J. Appl. Prob. **35,** 1003–1006 (1998)
Printed in Israel

ON CRITICAL VACCINATION COVERAGE IN MULTITYPE EPIDEMICS

TOM BRITTON,* *Uppsala University*

Abstract

Consider an epidemic outbreak in a large population resulting in an overall proportion infected $\tilde{p}$. The proportion needed to be vaccinated in order to prevent the epidemic, the critical vaccination coverage v_c, depends on individual and social heterogeneities in the population. In the present note it is shown that v_c is larger if, as is likely, individuals differ in terms of susceptibility than if they are equally susceptible.

Keywords: Multitype epidemics; susceptibility; basic reproduction number; critical vaccination coverage

AMS 1991 Subject Classification: Primary 92D30

1. Introduction

The significance of heterogeneities, both individual and social, have received much attention recently in the mathematical and statistical literature on epidemics of infectious diseases, e.g. Ball *et al.* (1997), Becker and Utev (1998). In the present note we focus on individual heterogeneities caused by varying susceptibility. A question of both practical and theoretical relevance is then to compare a heterogeneous population with a homogeneous population. In order to make this comparison meaningful the two populations have to be calibrated. Such comparisons have previously been based on a probabilistic viewpoint by calibrating certain model parameters of the two populations. In the present paper we adopt the approach of a statistician, an approach introduced by Becker and Utev (1998). Suppose an epidemic outbreak resulting in an overall proportion $\tilde{p}$ infected has occurred. Given these data, under which scenario, that of varying or equal susceptibility, is the appropriate estimate of the critical vaccination coverage v_c, i.e. the proportion it is necessary to vaccinate in order to surely prevent an epidemic, higher? It is shown that varying susceptibility always gives a higher v_c than if individuals are assumed equally susceptible. The result holds if vaccinees are drawn at random. If instead the vaccinees are selected in an optimal way there is no general conclusion: which population has the higher v_c depends on how susceptibility varies over the population.

2. The model and previous results

The model now defined is an SIR (susceptible–infectious–removed) model where individuals are grouped according to their susceptibility. The model was first analysed in detail by Ball (1985). Consider a population consisting of n individuals (initially immune individuals are neglected and not counted). Each individual is categorized as one of k types, labelled $1, \ldots, k$; n_i denotes the number of i-individuals and $\pi_i = n_i/n$ the corresponding proportion.

Received 9 March 1998; revision received 23 June 1998.
* Postal address: Department of Mathematics, Uppsala University, Box 480, S-751 06 Uppsala, Sweden.
Email address: tom.britton@math.uu.se

At any time t a susceptible i-individual gets infected at rate $\lambda_i \bar{I}(t)$, where $\bar{I}(t)$ denotes the proportion of infectious individuals at t. An individual who gets infected becomes infectious, possibly after a latency period with arbitrary distribution, and remains so for a period with distribution F (assumed to have mean 1 without loss of generality) after which the individual recovers and becomes immune. All infections and infectious periods are defined to be mutually independent. The epidemic starts with a few infectious individuals and goes on until there are no infectious individuals in the population: the final state of the epidemic. The model parameters $\{\lambda_i\}$ are called *susceptibilities* and the arithmetic mean susceptibility is denoted by $\bar{\lambda} = n^{-1} \sum_j n_j \lambda_j = \sum_j \pi_j \lambda_j$.

Let $\tilde{\boldsymbol{p}} = (\tilde{p}_1, \ldots, \tilde{p}_k)$ denote the proportions infected of each type at the end of the epidemic. Ball (1985) showed that, in case of a major outbreak, $\sqrt{n}(\tilde{\boldsymbol{p}} - \boldsymbol{p})$ converges to a Gaussian zero mean random vector with specified variance matrix as $n \to \infty$ (keeping $\boldsymbol{\pi}$ fixed). The deterministic vector $\boldsymbol{p} = \{p_i\}$ is the unique positive solution to the equations

$$1 - p_i = \mathrm{e}^{-\lambda_i \bar{p}}, \qquad i = 1, \ldots, k, \qquad \text{where } \bar{p} = \sum_j \pi_j p_j. \tag{1}$$

There exists a positive solution to (1) iff the basic reproduction number $R_0 = \sum_j \pi_j \lambda_j > 1$; otherwise there will be a minor outbreak with probability 1. In the present paper we are concerned only with major outbreaks, so we assume $R_0 > 1$.

