

ESTIMATING VACCINE EFFICACY UNDER THE HETEROGENEITY OF VACCINE ACTION IN A NONRANDOMLY MIXING POPULATION

M. Shafiqur Rahman

Department of Statistical Science, University College London, London,
United Kingdom and Institute of Statistical Research and Training,
University of Dhaka, Dhaka, Bangladesh

Vaccines often have heterogeneous actions because of possible variation in the immune systems of hosts. One must consider such heterogeneity of vaccine action when developing a vaccine efficacy parameter. Addressing this issue the summary model of vaccine action has been proposed in the literature to estimate vaccine efficacy in a randomly mixing population. However, nonrandom mixing is common, particularly in a small-group-mixing population. This article extends the summary model of vaccine action to such a nonrandomly mixing population. The interpretation and estimation of the summary vaccine efficacy were discussed in light of other two models of vaccine action: the leaky and all-or-nothing model. Vaccine efficacy under all models is defined as the relative reduction in transmission probability due to vaccine. Estimation of the transmission probabilities is described based on a deterministic epidemic model of an acute transmitted disease. This article further discusses, based on the above vaccine models, the estimation of vaccination coverage required to control epidemic. Methods are illustrated using data simulated by considering different patterns of mixing and vaccine action. Results confirm that the summary model performs better than other two models when vaccine action is heterogeneous.

Key Words: Heterogeneous vaccine action; Nonrandom mixing population; Vaccine efficacy.

1. INTRODUCTION

The design and implementation of a vaccination program for the control of infectious diseases requires a consistent estimate of vaccine efficacy. Accordingly, estimation of vaccine efficacy has received much attention in the literature. Most such literature often ignores the fact that the immune response tends to vary among the hosts, and thus the vaccines have heterogeneous action. Halloran et al. (1992) pointed out that one should consider such heterogeneity of vaccine action when developing a vaccine efficacy parameter. To address this issue, some studies (Halloran et al., 1996; Longini and Halloran, 1996; Becker and Utev, 2002), including Halloran et al. (1992) discussed the estimation of vaccine efficacy in a randomly mixing population. However, nonrandom mixing is common in

© M. Shafiqur Rahman

Received September 12, 2010; Accepted July 8, 2011

Address correspondence to M. Shafiqur Rahman, PhD, Institute of Statistical Research and Training, University of Dhaka, Dhaka-1000, Bangladesh; E-mail: shafiq@isrt.ac.bd

practice; for example, in a population stratified into different small groups such as households, where within-group mixing is usually higher than between groups and thus the overall mixing is nonrandom. Therefore, two important issues one needs to consider when estimating vaccine efficacy in a small-group-mixing population are heterogeneous vaccine action and nonrandom mixing.

A few studies have estimated vaccine efficacy in a nonrandomly mixing population (Haber et al., 1991a, 1995; Halloran et al., 1996; Longini and Halloran, 1996; Davis et al., 2006); however, these studies have ignored the heterogeneity of vaccine action across the vaccinated strata. Moreover, some other studies (Becker and Starczak, 1997; Ball and Lyne, 2002; Ball et al., 2004a,b; Ball and Lyne, 2006; Pellis et al., 2009) have modeled the epidemic in such a stratified population and discussed the estimation of optimal vaccination coverage under models that do not allow for the heterogeneity of vaccine action. The present work is designed to estimate vaccine efficacy under the heterogeneity of vaccine action in a population that is stratified into households, for a particular outbreak of acute, directly transmitted, infectious disease. In addition, this article focuses on estimation of optimal vaccination coverage, a fraction of population that needs to be vaccinated to control epidemic.

The article is organized as follows. In section 2, we briefly discuss some existing measures of vaccine efficacy under the heterogeneity of vaccine action. A deterministic model of an acute transmitted disease in a nonrandomly mixing population is discussed in section 3. Estimation of vaccine efficacy in a stratified population is described in section 4, and section 5 discusses the variance approximation of the estimator. Section 6 discusses the estimation of vaccination coverage required to control epidemic. For illustration the methods, a simulation study is described in section 7. Finally, section 8 includes a brief discussion and conclusion of the paper.

2. MEASURES OF VACCINE EFFICACY UNDER THE HETEROGENEITY OF VACCINE ACTION

Measures of protective vaccine efficacy (VE) depend on the nature of the host's response to the vaccine. Greenwood and Yule (1915) first recognized the VE estimation problem and discussed the need to model the possible heterogeneity in host susceptibility in both vaccinated and unvaccinated individuals. Since then, some researchers have modeled unmeasured heterogeneity through stratification (Smith et al., 1984; Halloran et al., 1992, 1996; Longini and Halloran, 1996), while others have modeled it by considering the susceptibility to follow probability distributions (Brunet et al., 1993; Struchiner et al., 1995). Smith et al. (1984) defined two models of vaccine action: model 1 assumes that the vaccine equally reduced the probability of infection given exposure to infection in all of the vaccinated individuals whereas model 2 assumes that the vaccine may give complete protection to a fraction of the vaccinated while the remaining fraction receive no protection. These models are increasingly being used in practice.

Haber et al. (1991b) obtained expressions for the protective effects of two vaccine mechanisms of Smith et al. (1984) based on a deterministic epidemic model of an acute directly transmitted disease, where vaccine efficacy was defined as the relative reduction in the transmission probability due to the vaccine. For a randomly

mixing population, the transmission probabilities were estimated from the observed attack rates in the vaccinated and unvaccinated by solving their proposed model. These methods have been extended further to a nonrandomly mixing population (Haber et al., 1991a, 1995). In the subsequent year, Halloran et al. (1992) described models 1 and 2 of Smith et al. (1984) as the “leaky” and “all-or-nothing” model, respectively and indicated that in the case of the all-or-nothing model, there are two vaccinated strata, each with different vaccine actions. Based on this premise, Halloran et al. (1992) discussed several models of vaccine action and derived a general expression for a summary vaccine efficacy parameter, where vaccine efficacy was defined as the relative reduction in susceptibilities to infection due to vaccine. Extending the results of Haber et al. (1991b) for a randomly mixing population, the authors discussed the interpretation and estimation of summary vaccine parameter in light of the leaky model and all-or-nothing model.

The present article extends the summary vaccine efficacy parameter to a nonrandomly mixing population that is stratified into households, for an outbreak of an acute directly transmitted disease. For comparative purposes, the leaky and all-or-nothing models of Smith et al. (1984) are briefly discussed. Vaccine efficacy under all models is defined as the relative reduction in the transmission probability due to vaccine. The transmission probability to a susceptible person can be defined as the probability that this person will become infected by a single infected person during one short unit of time (Haber et al., 1991b). Estimation of the transmission probabilities from the observed attack rates is discussed based on a deterministic model.

