Maximum likelihood estimation of influenza vaccine effectiveness against transmission from the household and from the community

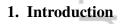## Kylie E. C. Ainslie[a], Michael J. Haber[a]*, Ryan E. Malosh[b], Joshua G. Petrie[b], and Arnold S. Monto[b]

Influenza vaccination is recommended as the best way to protect against influenza infection and illness. Due to seasonal changes in influenza virus types and subtypes a new vaccine must be produced, and vaccine effectiveness (VE) must be estimated, annually. Since 2010, influenza vaccination has been recommended universally in the U.S., making randomized clinical trials unethical. Recent studies have employed a monitored household cohort study design to determine separate VE estimates against influenza transmission from the household and community.

We developed a probability model and accompanying maximum likelihood procedure to estimate vaccine-related protection against transmission of influenza from the household and the community.

Using agent-based stochastic simulations, we validated that we can obtain maximum likelihood estimates of transmission parameters and VE close to their true values. Sensitivity analyses to examine the effect of deviations from our assumptions were conducted. We used our method to estimate transmission parameters and VE from data from a monitored household study in Michigan during the 2012-13 influenza season and were able to detect a significant protective effect of influenza vaccination against community-acquired transmission. Copyright © 2017 John Wiley & Sons, Ltd.

Keywords: vaccine effectiveness, influenza, maximum likelihood, household, observational studies

1. Introduction

Influenza vaccination is recommended every season due to changes in influenza virus types, subtypes, and phenotypes from one season to the next. The variation in the influenza virus requires the production of a new vaccine each season, thus vaccine effectiveness (VE) must be estimated each year [1]. The concept of vaccine effectiveness is based on comparing the probability of illness of a vaccinated person to that of an unvaccinated person, i.e., it measures the benefit of vaccination for a single individual. VE is defined as one minus the risk ratio (RR), where risk is defined as the probability of becoming

[a]*Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322*
[b]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI 48109*
Correspondence to: Michael Haber, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322. E-mail: mhaber@emory.edu

infected and ill throughout the influenza season. In this work we use the term 'vaccine effectiveness' rather than 'vaccine efficacy' because the former is estimated using observational studies, while the latter is estimated from a randomized trial.

Placebo controlled randomized clinical trials can no longer be used to assess influenza VE in the U.S. due to the recommendation that all individuals older than 6 months be vaccinated [2]. As a result, observational studies have been increasingly used to assess the benefit of influenza vaccination. Most commonly, observational studies on unrelated individuals have been used to estimate VE against influenza illness requiring outpatient medical care [3, 4, 5, 6]. The household unit has been shown to play an important role in the transmission of influenza [7, 8]. Additionally, household data has been shown to provide more robust estimates of VE than data consisting of unrelated individuals [9], thus recent VE studies have employed a monitored household (MH) cohort design [6, 10, 11].

In a MH study, entire households are enrolled and closely monitored over the course of the study period. Whenever a participant has an acute respiratory illness (ARI), s/he has to report to study personnel who arrange for a swab to be taken and tested for influenza infection. Despite being expensive and logistically complex, MH studies are attractive to assess influenza VE because they allow for the observation of time of influenza disease and vaccination as well as allow for the estimation of VE against community-acquired and household-acquired influenza [6, 10]. An additional advantage of a MH study is that it allows for the examination of symptomatic influenza of any severity regardless of whether participants sought medical care [3, 6, 10]. Other commonly used study designs (ordinary cohort, in which independent individuals are followed rather than households, case-control, and test-negative [12, 13]) are only able to capture individuals infected with influenza who seek medical care. These studies are prone to bias as many people infected with influenza may not seek medical care, and those who seek medical care might not represent the entire population.

Statistical methods have been developed to estimate influenza VE from household data, first from final data in which influenza infection was assessed after the season by serological testing [14, 15, 16] and more recently from time-to-event data [6, 9, 10, 17, 18]. Longini and Koopman [14] developed a probability model and maximum likelihood procedure for the separate estimation of influenza transmission parameters in the household and community from final count data. Haber *et al.* [15] extended this model to assess the impact of risk factors on influenza transmission. It has been shown that the use of time-to-event data produces VE estimates with smaller bias compared to estimates produced using final data [9]. Halloran *et al.* [17] developed a framework to estimate VE from time-to-event household data using the secondary attack rate (SAR), but did not account for community transmission. Davis and Haber [9] incorporated temporal information into VE estimates using survival models as a method to estimate influenza transmission probabilities from community and household contacts. Neither of these approaches allow for the estimation of source-specific VE, which is the focus of this work. Ohmit *et al.* [10] estimate source-specific VE using the Cox proportional hazards model, but make assumptions about the source of infection based on viral type/subtype and the timing of infection, as infection source cannot be directly observed.

We present a probability model and accompanying maximum likelihood procedure to first estimate source-specific transmission parameters and then to estimate vaccine-related protection against transmission of influenza from the household and from the community from time-to-event household data. Our approach does not require the source of infection to be known and incorporates temporal information into VE estimates. Additionally, our model allows vaccination to occur during the study and does not assume household VE is equal to community VE, thus providing a framework to estimate VE against influenza infection separately from each source. However, for interpretability, VE needs to be estimated assuming all vaccinated individuals are vaccinated over the entire study period. To assess VE, we use symptomatic influenza, defined as laboratory-confirmed infection with the influenza virus that develops into an acute respiratory illness (ARI), as our outcome of interest. We perform a simulation study to evaluate our model and then apply our model to data from the Household Influenza Vaccine Effectiveness (HIVE) study. The HIVE study has been established in Ann Arbor, MI as an ongoing, longitudinal MH cohort to better characterize the impact of households on influenza transmission [10, 18, 19].

2. Methods

We consider a population composed of households of varying sizes as in a MH study. There are potentially two different sources of influenza exposure: (1) exposure to other infected household members and (2) exposure to infected individuals in the larger community. We define the study period as a single influenza season.
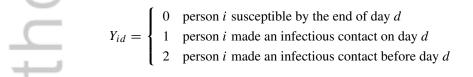
We make several important model assumptions: *(1)* Each member of the study population belongs to a household. *(2)* Persons are only classified by their household membership (i.e., there are no other stratifying variables or covariates). *(3)* Each person makes daily contacts with each member of their household and with randomly selected community contacts. *(4)* There is random mixing within the household and among community members. *(5)* A person can only be infected once during the study. Thus, once a person is infected with influenza s/he is removed from the at risk population for the remainder of the study. *(6)* Asymptomatic influenza cases - persons infected with influenza, but do not develop an ARI - have a very small probability of transmitting influenza to others (and therefore, are not accounted for in our model). *(7)* The per-contact transmission probabilities within a household and among the community for vaccinees and non-vaccinees remain constant throughout the study. *(8)* The vaccine provides reduction in transmission probability (leaky vaccine model [20, 21]) and only affects susceptibility to influenza. *(9)* The length of the latent and infectious periods are constant and known.