An interesting question is of course to compare the final size, or equivalently the final overall proportion infected, for different parameter choices. In particular it is interesting to compare a heterogeneous population with a homogeneous one. Ball (1985) showed that a homogeneous population with susceptibility λ (i.e. $\lambda_1 = \ldots = \lambda_k = \lambda$) always results in a larger final size than any heterogeneous population with arithmetic mean susceptibility $\bar{\lambda} = \lambda$. Andersson and Britton (1998) showed that a heterogeneous population has a higher average expected *resistance* than a homogeneous population with the same arithmetic mean susceptibility. They thus suggested an alternative calibration by assuming the same average expected resistance, which is equivalent to having the same *harmonic* mean susceptibility (the harmonic mean is defined by $\check{\lambda} = (\sum_j \pi_j / \lambda_j)^{-1}$). They then compared the law-of-large-number limit of the final size, within the class of populations having the same harmonic mean susceptibility $\check{\lambda}$. It was found that the largest outbreak occurs in a heterogeneous population for $\check{\lambda} < 2/(1-\mathrm{e}^{-2}) \approx 2.31$ but in the homogeneous population when $\check{\lambda} \geq 2/(1-\mathrm{e}^{-2}) \approx 2.31$.

3. Results

A different approach is the calibration a statistician or health practitioner would use, a method previously suggested by Becker and Utev (1998) when considering social heterogeneities. The approach is as follows. Suppose an epidemic outbreak in a large population resulted in a proportion $\tilde{p}$ of infected individuals. A question of practical relevance is then: how large a proportion of randomly chosen individuals would it have been necessary to vaccinate (assuming a perfect vaccine) in order to surely prevent the epidemic, or equivalently to prevent future outbreaks in the present or similar populations? That is, what is the critical vaccination coverage $v_{\mathrm{c}}^{\mathrm{ran}}$? The answer to this question depends on whether or not individuals are equally susceptible. The calibration is thus that the observed overall proportion infected $\tilde{p}$ is fixed. For this fixed proportion $\tilde{p}$ we will compare $v_{\mathrm{c}}^{\mathrm{ran}}$ for different assumptions on heterogeneity in terms of susceptibility.

Suppose that the population could be divided into k different groups according to their susceptibility, assuming (approximately) equal susceptibility within each group, and that the

observed proportion infected in group i was $\tilde{p}_i$, $i = 1, \ldots, k$, satisfying $\sum_i \pi_i \tilde{p}_i = \tilde{p}$. For the epidemic model defined in the previous section, the reproduction number after a randomly chosen proportion v is vaccinated is simply

$$R_v = (1 - v)R_0 = (1 - v) \sum_j \pi_j \lambda_j$$

and v_c^{ran} is the smallest v for which $R_v \leq 1$. This means that

$$v_c^{ran} = 1 - \frac{1}{R_0} = 1 - \left(\sum_j \pi_j \lambda_j \right)^{-1}.$$

From the results stated in Section 2 it follows that $\tilde{p}$ converges in probability to $\bar{p} = \sum_i \pi_i p_i$ as $n \to \infty$, where $\{p_i\}$ is the solution to (1). This implies that λ_i may be estimated consistently by $\hat{\lambda}_i = -\ln(1 - \tilde{p}_i)/\tilde{p}$ and hence also that v_c^{ran} may be estimated consistently by

$$\hat{v}_c^{ran} = 1 - \left(\tilde{p} \Big/ \sum_j \pi_j [-\ln(1 - \tilde{p}_j)] \right).$$

From now on we will assume a large population so that the randomness, which is of order $1/\sqrt{n}$, can be neglected and instead of writing $\tilde{p}_i$ we write the corresponding large population limit p_i. We have the following theorem relating v_c^{ran} for various populations.

Theorem 1. *Among all populations, possibly with varying susceptibility, having final proportion infected $\bar{p}$, v_c^{ran} is minimized for the homogeneous population.*

Proof. The critical vaccination coverage v_c^{ran} has been shown to satisfy

$$v_c^{ran} = 1 - \left(\bar{p} \Big/ \sum_j \pi_j [-\ln(1 - p_j)] \right). \tag{2}$$

In case of a homogeneous population all p_j's are equal (or equivalently the number of subgroups k equals 1). Then the right-hand side of equation (2) reduces to $1 - \bar{p}/[-\ln(1 - \bar{p})]$. The theorem is hence proven if we can show that

$$1 - \left(\bar{p} \Big/ \sum_j \pi_j [-\ln(1 - p_j)] \right) \geq 1 - \frac{\bar{p}}{[-\ln(1 - \bar{p})]}.$$

But this is equivalent to showing

$$\sum_j \pi_j [-\ln(1 - p_j)] \geq -\ln(1 - \bar{p})$$

which follows from the convexity of $f(x) = -\ln(1 - x)$, and the inequality is strict unless $p_i = \bar{p}$ for all i's.