3. A DETERMINISTIC MODEL OF AN ACUTE TRANSMITTED DISEASE IN A NONRANDOMLY MIXING POPULATION

In this section, the deterministic model proposed by Haber et al. (1995) is adapted to model the transmission of a infectious disease in a population stratified into households. Let M_k be the number of households of sizes k ($k = 1, 2, \dots, L$), and therefore, $M = \sum_{k=1}^L M_k$ is the total number of households in the community. Assume that there is a stronger homogeneous mixing among the members within a household than for members between households. Let λ_{ij} ($i, j = 1, \dots, M$) be the number of contacts with members of household j that a person from household i makes per unit of time, and let $\lambda_i = \sum_{j=1}^M \lambda_{ij}$. The proportion of contacts made by a person from household i with persons of household j is $\rho_{ij} = \lambda_{ij}/\lambda_i$. There are certain constraints on the ρ_{ij} inherent in their meaning:

$$\rho_{ij} = \rho_{ji}, \quad \rho_{ii} = 1, \quad \text{and} \quad 0 \leq \rho_{ij} \leq 1.$$

Let β_i be the probability that a contact between a susceptible from household i and an infective results in transmission of infection. Assume that this probability does not depend on the household of the infected person. Further assume that the infection can be described by the SIR (susceptible \rightarrow infective \rightarrow removed) model (Bailey, 1975). In this model, a susceptible person who becomes infected remains infectious to others for a certain time period and then becomes immune.

As households are small in size, we measure all components of the SIR deterministic model in terms of proportion instead of the actual numbers (Becker

and Utev, 1997; Ball and Lyne, 1999). Let $S_i(t)$, $I_i(t)$, and $R_i(t)$ denote the proportions of susceptible, infected, and naturally immune persons, respectively in household i at time t , where $S_i(t) + I_i(t) + R_i(t) = 1$. Assume that the total number of persons in household i is fixed over the time of epidemic. Then the deterministic description of disease transmission is given by:

$$\frac{dS_i(t)}{dt} = -S_i(t)\beta_i\lambda_i \left(\sum_{j=1}^M \rho_{ij}I_j(t) \right) \tag{1}$$

$$\frac{dI_i(t)}{dt} = S_i(t)\beta_i\lambda_i \left(\sum_{j=1}^M \rho_{ij}I_j(t) \right) - I_i(t)/\tau \tag{2}$$

$$\frac{dR_i(t)}{dt} = I_i(t)/\tau, \tag{3}$$

where τ is the average length of an exponentially distributed infectious period.

Substituting $I_j(t)$ from equation (3) into equation (1) yields

$$\frac{1}{S_i(t)} \frac{dS_i(t)}{dt} = -\tau\beta_i\lambda_i \sum_{j=1}^M \left(\rho_{ij} \frac{dR_j(t)}{dt} \right) \tag{4}$$

Let T be the time at which the observation is made after the end of the outbreak. Integrating both sides of equation (4) from 0 to T and exponentiating, we have

$$\frac{S_i(T)}{S_i(0)} = \exp \left[-\tau\beta_i\lambda_i \sum_{j=1}^M (\rho_{ij}(R_j(T) - R_j(0))) \right] \tag{5}$$

Let $A_i = [S_i(0) - S_i(T)]/S_i(0)$ denote the observed attack rate in household i at time T , and let $Z_i = R_i(T) - R_i(0)$ be the proportion of persons who recovered from the disease during $(0, T)$. Using these values, equation (5) becomes

$$A_i = 1 - \exp[-\tau\beta_i D_i] \tag{6}$$

where $D_i = \lambda_i \sum_{j=1}^M \rho_{ij}Z_j$. Therefore, the transmission probabilities can be estimated as

$$\hat{\beta}_i = \frac{-\ln(1 - A_i)}{\tau D_i}, \quad i = 1, \dots, M_k \tag{7}$$

The $\hat{\beta}_i$'s for all households of size k are equal, because all D_i 's as well as A_i 's for those households are equal. Therefore, the transmission probability for a household of size k can be estimated by taking the average over transmission probabilities obtained for all households of size k as

$$\hat{\beta}_k = -\frac{1}{M_k} \sum_{i=1}^{M_k} \frac{\ln(1 - A_i)}{\tau D_i} = \frac{-\ln(1 - A_k)}{\tau D_k}, \quad k = 1, \dots, L. \tag{8}$$

4. ESTIMATION OF VACCINE EFFICACY

Let us assume that vaccination is not random, but that all of the individuals belonging to the same household are either vaccinated or not. Let v be the vaccination status, where $v = 0$ if unvaccinated, and $v = 1$ if vaccinated. Incorporating the vaccination status, equation (8) can be written as

$$\hat{\beta}_{k,v} = \frac{-\ln(1 - A_{k,v})}{\tau D_{k,v}}, \quad v = 0, 1, \quad k = 1, \dots, L \quad (9)$$

4.1. Leaky Vaccine Model

Under the definition of the leaky vaccine model, the vaccine efficacy in a household of size k can be estimated as

$$\widehat{VE}_k^{(M1)} = 1 - \frac{\hat{\beta}_{k,1}}{\hat{\beta}_{k,0}} = 1 - \frac{D_{k,0} \ln(1 - A_{k,1})}{D_{k,1} \ln(1 - A_{k,0})} \quad (10)$$

If vaccination does not affect the number of contacts made by an individual and consequently the distribution of an individual's contacts, then $D_{k,0} = D_{k,1}$, and therefore, equation (10) becomes

$$\widehat{VE}_k^{(M1)} = 1 - \frac{\ln(1 - A_{k,1})}{\ln(1 - A_{k,0})} \quad (11)$$

If the number of contacts made by an individual depends on his/her vaccination status, then (11) provides a biased estimate of true $VE_k^{(M1)}$, because the assumption of equal exposure to infection of vaccinated and unvaccinated individuals is violated (Haber et al., 1995). In practice, it is likely that the vaccinated individuals have more contacts than unvaccinated individuals, because they may feel that they are protected from infection (Fine and Clarkson, 1987; Comstock, 1990). Therefore, it is wise to use equation (10) rather than equation (11) to estimate vaccine efficacy in a population stratified into households.

4.2. All-or-Nothing Vaccine Model

Under this vaccine model, a fraction α_i of vaccinees in household i is completely protected by the vaccine while the remaining fraction $(1 - \alpha_i)$ receive no protection but have equal transmission probabilities to those of unvaccinated persons from household j . Let $\beta = \beta_{i,1} = \beta_{i,0}$ be the common value of such transmission probabilities. Therefore, using equation (6) the observed attack rates for the unvaccinated and vaccinated household of size k can be written as:

$$A_{k,0} = 1 - \exp[-\tau\beta D_{k,0}] \quad (12)$$

$$A_{k,1} = (1 - \alpha_k) \left(1 - \exp[-\tau\beta D_{k,1}] \right) \quad (13)$$

Under the vaccine action as in the all-or-nothing model, it has been shown in Haber et al. (1991b) that $VE_k = \alpha_k$. Solving equations (12) and (13) for α_k we have the estimated vaccine efficacy for the household of size k as:

$$\widehat{VE}_k^{(M2)} = \hat{\alpha}_k = 1 - \frac{A_{k,1}}{1 - (1 - A_{k,0})^{D_{k,1}/D_{k,0}}} \tag{14}$$

If vaccination does not affect the number of contacts and so the distribution of contacts made by an individual, equation (14) reduces to:

$$\widehat{VE}_k^{(M2)} = \hat{\alpha}_k = 1 - \frac{A_{k,1}}{A_{k,0}} \tag{15}$$

The estimator (15) underestimates the true $VE^{(M2)}$ if vaccination alters an individual’s contact rate in practice. Moreover, the estimators (10) and (15) under the leaky and all-or-nothing vaccine models, respectively, are similar to those discussed in Haber et al. (1995) for a non-randomly mixing population that is partitioned into two large groups.

