Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Shawnee Anderson Date

Modeling the preventive effectiveness of Influenza vaccination

By

Shawnee M. Anderson
Masters of Science
Biostatistics

_____

Michael Haber, Ph.D
Advisor

_____

Yijuan Hu, Ph.D
Committee Member

_____

Lance Waller, Ph.D
Chair, Department of Biostatistics
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Modeling the preventive effectiveness of Influenza vaccination

By

Shawnee M. Anderson
Bachelors of Arts, Applied Mathematics & Statistics

Advisor: Michael Haber, Ph.D

An abstract of
A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science in Biostatistics

2015

Abstract

Modeling the preventive effectiveness of Influenza vaccination
By Shawnee M. Anderson

Vaccine effectiveness (VE), the observed relative reduction in risk associated with vaccination, is often used to measure the performance of vaccine or overall vaccination program. It gives the *direct* benefit of vaccination provided to an individual. However, as a population measure, it fails to account for the indirect effects of vaccination, resulting from the reduced number of infectious individuals in the population at any time point. Population vaccine effectiveness (PVE), the reduction in incidence for a population that has a vaccination program compared to the same (or similar) unvaccinated population, provides a way for public health officials to quantify the overall population benefit of vaccination. PVE however is difficult to estimate given the need for data on the population in the absence of vaccination. This thesis will provide, through use of a stochastic simulation program, a means of modeling PVE through data on vaccine coverage and observed vaccine effectiveness.

Modeling the preventive effectiveness of Influenza vaccination

By

Shawnee M. Anderson
Bachelors of Arts, Applied Mathematics & Statistics

Advisor: Michael Haber, Ph.D

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science in Biostatistics

2015

Contents

List of Tables

List of Figures

A Introduction

A.1 Background

The 2014 measles outbreak at a Disneyland amusement park in Southern California, brought a renewed focus on the role that vaccination plays in the control of infectious disease outbreaks either through eradication or mitigation. From a public health perspective, the evaluation of disease intervention programs, which include vaccination, helps to determine where, when and to whom resources should be delivered to in order to proactively combat disease spread or in response to an active disease outbreak. For the purposes of this thesis, evaluation of a vaccination program will focus only on the performance of a vaccine on the entire population. The question then becomes, how do we go about measuring the performance of a vaccine.

A.2 Vaccine Efficacy

The most basic measure of the performance of a vaccine, vaccine efficacy is defined as $1-\frac{r_{uv}}{r_v}$, where r_{uv} and r_v are the transmission probabilities in the unvaccinated and vaccinated respectively. In this definition we see that vaccine efficacy is an individual level measure quantifying the individual benefit a vaccine provides in the reduction of risk or transmissability and or suscpetibility (taking into account vaccinations that do not confer immunity) within those who are vaccinated. This allows for straight-forward comparisons of different vaccines performance easy to do.

A.3 Vaccine Effectiveness (VE)

In a vaccinated population, we can estimate the direct effect of vaccination by looking at the observed VE defined as $1-\frac{AR_v}{AR_u}$, where AR_v and AR_u are the attack rates in both the vaccinated and unvaccinated population groups respectively. *Vaccine effectiveness* then compares a vaccinated and an unvaccinated person in the same population, and is used to quantify disease risk. Estimation of this measure depends on the total population of each vaccination subgroup, as well as reliable informationon on symptomatic disease incidence in the population. For large populations (cities, counties, etc.) this can be quite difficult to obtain, and thus arriving at unbiased

estimates of the attributable risk of an individual not vaccinating may not be possible.

A.4 Indirect Effects

Because a large number of biological, behavioral and demographic changes due to vaccination are unknown or not easily measured, specifically including these indirect effects in disease models is difficult. Instead of accounting for each of these sources we assume one underlying unobserved process that acts to modify risk of infection in the vaccinated.

Suppose we have two individuals from a closed population, Tomerot and Chombi, brothers who make at least one effective contact with one another each day. Note that an *effective contact* is defined as contact sufficient enough for disease transmission; the definition of contact is entirely driven by the disease or virus of interest, i.e. intercourse for the transmission of HIV and physical proximity (less than six feet) of an individual infected with influenza . Suppose further that of the two brothers, Tomerot is the only one vaccinated. By virtue of Tomerot's vaccination status and reduced risk of infection or transmission, Chombi's overall risk of infection or transmission is changed as well. This change in Chombi's disease risk is the *indirect effect of vaccination*, the effect of a vaccinated individual on another's disease outcome. Such an indirect effect need not come solely from the direct benefit of a reduction in transmisability of a vaccinated individual. Behavioral changes made by vaccinated individuals can confer a protective effect on unvaccinated indivuals as well. If as a result of vaccination, a vaccinated individual chooses to reduce or increase the number of effective contacts they make with the population, this will in turn reduce or increase the overall number of infectives, indirectly reducing or increasing the disease risk for all other susceptible individuals.

