. The proportion of individuals that are eventually infected is smaller than the proportion of households, but increases if households are larger or within-household transmission is more intense.

We also considered the effort required to control an emergent epidemic by identifying cases, isolating them and placing their household under quarantine. We showed that in communities composed of small households a fairly modest, and likely achievable, case detection probability is sufficient to halt or prevent an epidemic, even without household quarantine. In larger households this critical detection probability is much higher. Without quarantine, each additional person in the household makes epidemic control much more challenging to achieve. However, combining case detection and isolation with effective quarantine of the whole household can greatly reduce the critical detection probability, even in communities of large households with intense within-household transmission. This results indicates that, when an infectious case is detected in a large household other members of the household are likely to be infected, but not yet infectious, and preventing further transmission by these individuals is a crucial component of epidemic control.

The intrinsic growth rate of an epidemic can be estimated directly from incidence data. The basic or household reproduction number may be inferred from the data, or the intrinsic growth rate, with an assumed transmission model (e.g. Fraser, 2007). We have shown that, when a significant part of transmission occurs within households, the details of this model are important. The same intrinsic growth rate is consistent with a range of within-household and between-household transmission rate pairs, which depend on the household size of the community. If the intrinsic growth rate is invariant, the basic reproduction number R_0 is not very sensitive to the balance of transmission within and between households, or the household size; the intrinsic growth rate and basic reproduction number are approximately proportional. However, the household reproduction number is more sensitive. In general the basic reproduction number may be useful because it indicates the proportion of new infections $(1-1/R_0)$ that need to be prevented in order to halt epidemic growth. However, control strategies may not prevent infections directly. We have shown that the control effort associated with case detection, isolation and household quarantine can be sensitive to assumptions about the transmission balance and household size, even if the intrinsic growth rate (and so the basic reproduction number) is invariant. In comparison to a model with household structure and within-household transmission, an unstructured model underestimates the critical case detection probability if quarantine is ineffective, but slightly overestimates it if quarantine is effective.

Our analysis shows that the demographic structure of a community can be a critical factor influencing epidemic risk and control. If a significant component of transmission occurs within-households, large households act as amplifiers, the risk of an epidemic is greater, and the number of cases grows faster. The mean household sizes in the West African countries afflicted by the 2014 Ebola epidemic are between 5 and 6.3, among the highest in the world. For comparison the mean household size in the 34 OECD countries is 2.6 (OECD, 2014). For infectious diseases such as Ebola where the majority of transmission is among family members, putting aside any asymmetries of healthcare infrastructure and resources, demography alone can determine upon which side of the epidemic threshold a population finds itself.

Acknowledgments

This paper was improved by insightful and constructive criticism from two anonymous reviewers.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jtbi.2015.11.025.

References

Aylward, B., Barboza, P., Bawo, L., Bertherat, E., Bilivogui, P., Blake, I., Brennan, R., Briand, S., Chakauya, J.M., Chitala, K., Conteh, R.M., Cori, A., Croisier, A., Dangou, J.M., Diallo, B., Donnelly, C.A., Dye, C., Eckmanns, T., Ferguson, N.M., Formenty, P., Fuhrer, C., Fukuda, K., Garske, T., Gasasira, A., Gbanyan, S., Graaff, P., Heleze, E., Jambai, A., Jombart, T., Kasolo, F., Kadiobo, A.M., Keita, S., Kertesz, D., Kone, M., Lane, C., Markoff, J., Massaquoi, M., Mills, H., Mulba, J.M., Musa, E., Myhre, J., Nasidi, A., Nilles, E., Nouvellet, P., Nshimirimana, D., Nuttall, I., Nyenswah, T., Olu, O., Pendergast, S., Perea, W., Polonsky, J., Riley, S., Ronveaux, O., Sakoba, K., Santhana Gopala Krishnan, R., Senga, M., Shuaib, F., Van Kerkhove, M.D., Vaz, R., Wijekoon Kannangarage, N., Yoti, Z., 2014. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. New Engl. J. Med. 371 (October (16)), 1481–1495.