2.1. Probability Model

In real data, it is difficult to ascertain the source of infection for each individual. Recent studies have attributed the source of infection to a household member if the influenza type/subtype are the same and the secondary case was identified within a short time period from the index case [6, 10, 11, 18]. However, it is not possible to actually observe source of infection (except in challenge studies) thus, it is important to develop a probability model that can be used when source of infection is unknown.

Below we present a probability model for estimating VE for household-acquired and community-acquired influenza. To accomplish this, we first estimate influenza transmission parameters within the household and in the community using a maximum likelihood procedure. Then, using these transmission parameter estimates, we estimate source-specific influenza VE. Table 1 defines the parameters used in the model.

Consider a susceptible person i on day d. Let Y_{id} denote the infection status for that person, where

$$Y_{id} = \begin{cases} 0 & \text{person } i \text{ susceptible by the end of day } d \\ 1 & \text{person } i \text{ made an infectious contact on day } d \\ 2 & \text{person } i \text{ made an infectious contact before day } d \end{cases}$$

Let v_{id} denote her/his vaccination status ($v_{id} = 0, 1$ for unvaccinated and vaccinated, respectively). We define $\beta_{v_{id}}$ as the daily transmission probability to that person from a single household contact; similarly, we define $\gamma_{v_{id}}$ as the daily transmission probability of infection from the community to that person when everyone else is infectious. Since v_{id} can only take on the values 0, 1 for any person on any day, we have four transmission parameters: $\beta_0, \beta_1, \gamma_0, \gamma_1$.

We assume that the latent period, the time between an individual getting infected and becoming infectious, begins on the day after the infectious contact was made. An infected person becomes *infectious* $L + 1$ days after making an infectious contact and remains infectious for I days, where L and I are the length of the latent and infectious periods, respectively. For example, if an infectious contact is made on day 1 then, the latent period lasts two days (days 2 and 3) and the infectious period lasts four days (days 4-7). After the duration of the infectious period, the person recovers and remains immune for the rest of the study. It is usually assumed that the day of becoming infectious is the day of onset of symptoms, i.e., the length of the incubation period equals the length of the latent period. During our estimation process,

we observe the first day of the infectious period and determine the day of the infectious contact by subtracting $L + 1$. We let $p(d)$ denote the prevalence of influenza infection on day d, defined as the proportion of the population who is infectious [22], and let m_{id} be the number of infectious persons in the household of person i on day d. When determining the probability that a person is infected from community contacts on day d, we multiply $gamma_{v_{id}}$ by the proportion of infectious individuals in the population that that day, $p(d)$.

Let $\pi_{ijd} = \mathbb{P}(Y_{id} = j | Y_{i(d-1)} = 0)$ denote the conditional probability that person i has infection status j on day d, given that s/he was susceptible on the previous day, $j = 0, 1, 2$. Let $\psi_{ijd} = \mathbb{P}(Y_{id} = j)$ denote the unconditional probability that person i has infection status j on day d, $j = 0, 1, 2$. All of the probabilities involving individual persons are conditioned on the vaccination history of individual i, $\mathbf{V}_i = (V_{i1}, \ldots, V_{iD})$, $i = 1, \ldots, N$.

The conditional probabilities π_{ijd} can be written as follows:

$$
\begin{array}{rclcl}
\pi_{i0d} & = & \mathbb{P}(Y_{id} = 0 | Y_{i(d-1)} = 0) & = & (1 - \beta_{v_{id}})^{m_{id}}(1 - \gamma_{v_{id}} \cdot p(d)) \\
\pi_{i1d} & = & \mathbb{P}(Y_{id} = 1 | Y_{i(d-1)} = 0) & = & 1 - \left[(1 - \beta_{v_{id}})^{m_{id}}(1 - \gamma_{v_{id}} \cdot p(d)) \right] \\
\pi_{i2d} & = & \mathbb{P}(Y_{id} = 2 | Y_{i(d-1)} = 0) & = & 0
\end{array}
$$

where, given person i was susceptible on day $d - 1$, π_{i0d} is the probability of person i escaping infection on day d, π_{i1d} is the probability of person i becoming infected on day d, and π_{i2d} is the probability that person i was infected on a previous day. Under assumption *(5)* $\pi_{i2d} = 0$.

The unconditional probability of person i having infection status j ($j = 0, 1$) on day d is defined as

$$
\psi_{ijd} = \mathbb{P}(Y_{id} = j) = \mathbb{P}(Y_{id} = j | Y_{i(d-1)} = 0) \cdot \mathbb{P}(Y_{i(d-1)} = 0) + \mathbb{P}(Y_{id} = j | Y_{i(d-1)} > 0) \cdot \mathbb{P}(Y_{i(d-1)} > 0).
$$

By assumption (5), $\mathbb{P}(Y_{id} = j | Y_{i(d-1)} > 0) = 0$, thus

$$
\begin{array}{rcl}
\mathbb{P}(Y_{id} = j) & = & \mathbb{P}(Y_{id} = j | Y_{i(d-1)} = 0) \cdot \mathbb{P}(Y_{i(d-1)} = 0) \\
& = & \pi_{ijd} \cdot \psi_{i0(d-1)} \\
& = &
\end{array}
$$

For example, the probability that person i, who was effectively vaccinated on day 2 gets infected on day 3 is

$$
\begin{array}{rcl}
\mathbb{P}(Y_{i3} = 1) & = & \pi_{i13} \cdot \psi_{i02} \\
& = & \pi_{i13} \cdot \pi_{i02} \cdot \pi_{i01} \\
& = & \left\{ 1 - \left[(1 - \beta_1)^{m_{i3}}(1 - \gamma_1 \cdot p(3)) \right] \right\} \left\{ (1 - \beta_1)^{m_{i2}}(1 - \gamma_1 \cdot p(2)) \right\} \left\{ (1 - \beta_0)^{m_{i1}}(1 - \gamma_0 \cdot p(1)) \right\}
\end{array}
$$

where the value of v_{id} changes from 0 to 1 on day 2.

