Math. Model. Nat. Phenom.
Vol. 7, No. 3, 2012, pp. 147–167
DOI: 10.1051/mmnp/20127310

Epidemiological Models With Parametric Heterogeneity: Deterministic Theory for Closed Populations

A.S. Novozhilov[1*]

Applied Mathematics–1, Moscow State University of Railway Engineering
Obraztsova 9, bldg. 9, Moscow 127994, Russia

Abstract. We present a unified mathematical approach to epidemiological models with parametric heterogeneity, i.e., to the models that describe individuals in the population as having specific parameter (trait) values that vary from one individuals to another. This is a natural framework to model, e.g., heterogeneity in susceptibility or infectivity of individuals. We review, along with the necessary theory, the results obtained using the discussed approach. In particular, we formulate and analyze an SIR model with distributed susceptibility and infectivity, showing that the epidemiological models for closed populations are well suited to the suggested framework. A number of known results from the literature is derived, including the final epidemic size equation for an SIR model with distributed susceptibility. It is proved that the bottom up approach of the theory of heterogeneous populations with parametric heterogeneity allows to infer the population level description, which was previously used without a firm mechanistic basis; in particular, the power law transmission function is shown to be a consequence of the initial gamma distributed susceptibility and infectivity. We discuss how the general theory can be applied to the modeling goals to include the heterogeneous contact population structure and provide analysis of an SI model with heterogeneous contacts. We conclude with a number of open questions and promising directions, where the theory of heterogeneous populations can lead to important simplifications and generalizations.

Keywords and phrases: SIR model, heterogeneous populations, distributed susceptibility, final epidemic size, heterogeneous contact structure, power law transmission function

Mathematics Subject Classification: 34C20, 34G20, 92D30

1. Introduction

The real-world populations are *heterogeneous*. The populations consist of individuals, and all the individuals are different. This is a basic fact, which does not require any proof. Individuals can differ in their age, spatial location, social habits, genome compositions, etc. Inasmuch as we aim to model the dynamics of interacting populations, it is an obvious condition to include this heterogeneous structure into the mathematical models. There are different kinds of population heterogeneity; here we only inspect

*Corresponding author. E-mail: anovozhilov@gmail.com

© EDP Sciences, 2012

Article published by EDP Sciences and available at http://www.mmnp-journal.org or http://dx.doi.org/10.1051/mmnp/20127310
https://doi.org/10.1051/mmnp/20127310 Article published by EDP Sciences and available at http://www.mmnp-journal.org or http://dx.doi.org/10.1051/mmnp/20127310

heterogeneity in population parameters (such as, e.g., susceptibility to a specific disease); the parameters are considered as an inherent and invariant property of individuals, whereas the parameter values can vary between individuals. We call such heterogeneity *parametric heterogeneity* ([23]). Taking into account the parametric heterogeneity yields important changes for the population dynamics. Most importantly, it means that both *evolutionary* and *ecological* aspects of the dynamics have to be accounted for. In other words, we must follow not only the total population numbers (ecological aspects) but also the changes of the population structure (evolutionary aspects), which are described by the transformation of the parameter distribution in the population with time.

The simplest way to include the parametric heterogeneity into mathematical models is to divide populations into subgroups, such that each subgroup has its own specific parameter value. By increasing the number of groups we can assume that there is a continuous distribution of the parameter in the population.

To illustrate the mathematical approach to the parametric heterogeneity and present the notations used throughout the text, we start with the simplest possible mathematical model: the Malthus equation

$$\dot{N} = mN, \tag{1.1}$$

where $N = N(t)$ is the total population size at time t. Equation (1.1) has the solution $N(t) = N(0)e^{mt}$ for the given initial condition $N(0)$.

Equation (1.1) has parameter m, which is the per capita growth rate (and is often called the Malthusian parameter). Model (1.1) can be used to describe, e.g., the growth of a bacterial colony. Employing equation (1.1) as a mathematical model to describe the changes of the population size notwithstanding, it is customary assumed that this parameter m is the same for any individual in the population. This can be hardly true for any realistic population. The simplest way to describe the population which has different values of m is to suppose that the total population consists of k subpopulations, each of which has its own parameter value; i.e., we replace (1.1) with the following system of equations:

$$\dot{N}_i = m_i N_i, \quad i = 1, \dots, k, \tag{1.2}$$

where now the total population size is $N(t) = \sum_i N_i(t)$. There are k initial conditions for (1.2): $N_i(0)$, $i = 1, \dots, k$. Is it possible to describe the behavior of the total population size, whose evolution is governed by (1.2), using only one equation? Summing all the equations in (1.2), we obtain

$$\dot{N} = \sum_i m_i N_i = N \sum_i m_i \frac{N_i}{N} = \mathsf{E}_t[m]N = \bar{m}(t)N, \tag{1.3}$$

where a natural notation used for the mean parameter value in the population at the time moment t:

$$\bar{m}(t) = \mathsf{E}_t[m] = \sum_i m_i p_i(t), \quad p_i(t) = \frac{N_i(t)}{N(t)}, \quad \sum_i p_i(t) = 1.$$

Equation (1.3) is very similar to (1.1), the difference is that it depends on the current parameter distribution in the population. If it is possible to find $\mathsf{E}_t[m]$, then the general solution to (1.3) is simply

$$N(t) = N(0) \exp\{\int_0^t \mathsf{E}_\tau[m]\, d\tau\} \tag{1.4}$$

for the initial population size $N(0)$.

An important disadvantage of the subgroup approach is that the heterogeneity within each group is not taken into account if a fixed number of groups is considered. As it was discussed, in case $k \to \infty$, we can conjecture that there is a limiting continuous parameter distribution. For the Malthus equation (1.1) this means that we suppose that each individual is characterized by its own parameter (trait) value ω,

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

such that the density of individuals with the given parameter value is $n(t,\omega)$, and the total population size is $N(t)=\int_\Omega n(t,\omega)\,d\omega$, where Ω is the set of the trait values. Now (1.3) is replaced with

$$\partial_t n(t,\omega)=m(\omega)n(t,\omega), \tag{1.5}$$

where $\partial_t\equiv\frac{\partial}{\partial t}$. For (1.5) we need the initial condition $n(0,\omega)$. Using the notation

$$p(t,\omega)=\frac{n(t,\omega)}{N(t)}$$

for the probability density function (pdf) of the current parameter distribution and the usual notation for the mean parameter value

$$\bar{m}(t)=\mathsf{E}_t[m]=\int_\Omega m(\omega)p(t,\omega)\,d\omega$$

we find that the solution to (1.5) is again given by (1.4). If one can calculate $\mathsf{E}_t[m]$ then the solution to (1.5) is found.

An important remark is worth spelling out. All the models we consider are deterministic. We use the probability theory language to describe evolution of the parameter distributions, however no stochastic effects are included into our models.

Models of the form (1.5) are infinite dimensional systems that describe the evolution of the parameter distribution with time along with the total population size. Such (and more complex) models were treated in [3,5] from the general point of the theory of differential equations in infinite dimensional spaces. However, it turns out that many such models can be reduced to low dimensional systems of ordinary differential equations (ODEs) [37]. The theory of heterogeneous populations, outlined in [37], provides the conditions when the mean parameter value $\mathsf{E}_t[\cdot]$ can be effectively calculated at any time moment using only the knowledge of the initial parameter distribution $p(0,\omega)$. Therefore the heterogeneous models, of which (1.5) is the simplest example, can be studies analytically. The examples include the analysis of heterogeneous Malthus equation (1.5) [35, 36], Lotka–Volterra system [50] (see also [4]), tumor cell dynamics [40], and the replicator equation [39].

A very rich field for the mathematical models with parametric heterogeneity is theoretical epidemiology, where it is natural to assume that individuals vary with respect to their susceptibility to a disease, infectivity to pass a pathogen, contact number, etc. (see, e.g., [6, 14–16, 19, 20, 24–26, 49, 60]). For many models from the cited literature, the general approach of the theory of heterogeneous populations yields important analytical results [51,52]. It is the main goal of the present manuscript to review and generalize these results.

The rest of the paper is organized as follows. In Section 2 we present the necessary mathematical facts from the theory of heterogeneous populations with parametric heterogeneity. Section 3 is devoted to formulation of various mathematical models of the epidemic spread that include parametric heterogeneity. In Section 4 the machinery from Section 2 is applied to the models from Section 3. Finally, Section 5 is devoted to discussion of some open problems and conclusions.

2. Mathematical theory of heterogeneous populations

In this section we present the mathematical development of the theory of heterogeneous populations with parametric heterogeneity as it is required for the subsequent application to the mathematical models of the epidemic spread. We aim to present neither the most abstract possible formulations for such models (dubbed *systems with inheritance*, see [28, 29]) nor the most general form of the dynamical systems that can be analyzed by similar means [37]. We refer the reader to the literature for a notably wider account of the necessary theory [28, 37, 38], and remark that our presentation has a primary goal to make the exposition self-contained.

Let us consider two interacting populations, each of which is characterized by its own parameter (trait) value at any given time moment, such that we can speak of the parameter distribution in each population

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

(the populations are *heterogeneous* with respect to a given parameter). Denote the densities of the populations as $n_1(t,\omega_1)$ and $n_2(t,\omega_2)$. This implies that the total population sizes are given by

$$N_1(t)=\int_{\Omega_1} n_1(t,\omega_1)\,d\omega_1,\quad N_2(t)=\int_{\Omega_2} n_2(t,\omega_2)\,d\omega_2,$$

respectively. Here Ω_i, $i=1,2$ are the sets of admissible parameter values, and we suppose that these sets are such that the corresponding integrals are always well defined (in fact, in the examples we usually suppose that Ω_i, $i=1,2$ are intervals of $\mathbb{R}^1$, but the general theory does not require it). An obvious generalization of the considered situation is that k populations are considered, some of which could be characterized by their own vector-parameters. Such generalization would require only additional notation, and therefore we deal with, without loss of generality, the case of two populations. For these two populations the pdfs are given by

$$p_i(t,\omega_i)=\frac{n_i(t,\omega_i)}{N_i(t)},\quad i=1,2,$$

such that $p_i(t,\omega_i)\geq 0$, $\int_{\Omega_i} p_i(t,\omega_i)\,d\omega_i=1$, $i=1,2$ for any $t>0$.