From the description of both vaccine efficacy and effectiveness, we can easily see that these are individual level measures that quantify a vaccine's direct performance. For pharmaceutical companies and consumers making vaccination decisions for themselves or those in their care, measures are of great importance in determining the relative "worth" of vaccination whether in the form of vaccine production cost or risk associatied with infection. Unfortunately, VE is not a good measure of the overall

performance of a vaccination program as it does not account for indirect effects (as described above), and thus says nothing about the reduction in disease incidence for a population, only providing a measure of the reduction in infection risk for an individual. Therefore, they do not provide public health officials with an objective way to measure the vaccines performance on the population as a whole or against other alternative vaccination schemes. In order to determine the population level impact of a vaccination program while also accounting for the indirect effects of vaccination, it is neccessary to observe its performance in a population in which vaccination does not exist.

A.5 Population Vaccine Effectiveness (PVE)

Population vaccine effectiveness (PVE) or vaccination impact defined as $1 - \frac{AR_p}{AR_n}$, where AR_p is the attack rate in the population of interest, and AR_n is the expected attack rate in the same population when no one is vaccinated. With this definition, PVE gives the risk associated with a population not having implemented a vaccination program. For a disease for which there is no previous vaccination, incidence data from previous outbreaks or seasons provide data on the unvaccinated population, and an estimate of PVE is made after introduction of a vaccination program. However, for a vast majority of diseases, vaccination programs already exist making estimation of PVE difficult if not ethically impossible. In these cases, estimation of PVE may occur using a *surrogate* unvaccinated population: a population with similar demographic and mixing patterns that has no vaccination program for the disease of interest. Alternatively, PVE can be estimated from stochastic models for the spread of a disease in a population.

A.6 Thesis Goal

This thesis will seek to model the population effectiveness of vaccination, specifically vaccination against seasonal influenza vaccination. Influenza vaccination presents several unique challenges in PVE estimation that are not found with other disease vaccination programs. As the name suggests, seasonal influenza demonstrates an incredible amount of heterogeneity in yearly viral strains making assessment of long-term vaccine effectiveness difficult. Because an influenza strain directly impacts

both the onset and intensity of a "flu season", estimates of vaccine effectiveness may be biased depending on the severity of flu strain pre and post-implementation of a vaccination program. Further complications arise from the relatively weak performance of an influenza vaccine program in comparison to other programs; pneumococcal conjugate vaccines have an observed effectiveness of 75 to 95% in comparison to vaccines for seasonal influenzas effectiveness of at most 60% in lab confirmed influenza virus infection and even as low as 30%. This can make it difficult to measure the expected reduction in incidence of influenza post implementation of a vaccination program. The *WHO Field Guide for the Evaluation of Influenza Vaccine Effectiveness* suggests several years of flu survailance pre and post-implemmentation in order to minimize the risk of biased PVE estimates.

In this thesis, we will present a model to (a) estimate the population effectiveness of a vaccination program in a heterogeneous population, and (b) explore factors that affect influenza population vaccine effectiveness. We will accomplish this by using SIMFLU (described in the next section), a detailed stochastic agent-based program that simulates an influenza outbreak in a population, where each member of the population belongs to a specific age group and has a vaccination status. SIMFLU allows us to modify various vaccination and demographic parameters covering several outbreak scenarios and population compositions. Simulation output will provide incidence on a population of 10,000 people under two population compositions: a homogenous population and an age-stratified (children and adults) population. With this output we will investigate the association between PVE and demographic and vaccination covariates, specifically vaccine coverage and observed vaccine effectiveness. This will help predict population level effectiveness of future vaccination strategies.

B Methods

B.1 Simulation

Estimation of population vaccination effectiveness is important in that it allows for estimation of the total population that must be vaccinated to eliminate or mitigate

disease spread in a population. By modeling incidence against vaccine coverage and observed vaccine effectiveness we obtain a surrogate estimate for PVE.

$$PVE = 1 - \frac{AR_p}{AR_n} \tag{1}$$

where AR_p is the observed symptomatic attack rate (incidence) in a flu season and AR_n is the attack rate in a population in which vaccination does not exist. Estimation of population PVE is challenging, as it requires an estimate of the expected incidence of the disease without vaccination in the same or similar population. To address this problem, we used the construction of a detailed agent-based infectious disease model where each person is assigned to an age-group stratum and given a vaccination status. Estimation of model paramters will be carried out by a stochastic simulation program, SIMFLU, where we simulate a disease outbreak in a partially vaccinated population. AR_p and AR_n wil be obtained from simulations results, where AR_n is obtained from the simulation where vaccination coverage is set to 0%.