The mathematical explanation for the above result differs from the corresponding result given by Ball (1985). Ball's result relies on the fact that the harmonic mean (of the λ_i's) is dominated by the arithmetic mean and the present result on the fact that the *geometric* mean (this

time of the p_i's) is dominated by the arithmetic mean. Further, the two methods of calibration between two populations do not always point in the same direction: an epidemic with smaller overall proportion infected may indeed have a higher vaccination coverage. For example, an outbreak in a population of two equally frequent types resulting in 40% infected, 20% infected in one type and 60% in the other, has $v_c^{\text{ran}} = 1 - 0.4/(0.5[-\ln(0.8)] + 0.5[-\ln(0.4)]) = 29.8\%$; this is greater than than the critical vaccination coverage of a homogeneous population with 50% infected, for which $v_c^{\text{ran}} = 1 - 0.5/[-\ln(0.5)] = 27.9\%$.

4. Discussion

A practical consequence of the theorem is that a vaccination strategy assumed to prevent future outbreaks may fail to do so if heterogeneities in susceptibility have been neglected. This may in particular happen when heterogeneity in susceptibility is (at least partially) caused by unobservable factors, for example in the immune system, in which case the p_i's would not be observed.

If the heterogeneities are caused by observable factors, i.e. the p_i's are observed, a vaccination strategy can do better than picking individuals to vaccinate at random. If a proportion v_i is vaccinated in group i, $i = 1, \ldots, k$, then the resulting reproduction number is $R_v = \sum_i \pi_i(1 - v_i)\lambda_i$ and the vaccination program surely prevents an outbreak iff $R_v \leq 1$. It is easy to show that the optimal vaccination strategy, the strategy with smallest coverage which surely prevents future epidemics, is the *top-to-bottom* strategy, i.e. to vaccinate all individuals in the groups with highest susceptibility 'down to' the group such that the resulting R_v equals 1. There is no monotonicity result similar to Theorem 1 when employing this vaccination strategy: for any fixed $\bar{p}$, whether the optimal vaccination coverage v_c^{opt} is larger in a homogeneous or a heterogeneous population, depends on the π_i's and p_i's. A homogeneous population has $v_c^{\text{opt}} = 1 - \bar{p}/[-\ln(1 - \bar{p})]$. It is not hard to show that if δ is small enough then the two-type population with $\pi_1 = \pi_2 = 0.5$, $p_1 = \bar{p} - \delta$ and $p_2 = \bar{p} + \delta$ has smaller v_c^{opt} whereas a two-type population with $\pi_1 = 1 - \bar{p}(1 - \delta)$, $\pi_2 = \bar{p}(1 - \delta)$, $p_1 = \bar{p}(1 - (1 - \delta)^2)/(1 - \bar{p}(1 - \delta))$ and $p_2 = 1 - \delta$ has larger v_c^{opt}, and in both examples the overall proportion infected $\bar{p}$ remains unchanged.

In the present note heterogeneities other than those caused by varying susceptibility, such as varying infectivity and heterogeneity due to non-uniform mixing, have been neglected. The conclusion of the present note resembles results in Becker and Utev (1998) treating other heterogeneities. The general rule seems to be: if vaccinations are allocated randomly in the population then heterogeneities imply a higher vaccination coverage, but if vaccinations are allocated optimally there is no unique answer. Of course, more investigation is needed to find in which situations the statement holds and, possibly, when it fails.

References

ANDERSSON, H. AND BRITTON, T. (1998). Heterogeneity in epidemic models and its effect on the spread of infection. *J. Appl. Prob.* **35**, 651–661.

BALL, F. G. (1985). Deterministic and stochastic epidemics with several kinds of susceptibles. *Adv. Appl. Prob.* **17**, 1–22.

BALL, F. G., MOLLISON, D. AND SCALIA-TOMBA, G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Prob.* **7**, 46–89.

BECKER, N. G. AND UTEV, S. (1998). The effect of community structure on the immunity coverage required to prevent epidemics. *Math. Biosci.* **147**, 23–39.