To describe the dynamics of the densities of the populations under study, we start with the general model

$$\begin{aligned}\frac{\partial_t n_1(t,\omega_1)}{n_1(t,\omega_1)}&=F_1\big(n_1(t,\omega_1),n_2(t,\omega_2)\big),\\ \frac{\partial_t n_2(t,\omega_2)}{n_2(t,\omega_1)}&=F_2\big(n_1(t,\omega_1),n_2(t,\omega_2)\big),\end{aligned}\tag{2.1}$$

where F_1, F_2 are given functions (the per capita growth rates of the interacting populations) and $\partial_t\equiv\frac{\partial}{\partial t}$. For (2.1) the initial conditions are $n_1(0,\omega_1)$ and $n_2(0,\omega_2)$.

It is the form of F_1 and F_2 that defines the dynamics of interacting populations, and to apply the theory of heterogeneous populations, as presented in, e.g., [37], this form must satisfy some additional requirements. In particular, we adopt that

$$F_i\big(n_1(t,\omega_1),n_2(t,\omega_2)\big)=f_i(\mathbf{v})+\varphi(\omega_i)g_i(\mathbf{v}),\quad i=1,2,\tag{2.2}$$

where f_i, g_i, φ_i, $i=1,2$ are given functions,

$$\mathbf{v}=(N_1,N_2,\bar\varphi_1(t),\bar\varphi_2(t)),\tag{2.3}$$

and

$$\bar\varphi_i(t)=\mathsf{E}_t[\varphi_i]=\int_{\Omega_i}\varphi_i(\omega_i)p_i(t,\omega_i)\,d\omega_i,\quad i=1,2.\tag{2.4}$$

In words, we require that the per capita growth rates depend explicitly only on the parameters, total population sizes and the mean values of the functions of the distributed parameters. Assume that we are interested only in the dynamics of the total population sizes. Integrating the first equation in (2.1) with respect to ω_1 and the second one with respect to ω_2, and using (2.2), (2.4), we obtain

$$\dot N_i=N_i\big(f_i(\mathbf{v})+\bar\varphi_i(t)g_i(\mathbf{v})\big),\quad i=1,2.\tag{2.5}$$

Inasmuch as $\mathbf{v}$ depends only on the total population sizes and $\bar\varphi_i(t)$, $i=1,2$, the dynamics of the populations sizes can be found if $\bar\varphi_i(t)$, $i=1,2$ are known. We remark that $\bar\varphi_i(t)$, $i=1,2$ depend on the current parameter distributions, which actually has to be found. The remarkable fact is that it is possible to calculate $\bar\varphi_i(t)$, $i=1,2$ for any time moment if the initial parameter distributions

$$p_i(0,\omega_i)=\frac{n_i(0,\omega_i)}{N_i(0)},\quad i=1,2$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

are known and two additional differential equations are allowed. This is the main result for our presentation. To give it a precise statement, introduce the notations

$$\mathsf{M}_i(t,\lambda) = \int_{\Omega_i} e^{\lambda\varphi_i(\omega_i)} p_i(t,\omega_i)\, d\omega_i, \quad i = 1,2, \tag{2.6}$$

for the moment generating functions (mgfs) of $\varphi_i(\omega_i)$, $i = 1,2$ respectively at any time moment. The mgf, it it exists, defines uniquely the given probability distribution. In the mgfs for the initial distributions we occasionally suppress the dependence on time:

$$\mathsf{M}_i(\lambda) \equiv \mathsf{M}_i(0,\lambda), \quad i = 1,2.$$

It is a basic fact that if $\mathsf{M}_i(t,\lambda)$, $i = 1,2$ are known then the mean values of $\varphi_i(\omega_i)$ can be found by simple differentiation:

$$\bar{\varphi}_i(t) = \mathsf{E}_t[\varphi_i] = \partial_\lambda \mathsf{M}_i(t,\lambda)|_{\lambda=0}, \quad i = 1,2.$$

Now we can state the following main

Theorem 2.1. *Let the dynamics of two interacting populations be described by system* (2.1) *with* (2.2), (2.3), (2.4). *Consider auxiliary variables* $q_i(t)$, $i = 1,2$ *that satisfy the following differential equations*

$$\dot{q}_i = g_i(\mathbf{v}), \quad q_i(0) = 0, \quad i = 1,2, \tag{2.7}$$

where g_i, $i = 1,2$ *are as in* (2.2). *Then*

$$\mathsf{M}_i(t,\lambda) = \frac{\mathsf{M}_i(0,\lambda + q_i(t))}{\mathsf{M}_i(0,q_i(t))} = \frac{\mathsf{M}_i(\lambda + q_i(t))}{\mathsf{M}_i(q_i(t))}, \quad i = 1,2. \tag{2.8}$$

Proof. From the first equation of (2.1) with (2.2) we have

$$\begin{aligned} n_1(t,\omega_1) &= n_1(0,\omega_1)\exp\{\int_0^t f_1(\mathbf{v})\, dt + \varphi_1(\omega_1)\int_0^t g_1(\mathbf{v})\, dt\} = \\ &= n_1(0,\omega_1)\exp\{\int_0^t f_1(\mathbf{v})\, dt + \varphi_1(\omega_1) q_1(t)\}. \end{aligned} \tag{2.9}$$

Using (2.9), we find

$$\begin{aligned} N_1(t) &= \int_{\Omega_1} n_1(t,\omega_1)\, d\omega_1 = \\ &= \exp\{\int_0^t f_1(\mathbf{v})\, dt\} \int_{\Omega_1} n_1(0,\omega_1)\exp\{\varphi_1(\omega_1)q_1(t)\}\, d\omega_1 = \\ &= \frac{1}{N_1(0)}\exp\{\int_0^t f_1(\mathbf{v})\, dt\} \int_{\Omega_1} p_1(0,\omega_1)\exp\{\varphi_1(\omega_1)q_1(t)\}\, d\omega_1 = \\ &= \frac{1}{N_1(0)}\exp\{\int_0^t f_1(\mathbf{v})\, dt\}\mathsf{M}_1(0,q_1(t)). \end{aligned} \tag{2.10}$$

Using the definition (2.6) for mgf,

$$\begin{aligned} \mathsf{M}_1(t,\lambda) &= \int_{\Omega_1} \exp\{\lambda\varphi_1(\omega_1)\} p_1(t,\omega_1)\, d\omega_1 = \\ &= \frac{1}{N_1(t)} \int_{\Omega_1} \exp\{\lambda\varphi_1(\omega_1)\} n_1(t,\omega_1)\, d\omega_1 = \text{from (2.9)} \\ &= \frac{\int_0^t f_1(\mathbf{v})\, dt}{N_1(0)N_1(t)} \int_{\Omega_1} \exp\{(\lambda + q_1(t))\varphi_1(\omega_1)\} p_1(0,\omega_1)\, d\omega_1 = \\ &= \frac{\int_0^t f_1(\mathbf{v})\, dt}{N_1(0)N_1(t)} \mathsf{M}_1(0,\lambda + q_1(t)). \end{aligned} \tag{2.11}$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

Finally, applying (2.10) and (2.11) we obtain (2.8). The same calculations are valid for the second population. □

Remark 2.2. Theorem 2.1 allows to reduce system (2.1) with (2.2) to the four dimensional system of ODEs (2.5), (2.7), where the mean parameter values are

$$\bar{\varphi}_i(t) = \mathsf{E}_t[\varphi_i] = \partial_\lambda \ln \mathsf{M}_i(\lambda)|_{\lambda=q_i(t)}\,, \quad i = 1, 2. \tag{2.12}$$

It is a simple matter (see, e.g., [37]) to prove that the mean parameter values satisfy the equations

$$\dot{\bar{\varphi}}_i(t) = g_i(\mathbf{v})\sigma_i^2(t), \quad i = 1, 2, \tag{2.13}$$

where $\sigma_i^2(t)$, $i = 1, 2$ are the current variances of $\varphi_i(\omega_i)$, $i = 1, 2$.

Remark 2.3. Using Theorem 2.1 we find that the solution for the total population size of the heterogeneous Malthus equation (1.5) is given explicitly as

$$N(t) = N(0)\mathsf{M}(t),$$

where $\mathsf{M}(t)$ is the mgf of the initial parameter distribution $p(0, \omega)$. This solution was studied in detail in [34, 35].

Remark 2.4. Theorem 2.1 shows that the mgfs at any time moment can be found from the mgfs at the initial moment using (2.8). This theorem also shows that by reducing the original infinite dimensional system (2.1) with (2.2) to the system of ODEs (2.5), (2.7) we actually do not loose any information because, as far as the functions $q_i(t)$, $i = 1, 2$ are known, the densities $n_i(t, \omega_i)$, $i = 1, 2$ at any time moment can be found from the total population sizes and the time dependent probability distributions $p_i(t, \omega_i)$, $i = 1, 2$, which can be inferred from the corresponding mgfs. For the following exposition we do not require explicit formulas for the evolution of distributions, and refer the reader to [37], where these results can be found.

Remark 2.5. Theorem 2.1 provides an analytical proof of (2.8). It is interesting to note that there is a simple probabilistic proof of a similar formula [1, 2].

Consider the so-called proportional *frailty model* [2]: it is assumed that the hazard rate of an individual is given as the product of an individual specific quantity Z and a basic rate $\alpha(t)$:

$$\alpha(t|Z) = \alpha(t)Z. \tag{2.14}$$

Here Z plays the role of the parameter distributed in the population. Given Z the probability of surviving up to time t is

$$S(t|Z) = e^{-ZA(t)}, \quad A(t) = \int_0^t \alpha(\tau)\, d\tau.$$

The population survival function is therefore

$$S(t) = \mathsf{P}[T > t] = \mathsf{E}[e^{-ZA(t)}] = \mathsf{M}(-A(t)). \tag{2.15}$$

The frailty distribution in the population surviving at time t can be found as follows (here $\mathsf{I}(T > t)$ is the indicator function):

$$\mathsf{M}(t, \lambda) = \mathsf{E}_t[e^{\lambda Z}] = \mathsf{E}[e^{\lambda Z}|T > t] = \frac{\mathsf{E}[e^{\lambda Z}\mathsf{I}(T > t)]}{\mathsf{P}[T > t]} = \\ = \frac{\mathsf{E}[e^{\lambda Z - ZA(t)}]}{\mathsf{E}[e^{-ZA(t)}]} = \frac{\mathsf{M}(\lambda - A(t))}{\mathsf{M}(-A(t))}\,. \tag{2.16}$$

Using $q(t) = -A(t)$ in (2.16) we obtain (2.8).