Ball, FG., 1996a. Threshold behaviour in stochastic epidemics among households. In: Heyde, C.C., Prohorov, Y.V., Pyke, R., Rachev, S.T. (Eds.), Athens Conference on Applied Probability and Time Series. Applied Probability. Lecture Notes in Statistics , vol. 114. Springer-Verlag, New York.

Ball, FG., 1996b. Reproduction numbers and critical immunity levels for epidemics in a community of houseoholds. In: Heyde, C.C., Prohorov, Y.V., Pyke, R., Rachev, S.T. (Eds.), Athens Conference on Applied Probability and Time Series, Applied Probability. Lecture Notes in Statistics vol. 114. Springer-Verlag, New York.

Ball, F.G., 1999. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. Math. Biosci. 156 (March (1–2)), 41–67.

Ball, F.G., Pellis, L., Trapman, P., 2014. Reproduction Numbers for Epidemic Models With Households and Other Social Structures II: Comparisons and Implications for vaccination. Unpublished manuscript arXiv:1410.4469v.

Becker, N.G., Dietz, K., 1995. The effect of the household distribution on transmission and control of highly infectious diseases. Math. Biosci. 127, 207–219.

Black, A.J., House, T., Keeling, M.J., Ross, J.V., 2013. Epidemiological consequences of household-based antiviral prophylaxis for pandemic influenza. J. R. Soc. Interface 10 (April (81)), 20121019.

Caswell, H., 2000. Matrix Population Models. Sinauer Associates.

Diekmann, O., Heesterbeek, H., Britton, T., 2013. Mathematical Tools for Understanding Infectious Disease Dynamics. Princeton University Press.

Faye, O., Boelle, P.Y., Heleze, E., Faye, O., Loucoubar, C., Magassouba, N., Soropogui, B., Keita, S., Gakou, T., Bah, E.H., Koivogui, L., Sall, A.A., Cauchemez, S., 2015. Chains of transmission and control of Ebola virus disease in Conakry, Guinea, in 2014: an observational study. Lancet Infect. Dis. (January)

Fraser, C., 2007. Estimating individual and household reproduction numbers in an emerging epidemic. PLoS One 2 (8), e758.

Goldstein, E., Paur, K., Fraser, C., Kenah, E., Wallinga, J., Lipsitch, M., 2009. Reproductive numbers, epidemic spread and control in a community of households. Math. Biosci. 221 (September (1)), 11–25.

House, T., Keeling, M.J., 2008. Deterministic epidemic models with explicit household structure. Math. Biosci. 213 (May (1)), 29–39.

Institut National de la Statistique and ICF International. Enquete Demographique et de Sante et a Indicateurs Multiples 2012, 2012. URL ⟨http://www.dhsprogram.com/pubs/pdf/FR280/FR280.pdf⟩.

Liberia Institute of Statistics and Geo-Information Services and ICF International. Liberia Demographic and Health Survey 2013, 2013. URL ⟨http://www.dhsprogram.com/pubs/pdf/FR291/FR291.pdf⟩.

OECD. OECD Family Database. SF1.1: Family size and household composition, 2014. URL ⟨http://www.oecd.org/els/soc/SF1_1_Family_size_and_composition.pdf⟩.

Pellis, L., Ball, F., Trapman, P., 2012. Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R0. Math. Biosci. 235 (January (1)), 85–97.

Ross, J.V., House, T., Keeling, M.J., 2010. Calculation of disease dynamics in a population of households. PLoS One 5 (3), e9666.

Statistics Sierra Leone and ICF International. Sierra Leone Demographic and Health Survey 2013, 2013. URL ⟨http://www.dhsprogram.com/pubs/pdf/FR297/FR297.pdf⟩.

Please cite this article as: Adams, B., Household demographic determinants of Ebola epidemic risk. J. Theor. Biol. (2015), http://dx.doi.org/10.1016/j.jtbi.2015.11.025