Epidemiol. Infect. (2004), **132**, 927–938. © 2004 Cambridge University Press
DOI: 10.1017/S0950268804002560 Printed in the United Kingdom

A stochastic model of vaccine trials for endemic infections using group randomization

T. W. RIGGS AND J. S. KOOPMAN*

Department of Epidemiology, University of Michigan, 611 Church Street, Room 203, Ann Arbor, MI 48109, USA

(*Accepted 13 March 2004*)

SUMMARY

To clarify the determinants of vaccine trial power for non-typable *Haemophilus influenzae*, we constructed stochastic SIS models of infection transmission in small units (e.g. day-care centres) to calculate the equilibrium distribution of the number infected. We investigated how unit size, contact rate (modelled as a function of the unit size), external force of infection and infection duration affected the statistical power for detection of vaccine effects on susceptibility or infectiousness. Given a frequency-dependent contact rate, the prevalence, proportion of infections generated internally and the power to detect vaccine effects each increased slightly with unit size. Under a density-dependent model, unit size had much stronger effects. To maximize information allowing inference from vaccine trials, contact functions should be empirically evaluated by studying units of differing size and molecular methods should be used to help distinguish internal *vs*. external transmission.

INTRODUCTION

The effect of vaccines on reducing transmission is of increasing interest in a variety of settings. We have demonstrated that vaccines for non-typable *Haemophilus influenzae* (NTHi) will have considerably greater benefit if they prevent transmission than if they only prevent disease given transmission [1]. To gain insight into the determinants of the power of vaccine trials to detect such effects, we constructed models of transmission in small units like day-care centres (DCCs). We simplified our model by assuming that the only source of immunity is from vaccines. A key insight is the role of contact functions (i.e. the contact rate as a function of the number of individuals in the unit) on such assessment [2]. Our model analysis makes clear that contact function is central to detection of vaccine effects on transmission.

The relationship between the contact rate and the number of individuals within the unit is often assumed without much critical evaluation. At one extreme, the number of contacts by an individual per unit time is independent of unit size. Transmission depends only on the proportion of infected and susceptible individuals within the unit; it is the same whether there is 1 infected among 5 or 5 infected among 25 individuals. That extreme is called frequency-dependent contact. At another extreme, the contact rate increases in proportion to the number of individuals within the unit. That extreme is called density-dependent contact. We formulate a contact function that can be adapted to either of these extremes, or to allow intermediate forms. For example, the contact rate could be density dependent within a small unit but approach frequency-dependent contact within a larger unit. We will also demonstrate how the contact function influences the power to detect vaccine effects on transmission.

The statistical power to detect an effect is always a major concern in designing vaccine trials. Power depends on the number of individuals in the trial, the difference between the vaccinated and the

* Author for correspondence.
(Email: jkoopman@umich.edu)

Table 1. *Summary of terms*

Term	Meaning	Comments
P[I]	Probability of exactly 'I' infected in a unit	A distribution of N + 1 mutually exclusive probability states
N	Unit size	
λ_{out}	Coefficient for outside force of infection	Acts on all levels except P[N] to convert susceptible to infected
γ	Recovery rate	Acts on all levels except P[0] to recover susceptible from infected
C_d	Contact rate per individual	Contact rate per individual per time per contactee
$C_{\max}$	Potential contact rate for a unit of unlimited size	Contact rate per individual per time
$C_N = \frac{C_d(\text{N}-1)}{1+(C_d/C_{\max})(\text{N}-2)}$	Contact rate as a function of unit size	Requires estimates of C_d and $C_{\max}$
$\lambda_{\text{in}} = C_N\,\text{I}/(\text{N}-1)$	Inside force of infection	Acts on all levels except P[0] and P[N], to generate a new infection from inside the unit
$0 \leqslant \sigma \leqslant 1$ unvaccinated state: $\sigma = 1$	Susceptibility effect	Decreased probability of acquiring infection post vaccination
$0 \leqslant \kappa \leqslant 1$ unvaccinated state: $\kappa = 1$	Infectiousness effect	Decreased probability of transmitting infection post vaccination
$0 \leqslant 1/\delta \leqslant 1$ unvaccinated state: $\delta = 1$	Duration effect	Decreased length of infectiousness period $(1/\delta)$ after vaccination
$\text{Incidence}_{\text{out}} = \sum_{\text{I}=0}^{\text{I}=\text{N}-1} \lambda_{\text{out}}(\text{N}-\text{I})\text{P[I]}$	Incidence from external forces	Rate of infection generation from P[0, …, N–1] to next level, summed across all levels
$\text{Incidence}_{\text{in}} = \sum_{\text{I}=1}^{\text{I}=\text{N}-1} \lambda_{\text{in}}(\text{N}-\text{I})\text{P[I]}$	Incidence from internal forces	Rate of infection generation from P[1, …, N − 1] to next level, summed across all levels
$\text{Prevalence}_{\text{out}} = \frac{100}{\text{N}(\gamma)}\text{Incidence}_{\text{out}}$	Prevalence from external forces	Percentage of persons in the unit infected from outside the unit
$\text{Prevalence}_{\text{in}} = \frac{100}{\text{N}(\gamma)}\text{Incidence}_{\text{in}}$	Prevalence from internal forces	Percentage of persons in the unit infected from within the unit
$\text{Prevalence}_{\text{all}} = \frac{100}{\text{N}}\sum_{\text{I}=0}^{i=\text{N}} \text{I*P[I]}$	Total prevalence	Percentage of persons in the unit who are infected

unvaccinated groups and the variability within each group, i.e. the probability distributions for the numbers of infected individuals within the vaccinated and the unvaccinated groups are required. Our stochastic model of SIS (susceptible → infected → susceptible) transmission derived the equilibrium probability distribution of the number infected in the unit. This method allowed us to calculate the statistical power of vaccine trials while varying the vaccine effect, the unit size, the external force of infection and through the contact function, the internal force of infection. Our model separated the internal *vs.* external forces and allowed assessment of how their relative magnitude influenced the power to detect vaccine effects on infectiousness or susceptibility. Although the immune response to *H. influenzae* is complex, a previous mathematical analysis [3] suggested that an SIS model fits the data reasonably well.