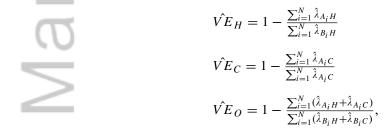### 2.2. Maximum Likelihood Procedure

Each person's contribution to the likelihood function depends on whether or not s/he became infected during the study and on the day of infection, if infected. If person i got infected on day d then his/her contribution to the likelihood function is: $L_i = \psi_{i1d}$. If person i did not become infected by the last day of the study D, $L_i = \psi_{i0D}$, the probability of escaping infection throughout the study. The overall likelihood is $L(\beta_0, \beta_1, \gamma_0, \gamma_1 | \text{data}) = \prod_{i=1}^{N} L_i$, where N is the number of study participants. We assume that persons are (conditionally) independent because our probabilities condition on the daily number of infected persons in the household and the daily prevalences of infection in the community. Finally, maximum likelihood estimates (MLEs) of the transmission parameters, $\hat{\beta}_0, \hat{\beta}_1, \hat{\gamma}_0, \hat{\gamma}_1$, are obtained by maximizing $L(\beta_0, \beta_1, \gamma_0, \gamma_1 | \text{data})$. Likelihood optimization was performed using a limited-memory modification of the BFGS quasi-Newton method [23] with a lower bound of 0 and an upper bound of 1 using the R function optim [24]. Standard errors (SEs) of transmission parameter estimates were obtained from the Hessian matrix from the maximization procedure,

empirically from simulations by taking the standard deviation of all simulation estimates, and by parametric bootstrap. Using the transmission parameter MLEs, we can estimate the distribution of Y_{id} for every (i, d) by plugging the parameter estimates into the equations for π_{ijd} and ψ_{ijd}.

We estimate VE using a two-step process. First, we estimate the transmission probabilities from the likelihood function in which a person's actual day of vaccination, before or during the study, is used. Second, using the estimated transmission probabilities from Step 1, we estimate VE by comparing the probability of becoming infected during the entire study between persons who became effectively vaccinated prior to the study and persons who remained unvaccinated throughout the study. In this way, the estimates of VE do not depend on the times of vaccination. For each person i, we added to the population two dummy persons: person A_i who was effectively vaccinated before the onset of the study, and person B_i who remained unvaccinated throughout the study. These dummy persons make the same contacts with real persons (but not with other dummy persons) that correspond with the contacts of person i. They can become infected but they are unable to infect others. Therefore, the dummy persons do not affect the infection probabilities of the real persons (i.e., they do not modify the vaccine's indirect effects). The vaccination status of all real persons remained unchanged for the purpose of estimating VE. We define λ_{iH} and λ_{iC} as the probability that person i is infected from a household (H) or community (C) contact during the study, respectively. The MLEs of λ_{iH} and λ_{iC} are obtained by substituting the MLEs of our transmission parameters for the true parameters (see Supplemental Information for details). Due to the very small probability of coinfection we do not include the probability of being coinfected in estimates of VE.

We denote the seasonal VEs against household transmission, community transmission, and overall transmission by VE_H, VE_C, and VE_O, respectively. The estimates of VE are

$$\hat{VE}_H = 1 - \frac{\sum_{i=1}^{N} \hat{\lambda}_{A_i H}}{\sum_{i=1}^{N} \hat{\lambda}_{B_i H}}$$

$$\hat{VE}_C = 1 - \frac{\sum_{i=1}^{N} \hat{\lambda}_{A_i C}}{\sum_{i=1}^{N} \hat{\lambda}_{A_i C}}$$

$$\hat{VE}_O = 1 - \frac{\sum_{i=1}^{N} (\hat{\lambda}_{A_i H} + \hat{\lambda}_{A_i C})}{\sum_{i=1}^{N} (\hat{\lambda}_{B_i H} + \hat{\lambda}_{B_i C})},$$

SEs for VE estimates were obtained empirically and by parametric bootstrap.

2.3. Simulations

To assess the performance of our method and the accuracy of our maximum likelihood estimates, we developed a stochastic agent-based simulation program to simulate an influenza outbreak in a population with households. One simulation corresponded to one outbreak. Each simulation featured a susceptible population of 1000 individuals with 10 initially infected individuals. Each individual was assigned to a household. Households varied in size from one to twelve members. The proportion of households of each size were based on DeKalb County, GA census data [25]. The influenza season lasted three months. It was assumed that each person made 10 daily community contacts (under the assumption of random-mixing) and made daily contact with each person in their household. Two vaccination scenarios were assessed: (1) all vaccinations occurred prior to the study period and (2) vaccinations occurred prior to the study (25%), during the first month (15%), and during the second month (10%). In each vaccination scenario, 50% of the population was vaccinated. A person was considered effectively vaccinated 14 days after the receipt of the influenza vaccine. The following parameter values were used as the daily transmission probability from an infectious person to a susceptible person with vaccination status v (v=0, unvaccinated, v=1, vaccinated): $\alpha_0 = 0.04$ ($\gamma_0 = \alpha_0 \cdot 10 = 0.40$), $\alpha_1 = 0.01$ ($\gamma_1 = \alpha_1 \cdot 10 = 0.1$), $\beta_0 = 0.15$, and $\beta_1 = 0.075$.

MLEs and SEs of the transmission parameters were calculated for each simulation. The MLEs of the transmission parameters were used to estimate household VE, community VE, and overall VE. True VE was calculated using the true transmission parameter values under each vaccination scenario. For each simulation scheme, the bias of the VE estimates

was assessed. Bias was defined as the true VE subtracted from the estimated VE. Two hundred outbreak simulations were performed and fifty bootstrap simulations were performed for each outbreak simulation. Source specific VE, SE, and 95% confidence intervals were estimated for each simulation. The same assumptions we made for our model (see section 2) were used for the simulation program. The latent period was set to 2 days and the infectious period was set to 4 days [26].

2.4. Sensitivity Analyses

Sensitivity analyses were performed to assess the bias of VE estimates using the maximum likelihood procedure when several model assumptions were relaxed. Each sensitivity analysis was conducted independently. In the first sensitivity analysis, we relaxed assumption *(10)* (the latent and infectious periods fixed and known) and allowed the latent and infectious periods to follow a distribution. We assessed bias under two different scenarios: (1) the mean latent and infectious periods were correctly specified and (2) the mean latent and infectious periods were incorrectly specified. For scenario (1) the latent period was 1, 2, or 3 days with probabilities 0.25, 0.5, and 0.25, respectively. The infectious period was 3, 4, 5, or 6 days with probabilities 0.3, 0.5, 0.1, and 0.1, respectively. The mean latent and infectious periods were 2 and 4 days, respectively. For scenario (2) the latent period was 1, 2, or 3 days with probabilities 0.1, 0.1, and 0.8, respectively with a mean of 2.7 days. The infectious period was 3, 4, 5, or 6 days with probabilities 0.05, 0.1, 0.65, and 0.2, respectively with a mean of 5 days. The specified mean latent and infectious periods, used for the derivation of the MLEs, were 2 and 4 days, respectively. In the second sensitivity analysis, we assessed the bias of VE estimates when the prevalence in the cohort was allowed to differ from the prevalence in the overall population. An overall population of households, comprised of 2000 individuals, was simulated, from which households were selected to make up the study cohort of 1000 individuals. Households in the cohort were assumed to be a random sample of the households in the overall population. MLEs, SEs, and 95% confidence intervals were calculated for each simulation.