4.3. Summary Vaccine Model

The transmission probability defined in the earlier section is a function of susceptibility of the susceptible, and therefore, defining vaccine efficacy, under the summary model of vaccine action, based on relative transmission probabilities is equivalent to those based on relative susceptibilities (Halloran et al., 1992). According to the vaccine mechanism under the summary model, suppose a vaccinated household i is separated into two strata in terms of host response to the vaccine, where stratum 1 consists of a fraction α_i of vaccinees, and stratum 2 consists of the remaining fraction $1 - \alpha_i$. Accordingly, let $\beta_{i,1}^{(1)}$ and $\beta_{i,1}^{(2)}$ be the transmission probabilities among the vaccinees in strata 1 and 2, respectively. Let $\beta_{j,0}$ be the transmission probabilities among the members of unvaccinated household j . If the vaccine acts as in the summary model, then the following conditions hold true:

$$0 \leq \beta_{i,1}^{(1)} \leq \beta_{i,1}^{(2)}; \quad \beta_{i,1}^{(1)} \leq \beta_{i,1}^{(2)} \leq \beta_{j,0}$$

In this case, equations (1)–(5) are remain unchanged. Let $S_{i,1}^{(1)}$ and $S_{i,1}^{(2)}$ denote the proportions of vaccinated susceptible in strata 1 and 2, respectively in household i . Then equation (5) can be written for strata 1 and 2, respectively, as

$$\frac{S_{i,1}^{(1)}(T)}{S_{i,1}^{(1)}(0)} = \exp \left[-\tau \beta_{i,1}^{(1)} D_{i,1}^{(1)} \right], \quad i = 1, \dots, M_k, \quad k = 1, \dots, L \tag{16}$$

$$\frac{S_{i,1}^{(2)}(T)}{S_{i,1}^{(2)}(0)} = \exp \left[-\tau \beta_{i,1}^{(2)} D_{i,1}^{(2)} \right], \quad i = 1, \dots, M_k, \quad k = 1, \dots, L \tag{17}$$

Let us define the observed attack rates for vaccinated strata 1 and 2, respectively, as

$$A_{i,1}^{(1)} = [S_{i,1}^{(1)}(0) - S_{i,1}^{(1)}(T)]/S_{i,1}^{(1)}(0)$$

$$A_{i,1}^{(2)} = [S_{i,1}^{(2)}(0) - S_{i,1}^{(2)}(T)]/S_{i,1}^{(2)}(0)$$

As both strata are part of the same household, they have common mixing distribution and recovery rate in the interval $(0, T)$. In this case, $D_{i,1}^{(1)} = D_{i,1}^{(2)} = D_{i,1}$. Substituting these values in (16) and (17) and solving for the transmission probability for strata 1 and 2 we have:

$$\hat{\beta}_{i,1}^{(1)} = \frac{-\ln(1 - A_{i,1}^{(1)})}{\tau D_{i,1}} \tag{18}$$

$$\hat{\beta}_{i,1}^{(2)} = \frac{-\ln(1 - A_{i,1}^{(2)})}{\tau D_{i,1}} \tag{19}$$

Using the same argument that was used to derive equation (8), equations (18) and (19) can be replaced by their average estimates obtained over all households size k as

$$\hat{\beta}_{k,1}^{(1)} = \frac{-\ln(1 - A_1^{(1)})}{\tau D_{k,1}} \tag{20}$$

$$\hat{\beta}_{k,1}^{(2)} = \frac{-\ln(1 - A_1^{(2)})}{\tau D_{k,1}} \tag{21}$$

Similarly, α_i can be replaced by α_k .

4.4. Strata Are Identifiable and α Known

Suppose that α_k and two strata of vaccinated individuals are identifiable, and therefore, $A_{k,1}^{(1)}$ and $A_{k,1}^{(2)}$ are also known. Then by using equations (9), (20) and (21) the summary vaccine efficacy for a household of size k can be estimated as

$$\widehat{VE}_k = 1 - \frac{\alpha_k \hat{\beta}_{k,1}^{(1)} + (1 - \alpha_k) \hat{\beta}_{k,1}^{(2)}}{\hat{\beta}_{k,0}} \tag{22}$$

$$= 1 - \frac{D_{k,0}}{D_{k,1}} \frac{\alpha_k \ln(1 - A_{k,1}^{(1)}) + (1 - \alpha_k) \ln(1 - A_{k,1}^{(2)})}{\ln(1 - A_{k,0})} \tag{23}$$

If vaccination does not alter the mixing distribution, then $D_{k,0} = D_{k,1}$, and (23) becomes

$$\widehat{VE}_k = 1 - \frac{\alpha_k \ln(1 - A_{k,1}^{(1)}) + (1 - \alpha_k) \ln(1 - A_{k,1}^{(2)})}{\ln(1 - A_{k,0})} \tag{24}$$

Similar to the estimators (11) and (15) under the leaky and all-or-nothing models, respectively, the estimator (24) provides biased estimate of the true \overline{VE}_k if vaccination alters an individual's contact rate.

4.5. Strata Are Not Identifiable and α Unknown

If strata are not identifiable and α is unknown, then one can derive bounds for \overline{VE}_k as

$$VE_k^{(M2)} \leq \overline{VE}_k \leq VE_k^{(M1)}$$

where $VE_k^{(M2)}$ and $VE_k^{(M1)}$ are the vaccine efficacy under the all-or-nothing and leaky models, respectively. The upper bound assumes that all vaccinated individuals are equally affected by the vaccine, and the lower bound assumes that some of them are completely protected while the remainder have no protection. The proof of this inequality is given in the appendix.

5. VARIANCE OF THE VACCINE EFFICACY ESTIMATOR

We derive an expression for the variance of the vaccine efficacy estimators given in (10), (14), and (23) under the leaky, all-or-nothing, and summary vaccine models, respectively. In the case of the leaky vaccine model, the estimator (10) is a function of the simpler estimator (11) multiplied by the constant term $D_{k,0}/D_{k,1}$. That is, if \widehat{VE}_k^* denotes the simpler estimator in (11), then variance of the estimator in (10) can be expressed as

$$\text{Var}(\widehat{VE}_k^{(M1)}) = (D_{k,0}/D_{k,1})^2 \text{Var}(\widehat{VE}_k^*) \tag{25}$$

where the following approximation for the variance of \widehat{VE}_k^* can be derived as in Haber et al. (1991a) (when there is no subsampling):

$$\text{Var}(\widehat{VE}_k^*) \approx \left[\frac{S_{k,1}(0) - S_{k,1}(T)}{S_{k,1}(0)S_{k,1}(T)} + (1 - \widehat{VE}_k^*)^2 \frac{S_{k,1}(0) - S_{k,1}(T)}{S_{k,1}(0)S_{k,1}(T)} \right] \bigg/ \left[\ln \left(\frac{S_{k,0}(0)}{S_{k,0}(T)} \right) \right] \tag{26}$$

For the estimator (14) under the all-or-nothing model, the following variance approximation can be derived using the delta method (Haber et al., 1995):

$$\text{Var}(\widehat{VE}_k^{(M2)}) \approx \frac{1}{B_{k,0}^2} \left[\frac{A_{k,1}(1 - A_{k,1})}{S_{k,1}(0)} + \frac{(A_{k,1}/B_{k,0})^2 r_k^2 A_{k,0}(1 - A_{k,0})^{2r_k - 1}}{S_{k,0}(0)} \right] \tag{27}$$

where $r_k = D_{k,1}/D_{k,0}$, and $B_{k,0} = 1 - (1 - A_{k,0})^{r_k}$.

