Estimating vaccine efficacy using animal efficacy data

Ann Yellowlees*, Richard H. J. Perry

Quantics Biostatistics, West End House, 28 Drumsheugh Gardens, Edinburgh EH3 7RN, UK

Ann.yellowlees@quantics.co.uk

Abstract

Animal models are used to predict the effect of an intervention in humans. An example is the prediction of the efficacy of a vaccine when it is considered unethical or infeasible to challenge humans with the target disease to assess the effect of the vaccine on the disease in humans directly. In such cases, data from animal studies are used to develop models relating antibody level to protection probability in the animal, and then data from a study or studies in human subjects vaccinated with the proposed vaccine regimen are used in combination with the relevant animal models to predict protection in humans, and hence estimate vaccine efficacy.

We explain the statistical techniques required to provide an estimate of vaccine efficacy and its precision. We present simulated examples showing that precise estimation of the relationship between antibody levels and protection in animals, at levels likely to be induced in humans by the vaccine regimen, is key to precise estimation of the vaccine efficacy.

Because the confidence interval for the estimate of vaccine efficacy cannot be expressed in analytical form, but must be estimated from resampling, or bootstrapping, it is not possible to design studies with required power analytically. Therefore we propose that a simulation-based design of experiments approach using preliminary data is used to maximise the power of further studies and thus minimise the human and animal experimentation required.

Keywords: Animal rule, Vaccine efficacy, Bootstrap, Confidence interval, Simulation

1. Introduction

In order for a vaccine to be evaluated, its efficacy must be defined. Vaccine efficacy is defined as the percentage reduction in the incidence of a disease among people who have received a vaccine compared to the incidence in unvaccinated people. It quantifies how effective the vaccine could be given ideal circumstances.

The basic formula for vaccine efficacy (VE) is (Kohberger RC et al., 2008):

VE = [1- {probability of disease, vaccinated}/{ probability of disease, unvaccinated)] x 100% [1]

Ideally vaccine efficacy is measured in a randomised controlled trial, by giving one group of subjects a vaccine and comparing the incidence of disease in that group to another group of people who do not receive the vaccine. However, efficacy studies for some products may be unfeasible in human subjects because of statistical limitations (e.g., disease incidence is too low or variable) or because of feasibility or ethical considerations (e.g., life threatening disease or where alternative vaccines or therapeutics are available) (NIAID, 2010; FDA, 2014).

The FDA has published draft guidance (FDA, 2014) on 'Product development under the animal rule'. The Animal Rule states that FDA will rely on evidence from animal studies to provide substantial evidence of effectiveness only when all of four specific scientific criteria are met. For vaccines, it states in addition that: *Sponsors should develop an approach for bridging animal responses to humans by careful selection of appropriate immune markers.* This provides a framework for the evaluation of vaccine efficacy in the absence of human clinical efficacy data and is the subject of this paper.

An 'appropriate immune marker', otherwise known as a protective correlate, is one which is statistically related to a clinical endpoint, and is reasonably likely to predict the clinical endpoint. Once such a marker has been identified, the relationship between the marker and the clinical endpoint in animals is modelled using data from animal studies. The relationship can then be bridged to humans by the assumption that the marker is also predictive of the clinical endpoint in humans. The vaccine efficacy can then be estimated by combining the human data with the predictive animal model.

This approach has been tested and validated in pertussis vaccines (Kohberger RC et al., 2008), where both clinical data and serological data for a protective correlate were available. However in general there is a leap of faith in assuming that the same model relating serologic measures to a clinical endpoint in animals can be used to predict the same clinical endpoint in humans.

There are many scientific issues to be resolved in the selection of a protective correlate. In this article we restrict ourselves to the statistical aspects of the models used and the statistical techniques required to estimate vaccine efficacy, including the precision of such estimates, via protective correlates. We present a series of simulated examples to illustrate the methods and

provide insight into how appropriate designs and sample sizes can be estimated to achieve the required precision for estimates of vaccine efficacy.