Our simulations will be based on the risk of infection of a randomly selected member of the population on a given day. We model risk of infection on day d of an individual from stratum k of vaccination status $v = 0$ or 1, for unvaccinated or vaccinated respectively, as:

$$P_d(\text{inf}|k, v) = 1 - [1 - P_d(t|k, v)]^{c_k} \tag{2}$$

where c_k are the average number of effective contacts made per individual in age-stratum k per day; within each strata the c_k contacts are taken to be independent. Further $P_d(t|k, v)$ is the probability at time d of transmission to an individual from stratum k of vaccination status v in a single contact. Through repeated simulation of a full flu season (4 months for the purposes of this thesis), we obtain a large body of data on the incidence counts of influenza for the population.

B.2 SIMFLU simulation routine

SIMFLU is a stochastic simulation program designed by Dr. Michael Haber, and implemented in C++. The program takes a single input file containing output, demographic, disease and vaccination parameters. Output options include obtaining data at the individual level per day d (in this case we take as our unit of time a day in the outbreak) up to average overall disease incidence over total simulations $NSIM$. As well for each simulation we may obtain detailed information on vaccination and contacts by each individual per day over the course of the outbreak.

Demographic parameters allow for adjustment of the total number of age-strata (=k) as well as their size, distribution of contacts made by an individual across all k stratums and total number of contacts (= c_k) made by an individual in stratum k. We may also adjust monthly vaccination coverage for each of our k strata; vaccination of an individual is assumed to occur in the month prior because the vaccine becomes effective two weeks after vaccination (i.e. a person vaccinated in month 2 is said to be fully vaccinated during month 3).

Characteristics of disease that we are able to adjust in the simulation include length of latent and infectious periods (measured in days), probability of illness given infection, relative infectiousness of an asymptomatic individual and transmission probabilities for each month for each vaccination status across all k strata.

An infected individual who has recovered is considered immune to infection; s/he can however have continued contact with individuals in the population. We set the latent period, defined as the period between infection and infectiousness, at 2 days. Further, we set the infectious period - the total time in which an individual may transmit the diease - at 4 days. Infected individuals exist in one of two states: symptomatic or asymptomatic, coded as states 3 or 2 with symptomatic infection occuring with probability 0.67. For the purposes of this thesis, we assume that asymptomatic individuals have their probability of disease transmission reduced by 40%.

B.3 Description of the one and two stratum details

We will consider two population scenarios for use in the SIMFLU program.

One Stratum: In the one stratum setup, we consider a homogenous population with random mixing. The total population is fixed at 10,000 individuals who each make on average 10 effective contacts a day. We vaccinate a proportion of this population prior to an influenza outbreak and at no other point in the season. We vary vaccine coverage from 40% to 95% in increments of 5%. It is assumed that our vaccine does not confer total protection against infection or transmission, the so-called "leaky vaccine". We set disease transmission at 0.06 per effective contact for unvaccinated individuals. We wish to look at a range of efficacy values for influenza and thus set transmission for the vaccinated at 0.01, 0.02 and 0.03 (adhering to high, medium and low levels of vaccine efficacy) yielding 36 vaccine coverage and efficacy scenarios.

Finally, to start the outbreak we set 30 individuals in the population as infecteds and in the infectious period.

Two Strata: The two strata setup will be similar to the one stratum case except for one change: we now have two age-strata of equal size, representing children and adults, and we will only vaccinate children. Demographic details and characteristics of disease remain the same for both age-strata. Under this setup, we wish to see how PVE for adults is affected by the vaccine coverage of children in stratum one. While unlikely, we still adhere to to the random mixing setup as in the one stratum case.

B.4 Estimation of PVE

We run 100 simulations using SIMFLU for each of our 39 vaccine coverage and efficacy scenarios. Original data from the SIMFLU output include information on incidence of influenza for each stratum broken down by vaccination, symptomatic and overall population status for each of the 3,900 simulations across efficacy and coverage scenarios. These data are reformatted in the R statistical analysis program, version 3.2.2, with the final dataset being a complete record of the 3,900 simulations. Each data point corresponds to a single simulation and contains: vaccine efficacy and coverage, total unvaccinated and vaccinated population, symptomatic and asymptomatic incidence in the unvaccinated and vaccinated population, overall incidence in the population, observed vaccine effectiveness and preventive vaccine effectiveness.

Sample mean and standard errors for the overall and efficacy subgroups are listed in Table 1. We next fit a Poisson log-link model to the count data for each simulation's symptomatic incidence, and then fit negative binomial model allowing for non-equal mean and variance that may provide a better fit to the data.