In the summary measure of vaccine efficacy, the estimator in (23) can be expressed as

$$\widehat{VE}_k = 1 - \frac{D_{k,0}}{D_{k,1}} \left[1 - \widehat{VE}_k^{(2)} + \alpha_k \left(\widehat{VE}_k^{(1)} - \widehat{VE}_k^{(2)} \right) \right] \tag{28}$$

where $\widehat{VE}_k^{(1)} = 1 - \frac{\ln(1-A_{k,1}^{(1)})}{\ln(1-A_{k,0}^{(1)})}$ and $\widehat{VE}_k^{(2)} = 1 - \frac{\ln(1-A_{k,1}^{(2)})}{\ln(1-A_{k,0}^{(2)})}$ are the vaccine efficacy estimators in vaccinated strata 1 and 2, respectively. As strata 1 and 2 are independent, $\widehat{VE}_k^{(1)}$ and $\widehat{VE}_k^{(2)}$ are independent. Therefore, the variance of the summary vaccine estimator can be approximated as

$$\text{Var}(\widehat{VE}_k) \approx \left(\frac{D_{k,0}}{D_{k,1}}\right)^2 \left[\text{Var}(\widehat{VE}_k^{(2)}) + \alpha_k^2 \left(\text{Var}(\widehat{VE}_k^{(1)}) + \text{Var}(\widehat{VE}_k^{(2)}) \right) \right]. \tag{29}$$

By definition, both $\widehat{VE}_k^{(1)}$ and $\widehat{VE}_k^{(2)}$ are closely related to the simpler estimator (11) of vaccine efficacy in the leaky vaccine model. Therefore, using the same approach applied in estimating $\text{Var}(\widehat{VE}_k^*)$ given in (26) the variances of $\widehat{VE}_k^{(1)}$ and $\widehat{VE}_k^{(2)}$ can be obtained, respectively, as:

$$\begin{aligned} \text{Var}(\widehat{VE}_k^{(1)}) &\approx \left[\frac{S_{k,1}^{(1)}(0) - S_{k,1}^{(1)}(T)}{S_{k,1}^{(1)}(0)S_{k,1}^{(1)}(T)} + (1 - \widehat{VE}_k^{(1)})^2 \frac{S_{k,1}^{(1)}(0) - S_{k,1}^{(1)}(T)}{S_{k,1}^{(1)}(0)S_{k,1}^{(1)}(T)} \right] \bigg/ \left[\ln \left(\frac{S_{k,0}(0)}{S_{k,0}(T)} \right) \right] \\ \text{Var}(\widehat{VE}_k^{(2)}) &\approx \left[\frac{S_{k,1}^{(2)}(0) - S_{k,1}^{(2)}(T)}{S_{k,1}^{(2)}(0)S_{k,1}^{(2)}(T)} + (1 - \widehat{VE}_k^{(2)})^2 \frac{S_{k,1}^{(2)}(0) - S_{k,1}^{(2)}(T)}{S_{k,1}^{(2)}(0)S_{k,1}^{(2)}(T)} \right] \bigg/ \left[\ln \left(\frac{S_{k,0}(0)}{S_{k,0}(T)} \right) \right] \end{aligned}$$

The preceding expression for the variance of the summary vaccine estimator is appropriate when vaccinated strata are identifiable and α_k is known. When α_k is unknown and the strata are not identifiable, the approximate variance of this estimator can be obtained by taking an average over $\text{Var}(\widehat{VE}_k^{(M1)})$ and $\text{Var}(\widehat{VE}_k^{(M2)})$. All of these variance expressions derived thus far can be used to obtain confidence intervals for the corresponding true vaccine efficacy.

6. ESTIMATION OF THE VACCINATION COVERAGE

For designing an effective vaccination program, it is important to estimate the vaccination coverage required to control an epidemic. In this section, we discuss estimation of such threshold parameter based on the deterministic approach discussed in section 3. We derive a threshold value f_k^* for the fraction f_k of vaccinated households of size k such that for $f_k \geq f_k^*$ there will be no epidemic, and for $f_k \leq f_k^*$ there will be an epidemic (Kermack and McKendrick, 1927; Haber et al., 1991b).

Let us define the dynamic variable $I_k(t) = I_{k,0}(t) + I_{k,1}(t)$, the total number of infectives at time t in all households of size k . Then taking differentiation we have

$$\frac{dI_k(t)}{dt} = \frac{dI_{k,0}(t)}{dt} + \frac{dI_{k,1}(t)}{dt} \tag{30}$$

Using equation (2) and $I_k(t) = I_{k,0}(t) + I_{k,1}(t)$, equation (30) can be rewritten as

$$\frac{dI_k(t)}{dt} = kM_{k,0}S_{k,0}(t)\beta_{k,0}H_{k,0}(t) + kM_{k,1}S_{k,1}(t)\beta_{k,1}H_{k,1}(t) - I_k(t)/\tau \tag{31}$$

where $M_{k,0}$ and $M_{k,1}$ are the number of unvaccinated and vaccinated households of size k , respectively, and $M_{k,0} + M_{k,1} = M_k$. Also $H_{k,v}(t) = \frac{1}{M_k} \left[\sum_{i=1}^{M_k} \lambda_{i,v} \left(\sum_{j=1}^M \rho_{ij,v} I_j(t) \right) \right]$, $v = 0, 1$. It has been discussed in the literature that $I_k(t)$ is either strictly decreasing (in the case of no epidemic) or strictly increasing to a maximum and strictly decreasing thereafter (in the case of an epidemic) (Hethcote, 1976; Haber et al., 1991b). Therefore, it can be proved that there will not be an epidemic if $\left. \frac{dI_k(t)}{dt} \right]_{t=0} \leq 0$.

6.1. When the Leaky Model is Operating

Let us consider the leaky model is operating in practice, and assume that no one in the population has naturally acquired immunity. Evaluating equation (31) at $t = 0$, and using the initial condition $S_{k,0}(0) = S_{k,1}(0) = 1$, we have

$$\left. \frac{dI_k(t)}{dt} \right]_{t=0} = kM_{k,0}\beta_{k,0}H_{k,0}(0) + kM_{k,1}\beta_{k,1}H_{k,1}(0) - I_k(0)/\tau \tag{32}$$

where $I_k(0) = 0^+$ is very small number.