2. Materials and methods

The first step in the process of estimating vaccine efficacy is the selection of an appropriate immune marker. This was discussed at a workshop hosted by the National Institute of Allergy and Infectious Diseases in 2009, entitled 'Utilization of serologic assays to support efficacy of vaccines in nonclinical and clinical trials: Meeting at the Crossroads' (NIAID, 2010). There, the term "correlate" was defined as a variable that is statistically related to a clinical endpoint, and a "protective correlate" was defined as a correlated variable that, based on additional evidence, is reasonably likely to be causally related to a clinical endpoint. (The term "surrogate" was avoided as its definition indicates that the measured variable (serologic data) can totally explain and replace a defined clinical endpoint. It is important to note that the animal rule does not require a surrogate (Fleming T, 2005).) We use this terminology, taking a protective correlate to be an appropriate immune marker for vaccine efficacy.

In the examples below, we assume that the serum level of an antibody, taken at a specific time (typically 14 days) after the last vaccination in a primary regimen, has been identified as a protective correlate. For simplicity we assume that unvaccinated animals and humans have 100% probability of disease. Therefore the expression for vaccine efficacy simplifies to:

$$VE = 1\text{- \{probability of disease, vaccinated\}} \qquad [2]$$

Figure 1 illustrates the steps in the estimation of vaccine efficacy.

Step 1 consists of an animal study designed to provide an estimate of the relationship between the antibody level in an animal, at a specific time point in relation to vaccination, and the probability of disease. The relationship between antibody level and disease - usually logistic regression - needs to be well characterised, so the animal study needs to have been designed to provide an accurate assessment of the relationship:

$$P = \frac{e^{a+b\ \log_{10} x}}{1+e^{a+b\ \log_{10} x}} \qquad [3]$$

where x is the antibody level, P is [1- probability of disease, given antibody level x], or the probability of protection (no disease), given antibody level x, and a and b are parameters estimated from the data. The parameters represent the horizontal location and the slope of the relationship respectively. The slope at the midpoint of the curve is $b/4$. The logistic curve in Figure 1, Step 1 illustrates the estimated relationship.

Step 2 is typically a human Phase 2 study including groups of subjects vaccinated according to the candidate vaccination regimens. This is designed to provide an estimate of the distribution of antibody levels, again at a specific time point in relation to vaccination, in a fully vaccinated

population. Typically the distribution is log normal – that is, the logarithm of the antibody level is normally distributed in the population

$$\log_{10}x \sim N(\mu, \sigma^2) \qquad [4]$$

The histogram in Figure 1, Step 2 shows a sampled distribution of $\log_{10}x$.