Poisson Model - One Stratum: To begin we consider a Poisson model with a log-link where our outcome is incidence of symptomatic influenza and independent variables include vaccine coverage and efficacy; we will refer to this as the *full model* in the remainder of this text:

$$\log(INC_s) = \beta_1 VCov_s + \beta_2 VE_s + \beta_3 VCov_s * VE_s + \text{offset} \tag{3}$$

where s indexes the simulation run and INC_s, $VCov_s$ and VE_s are the symptomatic incidence, vaccine coverage and observed vaccine effectiveness in the total population of the s^{th} simulation; since we wish to estimate $PVE = 1 - \frac{AR_p}{AR_n}$, where $AR_p = INC_s/10,000$ for all simulations in the one stratum case, we set the offset to $\log(10,000)$. We also fit a vaccine coverage only model, stratified on level of vaccine efficacy (low, medium or high):

$$\log(INC_{s_l}) = \beta_1 VCov_{s_l} + \text{offset} \tag{4}$$

where s_l is the s^{th} simulation in the $l = low, medium$ or $high$ vaccine efficacy subgroup.

Negative Binomial Model - One Stratum: Given the unequal sample variance and standard errors of our overall simulation data and efficacy subgroups, a negative binomial model would seem to be a better fit for the data. We will thus fit a negative binomial model with a log-link with outcome and independent variables as in the Poisson models (3) and (4).

Two Strata Set-up: Modeling for the two-strata scenario will proceed as in the one-stratum scenario for the coverage only Poisson and negative binomial model, once fitting the model with both coverage in children and vaccine effectiveness in adults and once for the model with coverage only.

We will run each of these 16 models with output produced by SIMFLU. All analysis will be carried out in SAS version 9.2. Model fit will be determined by Pearson chi-squared tests and AIC where appropriate.

C Results

Below we present results and findings for the full model and coverage only models; tables and figures are provided at the end of text in order to aid readability. We first note that over all scenarios and model formulation that the negative binomial model performed better than the Poisson model in terms of AIC and Pearson Chi-square, as expected. Thus we will focus our discussion of these results only on the negative binomial model. Further, for readabilities sake, we provide talbes and figures at the end of the text.

Table 1 gives summary statistics on vaccine effectiveness (VE), symptomatic incidence in the population (INC_p) and population vaccine effectiveness (PVE) for each of level of vaccine coverage; results are stratified by vaccine efficacy (low, medium and high). We notice that for low levels of efficacy and coverage, VE is always less than PVE in our simulated results. However at medium and high levels of efficacy, VE outpaces PVE when vaccine coverage is low. At 50% coverage when efficacy is medium or high, PVE sharply increases while VE remains relatively stable around the pre-set efficacy level.

Figures 1 & 2 show PVE and VE plotted against vaccine coverage for the one and two strata setups; again we stratify on level of vaccine efficacy. Here we see more clearly what the summary statistics bear out: there is a cross-over of size for VE and PVE at the medium and high levels of efficacy with the crossover occuring at low coverage for a vaccine with medium efficacy and at greater than half coverage for high efficacy. At low efficacy in a homogenous population, VE is always less than PVE. We can then see that accounting for direct *and* indirect effects of vaccination (VE versus PVE) in a homogenous population, population vaccine effectiveness gives a clearer view of the overall vaccines performance than when we rely only on the *direct* effect of vaccination only, i.e VE. While VE remains relatively stable (in particular in the two stratum scenario) we see that even when efficacy is low, at high levels of coverage

the reduction of incidence is at a minimum 75% in the population, a population performance masked by the individual performance of the vaccine.

C.1 Results from Models

Tables 2 gives covariate estimates for the full model and our efficacy stratified vaccine coverage models.

Unsurprisingly, all of our effects are significant; however we note the highly significant interaction term in both the one and two-strata formulations. We expect vaccine coverage to be associated with incidence in the population and therefore PVE, however its association is heavily modified by vaccine effectiveness, suggesting that stratification on level of efficacy is warranted. Tables 3 and 4 give parameter estimates for the vaccine coverage only model across the three efficacy levels for the one and two-strata case respectively. It is clear that as efficacy increases, the effect of vaccine coverage dramatically increases; as we saw in Figure 3, even at low efficacy there is a noticeable association between coverage and incidence, however the effect is most striking at the high efficacy level. Figures 3 & 4 plot each sample PVE curve against vaccine coverage for the one and two strata models on the same scale, so that we can more clearly see the impact of coverage on vaccine impact.