As assumed that there is one infective in the population at $t = 0$, so $I_k(0) = 1$. This implies the term $\sum_{j=1}^M \rho_{ij,v} I_j(0) = 1$, and hence, the term $H_{k,v}(0) = \lambda_{k,v}$. Moreover, the assumption on mixing due to vaccination is no longer valid before the vaccination starts, and therefore, $\lambda_{k,0} = \lambda_{k,1} = \lambda_k$. Assume that the transmission parameters ($\beta_{k,0}$ and $\beta_{k,1}$ or α or $\beta_{k,1}^{(1)}$ and $\beta_{k,1}^{(2)}$) are known from a previous study on a similar population. Substituting $I_k(0) = 1$, $H_{k,v} = \lambda_k$, $M_{k,0} = M_k(1 - f_k)$, and $M_{k,1} = M_k f_k$ in (32) we have

$$\left. \frac{dI_k(t)}{dt} \right]_{t=0} = \tau k M_k (1 - f_k) \lambda_k \beta_{k,0} + \tau k M_k f_k \lambda_k \beta_{k,1} - 1 \tag{33}$$

In a situation where there will be no epidemic, we have $\left. \frac{dI_k(t)}{dt} \right]_{t=0} \leq 0$. Therefore, solving equation (33) for f_k gives

$$f_k \geq \frac{\tau k M_k \lambda_k \beta_{k,0} - 1}{[\beta_{k,0} - \beta_{k,1}] \tau k M_k \lambda_k} \tag{34}$$

Since $\beta_{k,0} \geq \beta_{k,1}$ under the leaky model, the inequality (34) will be satisfied when $f_k^* \geq f_k$, and the threshold value f_k^* is given by

$$f_k^* = \frac{1 - 1/R_{0,k}}{1 - \beta_{k,1}/\beta_{k,0}} \tag{35}$$

where $R_{0,k} = \tau k M_k \lambda_k \beta_{k,0}$ is the basic reproduction number for the households of size k .

6.2. When the All-or-Nothing Model is Operating

In this model, we use $S_{k,1}(0) = (1 - \alpha_k)$ and $\beta_{k,0} = \beta_{k,1}$. Then an epidemic will not be occur if

$$\tau k M_k (1 - \alpha_k f_k) \lambda_k \beta_{k,0} \leq 1 \tag{36}$$

Solving equation (36) for f_k^* yields

$$f_k^* = \frac{1 - 1/R_{0,k}}{\alpha_k} \tag{37}$$

6.3. When the Summary Model is Operating

In the summary model, we use $S_{k,1}^{(1)}(0) = \alpha_k$, $S_{k,1}^{(2)}(0) = (1 - \alpha_k)$, and $\beta_{k,1}^{(1)} \leq \beta_{k,1}^{(2)} \leq \beta_{k,0}$. Then an epidemic will not be occur if

$$\tau k M_k (1 - f_k) \lambda_k \beta_{k,0} + \tau k M_k f_k \lambda_k \bar{\beta}_{k,1} \leq 1 \tag{38}$$

where $\bar{\beta}_{k,1} = \alpha_k \beta_{k,1}^{(1)} + (1 - \alpha_k) \beta_{k,1}^{(2)}$. Solving equation (38) for f_k^* yields

$$f_k^* = \frac{1 - 1/R_{0,k}}{1 - \bar{\beta}_{k,1}/\beta_{k,0}} \tag{39}$$

We can see that the denominator of f_k^* in the case of all models is the vaccine efficacy. Therefore, the equations (35), (37), and (39) can be replaced by a single equation as follows:

$$f_k^* = \frac{1 - 1/R_{0,k}}{Efficacy} \tag{40}$$

The threshold f_k^* depends on the vaccine efficacy and the $R_{0,k}$. We can see that f_k^* is positive only if $R_{0,k} > 1$ (which is the condition required to guarantee that an epidemic will occur in the absence of vaccination). From the previous study if we know $R_{0,k}$ for an acute directly transmitted disease, then f_k^* depends on the vaccine efficacy. If vaccine action is truly heterogeneous, then all models except the summary one provide a biased estimate of vaccine efficacy, which leads to bias in f_k^* . However, in practice, the model for the vaccine action is unknown; in this case a optimistic approach is to use the largest f_k^* (Haber et al., 1991b). Moreover, if $f_k^* > 1$, then it implies that an epidemic will occur regardless of the value of f_k , which could be possible if the disease is extremely infectious.

Finally, the overall fraction of households in the community that needs to be vaccinated can be estimated as follows:

$$F^* = \frac{1}{M} \sum_{k=1}^L M_k f_k^* \tag{41}$$

which has an interpretation similar to f_k^* .

7. SIMULATION STUDY

7.1. Simulation Design

A simulation study was conducted to illustrate the methods described in the earlier sections. We simulated an epidemic path of number of susceptibles, infectious, and recovered individuals based on the SIR epidemic process in a population partitioned into households. We considered a population of 500 households, where there were 5, 15, 35, 35, and 10% of households of sizes 1, 2, 3, 4, and 5, respectively; 50% of households were assumed to be vaccinated. We further assumed that initially there was one infective in the population from which the epidemic started, and that there were no naturally immune persons. The length of infectious period was assumed to be a random variable with mean $\tau = 5$.

In simulation, we considered two matrices of mixing distribution denoted by $\{\rho_{rk,v}\}$. In the first matrix (Table 1), there is higher mixing within households than between households, and an individual's contacts do not depend on vaccination status. In the second matrix, there is a higher mixing within households than between households, and individuals are more likely to contact individuals of the same vaccination status. In both matrices, the mixing distribution, $\rho_{rk,v}$, is proportionally related to the size of the household. However, the mixing distribution of all households of the same size is equal. One should note the symmetry condition $\rho_{rk,v} = \rho_{kr,v}$.

Table 1 Values of $\rho_{rk,v}$ used in simulation, where $v = 0$ for unvaccinated, $v = 1$ for vaccinated

		$k = 1$		$k = 2$		$k = 3$		$k = 4$		$k = 5$		
		v	0	1	0	1	0	1	0	1	0	1
$\rho_{rk,0} = \rho_{rk,1}$	$r = 1$	0	0.24	0.24	0.08	0.08	0.10	0.10	0.14	0.14	0.20	0.20
		1	0.24	0.24	0.08	0.08	0.10	0.10	0.14	0.14	0.20	0.20
	$r = 2$	0	0.08	0.08	0.30	0.30	0.12	0.12	0.16	0.16	0.23	0.23
		1	0.08	0.08	0.30	0.30	0.12	0.12	0.16	0.16	0.23	0.23
	$r = 3$	0	0.10	0.10	0.12	0.12	0.36	0.36	0.20	0.20	0.28	0.28
		1	0.10	0.10	0.12	0.12	0.36	0.36	0.20	0.20	0.28	0.28
	$r = 4$	0	0.14	0.14	0.16	0.16	0.20	0.20	0.42	0.42	0.32	0.32
		1	0.14	0.14	0.16	0.16	0.20	0.20	0.42	0.42	0.32	0.32
	$r = 5$	0	0.20	0.20	0.23	0.23	0.28	0.28	0.32	0.32	0.48	0.48
		1	0.20	0.20	0.23	0.23	0.28	0.28	0.32	0.32	0.48	0.48
$\rho_{rk,0} < \rho_{rk,1}$	$r = 1$	0	0.32	0.22	0.15	0.10	0.20	0.14	0.24	0.16	0.28	0.19
		1	0.22	0.42	0.10	0.19	0.14	0.29	0.16	0.31	0.19	0.33
	$r = 2$	0	0.15	0.10	0.36	0.16	0.20	0.18	0.25	0.19	0.27	0.21
		1	0.10	0.19	0.16	0.48	0.18	0.33	0.19	0.35	0.21	0.37
	$r = 3$	0	0.20	0.14	0.20	0.18	0.42	0.20	0.25	0.22	0.32	0.24
		1	0.14	0.29	0.18	0.33	0.20	0.56	0.22	0.39	0.24	0.40
	$r = 4$	0	0.24	0.16	0.25	0.19	0.25	0.22	0.48	0.25	0.30	0.26
		1	0.16	0.31	0.19	0.35	0.22	0.39	0.25	0.65	0.26	0.43
	$r = 5$	0	0.28	0.19	0.27	0.21	0.32	0.24	0.30	0.26	0.54	0.30
		1	0.19	0.33	0.21	0.37	0.24	0.40	0.26	0.43	0.30	0.75