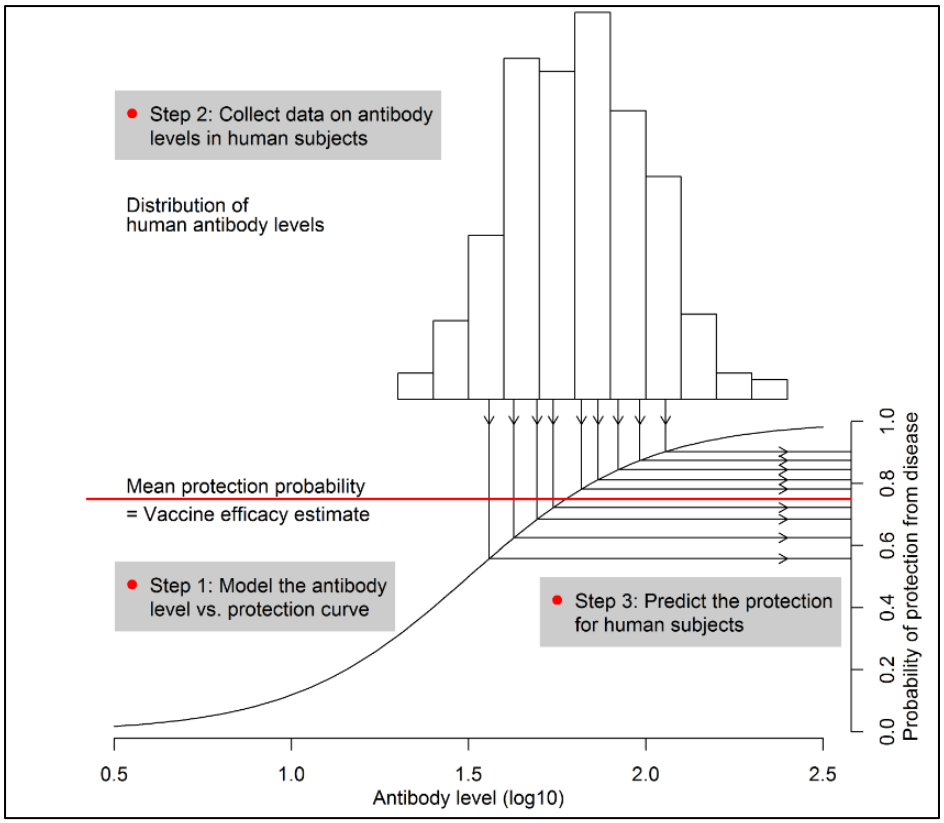

Figure 1: Steps in estimation of vaccine efficacy

Step 3 is to estimate the vaccine efficacy, and its precision, from the animal data obtained in Step 1 and the distribution estimated in Step 2.

2.1 Bootstrapping

The vaccine efficacy is estimated using Bayes' Theorem (Bayes T. and Price R., 1763) to combine the probability of protection (given the antibody level is x) with the likelihood of the antibody level being x. This is estimated from the data as follows:

1. For each human subject, calculate the estimated probability of protection given their antibody level x, from the fitted animal model.
2. Calculate the mean of these probabilities over all human subjects to provide the estimate of VE.

It is not possible to express the confidence limits for the vaccine efficacy estimate analytically (Kohberger RC et al., 2008). We use a two-stage bootstrap method as recommended by Kohberger, with the following two steps:

A. **Bootstrap of the logistic regression model**: Sample (with replacement) observations of antibody level and survival from the animal data. Estimate the relationship between antibody level and survival based on the sampled data.
B. **Bootstrap of the vaccine efficacy calculation:** Sample (with replacement) observations of antibody level from the human data. Apply the model from A to this data and estimate vaccine efficacy.

Steps A and B are repeated together many times (1,000 in our simulations) in order to estimate the distribution of vaccine efficacy. The distribution of estimates provides a percentile confidence interval for vaccine efficacy.

2.2 Simulation

In order to illustrate this procedure, and at the same time to explore the numbers of (a) animals and (b) human subjects needed to achieve adequate precision for the estimate of vaccine efficacy, we have carried out a simulation based on our experience of a particular model.

We assume that the underlying model for the probability of disease prevention is a logistic dose-response model defined as above, with a = -6 and b = 4. Therefore, for an animal or a human with antibody level of x, the probability of protection is $p_x = \exp(-6 + 4\log_{10}x)/(1 + \exp(-6 + 4\log_{10}x))$.

We also assume that the antibody level (x) achieved in vaccinated human subjects is normally distributed with mean 2.4 and variance 0.4 (on the log scale). In combination with the probability model above, this implies a vaccine efficacy of 88% (calculated by numerical integration; details available from the authors).

We simulate a pair of studies as follows:

1. Animal study with a total of N_a animals in 5 equally sized groups, mimicking an animal study with 5 dose groups. For each animal in group g (g = 1 to 5):
    a. Simulate an antibody level x from a log normal distribution: $\log_{10}x \sim N(\mu_{ag}, \sigma^2_a)$.
    b. Calculate the probability of protection given antibody level x:

$$P(x) = \exp(-6 + 4\log_{10}x)/(1 + \exp(-6 + 4\log_{10}x)) \qquad [5]$$

    c. Simulate a binary indicator, $I(x)$, for protection with $P(x)$ as the probability of protection.
2. Using the N_a simulated pairs $(x, I(x))$, estimate a logistic dose response relationship:

$$P_{hat}(x) = \exp(a + b\log_{10}x)/(1 + \exp(a + b\log_{10}x)) \qquad [6]$$

3. Human study with N_h human subjects. For each human subject:
    a. Simulate an antibody level y from the log normal distribution: $\log_{10} y \sim N(2.4, 0.4)$
    b. Calculate the probability of protection for each human subject using the estimated dose-response relationship:

$$P_{hat}(y) = \exp(a + b\log_{10} y)/(1 + \exp(a + b\log_{10} y)) \quad [7]$$

Estimate the vaccine efficacy from the pair of studies as the mean value of these probabilities of protection across the N_h human subjects.