METHODS

Model description

The contact process formulation

The generation of new infections within the unit is modelled as $C_N \cdot \text{S} \cdot \text{I}/(\text{N}-1)$ where S is the number of susceptible individuals, I is the number of infected, N is the unit size and C_N is the contact function (Table 1). We used the contact function

$$C_N = \frac{C_d(\text{N}-1)}{1+(C_d/C_{\max})(\text{N}-2)}$$

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

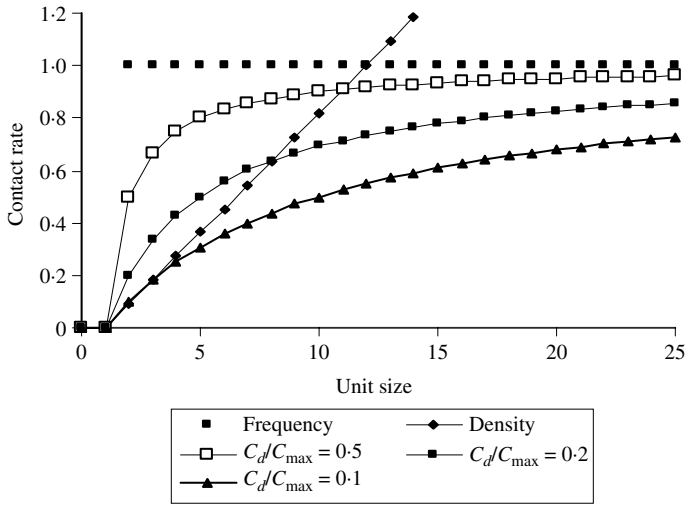

Fig. 1. The contact rate as a function of unit size (the contact function) is illustrated for frequency-dependent, density-dependent and intermediate formulations. The frequency-dependent contact rate is invariant to unit size; the density-dependent increases with unit size and the intermediate formulations initially rise, but approach a limiting value.

to describe the relationship between the contact rate and the number of individuals within the unit. C_d is the rate of contact per contactee, while $C_{\max}$ is the summed rate of contact with all other individuals. In the frequency-dependent model $C_d = C_{\max}$, while in the density-dependent model, $C_{\max} \gg C_d$. Figure 1 illustrates the form of C_N in density-dependent, frequency-dependent and intermediate formulations, when the latter each have the same value for $C_{\max}$. The density-dependent formulation results in a contact rate that increases linearly *vs.* the number of contactees $(N-1)$, the frequency-dependent formulation has a constant contact rate and the others approach $C_{\max}$ asymptotically.

Transmissions given system states

In a simple SIS model, an individual is either infected or susceptible, so the system is completely described by $N+1$ states, corresponding to $(I=0, \ldots, N)$ infected individuals within the unit. We defined the transmission model using the forward Kolmogorov equations [4, 5] to describe the transitions from each state to the adjacent states. There are three factors influencing the transitions: internal (λ_{in}) and external forces of infection (λ_{out}) acting upon each susceptible individual and a recovery rate (γ) acting upon each infected individual (see Table 1). Equilibrium is reached when the generation of new infections (from both the internal and external force of infection) is exactly balanced by the recovery of infected individuals to susceptible status, simultaneously for all P[I] $(I=0, \ldots, N)$ states (see Appendix).

Modelling vaccine effects

We model three potential vaccination effects: a susceptibility effect (σ), acting to decrease the probability of acquiring an infection, an infectiousness effect (κ), acting to decrease the probability of transmitting an infection and a duration effect (δ), acting to decrease the mean duration of the infectious state (see Appendix for further details of the model). Our trial design compares units where none is vaccinated with units where all are vaccinated. The model assumes that the community outside the vaccinated and unvaccinated units is itself unvaccinated and provides a constant external force of infection.

MODEL ANALYSIS

Among the unvaccinated units, there are five parameters (λ_{out}, γ, N, C_d, $C_{\max}$) that influence the number of infected individuals. We can characterize $C_{\max}$ through the ratio $C_d/C_{\max}$ and we show the distribution of the number of infected based on $\gamma=1$, while varying λ_{out}, N and C_d. One could derive another set of distributions based on resetting γ, but inspection of equations (A 1)–(A 3) shows that a rescaling of λ_{out} and C_d by multiplying all equations by the mean duration of infection $(1/\gamma)$ would make the relationships mathematically equivalent to those based on $\gamma=1$. Stated differently, the time unit is arbitrary, so if we adjust the time unit to one recovery period $(1/\gamma)$, then the external and internal transmission rates could each be rescaled to the mean recovery period.