We simulated a separate epidemic path under each of the vaccine models described in section 4. When the leaky model was operating, the transmission probabilities $\beta_{1,0} = \beta_{2,0} = \beta_{3,0} = \beta_{4,0} = \beta_{5,0} = 0.80$ for the unvaccinated group and $\beta_{1,1} = \beta_{2,1} = \beta_{3,1} = \beta_{4,1} = \beta_{5,1} = 0.20$ for the vaccinated group were used. These confirmed the true VE under the leaky model as 0.75. For the all-or-nothing model, the proportion of completely immune vaccinated individuals was considered as $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0.60$. The remaining proportion of vaccinees had similar transmission probabilities to the unvaccinated individuals as in the leaky model. Therefore, the true VE for the all-or-nothing model is 0.60. The transmission probabilities for the vaccinated strata 1 and 2 under the summary vaccine were $\beta_{1,1}^{(1)} = \beta_{2,1}^{(1)} = \beta_{3,1}^{(1)} = \beta_{4,1}^{(1)} = \beta_{5,1}^{(1)} = 0.22$ and $\beta_{1,1}^{(2)} = \beta_{2,1}^{(2)} = \beta_{3,1}^{(2)} = \beta_{4,1}^{(2)} = \beta_{5,1}^{(2)} = 0.30$, respectively. The transmission probabilities for the unvaccinated group were similar to those under the leaky model. Therefore, the true VE under the summary model is $1 - \frac{0.60 \times 0.22 + 0.40 \times 0.30}{0.80} = 0.685$. For each scenario, 500 simulations were performed, and the estimate of VE was taken from the average over 500 simulations. Models were compared by assessing 90% nominal coverage of confidence interval for the true vaccine efficacy parameter. Coverage was calculated as the percentage of simulations in which the true value was enclosed by the estimated confidence interval.

One of the main objectives of the study was to estimate vaccine efficacy under the heterogeneity of vaccine action in a nonrandomly mixing population. Among the already-described epidemic processes simulated based on three vaccine models, only the process under the summary model was generated considering the heterogeneity of vaccine action across the vaccinees. Using the outcomes of this epidemic process (where the true $VE = 0.685$), estimates of VE s with their corresponding coverage probabilities under each of the vaccine models were obtained, and results were compared to assess models' performance under the heterogeneity of vaccine action.

7.2. Results

Depending on mixing patterns, there are two estimators of VE under each of the vaccine models. When the individual vaccine model was operating as the true model, results of simulation based on the first ρ matrix indicate that both estimators of the model provided unbiased estimate with good coverage (Table 2). Based on the second ρ matrix, the estimators (10) in the leaky model, (14) in the all-or-nothing model, and (23) in the summary model provided accurate estimates of the corresponding true VE with good coverage. However, the estimators (11), (15), and (24) in the leaky, all-or-nothing, and summary models, respectively, underestimated the true VE and provided worse coverage, because these estimators ignore the change in mixing due to vaccination. Moreover, between the models of vaccine action, the summary and all-or-nothing models performed equally but slightly better than the leaky model, given that the respective model acted as the true model. This holds true under all types of mixing pattern considered.

To assess models' performance under the heterogeneity of vaccine action, further simulation was conducted under a common true model of heterogeneous vaccine actions (i.e., the summary model was operating as true model in this case),

Table 2 Estimates of vaccine efficacy, under all vaccine models, when individual model was operating as true model: mean estimate (\widehat{VE}), standard error (SE), and coverage (Cov %)

		Leaky model (true $VE = 0.75$)						All-or-nothing model (true $VE = 0.60$)						Summary model (true $VE = 0.685$)					
		Estimator (10)			Estimator (11)			Estimator (14)			Estimator (15)			Estimator (23)			Estimator (24)		
HH size		\widehat{VE}	SE	Cov	\widehat{VE}	SE	Cov	\widehat{VE}	SE	Cov	\widehat{VE}	SE	Cov	\widehat{VE}	SE	Cov	\widehat{VE}	SE	Cov
$\rho_{k,0} = \rho_{k,1}$	$k = 1$	0.787	0.034	69	0.771	0.039	65	0.621	0.022	73	0.625	0.024	67	0.694	0.022	74	0.667	0.025	71
	$k = 2$	0.775	0.026	72	0.763	0.033	69	0.619	0.020	75	0.614	0.023	72	0.692	0.023	73	0.668	0.021	74
	$k = 3$	0.738	0.020	77	0.747	0.031	74	0.616	0.018	77	0.620	0.027	78	0.689	0.018	79	0.679	0.017	77
	$k = 4$	0.741	0.019	83	0.732	0.022	81	0.608	0.012	88	0.591	0.020	88	0.683	0.010	89	0.683	0.011	86
	$k = 5$	0.746	0.019	86	0.741	0.021	85	0.604	0.013	91	0.596	0.021	87	0.604	0.013	91	0.692	0.021	87
$\rho_{k,0} < \rho_{k,1}$	$k = 1$	0.772	0.036	70	0.691	0.062	60	0.617	0.027	71	0.524	0.041	61	0.693	0.024	73	0.612	0.035	64
	$k = 2$	0.769	0.023	71	0.683	0.060	62	0.618	0.021	76	0.538	0.033	62	0.691	0.025	72	0.631	0.033	68
	$k = 3$	0.729	0.024	77	0.687	0.053	65	0.611	0.019	76	0.551	0.025	70	0.687	0.017	77	0.637	0.026	70
	$k = 4$	0.745	0.018	84	0.662	0.046	68	0.603	0.010	89	0.573	0.029	70	0.681	0.012	88	0.641	0.021	73
	$k = 5$	0.751	0.015	87	0.665	0.041	71	0.602	0.008	92	0.573	0.019	76	0.682	0.009	90	0.662	0.022	76

Table 3 Estimates of vaccine efficacy, for all vaccine models when summary model was operating as true model (true $VE = 0.685$): mean estimate (\widehat{VE}), and coverage (Cov %)

		Leaky model				All-or-nothing model				Summary model			
		Estimator (10)		Estimator (11)		Estimator (14)		Estimator (15)		Estimator (23)		Estimator (24)	
HH size		\widehat{VE}	Cov	\widehat{VE}	Cov	\widehat{VE}	Cov	\widehat{VE}	Cov	\widehat{VE}	Cov	\widehat{VE}	Cov
$\rho_{rk,0} = \rho_{rk,1}$	$k = 1$	0.723	63	0.725	62	0.644	63	0.640	62	0.694	74	0.667	71
	$k = 2$	0.724	65	0.720	65	0.645	64	0.642	63	0.692	73	0.668	74
	$k = 3$	0.720	66	0.718	64	0.646	65	0.645	65	0.689	79	0.679	77
	$k = 4$	0.719	67	0.718	67	0.648	65	0.648	65	0.683	89	0.683	86
	$k = 5$	0.718	68	0.715	68	0.651	67	0.651	66	0.604	91	0.692	87
$\rho_{rk,0} < \rho_{rk,1}$	$k = 1$	0.749	67	0.727	68	0.650	65	0.621	56	0.693	73	0.612	64
	$k = 2$	0.745	66	0.728	69	0.652	68	0.622	57	0.691	72	0.631	68
	$k = 3$	0.742	67	0.723	68	0.647	67	0.623	58	0.687	77	0.637	70
	$k = 4$	0.740	68	0.721	70	0.648	68	0.624	58	0.681	88	0.641	73
	$k = 5$	0.740	69	0.720	72	0.651	69	0.631	59	0.682	90	0.662	76

and estimates of VE under each of the vaccine models were obtained. Results (Table 3) of simulation based on the first matrix of mixing pattern show that the leaky and all-or-nothing models possessed a biased estimate; however, the summary model provided an unbiased estimate with good coverage. Simulation based on the second matrix indicates that the estimator ignoring the change in mixing due to vaccination still provided a biased estimate.