The variability of the estimate of VE for the simulated study pair is estimated using a two stage bootstrap method as described above.

We simulated data for study pairs for the 12 sets of parameter values (scenarios) shown in Table 1. The standard deviation of the $\log_{10}$ antibody levels for the animals was set to $\sigma_a = 0.1$ in each case.

Table 1: Scenarios and simulation parameters

Scenario	N_a	Percentile of protection for middle animal dose group	Geometric mean antibody level in each of 5 animal dose groups = 10^{μ}	N_h
1	50	50	8, 16, 32, 63, 126	50
2		80	18, 35, 70, 140, 281	
3		95	43, 876, 172, 344, 689	
4	100	50	8, 16, 32, 63, 126	50
5		80	18, 35, 70, 140, 281	
6		95	43, 876, 172, 344, 689	
7	50	50	8, 16, 32, 63, 126	100
8		80	18, 35, 70, 140, 281	
9		95	43, 876, 172, 344, 689	
10	100	50	8, 16, 32, 63, 126	100
11		80	18, 35, 70, 140, 281	
12		95	43, 876, 172, 344, 689	

For each set of parameter values, 1,000 study pairs were simulated.

3. Results

First we present the results of a single simulation from Scenario 1, to illustrate how the VE and its 95% confidence interval are calculated. We then expand this to explore, for the same scenario, how variable these results could be across repeated pairs of studies by looking at all 1,000 simulations.

Finally we compare the results across the 12 scenarios.

3.1 Calculation of the estimate of VE and its confidence interval

Figure 2 shows the data for a simulation with 50 animals, with mean log antibody level for the middle group providing 50% protection, and 50 human subjects.