Concluding, using the mathematical theory of heterogeneous populations with parametric heterogeneity, outlined in this section, it is possible to model communities of populations when each population is characterized by its own parameter value, and the per capita growth rates depend on this parameter and the average characteristics of the interacting populations.

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

3. Heterogeneous models in epidemiology: Model formulation

The modern mathematical epidemiology has its roots in now classical work by Kermack and McKendrick [21,43], where the total population of the constant size N was subdivided into three categories: susceptible individuals S that are prone to infection, infectious individuals I that transmit the disease, and removed individuals R that either acquire life-long immunity or die. In the simplest case, assuming that the period of infection is exponentially distributed with the mean $1/\gamma$ and the transmission process is described by mass-action kinetics [31,47] (this means that the contact rate is proportional to the total population size N), we have that the dynamics is described by

$$\begin{aligned} \dot{S} &= -\beta SI, \\ \dot{I} &= \beta SI - \gamma I, \\ \dot{R} &= \gamma I, \end{aligned} \tag{3.1}$$

where β is the transmission parameter that encompasses the information on the probability of a successful contact (i.e., the contact that results in infection) and the contact rate. From (3.1) it follows that $S(t)+I(t)+R(t) = N$ for any t, therefore the third equation is redundant and in the following we usually omit it. SIR model (3.1) includes two parameters that in reality vary from individual to individual. Therefore, it is important to take this heterogeneity into account.

3.1. Heterogeneous SIR model with distributed susceptibility to a disease

There are a number of studies in the literature which model heterogeneous susceptibility to a disease, with either a finite number of different susceptibility classes [7, 9, 13, 23, 27, 32, 33, 55, 57] or with a continuous distribution of susceptibility [14, 16, 24–26]. We present a formulation from [51] keeping in mind that both discrete and continuous distributions can be accommodated.

Let us denote $s(t,\omega)$ the density of the susceptible individuals having the trait value ω. The total size of the susceptible population is $S(t) = \int_\Omega s(t,\omega)\,d\omega$. Assuming the law of mass action, we obtain that the changes in the susceptible and infectious populations are described by

$$\begin{aligned} \partial_t s(t,\omega) &= -\beta(\omega)s(t,\omega)I(t), \\ \dot{I}(t) &= I(t)\int_\Omega \beta(\omega)s(t,\omega)\,d\omega - \gamma I(t). \end{aligned} \tag{3.2}$$

Or, using the notations

$$\bar{\beta}(t) = \mathsf{E}_t[\beta] = \int_\Omega \beta(\omega)p_s(t,\omega)\,d\omega, \quad p_s(t,\omega) = \frac{s(t,\omega)}{S(t)},$$

(3.2) can be rewritten as

$$\begin{aligned} \partial_t s(t,\omega) &= -\beta(\omega)s(t,\omega)I(t), \\ \dot{I}(t) &= \bar{\beta}(t)SI - \gamma I(t). \end{aligned} \tag{3.3}$$

For (3.3) the initial conditions are $s(0,\omega) = s_0(\omega) = S_0 p_s(0,\omega)$, $I(0) = I_0$.

Model (3.2), (3.3) can be also obtained from the general epidemic equation [19]

$$\partial_t s(t,\omega) = s(t,\omega)\int_\Omega\int_0^\infty A(\tau,\omega,\eta)\partial_t s(t-\tau,\eta)\,d\tau d\eta, \tag{3.4}$$

where $A(\tau,\omega,\eta)$ is the expected infectivity of an individual that was infected τ units ago while having trait value η towards to a susceptible with trait value ω. If we assume that $A(\tau,\omega,\eta) = \beta(\omega)f(\tau)$ for a given function f and set

$$I(t) = -\int_\Omega\int_0^t f(\tau)\partial_t s(t-\tau,\eta)\,d\tau d\eta,$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

then, after some algebra, we obtain

$$\begin{aligned}\partial_t s(t,\omega) &= -\beta(\omega)s(t,\omega)I(t),\\ \dot{I}(t) &= f(0)I(t)\int_\Omega \beta(\omega)s(t,\omega)\,d\omega - \int_\Omega\int_0^\infty f'(\tau)\partial_t s(t-\tau,\eta)\,d\tau d\eta.\end{aligned} \tag{3.5}$$

Letting $f(\tau) = e^{-\gamma\tau}$ reduces (3.5) to (3.3) (this can be shown to be the only case to end up with an ODE system, for this it is necessary and sufficient that f satisfies the equation $\dot{f} = -\gamma f$).

Remark 3.1. If in the general epidemic equation (3.4) it is assumed that $f(\tau) = \chi(T-\tau)$, where χ is the Heaviside function then from (3.5) it follows that

$$\begin{aligned}\partial_t s(t,\omega) &= -\beta(\omega)s(t,\omega)I(t),\\ \dot{I}(t) &= \bar{\beta}(t)SI - \bar{\beta}(t-T)S(t-T)I(t-T).\end{aligned} \tag{3.6}$$

Model (3.6) is the model with distributed susceptibility studied in [24–26].

3.2. Heterogeneous SIR model with distributed infectivity

Let $\beta(\omega)$ be the transmission parameter of an individual with the infectivity value ω, and $i(t,\omega)$ be the density of the infectious hosts with trait value ω at the time moment t, $I(t) = \int_\Omega i(t,\omega)\,d\omega$. We assume that the susceptible population is homogeneous. To describe the dynamics of the infectious population it is necessary to specify what trait value is assigned to a newly infected individual, which was infected by an infectious individual with the trait value η. Denoting $\psi(\omega,\eta)$ the pdf that prescribes the probability that a newly infected individual is assigned the trait value ω if infected by an individual with the trait value η, we obtain

$$\begin{aligned}\dot{S}(t) &= -\bar{\beta}(t)S(t)I(t),\\ \partial_t i(t,\omega) &= S(t)\int_\Omega \psi(\omega,\eta)\beta(\eta)i(t,\eta)\,d\eta - \gamma i(t,\omega),\end{aligned} \tag{3.7}$$

where now

$$\bar{\beta}(t) = \int_\Omega \beta(\omega)p_i(t,\omega)\,d\omega, \quad p_i(t,\omega) = \frac{i(t,\omega)}{I(t)}.$$

The initial conditions are $S(0) = S_0$, $i(0,\omega) = i_0(\omega) = I_0 p_i(0,\omega)$.

There are several natural choices for $\psi(\omega,\eta)$, the simplest of which is $\psi(\omega,\eta) = \delta(\omega-\omega')$, where δ is the delta-function. This option means that the newly infected individual acquires the trait value of the person by whom he was infected (this is equivalent to $\psi(\omega,\eta) = p_i(t,\omega)$, i.e., the trait values are assigned according to the current distribution of the infectivity). Using $\psi(\omega,\eta) = \delta(\omega-\omega')$ in (3.7) we obtain

$$\begin{aligned}\dot{S}(t) &= -\bar{\beta}(t)S(t)I(t),\\ \partial_t i(t,\omega) &= \beta(\omega)i(t,\omega)S(t) - \gamma i(t,\omega).\end{aligned} \tag{3.8}$$

Model (3.8) is very similar in form to (3.3). However, it should be clear that, in general, model (3.3) is much closer to reality than (3.8), which corresponds to the case when several different strains of an infection can be passed on.

3.3. Heterogeneous SIR model with distributed susceptibility and infectivity

Let us assume now that both the susceptibility and infectivity are distributed in the population experiencing the disease; this will generalize models (3.3) and (3.8).

Let $s(t,\omega_1)$ and $i(t,\omega_2)$ be the densities of the susceptible and infectious individuals respectively. For the following, simplifying, we assume that the traits of the two subpopulations are independent, i.e.,

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

$\beta(\omega_1,\omega_2)=\beta_1(\omega_1)\beta_2(\omega_2)$. The number of susceptible hosts with the trait value ω_1 infected by infectious individuals with the trait value ω_2 is given by

$$\beta_1(\omega_1)s(t,\omega_1)\beta_2(\omega_2)i(t,\omega_2).$$

Therefore, the total change in the infectious subpopulation with trait value ω_2 is

$$\beta_2(\omega_2)i(t,\omega_2)\int_{\Omega_1}\beta_1(\omega_1)s(t,\omega_1)\,d\omega_1,$$

assuming that $\psi(\omega,\eta)=\delta(\omega-\eta)$ (cf. (3.8)). A similar expression describes the change in the susceptible population. It is worth emphasizing that nothing else except for the standard law of mass action is supposed to formulate the terms for the change in susceptible and infectious subpopulations. Combining the above assumptions we obtain the following model:

$$\begin{aligned}\partial_t s(t,\omega_1) &= -\beta_1(\omega_1)s(t,\omega_1)\int_{\Omega_2}\beta_2(\omega_2)i(t,\omega_2)\,d\omega_2\\ &= -\beta_1(\omega_1)s(t,\omega_1)\bar\beta_2(t)I(t),\\ \partial_t i(t,\omega_2) &= \beta_2(\omega_2)i(t,\omega_2)\int_{\Omega_1}\beta_1(\omega_1)s(t,\omega_1)\,d\omega_1-\gamma i(t,\omega_2)\\ &= \beta_2(\omega_2)i(t,\omega_2)\bar\beta_1(t)S(t)-\gamma i(t,\omega_2),\end{aligned}\tag{3.9}$$

where

$$\bar\beta_1(t)=\mathsf{E}_t[\beta_1]=\int_{\Omega_1}\beta_1(\omega_1)p_s(t,\omega_1)\,d\omega_1,\quad \bar\beta_2(t)=\mathsf{E}_t[\beta_2]=\int_{\Omega_2}\beta_2(\omega_2)p_i(t,\omega_2)\,d\omega_2.$$

Model (3.9) is supplemented with the initial conditions $s(0,\omega_1)=S_0p_s(0,\omega_1)$, $i(0,\omega_2)=I_0p_i(0,\omega_2)$.

In the case $\gamma=0$ we obtain the heterogeneous SI model with distributed susceptibility and infectivity

$$\begin{aligned}\partial_t s(t,\omega_1) &= -\beta_1(\omega_1)s(t,\omega_1)\bar\beta_2(t)I(t),\\ \partial_t i(t,\omega_2) &= \beta_2(\omega_2)i(t,\omega_2)\bar\beta_1(t)S(t),\end{aligned}\tag{3.10}$$

for which the global dynamics is simple and is similar to the simplest homogeneous SI model, when $S(t)\to 0$, $I(t)\to N$, when $t\to\infty$.