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

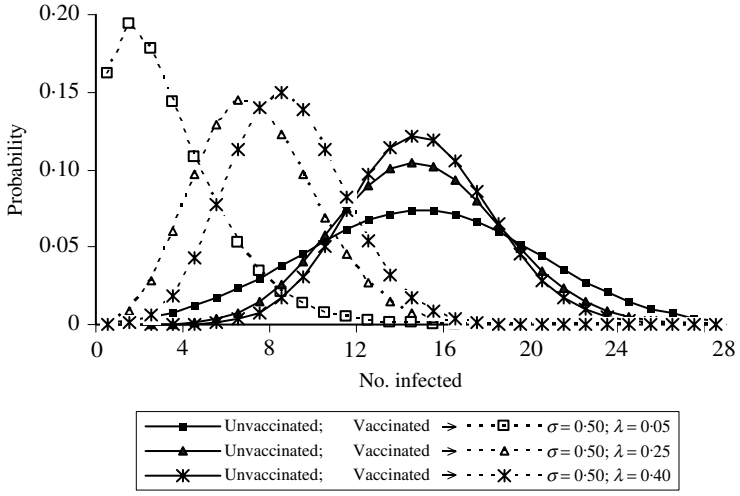

Fig. 2. The overlapping distributions for the expected number of infected individuals in an unvaccinated unit of 48 are contrasted with a vaccinated unit with a vaccine effect to reduce susceptibility by 50%. Mean prevalence in the unvaccinated group was 30%, with high, moderate and low external force of infection. At a lower outside force of infection, there is greater separation of the probability distributions for the unvaccinated and the vaccinated groups, and therefore more statistical power to detect a vaccine effect.

Prevalence attributable from inside *vs.* outside the unit

The primary outcome variable was the equilibrium probability distribution of the number infected within the unit. From this, the total prevalence and the infections generated externally or internally as defined in Table 1 were calculated. Since the primary outcome is a probability distribution, the prevalence and incidence are each described by an array of terms, corresponding to the frequency of each possible system state. The expectation of the number of infections within the unit provides a summary measure for the distribution. Since the mean duration of infection equals 1, the equilibrium prevalence and incidence values are equal. Because we are modelling the endemic prevalence of NTHi, we were most interested in the 10–50% prevalence range.

Table 2. *Infection prevalence and probability of disappearance of infection* vs. *unit size at the extremes of the contact function*

	Density dependent		Frequency dependent	
n	Prevalence	P[0]*	Prevalence	P[0]*
12	0·242	0·137	0·242	0·137
24	0·546	0·0002	0·255	0·017
36	0·696	10^{-9}	0·260	0·002
48	0·771	10^{-16}	0·263	0·0002

* P[0] = probability that no one is infected within the unit.
Prevalence = prevalence of infection within the unit.
This example assumes $\lambda_{out} = 0{\cdot}10$, $\gamma = 1$ for both models. In order to maintain a constant λ_{in} among units of 12, $C_d = C_{max} = 1{\cdot}0$ in the frequency-dependent model and $C_d = 1/11$, $C_{max} = \infty$ in the density-dependent model.

Calculation of power in vaccine trials

The calculation of power is illustrated in Figure 2 for three different outside forces of infection when the prevalence in each situation is 30%. To maintain 30% prevalence when the outside force of infection is increased, the transmission rate within the unit must be decreased. The distributions for the vaccinated and the unvaccinated groups are shown. The statistical power is calculated by first determining the threshold for the number of infected individuals in the unvaccinated distribution that represents the lowest 5th percentile. The cumulative distribution among the vaccinated group, that is less than this threshold, represents the statistical power of the trial. In the trials illustrated, there is less power under conditions of higher external force of infection. With a higher outside force of infection generating the 30% prevalence, the distribution of the number infected within the unvaccinated group has less variance, which would decrease power. However the power is increased, because when transmission rates within the unit are higher, vaccination reduces transmission within the unit more, and the overlap between the vaccinated and unvaccinated groups is greater.

In order to compare trials with units of differing sizes, we kept the total number in each arm of the trial

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

Table 3. *Infection prevalence and proportion of infections generated inside the unit* vs. *unit size and external force of infection* (λ_{out}) *for a frequency-dependent contact function*

	$n=12$		$n=24$		$n=48$	
λ	Prevalence	Internal proportion	Prevalence	Internal proportion	Prevalence	Internal proportion
0·10	0·242	0·686	0·255	0·709	0·263	0·719
0·15	0·299	0·648	0·309	0·665	0·314	0·672
0·20	0·343	0·616	0·351	0·63	0·355	0·636
0·25	0·378	0·589	0·384	0·6	0·387	0·605
0·30	0·408	0·565	0·413	0·574	0·416	0·578
0·35	0·433	0·543	0·438	0·551	0·44	0·555
0·40	0·456	0·524	0·46	0·53	0·462	0·534
0·45	0·476	0·506	0·48	0·512	0·481	0·515
0·50	0·495	0·49	0·498	0·495	0·499	0·498

$C_d = C_{max} = 1{\cdot}0$ in the frequency-dependent model.

Table 4. *Infection prevalence and proportion of infections generated inside the unit* vs. *unit size and external force of infection* (λ_{out}) *for a density-dependent contact function*

	$n=12$		$n=24$		$n=48$	
λ	Prevalence	Internal proportion	Prevalence	Internal proportion	Prevalence	Internal proportion
0·10	0·242	0·686	0·546	0·917	0·771	0·97
0·15	0·299	0·648	0·564	0·884	0·775	0·956
0·20	0·343	0·616	0·58	0·855	0·778	0·943
0·25	0·378	0·589	0·594	0·829	0·781	0·93
0·30	0·408	0·565	0·606	0·805	0·784	0·917
0·35	0·433	0·543	0·618	0·783	0·787	0·905
0·40	0·456	0·524	0·628	0·763	0·79	0·894
0·45	0·476	0·506	0·638	0·744	0·793	0·882
0·50	0·495	0·49	0·648	0·726	0·795	0·871

$C_d = 1/11$, $C_{max} = \infty$ in the density-dependent model, so that the $n=12$ units are identical to those in Table 3.

constant by comparing trials with four units of size 12, two units of size 24 and one unit of size 48. Assuming independence for the trials, we used the probability distribution from one trial to calculate the expected distribution of k replications by expanding the initial distribution to the kth power. This approach was used to derive the distributions from which the total prevalence, the prevalence attributable to internal and external forces and the power were calculated.