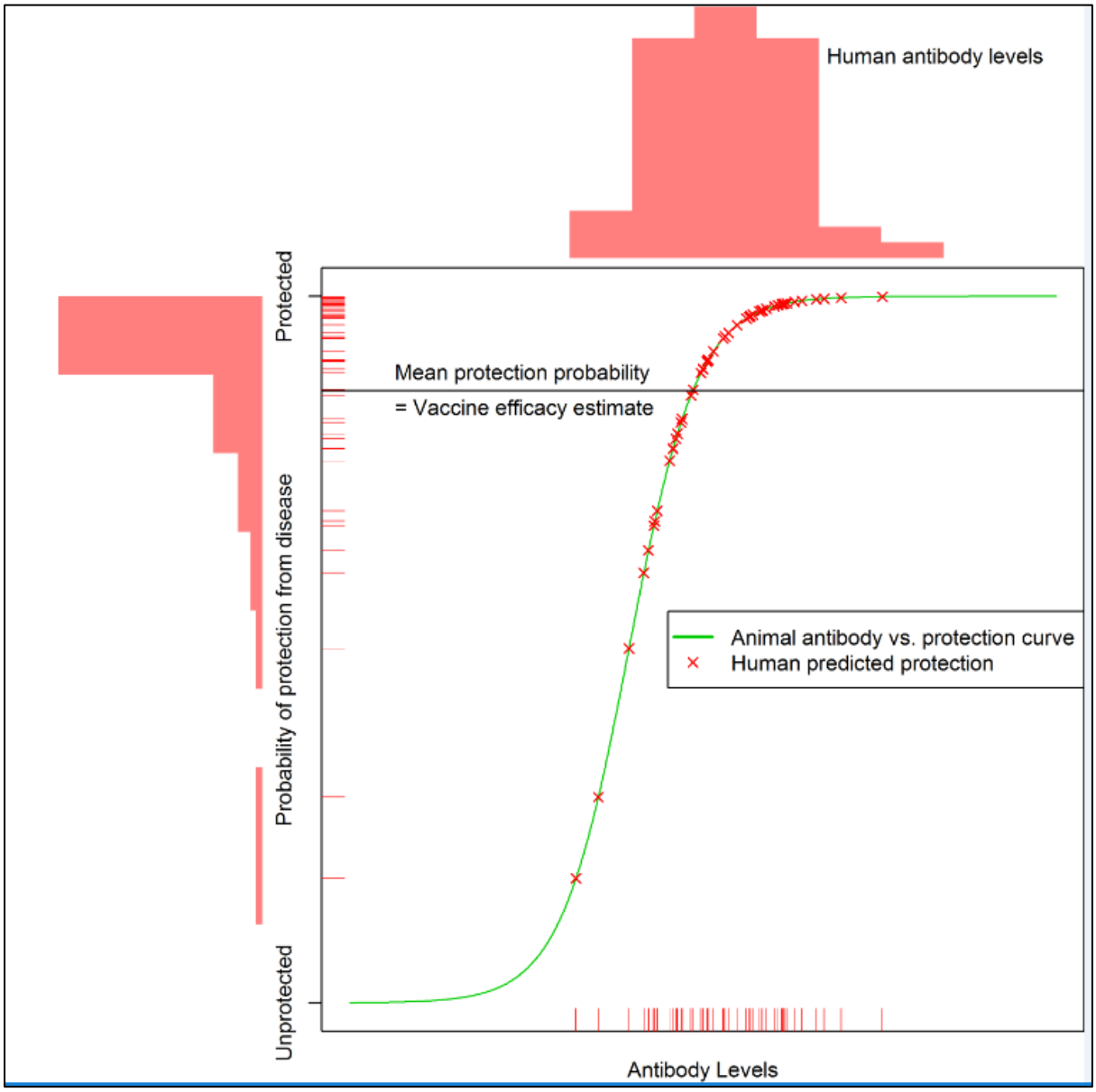

Figure 2: Relating animal antibody vs. protection curve to human disease protection via human antibody levels. Logistic fit and VE estimate: N_a = 50, N_h = 50, middle animal dose group at 50% protection

This particular example provided a vaccine efficacy estimate (and 95% confidence interval) of 87% (78%, 94%), compared with the true value (from the simulation parameters) of 88%.

3.2 Simulation results for Scenario 1: $N_a = 50$, $N_h = 50$, middle animal dose group at 50% protection

The 1,000 simulations provide a range of estimates. Figure 3 shows the 1,000 values of VE estimated from these simulations in a histogram (top of figure), the 1,000 values of the width of the 95% confidence interval in a histogram (left edge), and the relationship between these in a scatterplot. (The confidence interval is expressed as the absolute % points difference between the lower and upper 95% VE confidence limits.)