3.4. Heterogeneous SI model with heterogeneous contact structure

Above we discussed mainly about heterogeneity of the hosts: whether all susceptible individuals are of the same type with equal susceptibility, and whether all infectious individuals have equal ability to infect others. Another aspect of heterogeneity is the possible heterogeneous social contact network, which is the one of the central topics in mathematical epidemiology, e.g, [7, 11, 17, 42, 48]. It is difficult to apply the general theory of heterogeneous populations as presented in Section 2 to such models, however, there is a simple case, for which some results can be obtained.

Let us assume that $n(t,\omega)$ denotes the density of individuals in the population, which are making ω contacts on average. Every individual can be contacted by another individual, and the individuals differ in an average number of contacts. This situation is usually termed as *separable mixing*. If we denote r the probability of transmission the disease given a contact, then, the simplest SIR-model with separable mixing can be described by the following system:

$$\begin{aligned}\partial_t s(t,w) &= -r\omega s(t,\omega)\frac{\int_\Omega \omega i(t,\omega)d\omega}{\int_\Omega \omega s(t,\omega)\,d\omega+\int_\Omega \omega i(t,\omega)\,d\omega},\\ \partial_t i(t,w) &= -r\omega s(t,\omega)\frac{\int_\Omega \omega i(t,\omega)d\omega}{\int_\Omega \omega s(t,\omega)\,d\omega+\int_\Omega \omega i(t,\omega)\,d\omega}-\gamma i(t,\omega)\,.\end{aligned}\tag{3.11}$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

In case of SI model (3.10) ($\gamma = 0$) we have $s(t,\omega) + i(t,\omega) = n_0(\omega)$ for any t, and $n_0(\omega)$ is a given density which specifies probability density function of the contact distribution. Using the property $i(t,\omega) = n_0(\omega) - s(t,\omega)$, we obtain

$$\partial_t s(t,w) = -r\omega s(t,\omega)\left[1 - \frac{\int_\Omega \omega s(t,\omega)d\omega}{\int_\Omega \omega n_0(\omega)d\omega}\right]. \tag{3.12}$$

We conclude this section with an obvious statement that the list of possible epidemiological models with parametric heterogeneity can be made longer. What is important, however, that models (3.3), (3.6), (3.8), (3.9), (3.10), and (3.12) all are written in the form (2.1)–(2.4), which allows a unified treatment of these models within the mathematical framework outlined in Section 2.

4. Mathematical analysis of an SIR model

4.1. Analysis of heterogeneous SIR model with distributed susceptibility and infectivity

Model (3.9) comprises, as particular cases, models (3.3) and (3.8). Therefore, the main result is stated for this model (the case of distributed susceptibility alone was treated in [51]).

Theorem 4.1. *Heterogeneous SIR model* (3.9) *with distributed susceptibility and infectivity is equivalent to the following two-dimensional non-autonomous system of ODEs:*

$$\begin{aligned} \dot S &= -h_1(S)h_2(t,I),\\ \dot I &= h_1(S)h_2(t,I) - \gamma I, \end{aligned} \tag{4.1}$$

where

$$h_1(S) = S_0(\partial_\lambda \mathsf{M}_1^{-1}(0,\lambda)|_{\lambda = S/S_0})^{-1}, \tag{4.2}$$

$$h_2(t,I) = I_0 e^{-\gamma t}(\partial_\lambda \mathsf{M}_2^{-1}(0,\lambda)|_{\lambda = Ie^{\gamma t}/I_0})^{-1}, \tag{4.3}$$

and $\mathsf{M}_i^{-1}(0,\lambda)$, $i = 1,2$ *are the inverse functions to the mgfs of the initial distributions of susceptibility and infectivity respectively.*

Proof. According to Theorem 2.1, system (3.9) is equivalent to

$$\begin{aligned} \dot S &= -\bar\beta_1(t)\bar\beta_2(t)SI,\\ \dot I &= \bar\beta_1(t)\bar\beta_2(t)SI - \gamma I,\\ \dot q_1 &= -\bar\beta_2(t)I,\\ \dot q_2 &= \bar\beta_1(t)S, \end{aligned} \tag{4.4}$$

where

$$\bar\beta_i(t) = \partial_\lambda \ln \mathsf{M}_i(0,\lambda)|_{\lambda = q_i(t)}, \quad i = 1,2.$$

From (4.4) it follows that

$$\begin{aligned} \frac{d}{dt}\ln S &= \frac{d}{dt}\ln \mathsf{M}_1(0,q_1),\\ \frac{d}{dt}\ln I &= \frac{d}{dt}\ln \mathsf{M}_2(0,q_2) - \gamma, \end{aligned}$$

from which we have two first integrals

$$\frac{S}{S_0} = \mathsf{M}_1(0,q_1), \quad \frac{Ie^{\gamma t}}{I_0} = \mathsf{M}_2(0,q_2). \tag{4.5}$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

Since the mgfs are monotone functions, we can express q_1 and q_2 in (4.5) through S and I respectively. Putting the resulting expressions into (4.4) and remembering the inverse function theorem, we obtain (4.1) with (4.2), (4.3). □

Theorem 4.1 gives an important example of a bottom up approach in building mathematical models in epidemiology [18,30]. We start with a detailed model when each individual in the population has its own parameter value. This model, albeit still oversimplified and unrealistic, takes into account physiological structure of the population without any *ad hoc* assumptions on the dynamics of the total population sizes. The general theory of the heterogeneous populations with parametric heterogeneity allows to reduce the model to the population-level description, which, incidentally, still can be described as a classical SIR model, which contains non-linear and time-dependent transmission rates. This shows that a great deal of mathematical models, built from the first principles, are still valid and can be used for ecological predictions, because they can be shown to follow from detailed individual-based models.

Let us consider an explicit example (more details are given in [52]).

4.1.0.1. On the power law transmission function. The *transmission function* (the number of new infectious cases per time unit) is considered to be one of the major ingredients of the models describing spread of an infectious disease [19]. Historically, borrowing an analogy from the chemical kinetics, this function was supposed to be of a simple bilinear form, $\propto SI$, which means that the law of mass action is assumed to hold [31] (this is often called density dependent transmission function). If one assumes that the number of contacts is fixed for any individual and does not depend on the population size, than the transmission function takes the form $\propto SI/N$ (frequency dependent transmission function). If the population size constant these two transmission function yield the same predictions, whereas variable population size can produce different behaviors (e.g., [12, 54]). Other transmission functions are also possible, see [47]. Important point here is that most of the used nonlinear functions, that are somewhat intermediate between density and frequency dependent modes of transmission, are phenomenological and lack any mechanistic derivation.

One of most frequently used transmission function takes the form

$$T(S,I) = \beta S^p I^q, \quad p,q > 0. \tag{4.6}$$

This function is often called the power law transmission function. It was used in, e.g., [61,62] in the form $T(S,I) = \beta S^p I$. Severo [58] considered the full form (4.6), see also [44,45] for a detailed mathematical analysis of epidemiological models with the power law transmission function. Simulations [56,59] show that power law transmission function indeed improves an accuracy of the mean-field SIR models. Using Theorem 4.1 we can show that the power law transmission function can be inferred from the heterogeneous SIR models (3.9) and (3.10).

Let us assume that the initial distribution of susceptibility is a gamma-distribution with parameters k_1 and ν_1:

$$p_1(0,\omega) = \frac{\nu_1^{k_1}}{\Gamma(k_1)}\omega^{k_1-1}w^{-\nu_1\omega}, \quad \omega \geq 0, \quad k_1,\nu_1 > 0. \tag{4.7}$$

From (4.7) one has, assuming additionally that $\beta_1(\omega_1) = \omega_1$,

$$\mathsf{M}_1(\lambda) = \left(1 - \frac{\lambda}{\nu_1}\right)^{-k_1}, \tag{4.8}$$

therefore the expression in (4.2) is given by

$$h_1(S) = \frac{k_1}{\nu_1} S \left(\frac{S}{S_0}\right)^{1/k_1}. \tag{4.9}$$

An analogous expression can be obtained for $h_2(t,I)$ if it is postulated that the initial distribution of infectivity is also a gamma-distribution (4.7) with parameters k_2,ν_2 and $\beta_2(\omega_2) = \omega_2$. Expression (4.9) implies

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

Corollary 4.2. *The power law transmission function* (4.6) *with the heterogeneity parameters* $q = 1$, $p = 1+1/k_1$ *can be obtained as a consequence of the distributed heterogeneous SIR model* (3.3) *with distributed susceptibility if the initial distribution of susceptibility is the gamma-distribution with parameters* k_1, ν_1.

The power law transmission function (4.6) *with the heterogeneity parameters* $q = 1 + 1/k_2$, $p = 1 + 1/k_1$ *can be obtained as a consequence of the distributed heterogeneous SI model* (3.10) *with distributed susceptibility and infectivity if the initial distribution of susceptibility is the gamma-distribution with parameters* k_1, ν_1 *and the initial distribution of infectivity if the gamma-distribution with parameters* k_2, ν_2.

Remark 4.3. Using (2.8) and (2.12) it is straightforward to show that for the initial gamma distribution of susceptibility, the time-dependent distribution is also the gamma-distribution with parameters k_1 and $\nu_1 - q_1(t)$, where $q_1(t)$ is the solution of the corresponding auxiliary equation (2.7). That is, we have

$$\bar{\beta}_1(t) = \mathsf{E}_t[\beta_1] = \frac{k_1}{\nu_1 - q_1(t)}, \quad \sigma_1^2(t) = \frac{k_1}{(\nu_1 - q_1(t))^2}.$$

Note that for any time moment the coefficient of variation $cv = \sigma_1(t)/\bar{\beta}_1(t) = 1/\sqrt{k_1}$ and does not depend on t.

Remark 4.4. To analytically analyze models with parameter distributions it is useful to have families of distributions for which their mgfs are known, as in the case of the gamma-distribution. A very general family of distributions is the so-called power variance function distributions (PVF distributions) [2], which are defined through mgf:

$$\mathsf{M}(\lambda) = \exp\left[-\rho\left\{1 - \left(\frac{\nu}{\nu - \lambda}\right)^m\right\}\right], \tag{4.10}$$

with $\nu > 0$, $m > -1$, $m\rho > 0$.