For the household of larger size, all types of estimator provided comparatively more accurate estimates of the true value than those for the household of smaller size, which implies the usual problem related to sample size in getting a consistent estimate.

Estimates of VE under each of the vaccine models have different interpretations. For example, the estimate 0.600 under the all-or-nothing model can be interpreted as 60% of the vaccinated individuals begin completely protected from infection. The interpretation of the estimate 0.75 under the leaky model is that the transmission probability given exposure to one infective is reduced by 75%. The estimate under the summary model has, in terms of average, similar interpretation to those under the leaky model.

Table 4 Estimates of the vaccination coverage under all models (using estimates of VE from table 2 when $\rho_{rk,0} < \rho_{rk,1}$); the basic reproduction number was assumed at $R_0 = 2$

HH size	f^* (Leaky)	f^* (All-or-nothing)	f^* (Summary)
$k = 1$	0.664	0.833	0.727
$k = 2$	0.664	0.832	0.729
$k = 3$	0.664	0.834	0.729
$k = 4$	0.665	0.834	0.728
$k = 5$	0.666	0.834	0.728
Overall	0.664	0.834	0.728

In section 6, we discussed the estimation of the vaccination coverage, where the estimator f_k^* we derived is the function of the basic reproduction number R_0 and the vaccine efficacy. In this section, the methods were illustrated using the estimated vaccine efficacy from the simulation study and the basic reproduction number assumed at $R_0 = 2$. The estimated vaccination coverage under all models of vaccine action is presented in Table 4.

Under each model, there was no substantial difference among the household specific estimates of f_k^* ($k = 1, \dots, 5$). However, between the models, the largest overall estimate was observed for the all-or-nothing model, which was followed by the estimate for the summary model, while the smallest one was observed for the leaky model. We do not know which model is operating in practice. In this case, the optimistic approach is to use the largest one. Moreover, following others studies (Ball et al. 1999), it is also suggested that larger households should be vaccinated first, because members of a larger household are expected to be infected more rapidly than those of a smaller household because of the existence of higher mixing in larger household than smaller one. Otherwise, the epidemic will spread very fast.

8. DISCUSSION AND CONCLUSION

Estimation of vaccine efficacy parameters depends on several assumptions, particularly those related to population structure and their mixing patterns, the routes of transmission of the infection agent, and vaccine response (Haber et al., 1991a). Most published studies on vaccine efficacy assume random mixing throughout the population. However, nonrandom mixing is common in a population that is stratified into small groups such as households. Although some studies consider nonrandom mixing, the heterogeneity of vaccine response among the hosts is often ignored. This work discussed the estimation of vaccine efficacy under the heterogeneity of vaccine action in a nonrandomly mixing population, for an acute directly transmitted disease. To allow for the heterogeneity of vaccine action in a nonrandomly mixing population, three models of vaccine actions were discussed, particularly focusing on the summary model proposed by Halloran et al. (1992) for a randomly mixing population. Estimation and interpretation of the summary vaccine efficacy were discussed in light of other two models: the leaky and all-or-nothing models suggested in Smith et al. (1984). Vaccine efficacy under all models was defined assuming a vaccine that lowers the transmission probability to vaccinated individuals. A deterministic approach proposed in Haber et al. (1995) was applied in estimating the transmission probabilities in terms of a log transformation of final attack rates.

To illustrate the methods, a simulation study was conducted under each of the vaccine models, depending on whether vaccination along with stratification alters the mixing pattern or not. The estimators under each of the vaccine models ignoring the change in mixing pattern due to vaccination underestimate their corresponding true VE . The simulation results, based on all types of mixing pattern, further confirmed that all models of vaccine action performed almost equally when the individual model was operating as true model of vaccine action. These results, particularly under the leaky and all-or-nothing models, are analogous to those in Haber et al. (1995) for a nonrandomly mixing population that is partitioned into two large groups. However, population in practice consists of several small groups

such as households, where nonrandom mixing is obvious. This work provided information on estimating vaccine efficacy in such a small-group-mixing population.

One of the main objectives of the study was to estimate and assess performance of the vaccine models under the heterogeneity of vaccine action. This was accomplished by conducting a further simulation study based on a true model that considers heterogeneous vaccine actions. Results, from simulation based on all types of mixing pattern, confirm that the summary model performed better than both the leaky and all-or-nothing models. The reason is simple: The leaky and all-or-nothing models do not properly take into account the heterogeneity of vaccine action, while the summary model does. In practice, vaccine may have heterogeneous effects because of varying immune systems among the hosts (Halloran et al., 1992). Therefore, it is worthwhile to estimate vaccine efficacy based on the summary model. The estimator (23) we derived under the summary model of vaccine action for a nonrandomly mixing population has interpretation similar to those derived in Halloran et al. (1992) for a randomly mixing population. In addition, we derived an approximate expression for standard error of this estimator, which is useful to have an idea of its sampling distribution and to make an inferential statement. Further, following Halloran et al. (1992), a bound for the summary vaccine efficacy was derived when vaccinated strata was not identifiable, where the lower and upper bounds were represented by the vaccine efficacy under the all-or-nothing and leaky models, respectively.

This article further discussed, under the preceding models of vaccine action, the estimation of vaccination coverage required to control an epidemic. We derived a threshold value f_k^* that depends on the basic reproduction number R_0 and the vaccine efficacy, and that has interpretation similar to those in other studies (Becker and Starczak, 1997; Ball and Lyne, 2002, 2006; Pellis et al., 2009). The methods were illustrated by using the vaccine efficacy obtained from the simulation study and assuming $R_0 = 2$. The numerical result shows that the estimates of f_k^* were not identical for all models of vaccine action. In this case, the optimistic approach is suggested to use the largest value of f_k^* . This kind of information would be useful for health policymakers to design an appropriate vaccination program.

Because of the unavailability of real field data, this article was unable to provide an application of the methods discussed. Therefore, providing an application of these methods to real field data could be a future research project. Further, these methods could be extended to a nonrandomly mixing population that is stratified in two levels: by households, then by homestay and outgoing groups, where unequal mixing between outgoing and home-stay groups is possible. Moreover, it would be reasonable to consider the heterogeneity of vaccine response for each individual rather than for a group of individuals, as vaccine responses vary from individual to individual. In this case, vaccine efficacy could be estimated by using a gamma frailty model, which takes into account the individual heterogeneity.