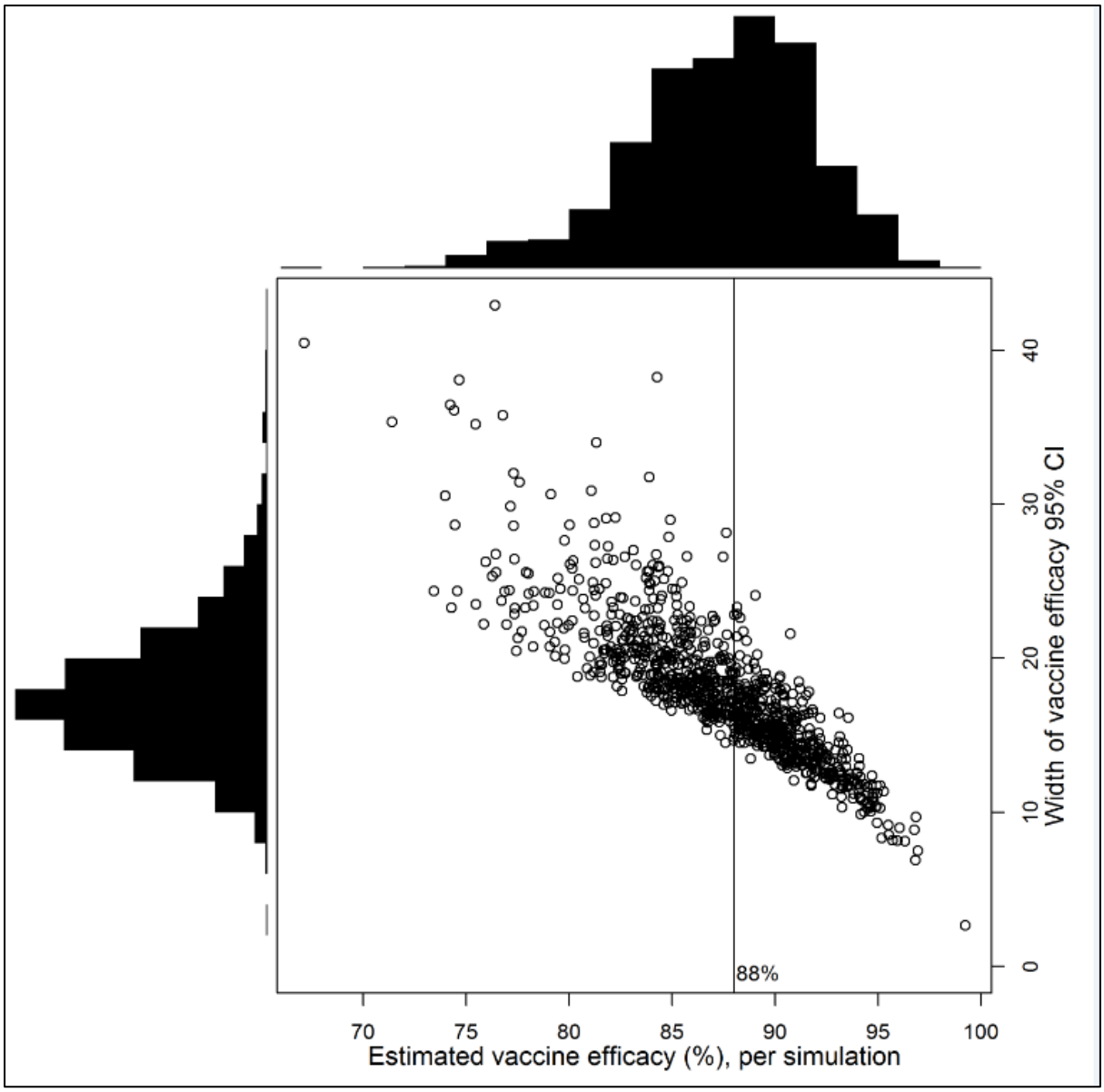

Figure 3: VE estimates and confidence intervals: $N_a = 50$, $N_h = 50$, middle animal dose group at 50% protection

The figure shows that the vaccine efficacy estimates are centred around the true value of 88%; 90% of the estimates lie in the interal (79%, 94%). The width of the confidence interval is centred around 18%, with 90% of the intervals no wider than 23%. In addition, the higher the estimated VE, the narrower the interval, as would be expected.

3.3 Comparison of study designs across scenarios

Figure 4 shows the values of VE estimates for each of the 12 scenarios. There is a suggestion that the scenario with the mean animal antibody level at the 80th percentile results in a median value closer to the true VE of 88%, however the differences between the different scenarios are relatively small. As expected, the simulations with smaller numbers of animals and humans are more variable.