2.5. A Real-Life Example

Our maximum likelihood approach was applied to a dataset from the HIVE study in Michigan during the 2012-13 influenza season designed to estimate household and community VE. The study population consisted of 321 households with a total sample size of 1426 members followed from October 2012 to May 2013. Households ranged in size from 4 to 10 members. Only households with at least 4 persons including at least 2 children were included in the study. At the onset of influenza-like symptoms, participants were instructed to contact study personnel, so a respiratory specimen could be collected. Specimens were tested for the presence of influenza virus by RT-PCR [6, 10, 18, 19]. Influenza infection was identified by RT-PCR in 111 individuals with influenza-like illness. Five individuals were infected with influenza twice [10], but only the first influenza infection was considered for our analysis. Index cases were assumed to be infected from the community and a household-acquired case was defined by transmission link to an index case within the household if both cases were caused by the same influenza type/subtype/lineage and if illness onset in the secondary case occurred 1-7 days after illness onset in the index case. Vaccination status was determined as previously described using a combination of medical records and state registry data. Adults and children aged 9-17 years old were considered effectively vaccinated 14 days after the receipt of the vaccine. Children under the age of 9 years old were considered effectively vaccinated 14 days after receipt of the second dose of the vaccine [10].

Influenza transmission parameters for vaccinated and unvaccinated individuals and VE against household-acquired and community-acquired transmission of influenza were estimated using our maximum likelihood approach. To avoid undefined values during maximum likelihood estimation due to a prevalence of zero within the cohort, values of zero were changed to 1/1426 (the total size of the cohort). Fifty parametric bootstrap simulations were performed to obtain SE estimates and 95% confidence intervals for transmission parameter and VE estimates. Simulations were performed using information from the study data, such as, transmission parameter estimates, proportion of households of each size and proportion of individuals vaccinated. Model adequacy was assessed by comparing the mean number of cases per

household size from 200 simulated outbreaks to the observed frequencies of cases per household size in the data. All simulations and analyses were performed in R 3.2.2 [24].

3. Results

3.1. Simulations

Mean transmission parameter and SE estimates from 200 simulations are shown in Table 2. When all vaccinations occurred prior to the study, our maximum likelihood procedure produced the following estimates: $\hat{\beta}_0 = 0.153$ (95% CI: 0.129, 0.177), $\hat{\beta}_1 = 0.078$ (95% CI: 0.064, 0.092), $\hat{\gamma}_0 = 0.429$ (95% CI: 0.380, 0.478), $\hat{\gamma}_1 = 0.118$ (95% CI: 0.092, 0.142). When vaccinations occurred prior to and during the study, $\hat{\beta}_0 = 0.156$ (95% CI: 0.132, 0.180), $\hat{\beta}_1 = 0.078$ (95% CI: 0.060, 0.096), $\hat{\gamma}_0 = 0.425$ (95% CI: 0.382, 0.468), $\hat{\gamma}_1 = 0.118$ (95% CI: 0.093, 0.143). Transmission estimates were similar between the two vaccination scenarios. All transmission parameter estimates were close to the true values ($\beta_0 = 0.15$, $\beta_1 = 0.075$, $\gamma_0 = 0.4$, $\gamma_1 = 0.10$). The greatest bias observed was in the estimate of γ_0 when vaccinations occurred prior to and during the study (Bias=0.031 corresponding to a relative bias of 7.75%); however, the estimates of γ_1 suffered from relative biases of 16% and 17% in vaccination scenarios (1) and (2), respectively (Table 2). SE estimates were calculated empirically, using the Hessian matrix from the maximum likelihood procedure, and from fifty parametric bootstraps. All three SE estimation methods produced similarly small SE estimates (Table 2).

Mean VE and SE estimates from 200 simulations are shown in Table 3. When all vaccinations occurred prior to the study, our maximum likelihood procedure produced the following estimates: $\hat{VE}_{MLH} = 0.478$ (95% CI: 0.358, 0.598), $\hat{VE}_{MLC} = 0.724$ (95% CI: 0.659, 0.789), and $\hat{VE}_{MLO} = 0.607$ (95% CI: 0.546, 0.668). When vaccinations occurred prior to and during the study, $\hat{VE}_{MLH} = 0.453$ (95% CI: 0.314, 0.592), $\hat{VE}_{MLC} = 0.719$ (95% CI: 0.652, 0.786), and $\hat{VE}_{MLO} = 0.602$ (95% CI: 0.535, 0.669). The greatest bias was observed in the estimate of VE against community-acquired influenza when all vaccinations occurred prior to the study, Bias=-0.024 (corresponding to a relative bias of 3.2%). SEs of the VE estimates were calculated empirically and via parametric bootstrap. Empirical SEs were very similar to bootstrap SEs. VE against household-acquired influenza consistently had the highest SE, while VE estimates against community-acquired and overall influenza were very similar. When vaccinations occurred during the study, the empirical SEs were slightly larger for all VE estimates than when all vaccinations occurred prior to the study. For all estimates of VE, we observed coverage probabilities of or greater than 95% with the exception of community VE when all vaccinations occurred prior to the study (Table 3).

3.2. Sensitivity Analyses

Table 4 shows the results of the sensitivity analyses performed when the assumption that the latent and infectious periods are fixed and known was relaxed. When the mean length of the latent and infectious periods were correctly specified and all vaccinations occurred prior to the study, $\hat{VE}_{MLH} = 0.474$ (95% CI: 0.348, 0.599), $\hat{VE}_{MLC} = 0.703$ (95% CI: 0.635, 0.770), and $\hat{VE}_{MLO} = 0.602$ (95% CI: 0.546, 0.657). When vaccinations occurred prior to and during the study, $\hat{VE}_{MLH} = 0.470$ (95% CI: 0.327, 0.612), $\hat{VE}_{MLC} = 0.699$ (95% CI: 0.630, 0.768), and $\hat{VE}_{MLO} = 0.600$ (95% CI: 0.531, 0.669). When the mean length of the latent and infectious periods were misspecified and all vaccinations occurred prior to the study, $\hat{VE}_{MLH} = 0.440$ (95% CI: 0.303, 0.577), $\hat{VE}_{MLC} = 0.679$ (95% CI: 0.617, 0.740), and $\hat{VE}_{MLO} = 0.589$ (95% CI: 0.541, 0.636). When vaccinations occurred prior to and during the study, $\hat{VE}_{MLH} = 0.442$ (95% CI: 0.311, 0.573), $\hat{VE}_{MLC} = 0.657$ (95% CI: 0.588, 0.726), and $\hat{VE}_{MLO} = 0.575$ (95% CI: 0.524, 0.626) (Table 4).