The mgf (4.10) describes a number of distributions. For instance, if $\rho \to \infty$ and $m \to 0$ in such a way that $\rho m \to k$, then (4.10) reduces to the mgf of the gamma-distribution (4.8). When $m > 0$ then (4.10) gives a compound Poisson distribution, which is the sum of independent gamma-distributions with the parameters ν and m, when the number of summand is Poisson distributed with expectation ρ. Such compound Poisson distribution has non-zero mass probability at zero, which means, in terms of susceptibility to a disease, that there is part of population of mass $\exp\{-\rho\}$ that is immune to the disease. When $m = -1/2$ and $\rho < 0$ we have an inverse gaussian (Wald) distribution. Other limiting distributions are possible, which gives a wide choice of heterogeneity distributions [2].

From (4.10) it can be shown that

$$\mathsf{E}[Z] = \frac{\rho m}{\nu}, \quad \mathsf{Var}[Z] = \frac{\rho m}{\nu}\frac{m+1}{\nu}.$$

Equation (2.8) implies that the time dependent parameter distribution of the models with parametric heterogeneity is still PVF distribution if the initial one is given by (4.10), with the following change in parameters:

$$\rho \to \rho\left(\frac{\nu}{\nu - q(t)}\right)^m, \quad \nu \to \nu - q(t), \quad m \to m.$$

Not all the distributions possess this "stability" property: e.g., the initial uniform distribution turns into truncated exponential distribution.

4.1.0.2. On the final epidemic size. A very important quantity of the SIR model (3.1) is the final epidemic size, which can be defined as the number of susceptible hosts that escape infection, and which we denote S_∞. From (3.1) it follows that in the homogeneous case this number can be found as the root to the equation

$$S_\infty = S_0 e^{\beta(S_\infty - N)/\gamma}. \tag{4.11}$$

Consider again the heterogeneous SIR model (3.3) with distributed susceptibility. Theorem 4.1 implies

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

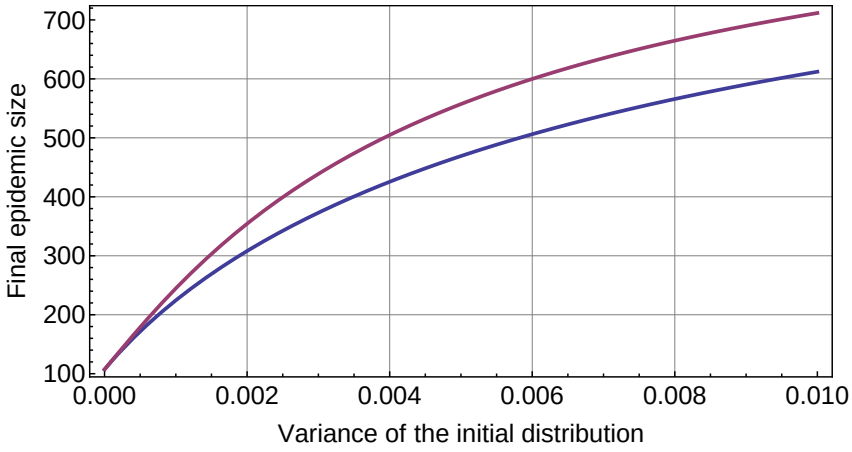


FIGURE 1. Final epidemic sizes versus the initial variance of the susceptibility distributions for two initial distributions: a gamma distribution (the top curve) and a Wald distribution (the bottom curve). All other parameters are the same for both cases

Corollary 4.5. *The final epidemic size of the heterogeneous SIR model with distributed susceptibility* (3.3) *can be found as the root to the equation*

$$S_\infty = S_0 \mathsf{M}\big((S_\infty - N)/\gamma\big). \tag{4.12}$$

Proof. From the first equation in (4.4) and recalling the third equation in (3.1) we find that

$$\frac{d}{dt}\ln S(t) = \frac{d}{dt}\ln \mathsf{M}(-R/\gamma),$$

from which, after integration and using the fact $R_\infty = N - S_\infty$, (4.12) follows. □

Remark 4.6. Equation (4.11) is a particular case of (4.12) if the initial distribution of susceptibility is the delta-function.

Remark 4.7. Equation (4.12) can also be obtained from the general epidemic equation (3.4), which shows that no assumptions on the distribution of epidemic length is required to deduce (4.12). A very careful mathematical analysis of this equation is given in [41], and we refer the reader to this reference for many intricate details. We note that for the first time equation (4.12) was written in [9] for a discrete distribution of susceptibility. It can be easily proved that the final epidemic size equation implies that the epidemic is the most severe in case of a totally uniform population; heterogeneity in susceptibility increases the final epidemic size S_∞ (see Fig. 1).

Consider a numerical example. Let us assume that we have two SIR model (3.3) with heterogeneous susceptibility, and the initial distributions are given by a gamma and Wald distributions respectively with the same initial mean values and the same variances. The final epidemic sizes are shown in Fig. 1 depending on the initial variances.

Figure 1 allows to make an important conclusion: to infer the final epidemic size for an SIR model with distributed susceptibility it is not enough to know only several first moments of the initial distribution. This conclusion is somewhat restrictive for a predictive use of such models; however, it also signifies that various approximation techniques can lead to erroneous conclusions. As an example consider the final epidemic size equation for (3.6), obtained in [26] with two different methods. First it was supposed that the susceptibility distribution is a gamma-distribution. The second approach was to consider an infinite

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

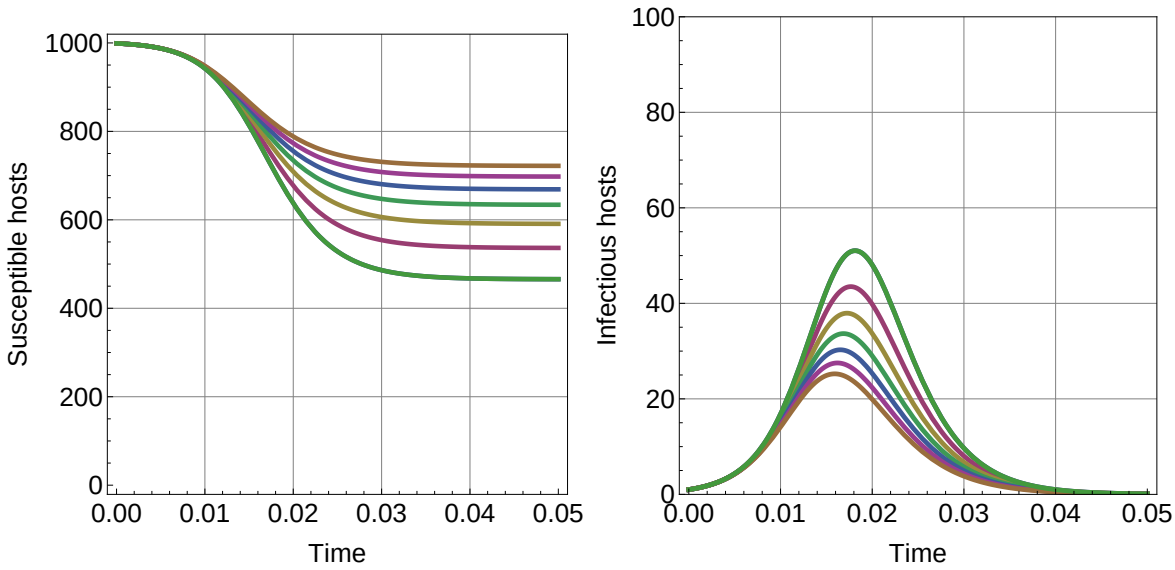


FIGURE 2. Solutions to the heterogeneous SIR model (3.3) with distributed susceptibility. The parameters are $S_0 = 999$, $I_0 = 1$, $\gamma = 700$, $\bar{\beta}_1(0) = 1$ and the initial distribution of susceptibility is the gamma distribution with $\sigma_1^2(0) = 0, 0.2, 0.4, 0.6, 0.8, 1, 1.2$ (from the bottom to top curves for the susceptible hosts, and the opposite direction for the infectious hosts)

dimensional system of ODE, which can be inferred from (3.6), for the moments of the corresponding distribution (this is a natural strategy for analyzing infinite dimensional dynamical systems of the form (2.2), see [49,60]). Inasmuch as it is impossible to solve a system with infinite number of equations, various techniques to close such systems exit. In particular, it was supposed in [26] that the coefficient of variation is constant. This assumption led to the same result as was obtained for the initial gamma distribution. Therefore, it was concluded that the exact form of the initial distribution is irrelevant because two seemingly different approaches lead to the same outcome. However, theory from Section 2 shows that the opposite is true. First, Remark 4.3 gives the explanation why two approaches in [26] turned out to be equivalent. And second, Fig. 1 provides indirect proof that the final epidemic size depends on the exact form of the underlying susceptibility distribution (see also [41] for exact mathematical statements).

Unfortunately, for the general case of the SIR model (3.9) with distributed susceptibility and infectivity we were not able to obtain a simple equation for the final epidemic size.

4.1.0.3. Numerical illustration. In Fig. 2 numerical solutions for (3.3) are shown for different initial variances of the parameter distribution and equal means, which confirms that the heterogeneity in susceptibility decreases the severity an an epidemic.

In Fig. 3 numerical solutions for system (3.8) are shown for different initial variances of infectivity and equal means. This figure shows that the distributed infectivity has an opposite effect on the severity of an epidemic comparing with the distributed susceptibility: the more heterogeneous the initial distribution of infectivity is, the smaller the number of susceptible hosts that escape the infection.

The last conclusion can be cast in an exact mathematical statement if one is interested only in the initial period of the epidemic.