D Discussion & Limitations

D.1 Discussion

By taking into account both direct and indirect effects of vaccination, population vaccine effectiveness is an attractive measure of the overall performance of a vaccination program for public health officials. However, because indirect effects are often part of an unobserveable random processes and estimation of PVE requires data on the same (or similar) population sans vaccination program, arriving at a reliable and unbiased estimate of PVE can be problematic. The simulation and analysis of simulated results in this thesis have tried to address this problem by describing the relationship between PVE with vaccine coverage and observable vaccine effectiveness in a population. From the results section vaccine coverage and observed vaccine

effectiveness are significantly associated with symptomatic incidence both in a population and a sub-group of that population (adults for the purposes of this thesis). Thus for the one and two-stratum case we can easily obtain estimates of preventive vaccine effectiveness.

Our visual and regression analysis showed clearly the role vaccine coverage plays in PVE. The vaccine coverage only model in particular demonstrates how, even at low levels of efficacy high vaccine coverage can lead to high levels of population protection against infection. Further, we have demonstrated that even with low vaccine effectiveness observed in a population following a flu season, the impact of vaccination trends upward with an increase in coverage. For example, from Table 2, in a one stratum population in which a vaccination with medium efficacy vaccine is implemented, while VE ranges from 0.52 to 0.63, PVE ranges from 0.46 to 0.98. The most common scenario for influenza vaccination puts coverage at 60% in the population; from Table 1 we see that at 60% vaccination, while individual effectiveness is at 0.64, population level protection is at 0.84.

D.2 Limitations

: Though we have shown a clear association between population vaccine effectiveness and vaccine coverage and observed vaccine effectiveness, use of this association for pre-planning purposes is a bit problematic. Observed vaccine effectiveness, one minus the ratio attack rates in the vaccinated and unvaccinated, is a function of total vaccinated and unvaccinated in the population as well as symptomatic attack rate *after* an outbreak. Unlike vaccine coverage, we can not fix observed vaccine effectiveness before an outbreak in order to achieve a coverage and effectivness combination for a certain level of population level vaccine effectiveness. However, this model formulation has utility in comparing the impact of vaccination of two competing vaccination programs with similar population vaccination levels.

We are further limited in the generalization of our results by the assumption of mixing between stratum in our two-stratum model. For the purposes of this thesis, we assumed random mixing in the two-strata model; an effective contact of an individual in stratum 1 or 2 was equally likely to be from either of our two strata. In our

formulation, the two strata referred to children and adults, thus this assumption is highly unlikely to be seen in any real world situation. As the scale of simulation increases, and the demographic stratification becomes finer, it will become neccessary to have clearer understanding of the mixing patterns among the various demographic groups we wish to represent in the model if we are to arrive at reliable and unbiased estimates of preventive vaccine effectiveness.

Finally, a more detailed meta-analysis of influenza disease and vaccine parameters is warranted. For this thesis, parameters were chosen in order to ensure an "outbreak" occured within our simulation so that we could observe the behavior of coverage and effectiveness on incidence. However in order to obtain more reliable estimates of incidence to aid in prediction and estimation of PVE, more robust estimates of parameters (including disease transmission and reduction in succeptibility due to vaccination) are needed.

		Low Efficacy				Medium Efficacy				High Efficacy			
VCOV %	Variable	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
0.20	VE	0.34	0.02	0.29	0.40	0.52	0.02	0.46	0.57	0.74	0.02	0.70	0.78
0.20	INCp	4572	79.51	4370	4769	4225	89.19	3991	4436	3798	99.78	3520	4011
0.20	PVE	0.41	0.01	0.39	0.44	0.46	0.01	0.43	0.49	0.51	0.01	0.48	0.55
0.25	VE	0.34	0.02	0.28	0.41	0.53	0.02	0.47	0.58	0.76	0.02	0.72	0.79
0.25	INCp	4395	86.49	4183	4587	3951	98.95	3743	4206	3382	114.51	3033	3591
0.25	PVE	0.43	0.01	0.41	0.46	0.49	0.01	0.46	0.52	0.56	0.01	0.54	0.61
0.30	VE	0.35	0.02	0.30	0.40	0.54	0.02	0.50	0.59	0.77	0.01	0.73	0.80
0.30	INCp	4224	97.07	3985	4550	3636	102.27	3390	3869	2959	96.08	2742	3293
0.30	PVE	0.46	0.01	0.41	0.49	0.53	0.01	0.50	0.56	0.62	0.01	0.58	0.65
0.35	VE	0.36	0.02	0.32	0.41	0.55	0.02	0.51	0.60	0.78	0.01	0.75	0.81
0.35	INCp	4004	109.71	3654	4220	3348	107.80	3096	3610	2485	130.59	2086	2776
0.35	PVE	0.48	0.01	0.46	0.53	0.57	0.01	0.54	0.60	0.68	0.02	0.64	0.73
0.40	VE	0.36	0.02	0.32	0.40	0.57	0.02	0.53	0.60	0.78	0.02	0.73	0.82
0.40	INCp	3823	112.09	3603	4089	3009	110.49	2749	3279	2026	163.60	1387	2403
0.40	PVE	0.51	0.01	0.47	0.54	0.61	0.01	0.58	0.65	0.74	0.02	0.69	0.82
0.45	VE	0.37	0.02	0.32	0.42	0.58	0.02	0.53	0.63	0.79	0.02	0.61	0.82
0.45	INCp	3606	99.70	3355	3843	2646	145.94	2148	2970	1451	249.74	70	1899
0.45	PVE	0.54	0.01	0.51	0.57	0.66	0.02	0.62	0.72	0.81	0.03	0.76	0.99
0.50	VE	0.38	0.02	0.33	0.43	0.59	0.02	0.54	0.64	0.80	0.02	0.75	0.86
0.50	INCp	3382	112.65	3002	3661	2200	158.74	1735	2497	873	241.28	283	1360
0.50	PVE	0.56	0.01	0.53	0.61	0.72	0.02	0.68	0.78	0.89	0.03	0.82	0.96
0.55	VE	0.39	0.02	0.34	0.43	0.61	0.02	0.56	0.66	0.80	0.03	0.67	0.85
0.55	INCp	3131	128.07	2888	3437	1743	239.82	919	2203	488	194.20	76.00	950
0.55	PVE	0.60	0.02	0.56	0.63	0.78	0.03	0.72	0.88	0.94	0.02	0.88	0.99