When the latent and infectious periods were allowed to follow a distribution, VE estimates were underestimated. When the mean latent and infectious periods were correctly specified the largest bias was observed in estimates of VE against community transmission with a relative bias of 6%. Larger bias was observed in all VE estimates when the mean length of the latent and infectious periods were misspecified with the largest bias (corresponding to a relative bias >11%) observed

in estimates of VE against community transmission when vaccination occurred prior to and during the study. Estimates of SE were similar when the mean latent and infectious periods were correctly specified and misspecified. Estimates of VE_{MLH} had the largest SE and the SE of VE_{MLC} and VE_{MLO} were similar (Table 4).

Table 5 shows the results of the sensitivity analysis when the cohort was selected from a simulated overall population. This simulation scenario allowed the prevalence of influenza infection in the cohort to differ from the prevalence in the overall population. When all vaccinations occurred prior to the study, $\hat{VE}_{MLH} = 0.480$ (95% CI: 0.344, 0.615), $\hat{VE}_{MLC} = 0.727$ (95% CI: 0.662, 0.791), and $\hat{VE}_{MLO} = 0.614$ (95% CI: 0.551, 0.677). When vaccinations occurred prior to and during the study, $\hat{VE}_{MLH} = 0.469$ (95% CI: 0.334, 0.604), $\hat{VE}_{MLC} = 0.722$ (95% CI: 0.648, 0.796), and $\hat{VE}_{MLO} = 0.607$ (95% CI: 0.544, 0.671) (Table 5). The biases of VE estimates when the cohort was selected from a larger overall population were very similar to the original simulations in which only the cohort was simulated (Tables 5 and 3, respectively).

3.3. A Real-Life Example

Estimates of transmission parameters from the HIVE study are shown in Table 6. The daily transmission probability from an infectious household contact to an unvaccinated susceptible is 0.013 (95% CI: 0.008, 0.019) and to a vaccinated susceptible is 0.013 (95% CI: 0.005, 0.021). The transmission rate from all daily infectious community contacts to an unvaccinated susceptible is 0.202 (95% CI: 0.159, 0.245) and to a vaccinated susceptible is 0.134 (95% CI: 0.081, 0.187) (Table 6). A transmission rate of 0.202 (0.134) means that, on average, 20.2% (13.4%) of unvaccinated (vaccinated) persons who make contacts with infectious persons on a given day will become infected. The lower estimated transmission rate in the community to vaccinated individuals compared to unvaccinated individuals suggests that there is at least a small protective effect of vaccination against community-acquired influenza infection. SE estimates were obtained using the Hessian matrix from the maximum likelihood procedure and using parametric bootstrap. The bootstrap SE estimates were similar to the Hessian matrix SE estimates (Table 6). Our simulated frequencies of cases per household size (Table **??**) were a good match to the observed frequencies (Table **??**) suggesting that our model captures the dependency between household size and attack rate.

Estimates of household VE, community VE, and overall VE are presented in Table 7. VE point estimates indicated significant protection against community-acquired influenza infection (0.336, 95% CI: 0.066, 0.606), and non-significant protection against household-acquired (0.052, 95% CI: -0.754 0.858) and overall (0.250, 95% CI: -0.019, 0.519) influenza infection (Table 7).

We compared our VE estimates to results found by Ohmit *et al.* [10] using unadjusted and adjusted hazard rate ratios (Table 8). The adjusted models adjusted for age in months and documentation of high-risk health status [10]. Our point estimate of VE against household-acquired influenza was substantially lower (by more than 0.25) than the estimates found by Ohmit *et al.* [10]. Our point estimate of VE against community-acquired influenza infection was slightly higher than the estimates in the original paper, and we were able to detect a significant protective effect of vaccination against overall influenza infection across all study participants. The original paper did not detect a significant protective effect of the vaccine from any source using either the unadjusted or adjusted model (Table 8). We estimated overall VE lower than both the unadjusted and adjusted estimates. Our method produced 95% confidence intervals that were slightly wider for household VE, but narrower than those reported in the original study for community and overall VE.

4. Discussion

We have presented a probability model and accompanying maximum likelihood procedure to estimate VE against household-acquired and community-acquired influenza infection from MH studies. Our method first estimates source-specific transmission parameters that characterize the daily probability of infection. We use these transmission parameters

Prepared using **simauth.cls**

to estimate the probability of influenza infection throughout the study and estimate VE against transmission of influenza from the household and from the community. Previous methods that estimate source-specific VE use final count data that do not take into account the time of infection [21, 14, 15]. Our approach improves upon these methods by incorporating time to event data, which allows for variation in influenza prevalence and timing of vaccination to be incorporated into estimates of VE. We used a stochastic agent-based simulation program to evaluate the bias and precision of our estimates.

Under our model assumptions, our method estimated the transmission parameters and VE close to the truth for two different vaccination scenarios (Tables 2 and 3). Transmission parameter estimates were very similar regardless of vaccination scenario. SEs were calculated empirically, using the Hessian matrix, and using a parametric bootstrap procedure. In all scenarios, bootstrap SEs were close to empirical SEs indicating that the bootstrap procedure performs well and is appropriate for the estimation of SE when analyzing real data (when estimation of SE empirically is not possible). Coverage probabilities of greater than or equal to 95% for all estimates of VE suggest that our method performs well. For estimates of VE against community-acquired influenza when all vaccinations occurred prior to the study, the coverage probability was slightly lower (92%) suggesting that our method produces confidence intervals that are too narrow for this estimate.

We developed a similar likelihood method under the assumptions that the source of infection (household or community) is known. We found that the estimates and their standard deviations were similar to those we obtained without this assumption. Hence, we conclude that knowing the source of infections does not substantially improve the VE estimates.

Results from our sensitivity analyses suggest that our maximum likelihood approach provided estimates of source specific VE with small bias and SE when the length of the latent and infectious periods are not constant and the mean is correctly specified. When the mean length of the latent and infectious periods were misspecified, our results suggest that caution should be used when using the maximum likelihood procedure to estimate community-acquired influenza when vaccination occurs during the study, as the estimates may be moderately biased if the mean length of the latent and infectious periods are misspecified. However, the misspecification of the mean length of the latent and infectious periods had little impact on SE of VE estimates.