Corollary 4.8. *Let $S_1(t)$, $S_2(t)$ be the solutions to* (3.3) *with the initial conditions that satisfy $\sigma_1^2(0) > \sigma_2^2(0)$ for the distribution of susceptibility, all other initial conditions being equal. Then there exists $\varepsilon > 0$ such that $S_1(t) > S_2(t)$ for all $t \in (0, \varepsilon)$.*

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

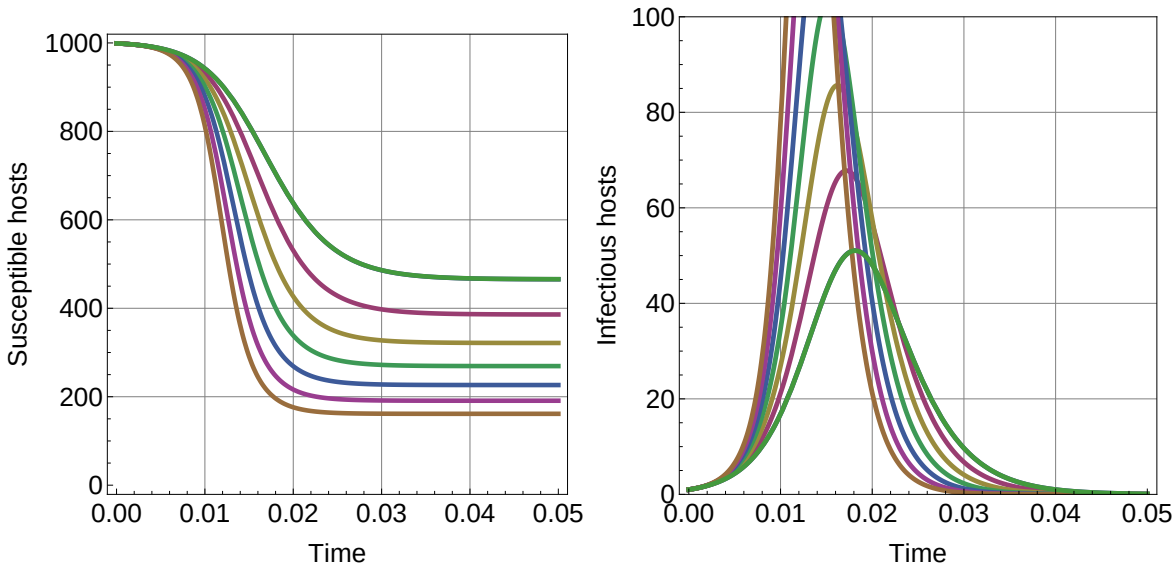


FIGURE 3. Solutions to the heterogeneous SIR model (3.8) with distributed infectivity. The parameters are $S_0 = 999$, $I_0 = 1$, $\gamma = 700$, $\bar{\beta}_1(0) = 1$ and the initial distribution of infectivity is the gamma distribution with $\sigma_2^2(0) = 0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03$ (from the top to bottom curves for the susceptible hosts, and the opposite direction for the infectious hosts)

Let $S_1(t)$, $S_2(t)$ be the solutions to (3.8) with the initial conditions that satisfy $\sigma_1^2(0) > \sigma_2^2(0)$ for the distribution of infectivity, all other initial conditions being equal. Then there exists $\varepsilon > 0$ such that $S_1(t) < S_2(t)$ for all $t \in (0, \varepsilon)$.

Proof. Using (2.13) and differentiating the first equation in (3.3) we obtain

$$S''(t) = I^2 S(\sigma^2(t) + \bar{\beta}(t)) - \bar{\beta}(t) I' S,$$

or, at the initial time moment, $S_1''(0) > S_2''(0)$, which proves the first part the corollary. The second part is proved in a similar way. □

Finally, for the general heterogeneous SIR model (3.9) with distributed infectivity and susceptibility some numerical results are shown in Fig. 4, where the initial distributions of susceptibility and infectivity are combined from those in Figs. 2 and 3. The conclusion from numerical calculations is that the interaction of the distributions of susceptibility and infectivity is nonlinear and cannot be predicted from knowing the final outcomes of separate epidemics (see also [8] for the discussion of the final epidemic size of the model (3.9)).

4.2. Heterogeneous SIR model with distributed contact number

Here we start with system (3.11):

$$\begin{aligned} \partial_t s(t,w) &= -r\omega s(t,\omega)\mu(t), \\ \partial_t i(t,w) &= -r\omega s(t,\omega)\mu(t) - \gamma i(t,\omega)\,, \end{aligned} \tag{4.13}$$

where

$$\mu(t) = \frac{\int_\Omega \omega i(t,\omega) d\omega}{\int_\Omega \omega s(t,\omega)\, d\omega + \int_\Omega \omega i(t,\omega)\, d\omega}\,.$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

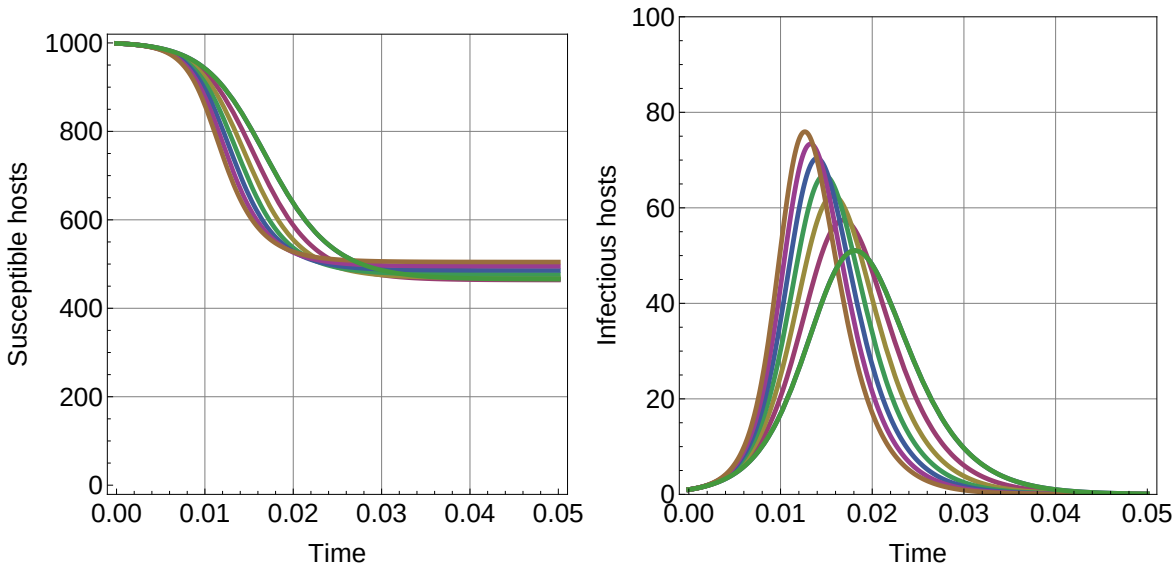


FIGURE 4. Solutions to the heterogeneous SIR model (3.9) with distributed infectivity and susceptibility . The parameters are $S_0 = 999$, $I_0 = 1$, $\gamma = 700$, $\bar{\beta}_1(0) = 1$ and the initial distribution of susceptibility is the gamma distribution with $\sigma_1^2(0) = 0, 0.2, 0.4, 0.6, 0.8, 1, 1.2$, the initial distribution of infectivity is the gamma distribution with $\sigma_2^2(0) = 0, 0.005, 0.01, 0.015, 0.02, 0.025, 0.03$

First we note that system (4.13) is not covered by Theorem 2.1, therefore, at least using the framework from Section 2, we cannot reduce this system to an ODE system. However, some results are still possible to obtain. In particular, we give a simple derivation of the expression for the initial growth of $\mu(t)$ (cf. [46]).

If we denote $K(t) = \int_\Omega \omega i(t,\omega)$, then

$$K'(t) = r\mathsf{E}[\omega^2]S\mu(t) - \gamma K(t).$$

Differentiating $\mu(t)$ yields

$$\begin{aligned}\mu'(t) &= \frac{K'(t)}{\int_\Omega \omega s(t,\omega)\,d\omega + \int_\Omega \omega i(t,\omega)\,d\omega} + \gamma\mu^2(t) = \\ &= \frac{r\mathsf{E}[\omega^2]S\mu(t)}{\int_\Omega \omega s(t,\omega)\,d\omega + \int_\Omega \omega i(t,\omega)\,d\omega} - \gamma\mu(t) + \gamma\mu^2(t) = \\ &= \mu(t)\left(\frac{r\mathsf{E}[\omega^2]S}{\mathsf{E}[\omega]S + K(t)} - \gamma\right) + \gamma\mu^2(t) = \\ &= \mu(t)\left(\frac{r\mathsf{E}[\omega^2]}{\mathsf{E}[\omega]} - \gamma\right) + \gamma\mu^2(t), \end{aligned} \tag{4.14}$$

where the last equality holds for $t \to 0$ because $S(t) \to N$ and $K(t) \to 0$. In words, we obtained the well known results [46] that initially the change in the number of susceptible hosts is proportional not to the mean number of contacts but to the mean number plus the coefficient of variation. This shows that the individuals who have a high number of contacts contribute disproportionally to the spread of an epidemic.

Corollary 4.8 implies that in case of distributed susceptibility or infectivity the initial phase of an epidemic can be determined by the first two moments. Here we prove that this is not so for model (4.13),

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

as it was shown in [60]. The proof uses Theorem 2.1, direct proof is given in [60]. We use the notation $L(t) = \int_\Omega \omega s(t,\omega)\, d\omega$.

Lemma 4.9. *Let $S_1(t)$, $S_2(t)$ be the solutions to model (4.13) with the initial conditions such that $\sigma_1^2(t) > \sigma_2^2(t)$ for the distribution of the contact number, all other initial conditions being equal. If $L(0) < K(0)$ then there exists $\varepsilon > 0$ such that $S_1(t) > S_2(t)$ for all $t \in (0, \varepsilon)$. If $L(0) > K(0)$ then the opposite holds.*

Proof. From the first equation in (4.13) it follows

$$\begin{aligned} S'' &= -r\mathsf{E}'[\omega]S\mu(t) - \mu(t)\mathsf{E}[\omega]S' - \mathsf{E}[\omega]S\mu'(t) = (\text{from (2.13) and (4.14)}) \\ &= r\sigma^2(t)\mu^2(t)S - \mu(t)\mathsf{E}[\omega]S\left[\frac{r\mathsf{E}[\omega^2]S}{L(t)+K(t)} - \gamma\right] + \ldots = \\ &= \sigma^2(t)\mu^2(t)S\left[1 - \frac{L(t)}{K(t)}\right] + \ldots, \end{aligned} \tag{4.15}$$

where dots denote terms that depend only on the first moment of the contact distribution.