Table 1: Summary statistics of VE, INCp and PVE over 100 simulations for each combination of vaccine covarage and efficacy; One Stratum

		Low Efficacy				Medium Efficacy				High Efficacy			
VCOV %	Variable	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
0.60	VE	0.40	0.02	0.36	0.44	0.62	0.02	0.57	0.68	0.78	0.04	0.62	0.85
0.60	INCp	2862	146.11	2476	3201	1240	248.95	562	1799	247	110.16	63	576
0.60	PVE	0.63	0.02	0.59	0.68	0.84	0.03	0.77	0.93	0.97	0.01	0.93	0.99
0.65	VE	0.41	0.02	0.36	0.46	0.63	0.03	0.54	0.69	0.76	0.06	0.43	0.85
0.65	INCp	2565	170.45	2019	2910	784	256.57	136.00	1502	142	66.46	35	362
0.65	PVE	0.67	0.02	0.63	0.74	0.90	0.03	0.81	0.98	0.98	0.01	0.95	0.99
0.70	VE	0.42	0.02	0.38	0.50	0.62	0.05	0.37	0.73	0.72	0.08	0.45	0.85
0.70	INCp	2191	255.71	1243	2704	421	204.31	74.00	1038	92	33.20	36	208
0.70	PVE	0.72	0.03	0.65	0.84	0.95	0.03	0.87	0.99	0.99	0.00	0.97	0.99
0.75	VE	0.43	0.02	0.37	0.50	0.63	0.05	0.36	0.71	0.70	0.08	0.50	0.86
0.75	INCp	1803	261.14	994	2280	257	120.74	54	624	70	21.15	31	160
0.75	PVE	0.77	0.03	0.71	0.87	0.97	0.02	0.92	0.99	0.99	0.00	0.98	1.00
0.80	VE	0.44	0.03	0.38	0.52	0.61	0.06	0.44	0.73	0.65	0.15	-0.44	0.85
0.80	INCp	1323	355.32	245	2014	157	70.11	44	364	54	15.59	22	95
0.80	PVE	0.83	0.05	0.74	0.97	0.98	0.01	0.95	0.99	0.99	0.00	0.99	1.00

Table 2: Table 1 (cont.): Summary statistics of VE, INCp and PVE over 100 simulations for each combination of vaccine covarage and efficacy; One Stratum

		Low Efficacy				Medium Efficacy				High Efficacy			
VCOV %	Variable	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
0.40	VE	0.49	0.03	0.42	0.55	0.74	0.02	0.70	0.80	0.86	0.01	0.83	0.90
0.40	INCp	1946	148.77	1652	2315	1765	146.27	1413	2096	1689	156.56	1181	2022
0.40	PVE	0.09	0.07	-0.08	0.23	0.17	0.07	0.02	0.34	0.21	0.07	0.05	0.45
0.45	VE	0.50	0.03	0.43	0.55	0.74	0.02	0.69	0.80	0.87	0.01	0.83	0.90
0.45	INCp	1896	130.15	1612	2184	1711	148.36	988	2042	1545	198.91	179	2017
0.45	PVE	0.11	0.06	-0.02	0.24	0.20	0.07	0.04	0.54	0.28	0.09	0.06	0.92
0.50	VE	0.50	0.03	0.42	0.57	0.74	0.02	0.70	0.80	0.87	0.04	0.46	0.90
0.50	INCp	1884	135.76	1569	2222	1639	142.59	1206	1983	1416	225.37	15	1796
0.50	PVE	0.12	0.06	-0.04	0.26	0.23	0.07	0.07	0.43	0.34	0.11	0.16	0.99
0.55	VE	0.51	0.03	0.44	0.58	0.75	0.02	0.69	0.80	0.88	0.02	0.83	0.92
0.55	INCp	1823	129.10	1483	2184	1517	200.20	317	1965	1213	254.68	188	1659
0.55	PVE	0.15	0.06	-0.02	0.30	0.29	0.09	0.08	0.85	0.43	0.12	0.22	0.91