Allowing the prevalence of influenza infection to differ in a randomly sampled cohort compared to the overall population had little impact on the bias of source-specific VE estimates (Table 5) indicating that the maximum likelihood approach is robust to differences in the prevalence of influenza between the study cohort and overall population. To investigate the bias of VE estimates when the sample population is a small fraction the overall population, we performed an additional sensitivity analysis in which we simulated an overall population of 10,100 people with a sample population of 1,000 people (results not shown). We saw no change in the amount of bias when the cohort was a smaller fraction of the overall population. Under the assumption that the cohort is a random sample of the overall population, we would expect the results to be similar regardless of the size of the overall population relative to the cohort.

It is well known that not all individuals infected with influenza develop an ARI [26]. However, little is known about the proportion of asymptomatic individuals in a given influenza season. One study estimated that approximately 67% of individuals infected with influenza develop symptoms [26], while other studies have estimated that as few as 23% [27] or as many as 84% [28] of influenza infected individuals develop symptoms. Despite the lack of symptoms, asymptomatic individuals are still infectious; however, less so than symptomatic individuals [26]. Little is known about the relative infectiousness of asymptomatic individuals compared to symptomatic individuals. While asymptomatic individuals are considered less infectious because they are not shedding as much virus as symptomatic individuals [29], asymptomatic may make more contacts while infectious than their symptomatic counterparts because they do not realize they are infected. Due to the many unknowns surrounding asymptomatic influenza infections, we did not include asymptomatic individuals in our sensitivity analyses. It will be important in future work to assess the impact of asymptomatic individuals in the population on VE estimates.

We applied our method to data from the HIVE study during the 2012-13 influenza season [10] (Table 7). Our VE point estimates for household and overall influenza infection were lower than the point estimates found in the original

study (Table 8) using unadjusted and adjusted hazard rate ratios. This difference in point estimates is likely due to the fact that our method does not require source of infection to be known, and in the case of the adjusted estimates, does not control for potential confounders, such as health status and age. Our estimate of community VE was similar to the original estimates, but we were able to detect a significant protective effect of vaccination against community-acquired influenza infection across all study participants (Table 7). Ohmit *et al.* found no significant protective effect of vaccination for overall, household-acquired, and community-acquired influenza infection across all study participants using the hazard rate ratio (Table 8) [10].

Our probability model makes many simplifying assumptions about influenza disease progression. In future work, we plan to relax some of our model assumptions to more realistically model the influenza disease process. First, we would like to allow for strata within the population, particularly age groups. Previous research indicates that transmission of influenza is different from child to child, child to adult, adult to child, and adult to adult [30]. Additionally, age has been identified as an important risk factor associated with influenza transmission in which children and the elderly are more susceptible to infection than young adults. Very young children and older adults are also more susceptible to complications from infection. The addition of strata requires additional assumptions about the contact patterns of individuals in the same stratum and between strata. Previous studies have found that contacts made by children and adolescents are more assortative than other age groups. The same study found that individuals aged 55 years and older had the least assortative contact patterns [31]. We plan to extend our method to incorporate these additional contact patterns. We also plan to use stratification to reduce confounding bias.

Second, we plan to extend our probability model for the all-or-none protection vaccination model in which a proportion of vaccinated individuals are completely protected from infection and the remaining vaccinated individuals are not protected at all [20, 21].

Finally, we plan to analyze data from different influenza seasons and different settings to better determine the effectiveness of influenza vaccination against household transmission compared to community transmission. One study during the 2010-2011 influenza season found that VE against household and community transmission were different [11]. However, more recent studies have not observed this difference in VE against household- and community-acquired influenza infection [6, 10]. Further research is required to elucidate the impact of contact dynamics within populations of households on influenza VE; however, our model and maximum likelihood procedure provide a framework to begin distinguishing influenza VE from different sources of infection.

5. Acknowledgements

References

1. Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations from the advisory committee on immunization practices (ACIP) - united states, 2013-2014. *MMWR 2013* 2013; **62**(7).

2. Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations from the advisory committee on immunization practices (ACIP). *MMWR 2010* 2010; **59**(No. RR-8).

3. Ohmit SE, Thompson MG, Petrie JG, Thaker SN, Jackson ML, Belongia EA, Zimmerman RK, Gaglani M, Lamerato L, Spencer SM, *et al.*. Influenza vaccine effectiveness in the 2011-2012 season: protection against each circulation virus and the effect of prior vaccination on estimates. *CID* 2014; **58**(1 February).

4. McLean HQ, Thompson MG, Sundaram ME, Kieke BA, Gaglani M, Murthy K, Piedra PA, Zimmerman RK, Nowalk MP, Raviotta JM, *et al.*. Influenza vaccine effectiveness in the United States during 2012-2013: variable protection by age and virus type. *JID* 2015; **211**:1529–1540.

5. Pierse N, Kelly H, Thompson MG, Bissielo A, Radke S, Huang QS, Baker MG, Turner N. Influenza vaccine effectiveness for hospital and community patients using control groups with and without non-influenza respiratory viruses detected, Auckland, New Zealand 2014. *Vaccine* 2016; **34**:503–509.

6. Ohmit SE, Petrie JG, Malosh RE, Johnson E, Truscon R, Aaron B, Martens C, Cheung C, Fry AM, Monto AS. Substantial influenza vaccine effectiveness in households with children during the 2013-2014 influenza season, when 2009 pandemic influenza A(H1N1 virus predominated. *JID* 2016; **213**(8):1229–1236.

7. Monto AS. Studies of the community and family: acute respiratory illness and infection. *Epidemiologic Reviews* 1994; **16**(2):351–373.

8. Cowling BJ, Ip DKM, Fang VJ, Suntarattiwong P, Olsen SJ, Levy J, Uyeki TM, Leung GM, Peiris JSM, Chotpitayasunondh T, *et al.*. Aerosol transmission is an important mode of influenza A virus spread. *Nat. Commun.* 2013; **4**(1935):doi:10.1038/ncomms2922.

9. Davis XM, Haber M. Estimating vaccine efficacy from household data observed over time. *Statist. Med.* 2004; **23**:2961–2974.

10. Ohmit SE, Petrie JG, Malosh RE, Fry AM, Thompson MG, Monto AS. Influenza vaccine effectiveness in households with children during the 2012-2013 season: assessments of prior vaccination and serologic susceptibility. *JID* 2015; **211**(10):1519–1528.

11. Ohmit SE, Petrie JG, Malosh RE, Cowling BJ, Thompson MG, Shay DK, Monto AS. Influenza vaccine effectiveness in the community and the household. *CID* 2013; **56**(10):1363–1369.

12. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* 2013; **17**:21652168.

