Mathematical Biosciences 270 (2015) 81–89

Contents lists available at ScienceDirect

Mathematical Biosciences

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

The influence of assumptions on generation time distributions in epidemic models

Åke Svensson*

Department of Mathematics, Stockholm university, SE-106 91 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 8 May 2015
Revised 7 October 2015
Accepted 8 October 2015
Available online 23 October 2015

Keywords:
Epidemic models
Kermack–McKendrick-model
Epidemic curve
Generation time distribution
Basic reproduction number

ABSTRACT

A simple class of stochastic models for epidemic spread in finite, but large, populations is studied. The purpose is to investigate how assumptions about the times between primary and secondary infections influences the outcome of the epidemic. Of particular interest is how assumptions of individual variability in infectiousness relates to variability of the epidemic curve. The main concern is the final size of the epidemic and the time scale at which it evolves. The theoretical results are illustrated by simulations.

1. Introduction

Epidemics are complex processes. The possibility for an infection to spread in a population is related both to medical-biological properties deciding the interplay between an infectious person and the infectious agent and to social factors involved in contacts between infectious and susceptible individuals.

In this paper we will consider assumptions about randomness. It is a common understanding that chance plays an important part in spread of infections. Epidemics in large populations are mass phenomena and we can expect that the influence of chance on overall properties will, due to some form of the theorem of large numbers, even out. If this is the case it is crucial to understand for which properties it is sufficient to consider mean properties and how, in that case, they are related to the stochastic properties of the infectious agent and the population. It is worth pointing out that it is well-known that randomness influences the outcomes of an epidemic even in large populations. An example is that it always is a positive probability that the spread stops early with only a few infected. Another random outcome is the time it takes for an epidemic to grow large.

The assumptions used to build a model of an epidemic have to be considered carefully. They should include features that are related to the phenomena under study. If the aim is, as it normally should be, restricted to a study of a few aspects it is also recommendable that the assumptions are as simple as possible. A consequence is that the model should only use assumptions that are important for the predictions of the model.

In this paper we will use a simple model for the spread of an infection to study the impact of some basic assumptions of how an infectious agent is transmitted. The aim is to describe human-to-human spread of an infection in a large closed population. The model used has a long history and is basically a stochastic version of the Kermack–MacKendrick model ([7]) applied to a finite population. It is described in Section 2. We will here follow the formulation and terminology of [10]. There are several treatments of models with similar structure, see e.g. [3].

The assumptions are related to how many persons an infected person may infect and when secondary infections occur. The times that elapses from a person is infected till he infects other persons play are important both in applied and theoretical studies of epidemic spread (see e.g. [4], [12], [6], and [11]). These times enter into the model studied here through the generation time distribution. In Section 3 different approaches to assumptions about this distribution are considered and Section 4 contains a discussion how a specific generation time density can be motivated.

The epidemic is assumed to start with the introduction of one (newly) infected person into the population. The focus of the study is how the assumptions are reflected in the appearance of the epidemic curve, which describes how many persons in a population that are infected at time t after the infection entered the population. The appearance of the epidemic curves are analysed using martingale theory in Section 5. Simulated epidemic curves are presented and discussed in Section 6. We will in particular be concerned with the

* Tel.: +468164551.
E-mail address: akes@math.su.se

http://dx.doi.org/10.1016/j.mbs.2015.10.006
0025-5564/

proportion of the population that finally will be infected and at which time scale the epidemic evolves.

In Section 7 we consider non-parametric estimates of basic parameters in the model based on one observed epidemic curve. Since the epidemic curve is based on times of infection that are seldom observed this may seem an unrealistic theoretical exercise. However, the possibility to estimate the parameters that define the model shows what can be recovered from an observation and thus also which assumptions have identifiable impacts on the predictions of the model.

2. A simple epidemic model

We will assume that the epidemic takes place in a closed, finite population with n members. At time $t = 0$, one newly infected person enters the population and starts the infectious spread.

The spread is assumed to depend on two, possible random, entities, λ and K. Here λ is a non-negative (random) number that decides the "total amount of infectivity" spread by an infected person. λ is related to how many secondary infections an infected person may cause. K is a (random) positive measure, with total mass 1, defined on $[0, \infty[$ and is related to how secondary infections are distributed in time. Observe that both λ and K are considered to be random and that they may be dependent. In Section 3 we give examples of how randomness in λ and K can occur.