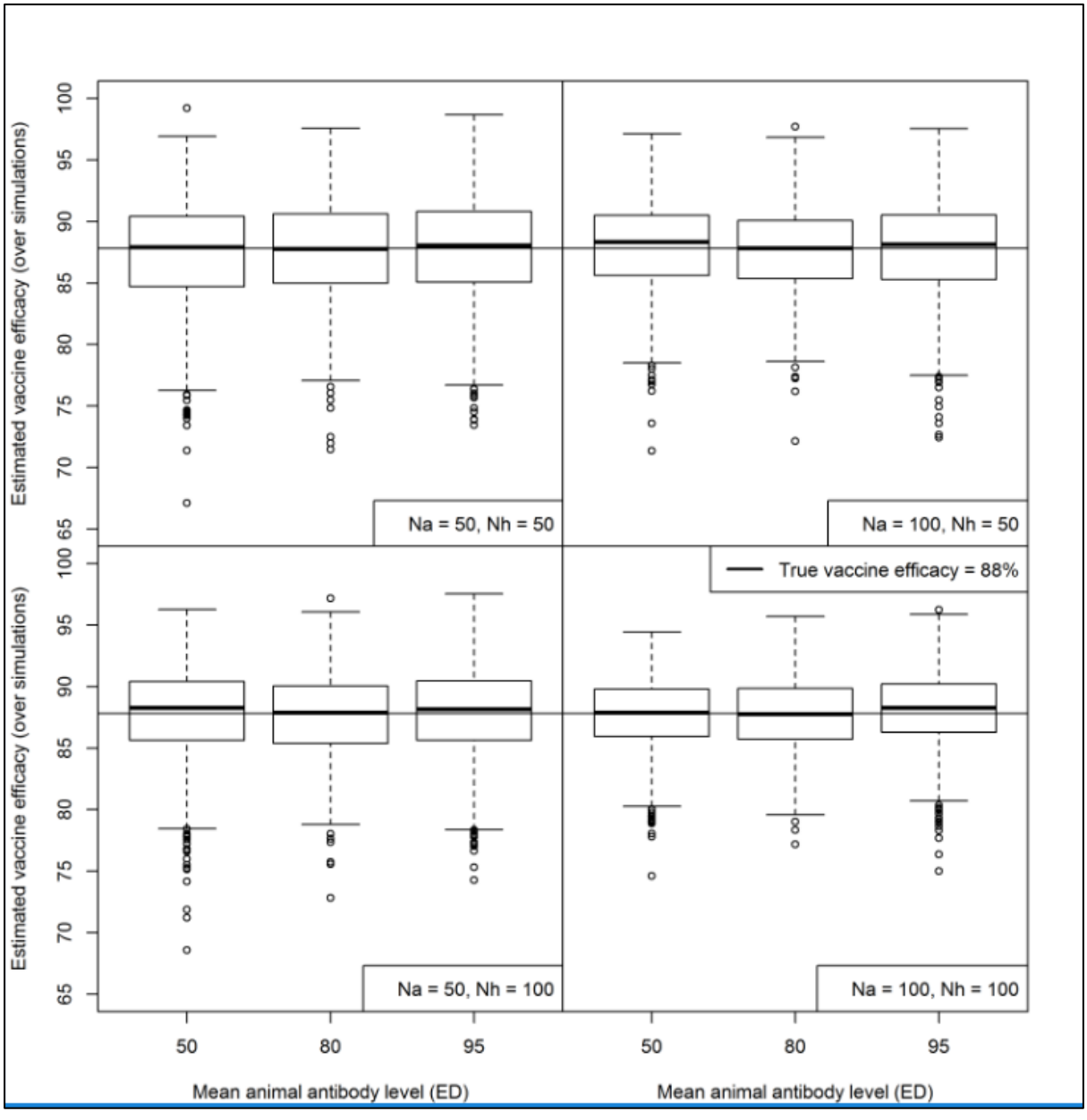

Figure 4: Box plots for VE estimates for each of the 12 scenarios

Figure 5 shows the ranges of the widths of the 95% confidence intervals for the VE estimates for each of the 12 scenarios.