13. Sullivan SG, Feng S, Cowling BJ. Potential of the test-negative design for measuring influenza vaccine effectiveness: a systematic review. *Expert Rev. Vaccines* 2014; **13**(12):1571–1591.

14. Longini IR, Koopman JS. Household and community transmission parameters from final distributions of infections in households. *Biometrics* 1982; **38**:115–126.

15. Haber M, Longini IM, Cotsonis G. Models for the statistical analysis of infectious disease data. *Biometrics* 1988; **44**:163–173.

16. Davis XM, Haber M. Estimation of vaccine efficacy from household data. CD-ROM 2001.

17. Halloran EM, Preziosi MP, Chu H. Estimating vaccine efficacy from secondary attack rates. *JASA* 2003; **98**(461):38–46.

18. Petrie JG, Ohmit SE, Cowling BJ, Johnson E, Cross RT, Malosh RE, Thompson MG, Monto AS. Influenza transmission in a cohort of households with children: 2010-2011. *PLoS One* 2013; **8**(9).

19. Malosh R, Ohmit SE, Petrie JG, Thompson MG, Aiello AE, Monto AS. Factors associated with influenza vaccine receipt in community dwelling adults and their children. *Vaccine* 2014; **32**:1841–1847.

20. Smith PG, Rodrigues LC, Fine PEM. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *Int. J. Epidemiol.* 1984; **13**(1):87–93.

21. Haber M, Longini IM, Halloran ME. Measures of the effects of vaccination in a randomly mixing population. *International Journal of Epidemiology* 1991; **20**:300–310.

22. Halloran ME. *Epidemiologic Methods for the Study of Infectious Diseases*, chap. Concepts of Transmission and Dynamics. Oxford University Press: New York, NY, 2001.

23. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. *SIAM J. Scientific Computing* 1995; **16**:1190–1208.

24. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2016. URL https://www.R-project.org/.

25. US Census Bureau. Cesus 2010, summary file 1, households and families: 2010 and qt-p11 2010. URL http://factfinder.census.gov.

26. Carrat F, Vergu E, Ferguson NM, Lemaitre M, Cauchemez S, Leach S, Valleron AJ. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. *Am J Epidemiol* 2008; **167**(7):775–785.

27. Hayward AC, Fragaszy EB, Wang L, Copas A, Edmunds WJ, Ferguson N, Goonetilleke N, Harvey G, Kovar J, Lim MSC, *et al.*. Comparative community burden and severity of seasonal and pandemic influenza: results of the Flu Watch cohort study. *Lancet Respir Med* 2014; **2**(6):445–454.

28. Leung NHL, Xu C, Ip DKM, Cowling BJ. The fraction of influenza virus infections that are asymptomatic: a systematic review and meta-analysis. *Epidemiology* 2015; **26**(6):862–872.

29. Foy HM, Cooney MK, Allan ID, Albrecht JK. Influenza B in households: virus shedding without symptoms or antibody response. *Am. J. Epidemiol.* 1987; **126**:506–515.

30. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* 2006; **164**(10):936–944.

31. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, *et al.*. Social contacts and mixing patters relevant to the spread of infectious diseases. *PLoS Medicine* 2008; **5**(3):381–391.

Table 1. Model Parameters

Parameter	Definition
i	Index over people, where $i = 1, \ldots, N$
j	Index of infection status on a given day: (0=escaped infection (susceptible), 1=made an infectious contact on this day, 2=made an infection contact before this day).
v_{id}	Vaccination status of person i on day d. Also denoted at v (0=unvaccinated, 1=vaccinated)
β_v	Daily transmission probability from an infectious household member to a susceptible with vaccination status $v = 0, 1$
γ_v	Daily transmission probability from infectious community contacts to a susceptible with vaccination status $v = 0, 1$
m_{id}	The number of infectious persons in the household of person i on day d
$p(d)$	The prevalence of influenza infection on day d
π_{ijd}	Conditional probability that person i has infection status j on day d given that s/he was susceptible on the previous day
ψ_{ijd}	Unconditional probability that person i has infection status j on day d
λ_{iH}	Probability that person i was infected from a household contact by the end of the study
λ_{iC}	Probability that person i was infected from a community contact by the end of the study

Table 2. Household and community transmission parameter estimates by vaccination status from 200 simulated influenza outbreaks

Scenario	Value	β_0	β_1	γ_0	γ_1
	True	**0.150**	**0.075**	**0.400**	**0.100**
	Estimate	**0.153** (0.129, 0.177)	**0.078** (0.064, 0.092)	**0.429** (0.380, 0.478)	**0.118** (0.092, 0.142)
All vaccinations (50%) occurred prior to the study	Bias	0.003	0.003	0.029	0.018
	SE (Empirical)	0.012	0.007	0.025	0.011
	SE (Hessian)	0.013	0.007	0.026	0.011
	SE (Bootstrap)	0.014	0.007	0.027	0.012
Vaccinations occurred prior to the study (25%), in month 1 (15%), and in month 2 (10%)	**Estimate** (95% CI)	**0.156** (0.132, 0.180)	**0.078** (0.060, 0.096)	**0.425** (0.382, 0.468)	**0.118** (0.093, 0.143)
	Bias	0.006	0.003	0.025	0.018
	SE (Empirical)	0.012	0.009	0.022	0.013
	SE (Hessian)	0.012	0.008	0.023	0.014
	SE (Bootstrap)	0.012	0.009	0.022	0.015

Transmission parameters (β_0-household, unvaccinated; β_1-household, vaccinated; γ_0-community, unvaccinated; and γ_1-community, vaccinated) and SEs. SEs were calculated empirically, using the Hessian matrix from the maximization procedure, and by parametric bootstrap.

Table 3. Maximum likelihood VE estimates against influenza infection in the household, community, and overall from 200 simulated influenza outbreaks

Scenario	Value	VE_{MLH}	VE_{MLC}	VE_{MLO}
All vaccinations (50%) occurred prior to the study	**True**	**0.477**	**0.746**	**0.607**
	Estimate (95% CI)	**0.467** (0.354, 0.580)	**0.721** (0.662, 0.779)	**0.601** (0.544, 0.659)
	Bias	-0.010	-0.025	-0.006
	SE (Empirical)	0.056	0.030	0.030
	SE (Bootstrap)	0.064	0.034	0.030
	Coverage Probability	0.960	0.920	0.960
Vaccinations occurred prior to the study (25%), in month 1 (15%), and in month 2 (10%)	**True**	**0.470**	**0.745**	**0.605**
	Estimate (95% CI)	**0.469** (0.339, 0.600)	**0.717** (0.645, 0.789)	**0.602** (0.540, 0.664)
	Bias	-0.001	-0.028	-0.003
	SE (Empirical)	0.067	0.037	0.032
	SE (Bootstrap)	0.069	0.038	0.033
	Coverage Probability	0.960	0.950	0.970

Household VE (VE_{MLH}), community VE (VE_{MLC}), overall VE (VE_{MLO}) and SEs. SEs were calculated empirically and by parametric bootstrap.