RESULTS

The effect of the contact function on prevalence and fraction of infections internally generated

Tables 2–4 illustrate the influence of the contact function and unit size on the endemic infection prevalence, the probability that no one within the unit is infected and the proportion of infections generated internally. Given a frequency-dependent contact function, as the unit size increases, the prevalence rises <10%, and the probability of the unit having no infected individuals dramatically decreases. Given a density-dependent contact function, as the unit size increases, the prevalence increases >200% and the probability of the unit having no infected individuals decreases much faster than under the frequency-dependent model. Equivalency of contact rates among units of size 12 was maintained by setting $C_d = C_{max} = 1{\cdot}0$ for the frequency-dependent model and $C_d = 1/11$ and $C_{max} = \infty$ for the density-dependent model.

For both the frequency-dependent and the density-dependent models, as the external force of infection

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

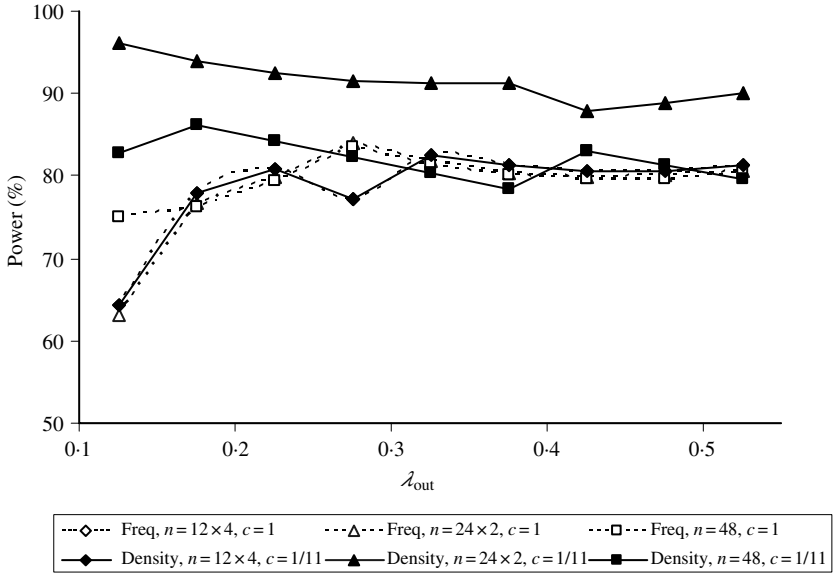

Fig. 3. The power to detect a 50% reduction in susceptibility is shown *vs.* the unit size, the outside force of infection and the contact function. The power to detect a susceptibility effect was dependent on the overall prevalence and the proportion of infections generated from inside *vs.* from outside the unit, which was in turn related to the contact function.

increased, the proportion of infections arising from transmission within the unit decreased for each unit size, with slightly higher proportions of internally generated infections among the larger units. For the density-dependent model, the proportion of internal infections increased with the unit size. Therefore, the contact function had a huge impact on the internal force of infection, the relationship between the internal and external force of infection and the total prevalence.

Estimating λ_{out}

Our model assumes a constant outside force of infection, independent of unit size. This could be verified in a vaccine trial by collecting data by unit size and including the frequency of the P[I] states, rather than a summary measure like the mean number infected across units. An estimate of the external force of infection could be made from rearrangement of equation (A 4) as $\lambda_{out}=(\gamma\cdot P[1])/(N\cdot P[0])$. Thus λ_{out} can be estimated once γ is known and its estimation is independent of the form of the contact function. Once λ_{out} and the probability distribution of the number of infected individuals within the unit are estimated, then the proportion of infections attributable to λ_{out} can be calculated. However, this estimate depends on a relatively high frequency of the P[0] and P[1] states (i.e. mostly in small units and low endemic levels of infection).

Comparison of vaccine effects in this model

A vaccine effect that reduces susceptibility by a proportion 'σ' will have a greater effect on reducing prevalence of infection than a vaccine that reduces infectiousness by the same proportion, since the former diminishes both λ_{out} and λ_{in}, while the latter selectively decreases λ_{in}. A vaccine effect on duration of infection ($1/\delta$) is algebraically equivalent to a susceptibility effect in this SIS model; that is, increasing the recovery rate by a factor of 2 is equivalent to reducing susceptibility by $\frac{1}{2}$ (see Appendix).

Power to detect a susceptibility effect

Figure 3 shows the power to detect a moderate effect on susceptibility ($\sigma=0{\cdot}50$, $\kappa=1{\cdot}0$) under the frequency-dependent and the density-dependent models for the various unit sizes and a variable external force of infection. In the frequency-dependent model, the power to detect a susceptibility effect increased as the external force increased and there was little difference comparing four replications of $n=12$, two replications of $n=24$ and a single trial of $n=48$. In the density-dependent model, the power to detect a susceptibility effect is more sensitive to the unit size than in the frequency-dependent model. As the unit size increases in the density-dependent model and the prevalence increases, there are fewer susceptible individuals in the unit and the power to detect a