From (4.15) lemma follows. □

As it was mentioned, Theorem 4.1 cannot be applied to (4.13). However, if we consider the case $\gamma = 0$, we still can reduce the heterogeneous model to ODE. We rewrite equation (3.12) in the form

$$\partial_t s(t,w) = -r\omega s(t,\omega)\left[1 - \frac{\bar\omega(t)S(t)}{C}\right], \tag{4.16}$$

where C is the number of contacts, which are made by the total population, $\bar\omega(t)$ is the average number of contacts made by one susceptible individual at time t. Theorem 4.1 implies that (4.16) is equivalent to the following ODE:

$$\dot S(t) = -rh(S)\left[1 - \frac{h(S)}{C}\right], \tag{4.17}$$

where $h(S)$ is given by (4.2). For example, if the initial distribution of the number of contacts of susceptible hosts is a gamma distribution with parameters k and ν, equation (4.17) takes the form

$$\dot S(t) = -r\frac{k}{\nu}\left[\frac{S}{S_0}\right]^{1/k} S\left(1 - \frac{k}{\nu}\left[\frac{S}{S_0}\right]^{1/k}\frac{S}{C}\right). \tag{4.18}$$

Numerical solutions of (4.17) for two parameter sets are given in Figs. 5 and 6. These figures show that initially the heterogeneous contact rates increase the speed of an epidemic; in the long term run, however, the presence of individuals who make a small number of contacts, slows the epidemic down.

5. Concluding comments

In this manuscript we reviewed and generalized several known results on the spread of epidemics in a closed heterogeneous population obtained with the help of the theory of heterogeneous populations with parametric heterogeneity. Our discussion was mainly about the SIR model (3.9) and its particular cases (3.3), (3.8), and (3.10). We also introduced the methods of the theory of heterogeneous populations to the models that take into account non-uniform contact structure of the population, noting that explicit results can be inferred only for a very limited case of the SI model (4.16).

While speaking of heterogeneity of the populations experiencing a disease, three different sources of heterogeneity can be accounted for. First, this is the heterogeneity in disease parameters that we termed the parametric heterogeneity, and we discussed this at length in the main text. Secondly, this is the heterogeneity of the social structure, which results in a different contact rates for different individuals. Finally, there is a third source of heterogeneity: the distribution of the infection length. For most

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

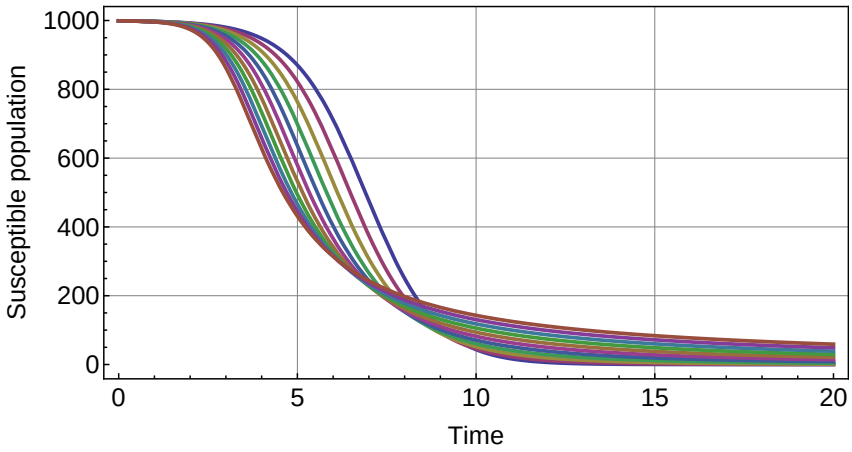


FIGURE 5. Solutions to the heterogeneous SI model (4.18) with distributed contact number. The parameters are $S_0 = 999$, $I_0 = 1, \bar{\omega}(0) = 1$ and the initial distribution of contact number is the gamma distribution with $\sigma^2(0) = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$

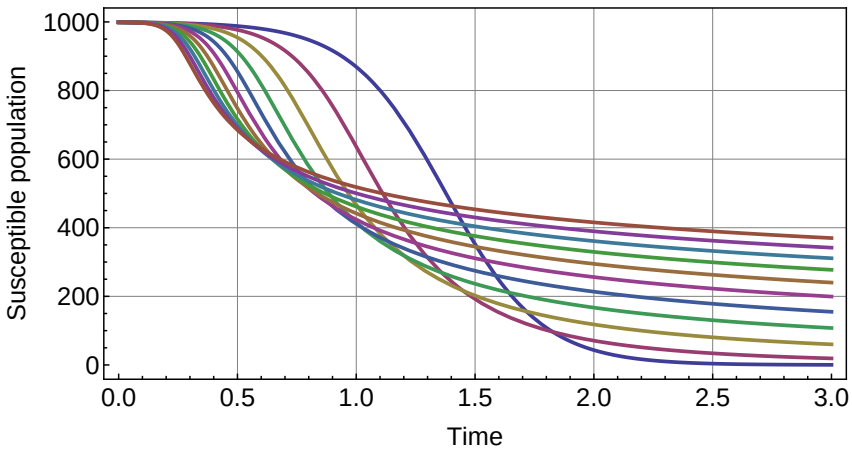


FIGURE 6. Solutions to the heterogeneous SI model (4.18) with distributed contact number. The parameters are $S_0 = 999$, $I_0 = 1, \bar{\omega}(0) = 5$ and the initial distribution of contact number is the gamma distribution with $\sigma^2(0) = 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100$

models in the text it was tacitly assumed that this distribution is exponential with the mean $1/\gamma$ for all the infectious individuals. This assumption is usually made to simplify the mathematics and does not correspond to real situations. Using the approach from Section 2 it is possible to slightly generalize our models by assuming that parameter γ is also distributed through the population. Proceeding along the lines of Section 3, we obtain

$$\begin{aligned} \dot{S} &= -\beta SI, \\ \partial_t i(t,\omega) &= \beta S i(t,\omega) - \gamma(\omega) i(t,\omega), \end{aligned} \tag{5.1}$$

which is equivalent, according to Theorem 2.1, to the system

$$\begin{aligned} \dot{S} &= -\beta SI, \\ \dot{I} &= \beta SI - \partial_\lambda \ln \mathsf{M}(\lambda)|_{\lambda=-t} I, \end{aligned} \tag{5.2}$$

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

with the mgf $\mathsf{M}(\lambda)$ of the initial distribution of γ. Systems (5.1) and (5.2) assume that the population consists of subpopulations, each of which has an exponentially distributed infection length, but this length varies from group to group according to the given initial distribution. A somewhat more interesting approach is to consider instead of (5.1) the following system

$$\begin{aligned} \dot{S} &= -\bar{\beta}(t)SI, \\ \partial_t i(t,\omega) &= \beta(\omega)Si(t,\omega) - \gamma(\omega)i(t,\omega), \end{aligned} \tag{5.3}$$

where $\beta(\omega)$ and $\gamma(\omega)$ are correlated. Model (5.3) is not covered by Theorem 2.1, but still can be tackled with the general theory from [37].

In the mathematical models considered in the text it was always assumed that the population size is closed. This assumption allowed to formulate the models in the form suitable for Theorem 2.1. It is a natural extension to consider models, when the demography and immigration processes are taken into account. This is very important because heterogeneous susceptibility of many diseases can be explained by a heritable genetic basis. For example, we can consider the simplest SIR model with distributed susceptibility and recruitment in the form

$$\begin{aligned} \partial_t s(t,\omega) &= \Lambda s(t,\omega) - \beta(\omega)s(t,\omega)I, \\ \dot{I} &= \bar{\beta}(t)SI - \gamma I. \end{aligned} \tag{5.4}$$

Model (5.4) can be reduced to an equivalent ODE system using Theorem 2.1. However, the long-term behavior of (5.4) is straightforward: the individuals with higher values of the susceptibility parameter will outcompete those with smaller ones. For the model to be realistic it is also required to include the stochastic hereditary element — mutations. That is, a more realistic counterpart of (5.4) is probably

$$\begin{aligned} \partial_t s(t,\omega) &= \Lambda \int_\Omega \varphi(\omega,\eta)s(t,\eta)\,d\eta - \beta(\omega)s(t,\omega)I, \\ \dot{I} &= \bar{\beta}(t)SI - \gamma I, \end{aligned} \tag{5.5}$$

where $\varphi(\omega,\eta)$ gives the probability that a parent with the parameter value η produces offspring with parameter value ω. Model (5.5), however, is not covered by Theorem 2.1. We also note that a very similar to (5.4) model was considered in [22], where, to solve the system, it was conjectured that the equation for the mean parameter value is (cf. (2.13))

$$\dot{\bar{\beta}}(t) = -I\sigma^2(0).$$

This means that is was implicitly assumed that the susceptibility distribution is normal, because the normal distribution is the one that satisfies the condition $\sigma^2(t) = \sigma^2(0)$.

Finally, all the models we discussed so far are deterministic. There is long history of studying stochastic SIR epidemics with multiple classes of susceptible and infectious individuals, e.g., [7, 9, 10, 57], which is usually centered around the asymptotic distributions of the final epidemic size. We announce here that some of the methods presented in this text can be used for studying stochastic models [53].

Acknowledgements. The research is supported in part by the Russian Foundation for Basic Re- search grant # 10-01-00374. The author supported by the grant to young researches from Moscow State University of Railway Engineering.

References

[1] O. O. Aalen. *Effects of frailty in survival analysis.* Statistical Methods in Medical Research, 3 (1994), No. 3, 227.
[2] O. O. Aalen, Ø. Borgan, H. K. Gjessing. Survival and event history analysis: a process point of view. Springer Verlag, 2008.

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

[3] A. S. Ackleh. *Estimation of rate distributions in generalized Kolmogorov community models*. Non-Linear Analysis, 33 (1998), No. 7, 729–745.

[4] A. S. Ackleh, D. F. Marshall, H. E. Heatherly. *Extinction in a generalized Lotka – Volterra predator–prey model*. Journal of Applied Mathematics and Stochastic Analysis, 13 (2000), No. 3, 287–297.

[5] A. S. Ackleh, D. F. Marshall, H. E. Heatherly, B. G. Fitzpatrick. *Survival of the fittest in a generalized logistic model*. Mathematical Models and Methods in Applied Sciences, 9 (1999), No. 9, 1379–1391.

[6] R. M. Anderson, R. M. C. May. Infectious diseases of humans: Dynamics and control. Oxford University Press, New York, 1991.

[7] H. Andersson and T. Britton. Stochastic epidemic models and their statistical analysis. Springer Verlag, 2000.

[8] V. Andreasen. *The final size of an epidemic and its relation to the basic reproduction number*. Bulletin of Mathematical Biology, (2011) in press.

[9] F. Ball. *Deterministic and stochastic epidemics with several kinds of susceptibles*. Advances in Applied Probability, 17 (1985), No. 1, 1–22.