In the following analysis we will tacitly assume that all measures and functions are regular enough to admit operations that simplify calculations, e.g. differentiation and exchange of the order of integration. Let $K(t) = K([0, t[)$. For simplicity we assume that there exist a density so that

$$K(t) = \int_0^t \kappa(s)ds. \tag{1}$$

The functions $\lambda\kappa(t)$ are referred to as infectiousness functions by Becker [3].

It is assumed that for a given infectious individual possibly infectious contacts, conditional on λ and K, occur according to a Poisson process. The number of contacts in the interval $I = [a, b[$, after infection is thus Poisson distributed with mean $\lambda K(I)$.

The contacted persons are chosen randomly in the population. An infectious contact results in a secondary case if the contact is taken with a susceptible person, i.e. a person that has not been previously infected.

There are several levels of randomness in the model. First there is a random variation of the infectiousness in an infected individual described by the random pair (λ, K). Then there is randomness in how this infectiousness results in possible infectious contacts. A consequence of the model is that the total number of possible infectious contacts taken by a random infected individual follows a mixed Poisson distribution i.e. it is Poisson distributed with the random mean λ.

In large populations when there are many infected individuals it is sometime (but not always) possible to, as least as a first approximation, disregard individual variations. Certain mean values are of particular interest. We will in the following have use of the basic reproduction number and the basic generation time density.

The basic reproduction number, R_0, is often defined as the mean number of secondary cases to an infected individual in a totally susceptible population. We will in this paper define it as the mean number of possible infectious contacts. In the class of models considered here the two definitions are equivalent. Thus

$$R_0 = \mathrm{E}(\lambda). \tag{2}$$

The expectation of random function $\lambda\kappa$, normalized to have total mass 1 is called **the basic generation time density** i.e.

$$g(t) = \frac{\mathrm{E}(\lambda\kappa(t))}{R_0}. \tag{3}$$

[12] discusses relations between epidemic models and demography where primary and secondary infections correspond to mothers and female offspring. In demography (which typically is concerned with large populations) the correspondence of $\mathrm{E}(\lambda\kappa(t))$ is often described as the rate of production of (female) offspring by a mother of age t.

We can also define **the basic generation time distribution**

$$G(t) = \int_0^t g(s)ds = \frac{\mathrm{E}(\lambda K(t))}{R_0}. \tag{4}$$

We will also consider the mean generation time

$$T_0 = \int_0^\infty tg(t)dt. \tag{5}$$

Later we will be concerned with the variability of the epidemic process. For this reason we introduce the variance function

$$V(t) = \frac{\mathrm{Var}(\lambda K(t))}{R_0^2}. \tag{6}$$

Note that

$$V_0 = V(\infty) = \frac{\mathrm{Var}(\lambda)}{R_0^2} \tag{7}$$

is the square of the coefficient of variance of the random variable λ.

3. Models of the generation time density

The generation time density plays an important part in the model. Its role is to explain the times between a primary infection and its secondary infections. We will consider two approaches to motivate assumptions about this density.

3.1. Models with non-random generation time density

A common assumption is that the infectivity of an infected persons develops in time without individual variation. This implies that the function $\kappa(t)$ is constant, i.e.

$$\kappa(t) = g(t), \tag{8}$$

The intensity of the Poisson process that generates possible infectious contacts of an infectious person at time t after infection is $\lambda g(t)$ where λ is a random variable.

3.2. Models with latent and infectious times

In SEIR-models it is assumed that an infection is followed by a period, called the latent period, during which the infected person do not transmit the infection. The latent period is then followed by an infectious period. Both the latent and infectious periods may have random individual duration. In this paper we shall, for simplicity, only consider models where the infectivity is assumed to be constant without individual variation, throughout the infectious time.

Let X be the duration of the latent period, Y the duration of the infectious period, and α the rate at which the possible infectious contacts are taken during the infectious period. Then

$$\lambda\kappa(t) = \alpha I(X < t \le X + Y). \tag{9}$$

Obviously, in this model, the number of secondary infections caused by an infected depends on the length, Y, of the infectious period. The total individual infectivity will be $\lambda = \alpha Y$.

If we normalize κ to be a density of a measure with total mass 1 we get

$$\kappa(t) = I(X < t \le X + Y)/Y. \tag{10}$$

The basic generation number equals

$$R_0 = \alpha\mathrm{E}(\mathrm{Y}), \tag{11}$$