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

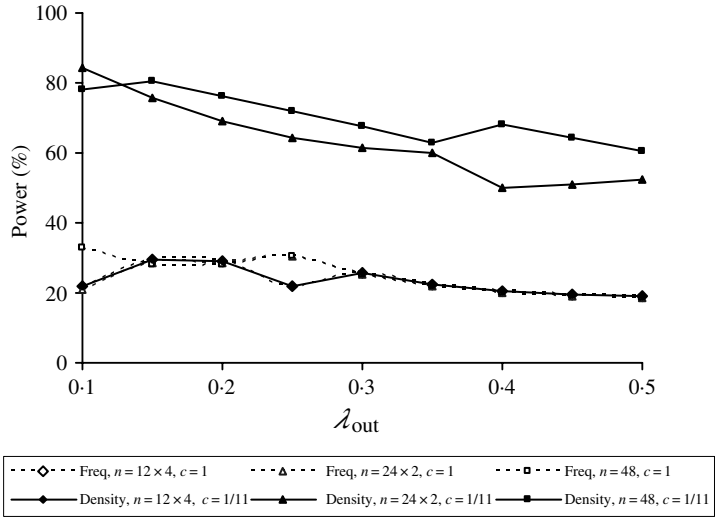

Fig. 4. The power to detect a 50% reduction in infectiousness is shown *vs.* the unit size, the outside force of infection and the contact function. The power to detect an infectiousness effect was generally less than the power to detect a susceptibility effect, but was more strongly related to the contact function and the unit size.

susceptibility effect is diminished (see two replications of $n=24$ *vs.* one trial of $n=48$ in Fig. 3).

Power to detect an infectiousness effect

The contact function has a much stronger influence on the power to detect an infectiousness effect ($\kappa=0{\cdot}5$, $\sigma=1{\cdot}0$), as shown by Figure 4. In the ranges examined, if the contact process was frequency dependent, then the unit size had only a modest effect on the power to detect a vaccine effect. If the contact function was density-dependent, then the power to detect an infectiousness effect was sensitive to the unit size and to the external force of infection, since the former will raise the proportion of infections generated inside the unit and the latter will lower it. So the power to detect an infectiousness effect is greatest at low levels of external force of infection and high contact rates generated within a larger unit size by the density-dependent contact function.

Power to detect an effect on susceptibility or infectiousness, given constant prevalence

Since our model is based on an endemic infection, we illustrated the power to detect vaccine effects on susceptibility or infectiousness at constant prevalence. With each value of λ_{out} we calculated the value of C_d to obtain a constant endemic prevalence (among the unvaccinated group) of 20, 30, 40, 50 and 60%. Figure 5 shows the power to detect a susceptibility effect ($\sigma=0{\cdot}50$, $\kappa=1{\cdot}0$) at those constant prevalence levels. At all prevalence values, the power decreases as the external force increases, but the slope is more negative at lower prevalence values. In the range illustrated, in order to obtain a reasonable value for the power of the trial, a prevalence of $\geqslant 0{\cdot}40$ is required. These curves do not show a monotonic decrease; there are discrete jumps in the value for the power. This is a consequence of the probability distribution having discrete values and the necessity of using a threshold where α has a value as close as possible to (but less than) 0·05. When the distribution changes slightly so that $\alpha<0{\cdot}05$ shifts to the next highest number of infected individuals, the power abruptly increases.

The power to detect an infectiousness effect ($\kappa=0{\cdot}50$, $\sigma=1{\cdot}0$) under the same condition of constant prevalence is shown in Figure 6. The power decreases as the external force increases across all levels of prevalence, but the slope of power *vs.* external force of infection is more negative than for the susceptibility effect. To detect an infectiousness vaccine effect, both a low external force and a reasonable overall prevalence are required. That is, most of the infections must be generated internally and their number must be sufficiently large in order to detect a significant difference given a vaccine infectiousness effect.

DISCUSSION

To capture all vaccine effects on transmission, one should study populations in settings where

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

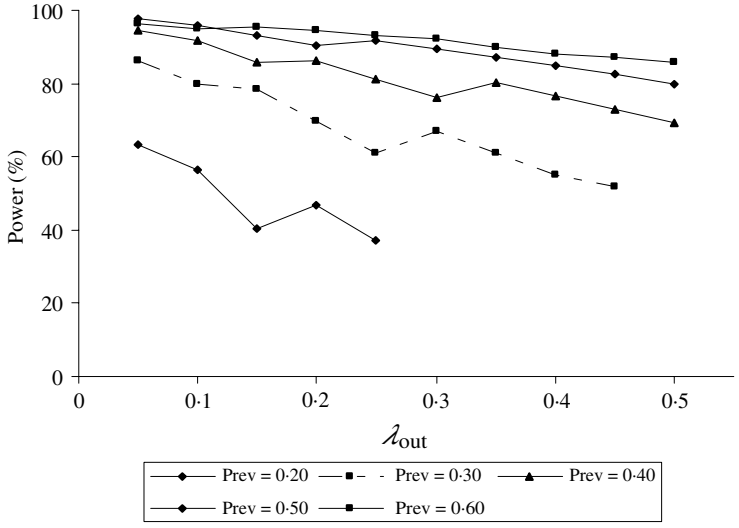

Fig. 5. The power to detect a 50% reduction in susceptibility is plotted *vs*. the outside force of infection at contours of constant prevalence. The abrupt jumps in the curve occur as a result of the probability distributions being discrete, rather than continuous.