[10] F. Ball and D. Clancy. *The final size and severity of a generalised stochastic multitype epidemic model*. Advances in Applied Probability, 25 (1993), No. 4, 721–736.

[11] S. Bansal, B. T. Grenfell, and L. A. Meyers. *When individual behaviour matters: homogeneous and network models in epidemiology*. Journal of Royal Sosciety Interface, 4 (2007), No. 16, 879–891.

[12] F. S. Berezovskaya, A. S. Novozhilov, G. P. Karev. *Population models with singular equilibrium*. Mathematical Biosciences, 208 (2007), No. 1, 270–299.

[13] B. Bonzi, A. A. Fall, A. Iggidr, G. Sallet. *Stability of differential susceptibility and infectivity epidemic models*. Journal of Mathematical Biology, 62 (2011), No. 1, 39–64.

[14] R. D. Boylan. *A note on epidemics in heterogeneous populations*. Mathematical Biosciences, 105 (1991), No. 1, 133–137.

[15] A. S. Bratus, A. S. Novozhilov, Platonov A. P. Dynamical systems and models in biology. Fizmatlit, 2010.

[16] F. A. B. Coutinho, E. Massad, L. F. Lopez, M. N. Burattini, C. J. Struchiner, R. S. Azevedo-Neto. *Modelling heterogeneities in individual frailties in epidemic models*. Mathematical and Computer Modelling, 30 (1999), No. 1, 97–115.

[17] L. Danon, A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross, M. C. Vernon. *Networks and the epidemiology of infectious disease*. Interdisciplinary Perspectives on Infectious Diseases, 28 (2011).

[18] A. M. De Roos, L. Persson. *Unstructured population models: Do population-level assumptions yield general theory?* In K. Cuddington and B. Beisner, editors, Ecological paradigms lost: Routes of theory change, pages 31–62. Academic Press, 2005.

[19] O. Diekmann, J. A. P. Heesterbeek. Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation, John Wiley, 2000.

[20] O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz. *On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations*. Journal of Mathematical Biology, 28 (1990), No. 4, 365–382.

[21] O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz. *The legacy of Kermack and McKendrick*. In D. Mollison, editor, Epidemic models: Their structure and relation to data, 95–115. Cambridge University Press, 1995.

[22] M.A. Duffy, L. Sivars-Becker. *Rapid evolution and ecological host–parasite dynamics*. Ecology Letters, 10 (2007), No. 1, 44–53.

[23] J. Dushoff. *Host heterogeneity and disease endemicity: A moment-based approach*. Theoretical Population Biology, 56 (1999), No. 3, 325–335.

[24] G. Dwyer, J. Dushoff, J. S. Elkinton, J. P. Burand, S. A. Levin. *Variation in susceptibility: Lessons from an insect virus*. In U. Diekmann, H. Metz, M. Sabelis, and K. Sigmund, editors, Adaptive dynamics of infectious diseases: In pursuit of virulence management, 74–84. Cambridge Univercity Press, 2002.

[25] G. Dwyer, J. Dushoff, J. S. Elkinton, S. A. Levin. *Pathogen-driven outbreaks in forest defoliators revisited: Building models from experimental data*. The American Naturalist, 156 (2000), No. 2, 105–120.

[26] G. Dwyer, J. S. Elkinton, J. P. Buonaccorsi. *Host heterogeneity in susceptibility and disease dynamics: Tests of a mathematical model*. The American Naturalist, 150 (1997), No. 6, 685–707.

[27] J. J. Gart. *The statistical analysis of chain-binomial epidemic models with several kinds of susceptibles*. Biometrics, 28 (1972), No. 4, 921–930.

[28] A. N. Gorban. *Selection theorem for systems with inheritance*. Mathematical Modelling of Natural Phenomena, 2 (2007), No. 4, 1–45.

[29] A. N. Gorban. *Self-simplification in Darwin's systems*. In A. N. Gorban and D. Roose, editors, Coping with complexity: Model reduction and data analysis, pages 311–340. Springer Verlag, 2010.

[30] A. Hastings. *Unstructured models in ecology: past, present, and future*. In K. Cuddington and B. E. Beisner, editors, Ecological paradigms lost: Routes of theory change, pages 9–30. Academic Press, 2005.

[31] J. A. P. Heesterbeek. *The law of mass-action in epidemiology: a historical perspective*. In K. Cuddington and B. E. Beisner, editors, Ecological paradigms lost: Routes of theory change, pages 81–104. Academic Press, 2005.

[32] S. Hsu Schmitz. *Effects of genetic heterogeneity on HIV transmission in homosexual populations*. In C. Castillo-Chavez, editor, Mathematical approaches for emerging and reemerging infectious diseases: Models, methods, and theory, pages 245–260. IMA, 2002.

[33] J. M. Hyman, J. Li. *Differential susceptibility epidemic models*. Journal of Mathematical Biology, 50 (2005), No. 6, 626–644.

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press

[34] G. P. Karev. *Heterogeneity effects in population dynamics*. Doklady Mathematics, 62 (2000), No. 1, 141–144.
[35] G. P. Karev. *Inhomogeneous models of tree stand self-thinning*. Ecological Modelling, 160 (2003), No. 1-2, 23–37.
[36] G. P. Karev. *Dynamics of inhomogeneous populations and global demography models*. Journal of Biological Systems, 13 (2005), No. 1, 83–104.
[37] G. P. Karev. *On mathematical theory of selection: continuous time population dynamics*. Journal of Mathematical Biology, 60 (2010), No. 1, 107–129.
[38] G. P. Karev. *Replicator equations and the principle of minimal production of information*. Bulletin of Mathematical Biology, 72 (2010), No. 5, 1124–1142.
[39] G. P. Karev, A. S. Novozhilov, F. S. Berezovskaya. *On the asymptotic behavior of the solutions to the replicator equation*. Mathematical Medicine and Biology, 28 (2011), No. 2, 89–110.
[40] G. P. Karev, A. S. Novozhilov, E. V. Koonin. *Mathematical modeling of tumor therapy with oncolytic viruses: Effects of parametric heterogeneity on cell dynamics*. Biology Direct, 1 (2006), No. 30, 19.
[41] G. Katriel. *The size of epidemics in populations with heterogeneous susceptibility*. Journal of Mathematical Biology, (2011), in press.
[42] M. J. Keeling, P. Rohani. Modeling infectious diseases in humans and animals. Princeton University Press, 2008.
[43] W. O. Kermack, A. G. McKendrick. *A Contribution to themathematical theory of epidemics*. Proceedings of the Royal Society of London. Series A, 115 (1927) No. 772, 700–721.
[44] W. M. Liu, H. W. Hethcote, S. A. Levin. *Dynamical behavior of epidemiological models with nonlinear incidence rates*. Journal of Mathematical Biology, 25 (1987), No. 4, 359–380.
[45] W. M. Liu, S. A. Levin, Y. Iwasa. *Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models*. Journal of Mathematical Biology, 23 (1986), No. 2, 187–204.
[46] R. M. May, R. M. Anderson. *The transmission dynamics of human immunodeficiency virus*. Proceedings of the Royal Society of London. Series B: Biological Sciences, 321 (1988), 565–607.
[47] H. McCallum, N. Barlow, J. Hone. *How should pathogen transmission be modelled?* Trends in Ecology & Evolution, 16 (2001), No. 6, 295–300.
[48] L.A. Meyers. *Contact network epidemiology: Bond percolation applied to infectious disease prediction and control*. Bulletin of American Mathematical Society, 44 (2007), 63–86.
[49] M. Nikolaou, V. H. Tam. *A New Modeling approach to the effect of antimicrobial agents on heterogeneous microbial populations*. Journal of Mathematical Biology, 52 (2006), No. 2, 154–182.
[50] A. S. Novozhilov. *Analysis of a generalized population predator–prey model with a parameter distributed normally over the individuals in the predator population*. Journal of Computer and System Sciences International, 43 (2004), No. 3, 378–382.
[51] A. S. Novozhilov. *On the spread of epidemics in a closed heterogeneous population*. Mathematical Biosciences, 215 (2008), No. 2, 177–185.
[52] A. S. Novozhilov. *Heterogeneous Susceptibles–Infectives model: Mechanistic derivation of the power law transmission function*. Dynamics of Continuous, Discrete and Impulsive Systems (Series A, Mathematical Analysis), 16 (2009), No. S1, 136–140.
[53] A. S. Novozhilov. *On the stochastic SIR model with heterogeneous susceptibility*. (2012), in preparation.
[54] A. S. Novozhilov, F. S. Berezovskaya, E. V. Koonin, G. P. Karev. *Mathematical modeling of tumor therapy with oncolytic viruses: Regimes with complete tumor elimination within the framework of deterministic models*. Biology Direct, 1 (2006), No. 6, 18.
[55] P. Rodrigues, A. Margheri, C. Rebelo, M. G. M. Gomes. *Heterogeneity in susceptibility to infection can explain high reinfection rates*. Journal of Theoretical Biology, 259 (2009), No. 2, 280–290.
[56] M. Roy, M. Pascual. *On representing network heterogeneities in the incidence rate of simple epidemic models*. Ecological Complexity, 3 (2006), No. 1, 80–90.
[57] G. Scalia-Tomba. *Asymptotic final size distribution of the multitype reed- frost process*. Journal of Applied Probability, 23 (1986), No. 3, 563–584.
[58] N. C. Severo. *Generalizations of some stochastic epidemic models*. Mathematical Biosciences, 4 (1969), 395–402.
[59] P. D. Stroud, S. J. Sydoriak, J. M. Riese, J. P. Smith, S. M. Mniszewski, P. R. Romero. *Semi-empirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing*. Mathematical Biosciences, 203 (2006), No. 2, 301–318.
[60] V. M. Veliov. *On the effect of population heterogeneity on dynamics of epidemic diseases*. Journal of Mathematical Biology, 51 (2005), No. 2, 123–143.
[61] E. B. Wilson, J. Worcester. *The law of mass action in epidemiology*. Proceedings of the National Academy of Sciences of the USA, 31 (1945), No. 1, 24–34.
[62] E. B. Wilson, J. Worcester. *The law of mass action in epidemiology, II*. Proceedings of the National Academy of Sciences of the USA, 31 (1945), No. 4, 109–116.

https://doi.org/10.1051/mmnp/20127310 Published online by Cambridge University Press