APPENDIX

Proof of the Inequality: $VE_k^{(M2)} \leq \overline{VE}_k \leq VE_k^{(M1)}$

We assumed that overall mixing in the population is nonrandom due to stratification and vaccination. Also we let all other assumptions described earlier

remain the same. We further assumed that α as well as the attack rates in vaccinated strata 1 and 2 are unknown. First, we prove that $VE_k^{(M1)} \geq \overline{VE}_k$,

$$\Rightarrow 1 - \frac{D_{k,0} \ln(1 - A_{k,1})}{D_{k,1} \ln(1 - A_{k,0})} \geq 1 - \frac{D_{k,0} \alpha_k \ln(1 - A_{k,1}^{(1)}) + (1 - \alpha_k) \ln(1 - A_{k,1}^{(2)})}{D_{k,1} \ln(1 - A_{k,0})}$$

Here $\ln(1 - A_{k,0})$ is negative for $0 \leq A_{k,0} \leq 1$. Therefore, we have to prove

$$\ln(1 - A_{k,1}) \geq \alpha_k \ln(1 - A_{k,1}^{(1)}) + (1 - \alpha_k) \ln(1 - A_{k,1}^{(2)}) \quad (42)$$

We assumed here that $\bar{A}_{k,1} = \alpha_k A_{k,1}^{(1)} + (1 - \alpha_k) A_{k,1}^{(2)}$ and $\bar{A}_{k,1} = A_{k,1}$. That is, the crude estimate is equal to the weighted average. However, the last inequality remains unchanged. It is obvious that $\ln(1 - A_{k,1})$ is a concave function as $\ln(1 - x)$ is concave for $x \in \mathfrak{R}$, and therefore, equation (42) is true, which implies that $VE_k^{(M1)} \geq \overline{VE}_k$. For details, see Halloran et al. (1992). Similarly, it can be proved that $VE_k^{(M2)} \leq \overline{VE}_k$. Finally, we have

$$VE_k^{(M2)} \leq \overline{VE}_k \leq VE_k^{(M1)}$$

Hence, if we cannot identify the strata and do not know α , we have an upper and lower bound for the true summary vaccine efficacy, \overline{VE}_k .

ACKNOWLEDGMENTS

I am very grateful to Sergey Utev, Pat Stone, anonymous referees, and associate editors for their valuable suggestions and comments that helped me strengthen the article.

REFERENCES

- Ball, F. G., Lyne, O. D. (2002). Optimal vaccination policies for stochastic epidemics among a population of households. *Mathematical Biosciences* 177–178:333–54.
- Ball, F. G., Britton, T., Lyne, O. D. (2004a). Stochastic multitype epidemics in a community of households: estimation and form of optimal vaccination schemes. *Mathematical Biosciences* 191(1):19–40.
- Ball, F. G., Britton, T., Lyne, O. D. (2004b). Stochastic multitype epidemics in a community of households: Estimation of threshold parameter R_* and secure vaccination coverage. *Biometrika* 91(2):345–62.
- Ball, F. G., Lyne, O. D. (2006). Optimal vaccination schemes for epidemics among a population of households, with application to variola minor in Brazil. *Statistical Methods in Medical Research* 15:481–497.
- Ball, F. G., Lyne, O. D. (2001). Stochastic multitype SIR epidemics among a population partitioned into households. *Advances in Applied Probability* 33:99–123.
- Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and its Applications*. London, UK: Griffin.
- Becker, N. G., Utev, S. (2002). Protective vaccine efficacy when vaccine response is random. *Biometrical Journal* 44:29–42.
- Becker, N. G., Utev, S. (1997). The effect of community structure on the immunity coverage required to prevent epidemics. *Mathematical Biosciences* 147:23–39.

- Becker, N. G., Starczak, D. N. (1997). Optimal vaccination strategies for a community of households. *Mathematical Biosciences* 139:117–132.
- Brunet, R. C. Struchiner, C. J., Halloran, M. E. (1993). On the distribution of vaccine protection under heterogeneous response. *Mathematical Biosciences* 116:111–125.
- Comstock, G. W. (1990). Vaccine evaluation by case-control or prospective studies. *American Journal of Epidemiology* 131:205–207.
- Davis, X. M., Waller, L. A., Haber, M. (2006). Estimating vaccine efficacy from outbreak size household data in the presence of heterogeneous transmission probabilities. *Journal of Biopharmaceutical Statistics* 16:499–516.
- Fine, P. E. M., Clarkson, J. A. (1987). Reflection on the efficacy of pertussis vaccines. *Review of Infectious Diseases* 9:866–883.
- Greenwood, M., Yule, U. G. (1915). The statistics of anti-typhoid and anticholera inoculations, and the interpretation of such statistics in general. *Proc. Royal Society of Medicine* 8:113–194.
- Haber, M., Longini, I. M., Jr., Halloran, M. E. (1991a). Estimation of vaccine efficacy in outbreaks of acute infectious disease. *Statistics in Medicine* 10:1573–1584.
- Haber, M., Longini, I. M., Jr., Halloran, M. E. (1991b). Measures of the effects of vaccination in a randomly mixing population. *International Journal of Epidemiology* 20:300–310.
- Haber, M., Longini, I. M. Jr., Halloran, M. E., Watelet, L. (1995). Estimation of vaccine efficacy in a non-randomly mixing population. *Biometrical Journal* 37:25–38.
- Halloran, M. E., Haber, M., Longini, I. M., Jr., Struchiner, C. J. (1991). Direct and Indirect Effects in Vaccine Efficacy and Effectiveness. *American Journal of Epidemiology* 133:456–463.
- Halloran, M. E., Haber, M., Longini, I. M., Jr. (1992). Interpretation and estimation of vaccine field efficacy under heterogeneity. *American Journal of Epidemiology* 136:328–343.
- Halloran, M. E., Longini, I. M., Jr., Struchiner, C. J. (1996). Estimability and interpretation of vaccine efficacy using frailty mixing models. *American Journal of Epidemiology* 144:83–97.
- Hethcote, H. W. (1976). Qualitative analyses of communicable disease models. *Mathematical Biosciences* 28:335–356.
- Kermack, W. O., McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proc. Royal Society of London Series A* 115:700–721.
- Longini, I. M., Jr., Halloran, M. E., Haber, M. (1993). Estimation of vaccine efficacy from epidemics of acute infectious agents under vaccine-related heterogeneity. *Mathematical Biosciences* 117:271–281.
- Longini, I. M., Jr., Halloran, M. E. (1996). A frailty mixture model for estimating vaccine efficacy. *Applied Statistics* 45:165–173.
- Pellis, L., Ferguson N. M., Fraser, C. (2009). Threshold parameters for a model of epidemic spread among households and workplaces. *Journal of Royal Society Interface* 6:979–987.
- Smith, C. E. G. (1970). Prospects for the control of infectious disease. *Proc. Royal Society of Medicine* 63:1181–1190.
- Smith P. G., Rodrigues, L. C., Fine, P. E. M. (1984). Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *International Journal of Epidemiology* 13:87–93.
- Struchiner, C. J., Brunet, R., Halloran M. E., Massad E., Azevedo-Neto, R. S. (1995). On the use of state-space models for the evaluation of health interventions. *Journal of Biological Systems* 3:851–65.