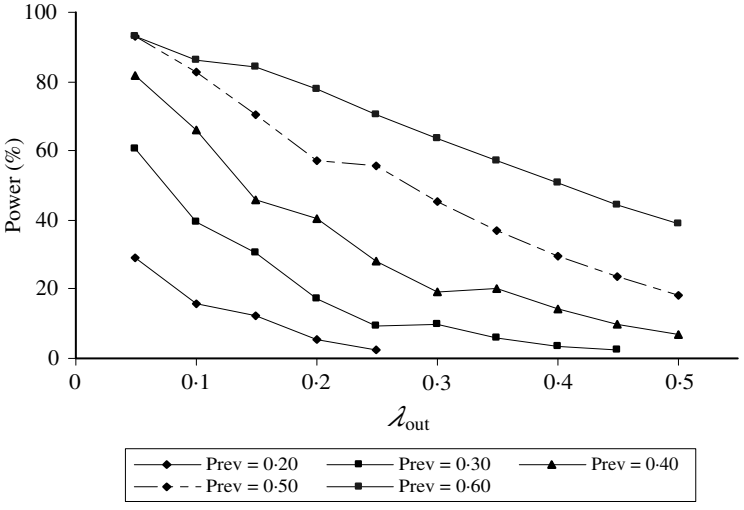

Fig. 6. The power to detect a 50% reduction of infectiousness is plotted *vs*. the outside force of infection at contours of constant prevalence. As the proportion of infections generated from within the unit decreases, there is less power to detect an infectiousness effect.

transmission occurs. The planning for such trials and the information to be extracted from them depends upon a correct formulation of the transmission process. Transmission has been summarized (e.g. ref. [6]) as occurring by 'mass-action', or proportional to $S \cdot I$, where S and I represent the numbers of susceptible and infected individuals present per unit area. Many others (e.g. refs [1, 7, 8]), have modelled the transmission rate as proportional to $S \cdot (I/N)$, dividing the product by the total population (N). The terminology 'density dependent' for the former and 'frequency dependent' for the latter has been suggested previously [2].

Contact rate as a function of unit size

The proper formulation of the contact function is central for modelling transmission dynamics within small units, since contact rates determine the internal force of infection. The density-dependent *vs*. frequency-dependent models represent extremes that will not apply across all unit sizes or with all infectious agents. The density-dependent model leads to a steady rise in the rate of infection transmission as the number of infected individuals within the unit increases, even if their proportion is fixed; this quickly generates high internal forces of infection.

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

The frequency-dependent model holds the transmission rate constant as unit size changes, and thus, the internal force is invariant to unit size. While the density-dependent model may be a reasonable description of the contact process for units with a small number of people, the process might saturate as the unit size increases. Particularly among small to moderately sized units (e.g. 5–40 individuals, as encountered in a DCC), the contact rate may vary considerably with N.

Using a contact function that is proportional to N for small values of N and reaches an asymptote for large values of N has been suggested previously [8–10]. McCallum et al. [2] expressed this as a transmission function

$$\frac{\mathrm{N}}{1+\varepsilon(\mathrm{N}-1)} \cdot \frac{\beta \cdot \mathrm{I} \cdot \mathrm{S}}{\mathrm{N}},$$

in which two parameters (ε, β) modify the contact rate. If $\varepsilon = 0$, the transmission is density dependent; if $\varepsilon = 1$, it is frequency dependent. While their formulation, and ours, each have two parameters, we suggest that ours is more easily understood as C_d and $C_{\max}$ relate to contact rates per potential contactee and contact rates independent of unit size, rather than the dimensionless constant, ε. By stipulating the value for ε and β in their model or of C_d and $C_{\max}$ in ours, one could construct a relationship of prevalence *vs.* unit size that is between the frequency-dependent and the density-dependent functions. There are more complex formulations between the extremes of density-dependent and frequency-dependent models [2] that are based on pair formation [11] or a power relationship [12] of the number of susceptible and infected individuals. One could also model a contact function that has a much steeper rate of rise for a given unit size by raising the $[\mathrm{N}-1]$ and $[\mathrm{N}-2]$ terms to a power, analogous to the Hill equation [13] in Michaelis–Menten dynamics. By empirical evaluation of prevalence and strength of internal force of infection as a function of unit size, the contact function could be estimated as frequency dependent, density dependent or intermediate.

Vaccine effects

Ideally, a vaccine confers complete protection (a 'sterilizing' effect) so that no matter how intense the internal or external force of infection, the vaccinated individual will not become infected [14]. Although some vaccines approach this ideal, many have relative effects on susceptibility, infectiousness, duration or pathogenicity. The success of *H. influenzae* type b vaccine [15, 16] has been cited as an example of the importance of infectiousness effects [17]. Pertussis vaccine is another example of an important infectiousness effect [18]. However, our study is not concerned with estimation of vaccine effects; we used hypothetical pure vaccine effects on susceptibility, infectiousness or duration to illustrate how the power to detect a difference in prevalence would be affected in different unit sizes and under scenarios of frequency- or density-dependent contact functions. A thorough discussion of this topic can be found in ref. [19], while a general framework for considering the data required to estimate transmission effects and to separate susceptibility from infectiousness effects is in ref. [20]. The present study design, where vaccination coverage levels are all *vs.* none, will not allow separation of susceptibility and infectiousness vaccine effects. Longini and colleagues propose a model for measuring both susceptibility and infectiousness effects in populations vaccinated at differing coverage levels [17]. All these approaches must specify a model of the contact process, but this is usually modelled as a constant contact rate (i.e. implicitly as a frequency-dependent contact function). Our study demonstrates that a vaccine trial design for an endemic infection requires more than the prevalence among unvaccinated *vs.* vaccinated groups to estimate the power of the trial. One must have some knowledge regarding the contact function in the range of unit sizes to be tested, the relative strength of internal *vs.* external force of infection acting on the units and the degree to which the vaccine effect relates to infectiousness or susceptibility.