Table 3: Summary statistics of VE, INCp and PVE over 100 simulations for each combination of vaccine covarage and efficacy; Two Strata

		Low Efficacy				Medium Efficacy				High Efficacy			
VCOV %	Variable	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
0.60	VE	0.50	0.02	0.45	0.56	0.75	0.02	0.70	0.90	0.88	0.01	0.84	0.92
0.60	INCp	1801	136.29	1460	2153	1354	248.90	12	1729	1022	262.97	112	1488
0.60	PVE	0.16	0.06	-0.01	0.32	0.37	0.12	0.19	0.99	0.52	0.12	0.30	0.95
0.65	VE	0.51	0.02	0.45	0.59	0.76	0.03	0.70	0.83	0.88	0.02	0.84	0.93
0.65	INCp	1726	165.67	729	2068	1123	302.32	35	1618	749	265.95	48	1176
0.65	PVE	0.19	0.08	0.03	0.66	0.47	0.14	0.24	0.98	0.65	0.12	0.45	0.98
0.70	VE	0.52	0.03	0.45	0.59	0.77	0.03	0.71	0.82	0.88	0.02	0.77	0.94
0.70	INCp	1656	185.63	853	2011	962	263.44	187	1543	565	278.61	27	1156
0.70	PVE	0.22	0.09	0.06	0.60	0.55	0.12	0.28	0.91	0.73	0.13	0.46	0.99
0.75	VE	0.52	0.03	0.42	0.59	0.77	0.06	0.33	0.88	0.88	0.04	0.67	0.94
0.75	INCp	1515	186.77	733	1863	714	301.71	14	1397	421	211.65	17	887
0.75	PVE	0.29	0.09	0.13	0.66	0.67	0.14	0.35	0.99	0.80	0.10	0.58	0.99
0.80	VE	0.53	0.03	0.47	0.66	0.77	0.05	0.46	0.87	0.88	0.04	0.75	0.94
0.80	INCp	1426	251.22	231	1796	575	268.55	30	1090	286	187.46	15	766
0.80	PVE	0.33	0.12	0.16	0.89	0.73	0.13	0.49	0.99	0.87	0.09	0.64	0.99
0.85	VE	0.54	0.05	0.38	0.71	NA	NA	NA	0.87	NA	NA	-Inf	0.94
0.85	INCp	1277	243.05	172	1779	462	254.78	28	1049	191	138.41	7	716
0.85	PVE	0.40	0.11	0.17	0.92	0.78	0.12	0.51	0.99	0.91	0.06	0.66	1.00
0.90	VE	0.56	0.05	0.42	0.66	NA	NA	NA	0.91	NA	NA	NA	1.00
0.90	INCp	1099	311.55	59	1644	304	225.89	17	938	131	92.18	12	428
0.90	PVE	0.48	0.15	0.23	0.97	0.86	0.11	0.56	0.99	0.94	0.04	0.80	0.99
0.95	VE	0.53	0.19	-1.21	0.86	NA	NA	NA	0.93	NA	NA	NA	0.96
0.95	INCp	931	322.46	34	1445	201	137.24	7	592	106	94.48	5	431
0.95	PVE	0.56	0.15	0.32	0.98	0.91	0.06	0.72	1.00	0.95	0.04	0.80	1.00

Table 4: Table 2 (cont.): Summary statistics of VE, INCp and PVE over 100 simulations for each combination of vaccine covarage and efficacy; Two Strata

	Low Efficacy			Medium Efficacy			High Efficacy		
Parameter	Estimate	std. error	p-value	Estimate	std. error	p-value	Estimate	std. error	p-value
intercept	-0.286	0.0127	<0.0001	0.738	0.0396	<0.0001	1.29	0.0355	<0.0001
VCov	-1.79	0.0239	<0.0001	-5.24	0.0756	<0.0001	-8.09	0.0678	<0.0001