Table 4. Estimates of VE when the latent and infectious periods are not constant from 200 simulated influenza outbreaks

Mean Length of Latent and Infectious Periods	Scenario	Value	VE_{MLH}	VE_{MLC}	VE_{MLO}
$\bar{L} = 2, \bar{I} = 4$	All vaccinations (50%) occurred prior to the study	**True**	**0.477**	**0.747**	**0.607**
		Estimate (95% CI)	**0.474** (0.349, 0.599)	**0.703** (0.635, 0.770)	**0.602** (0.546, 0.657)
		Bias	-0.003	-0.044	-0.005
		SE (Empirical)	0.064	0.035	0.028
		SE (Bootstrap)	0.074	0.031	0.032
	Vaccinations occurred prior to the study (25%), in month 1 (15%), and in month 2 (10%)	**True**	**0.471**	**0.745**	**0.606**
		Estimate (95% CI)	**0.470** (0.327, 0.612)	**0.699** (0.630, 0.768)	**0.600** (0.531, 0.669)
		Bias	-0.001	-0.046	-0.006
		SE (Empirical)	0.073	0.035	0.035
		SE (Bootstrap)	0.070	0.034	0.033
$\bar{L} = 2.7, \bar{I} = 5$	All vaccinations (50%) occurred prior to the study	**True**	**0.469**	**0.745**	**0.605**
		Estimate (95% CI)	**0.440** (0.303, 0.577)	**0.679** (0.617, 0.740)	**0.589** (0.542, 0.636)
		Bias	-0.029	-0.066	-0.016
		SE (Empirical)	0.070	0.031	0.024
		SE (Bootstrap)	0.074	0.028	0.027
	Vaccinations occurred prior to the study (25%), in month 1 (15%), and in month 2 (10%)	**True**	**0.462**	**0.743**	**0.604**
		Estimate (95% CI)	**0.442** (0.298, 0.559)	**0.657** (0.584, 0.720)	**0.575** (0.525, 0.623)
		Bias	-0.020	-0.086	-0.029
		SE (Empirical)	0.067	0.035	0.026
		SE (Bootstrap)	0.075	0.033	0.030

Bias of VE estimates were calculated allowing the latent and infectious periods to follow a distribution. VE_{MLH}, VE_{MLC}, and VE_{MLO} denote VE estimates using the maximum likelihood approach against household, community, and overall transmission, respectively. We considered two situations. First, the mean length of the latent ($\bar{L}$) and infectious ($\bar{I}$) periods were $\bar{L} = 2$ and $\bar{I} = 4$ (i.e., they were correctly specified). Second, $\bar{L} = 2.7$ and $\bar{I} = 5$ (i.e., they were misspecified). In both situations, during the estimation procedure, it was assumed that L=2 and I=4. VE estimates were obtained from 200 simulations under two different vaccination scenarios. SEs were calculated empirically and by parametric bootstrap.

Table 5. VE estimates from a random sample drawn from a larger population from 200 simulated influenza outbreaks

Scenario	Value	VE_{MLH}	VE_{MLC}	VE_{MLO}
All vaccinations (50%) occurred prior to the study	**True**	**0.476**	**0.746**	**0.605**
	Estimate	**0.480** (0.344, 0.615)	**0.727** (0.662, 0.791)	**0.614** (0.551, 0.677)
	Bias	0.004	-0.019	0.009
	SE (Empirical)	0.069	0.033	0.032
	SE (Bootstrap)	0.065	0.032	0.030
Vaccinations occurred prior to the study (25%), in month 1 (15%), and in month 2 (10%)	**True**	**0.469**	**0.745**	**0.604**
	Estimate	**0.469** (0.334, 0.604)	**0.722** (0.648, 0.796)	**0.607** (0.544, 0.671)
	Bias	0.000	-0.023	0.003
	SE (Empirical)	0.069	0.038	0.032
	SE (Bootstrap)	0.070	0.037	0.032

Bias of VE from a sample population randomly selected from a larger population. VE_{MLH}, VE_{MLC}, and VE_{MLO} denote VE estimates using the maximum likelihood approach against household, community, and overall transmission, respectively. True VE was calculated from the overall population. VE estimates were obtained from 200 simulations performed under two different vaccination scenarios. SEs were calculated empirically and by parametric bootstrap.

Table 6. Transmission parameter estimates from the HIVE Study

Value	β_0	β_1	γ_0	γ_1
Estimate (95% CI)	0.013 (0.008, 0.019)	0.013 (0.005, 0.021)	0.202 (0.159, 0.245)	0.143 (0.081, 0.187)
SE (Hessian)	0.003	0.003	0.032	0.026
SE (Bootstrap)	0.003	0.004	0.022	0.027

Maximum likelihood transmission parameter estimates for the HIVE Study (2012-13). Transmission parameters were defined as follows: β_0-household, unvaccinated; β_1-household, vaccinated; γ_0-community, unvaccinated; and γ_1-community, vaccinated.

Table 7. Maximum likelihood estimates of VE and confidence intervals from the HIVE Study

Value	VE_{MLH}	VE_{MLC}	VE_{MLO}
Estimate (95% CI)	0.052 (-0.754, 0.858)	0.336 (0.066, 0.606)	0.250 (-0.019, 0.519)
SE	0.411	0.138	0.137

Maximum likelihood VE estimates and 95% confidence intervals for the HIVE Study (2012-13). VE_{MLH}, VE_{MLC}, and VE_{MLO} denote VE estimates using the maximum likelihood approach against household, community, and overall transmission, respectively.

Table 8. VE estimates based on hazard rate ratios and confidence intervals from the HIVE Study reported by Ohmit *et al.* [10]

Model	VE_H	VE_C	VE_O
Unadjusted	0.31 (-0.73, 0.73)	0.27 (-0.13, 0.54)	0.30 (-0.09, 0.55)
Adjusted	0.37 (-0.73, 0.77)	0.30 (-0.09, 0.55)	0.43 (-0.18, 0.72)

VE point estimates and 95% confidence intervals from Ohmit *et al.* using both the unadjusted and adjusted hazard rate ratio. The adjusted models adjusted for age in months and documentation of high-risk health status [10]. VE_H, VE_C, and VE_O denote VE estimates against household, community, and overall transmission, respectively.