Observational data from DCCs

A likely setting for vaccine trials of NTHi focused on detecting transmission effects are DCCs. Attendance at a DCC is an important risk factor for colonization with *H. influenzae* [21], and siblings who attended DCCs were the principal source of colonization for younger infants who themselves were not yet enrolled [22]. However, there are few studies comparing transmission across DCC unit size. One study [23] demonstrated that prevalence of *H. influenzae* infection is related to DCC size; among units of size <10, prevalence was 32%, among units 10–50, the prevalence was 38% and among units >50, the

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

prevalence was 58%. In another study [24], the proportion of ampicillin-resistant strains of *H. influenzae* was related to the size of the DCC. In units having <20 children, 35% of strains were resistant, in units of 20–50 children, 34% were resistant and in units with >50 children, 63% were resistant. More importantly, when the bacterial DNA pattern was classified as to whether there was a shared pattern in the DCC, in units <20, there was never a shared pattern between two individuals. Among units of size 20–50, 32% shared a pattern with others and among the largest units, 52% shared a common pattern with others in the DCC. The authors conclude that the size of the DCC may affect the risk of transmission of ampicillin-resistant strains. In another study, [25], carriage rates for *H. influenzae* were 37% for children in a DCC compared with 11% among controls. DNA analyses of the *H. influenzae* matched in 38% of the DCC children but was only 4% in the controls. When the DCC children were examined serially, a child's newly acquired strain was the same as that previously found within the same DCC in 40% of cases. The size of the DCC appears to be an important determinant of the internal force of infection for *H. influenzae*; this must be understood in order to properly model transmission within DCCs.

Our analysis demonstrates the effect of the contact function and, therefore, unit size on the internal force of infection in small to moderately sized units. This unit size effect is supported by the clinical data available from DCCs of differing sizes. Further studies should be performed both to observe the effect of DCC size on the overall prevalence of *H. influenzae* and specifically to distinguish via DNA analyses those infections generated internally *vs.* externally. Accurate tracking of transmission by strain will allow better estimation of the outside *vs.* inside force of infection acting on the unit. As shown by Figures 3 and 4, the power to detect vaccine effects, especially those on infectiousness, is profoundly influenced by the proportion of infections generated from outside *vs.* inside the unit.

In this model, we compared units where everyone or no one in the unit was vaccinated and where the external force of infection was both constant and arose from a community that was not vaccinated. These simplifications helped us to highlight that under these conditions, vaccines reducing susceptibility would lower infection rates from both internal and external forces while vaccines reducing infectiousness would lower infection rates only if they were generated from internal forces. This model has no age-dependent structure in terms of prevalence, contact function or vaccine effects. Our conclusions may not apply to vaccine trials in populations having some degree of acquired immunity from natural infections. The next step is to extend these analyses to a more realistic model having both natural immunity and vaccine effects.

ACKNOWLEDGEMENTS

Supported by NIH grant no. 5 RO1 AI45168.

APPENDIX (see Table 1 for definitions of terms)

The Kolmogorov forward equations describing the transitions between the states are of three types:

$$dP[0]/dt = -\lambda_{out}\, NP[0] + \gamma P[1], \tag{A 1}$$

$$\begin{aligned} dP[I]/dt \text{ for } I = 1, \ldots, N-1 = {} & \lambda_{out}(N-I+1)P[I-1] \\ & -\lambda_{out}(N-I)P[I] + \frac{C_N(I-1)}{(N-1)}(N-I+1)P[I-1] \\ & -\frac{C_N I}{(N-1)}(N-I)P[I] \\ & -\gamma I P[I] + \gamma(I+1)P[I+1], \end{aligned} \tag{A 2}$$

$$\begin{aligned} dP[N]/dt = {} & \lambda_{out}\, P[N-1] + \frac{C_N}{(N-1)}(N-1)P[N-1] \\ & -\gamma N P[N]. \end{aligned} \tag{A 3}$$

For the vaccinated group, the equations are modified by multiplying each λ_{out} by σ, each C_N by $\sigma \cdot \kappa$ and each γ by $(1/\delta)$. The equivalence of reciprocal changes in susceptibility *vs.* duration effects can be appreciated by multiplying through by δ, and collecting factors: the duration effect $(1/\delta)$ is equivalent to a susceptibility effect $(\sigma \cdot \delta)$.

The three equations could also be summarized in matrix form as:

$dP/dt = AP$, where A is a square tridiagonal matrix whose non-zero elements are

supradiagonal elements:

$A_{I,I-1} = (N-I+1)[(I+1)C_N + \lambda_{out}]$

diagonal elements:

$A_{I,I} = -[I \cdot \gamma + (N-I)IC_N + (N-I)\lambda_{out}]$

subdiagonal elements:

$A_{I,I+1} = (I+1)\gamma \quad (0 \leqslant I \leqslant N).$

https://doi.org/10.1017/S0950268804002560 Published online by Cambridge University Press

The solution

We used Berkeley Madonna software [26] to numerically solve the set of differential equations, but an exact solution to the N+1 equations could be found by iteration, beginning with the P[0] state and substituting for the next state, with the constraints that $\sum_{I=0}^{I=N} P[I]=1$ and $\gamma>0$:

$$P[1]=\frac{\lambda_{out}NP[0]}{\gamma}, \quad \text{(A 4)}$$

$$P[2]=\frac{1}{2\gamma}[\lambda_{out}(N-1)+C_N]P[1], \quad \text{(A 5)}$$

In general, for [I, …, N−1], P[I+1] can be expressed in terms of the previous two states:

$$\begin{aligned}P[I+1]=&\frac{P[I]}{\gamma(I+1)}(\gamma(I)+\lambda_{out}(N-I)\\&+C_N I(N-I)/(N-1))\\&-\frac{P[I-1]}{(\gamma)(I+1)}(\lambda_{out}(N-I+1)\\&+C_N(I-1)(N-I+1)/(N-1))\end{aligned} \quad \text{(A 6)}$$

and

$$P[N]=\frac{1}{N\gamma}(\lambda_{out}+C_N)P[N-1]. \quad \text{(A 7)}$$

All of these can be expressed iteratively in terms of P[0] and since $\sum_{I=0}^{I=N} P[I]=1$, the coefficients of P[0] would sum to the normalizing factor, allowing calculation of P[0, …, N]. Therefore, an exact (non-ergodic) solution exists, showing that there is a unique set of equilibrium values for the P[I] states.