Table 5: effect estimates for negative binomial coverage only model; One Stratum

	Low Efficacy			Medium Efficacy			High Efficacy		
Parameter	Estimate	std. error	p-value	Estimate	std. error	p-value	Estimate	std. error	p-value
intercept	-0.241	0.032	<0.0001	1.03	0.067	<0.0001	1.49	0.067	<0.0001
VCov	-1.4	0.044	<0.0001	-4.17	0.093	<0.0001	-5.48	0.093	<0.0001

Table 6: effect estimates for negative binomial coverage only model; Two Strata

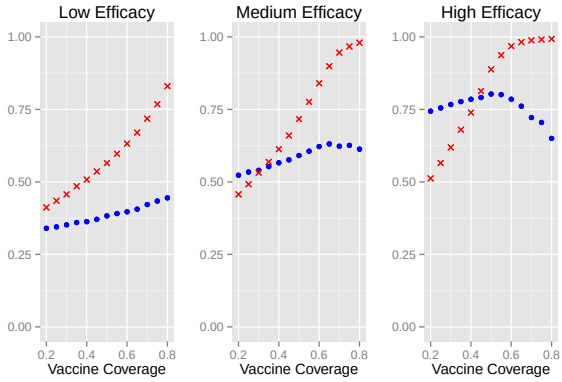

Figure 1: One Stratum: Vaccine effectiveness and population vaccine effectiveness by coverage; PVE is in red, VE in blue. Efficacy stratified

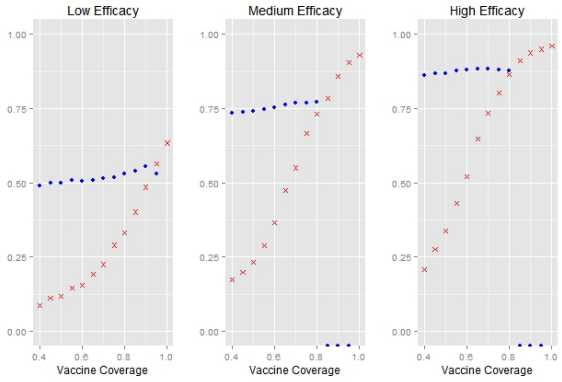

Figure 2: Two Strata: Vaccine effectiveness and population vaccine effectiveness by coverage; PVE is in red, VE in blue. Efficacy stratified

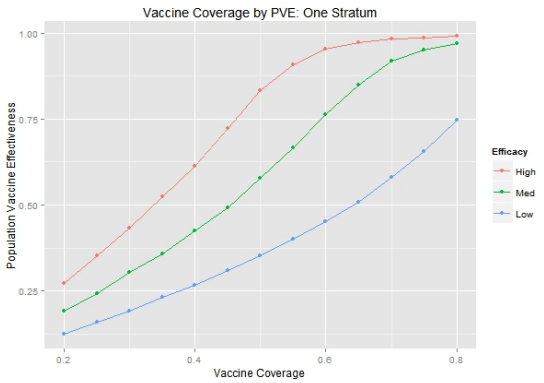

Figure 3: One Stratum: Population vaccine effectiveness by coverage; efficacy stratified

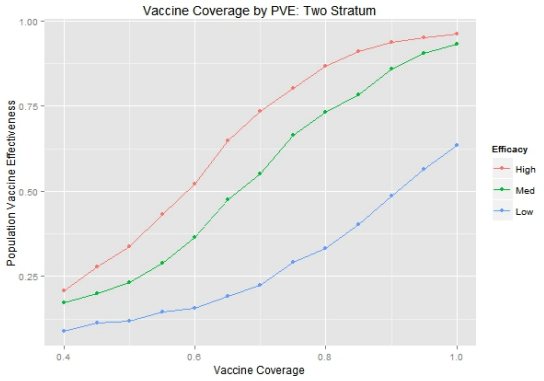

Figure 4: Two Stratum: Population vaccine effectiveness by coverage; efficacy stratified

E References

- Haber, M., Longini, I.M., and Halloran, M.E. (1991). Measures of the effects of vaccination in a randomly mixing population. International Journal of Epidemiology 20, 300-310.
- Halloran, M.E., Haber, M., Longini, I.M., and Struchiner, C.J. (1991). Direct and indirect effects in vaccine efficacy and effectiveness. American Journal of Epidemiology 133, 323-331.
- Haber, M., Watelet, L., and Halloran, M.E. (1995). On individual and population effectiveness of vaccination. International Journal of Epidemiology 24, 1249-1260.

- World Health Organization (14 August 2015). *WHO Field Guide for the evaluation of influenza vaccine effectiveness*. Retrieved from http://www.who.int/immunization/research/development/influenza$_f$*ield*$_g$*uide/en/*