The solution when there is no outside force of infection

In the special case of no outside force of infection ($\lambda_{out}=0$, $\gamma>0$), it can be shown by induction that the P[0] state will ultimately occur with probability = 1.

REFERENCES

1. Koopman JS, Lin X, Chick SE, Gilsdorf JR. Transmission model analysis of nontypeable Haemophilus influenzae. In: Sainfort F, Brandeau ML, Pierskalla WP, eds. Handbook of operations research/management science applications in health care. Dordrecht: Kluwer Academic Publishers, 2003 (in press).
2. McCallum H, Barlow N, Hone J. How should pathogen transmission be modelled? Trends Ecol Evol 2001; **16**: 295–300.
3. Coen PG, Heath PT, Barbour ML, Garnett GP. Mathematical models of *Haemophilus influenzae* type b. Epidemiol Infect 1998; **120**: 281–95.
4. Ross SM. Stochastic processes, 2nd edn. New York: John Wiley & Sons, 1996.
5. Mode CJ, Sleeman CK. Stochastic processes in epidemiology: HIV/AIDS, other infectious diseases and computers. River Edge, New Jersey: World Scientific Publishing Co., 2000.
6. Anderson RM, May RM. Population biology of infectious diseases: Part I. Nature 1979; **280**: 361–368.
7. Hethcote HW, ed. Three basic epidemiologic models. New York: Springer, 1989.
8. de Jong MCM, Diekmann O, Heesterbeek H. How does transmission of infection depend on population size? In: Mollison D, ed. Epidemic models: their structure and relation to data. Cambridge, UK: Cambridge Publishers, 1995: 84–94.
9. Diekmann O, Heesterbeek H. 'What is contact?' In: Levin S, editor-in-chief. Mathematical epidemiology of infectious diseases. New York: John Wiley & Son, Ltd., 2000: 153–169.
10. Dietz K, ed. Overall population patterns in the transmission cycle of infectious disease agents. New York: Springer-Verlag, 1982.
11. Heesterbeek JAP, Metz JAJ. The saturating contact rate in marriage and epidemic models. J Math Biol 1993; **31**: 529–539.
12. De Leo GA, Dobson AP. Allometry and simple epidemic models for microparasites. Nature 1996; **379**: 720–722.
13. Murray JD. Mathematical biology, 2nd edn. New York: Springer, 1993.
14. Smith PG, Rodrigues LC, Fine PE. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. Int J Epidemiol 1984; **13**: 87–93.
15. Ten Great Public Health Achievements – United States, 1900–1999. MMWR 1999; **48**: 241–248.
16. Takala AK, Santosham M, Almeido-Hill J, et al. Vaccination with Haemophilus influenzae type b meningococcal protein conjugate vaccine reduces oropharyngeal carriage of Haemophilus influenzae type b among American Indian children. Pediatr Infect Dis J 1993; **12**: 593–599.
17. Longini Jr IM, Sagatelian K, Rida WN, Halloran ME. Optimal vaccine trial design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. Stat Med 1998; **17**: 1121–1136.
18. Preziosi MP, Halloran ME. Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness. Vaccine 2003; **21**: 1853–1861.
19. Becker NG. Analysis of infectious disease data. Boca Raton, FL: CRC Publishers, 1989.
20. Rhodes PH, Halloran ME, Longini Jr IM. Counting process models for infectious disease data: distinguishing exposure to infection from susceptibility. J Roy Stat Soc B 1996; **58**: 751–762.
21. Principi N, Marchisio P, Schito GC, Mannelli S. Risk factors for carriage of respiratory pathogens in the

nasopharynx of healthy children. Ascanius Project Collaborative Group. Pediatr Infect Dis J 1999; **18**: 517–523.
22. Vives M, Garcia ME, Saenz P, et al. Nasopharyngeal colonization in Costa Rican children during the first year of life. Pediatr Infect Dis J 1997; **16**: 852–858.
23. Mohle-Boetani JC, Ajello G, Breneman E, et al. Carriage of Haemophilus influenzae type b in children after widespread vaccination with conjugate Haemophilus influenzae type b vaccines. Pediatr Infect Dis J 1993; **12**: 589–593.
24. Talon D, Leroy J, Dupont MJ, et al. Antibiotic susceptibility and genotypic characterization of Haemophilus influenzae strains isolated from nasopharyngeal specimens from children in day-care centers in eastern France. Clin Microbiol Infect 2000; **6**: 519–524.
25. Peerbooms PG, Engelen MN, Stokman DA, et al. Nasopharyngeal carriage of potential bacterial pathogens related to day care attendance, with special reference to the molecular epidemiology of Haemophilus influenzae. J Clin Microbiol 2002; **40**: 2832–2836.
26. Macey R, Oster G, Zahnley T. Berkeley Madonna, version 8.0. Berkeley, CA: University of California, 2000.