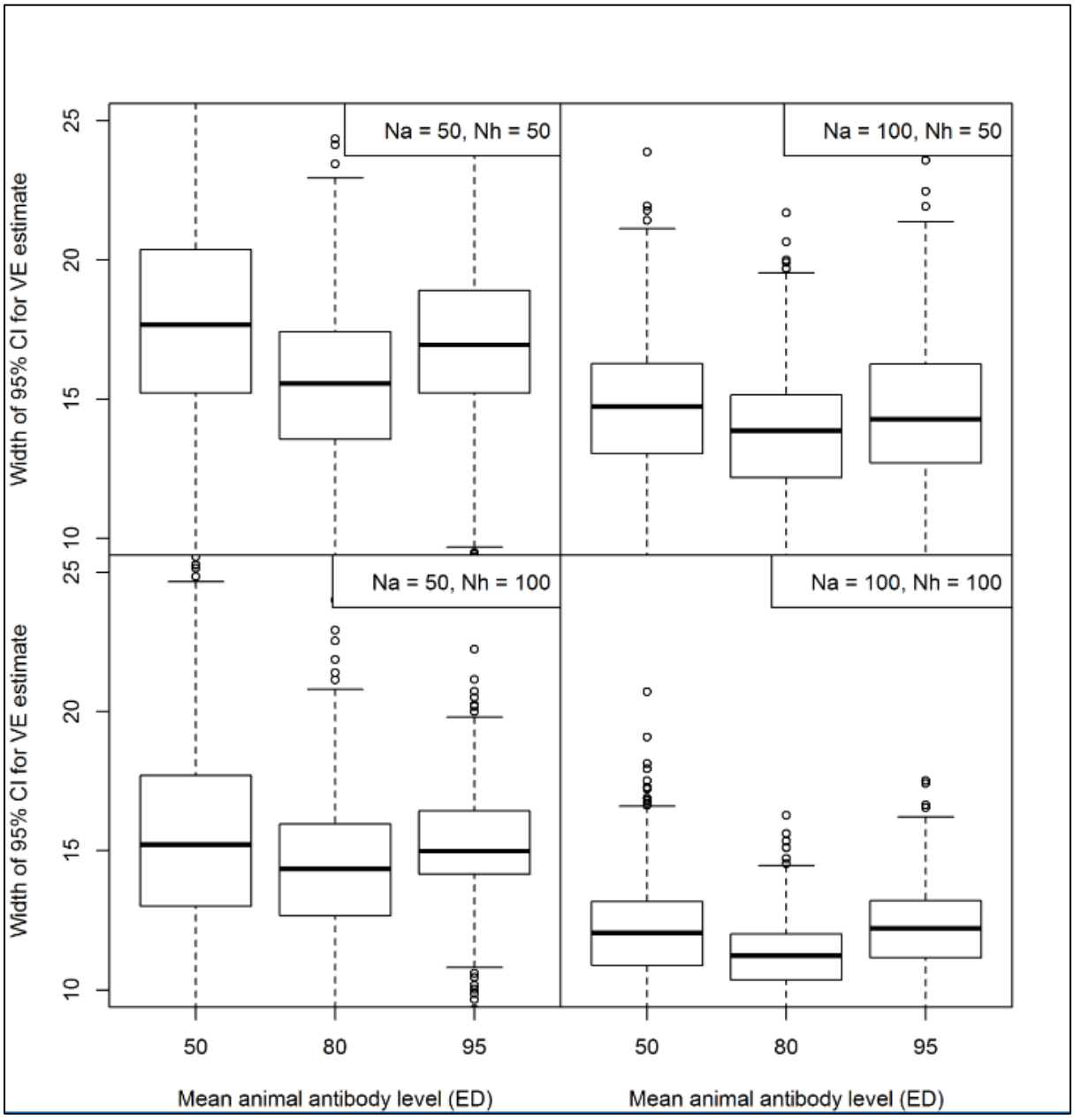

Figure 5: Box plots for VE estimates for each of the 12 scenarios

In each case, the confidence intervals are narrower when the mid-dose group for the animals is at the 80th percentile than either at the 50th percentile or the 95th percentile. This is due to a combination of the facts that (a) the relationship needs to be well estimated for the values seen for the human subjects, towards the higher end, and (b) for good estimation of the logistic curve, values

on both sides of the midpoint are required. Thus, when the animals are centred aound the 50th percentile, there is less information about the upper end of the curve than for the other two dose groupings, but when the animals are centred too high, the precision of the estimation of the curve is also compromised.

3.4 Optimising animal and human subject numbers

During early vaccine development the initial experiments to create the animal antibody Vs protection curve may not result in the mean animal antibody level being around the optimal value giving 80% protection seen in this simulation. To optimise the estimate of the VE, is it then better to conduct further animal experiments, improving the primary model fit, or to add to the information on human antibody levels, to improve the final estimate of the protection provided to humans?

To illustrate how this question can be answered we take the example of a preliminary study where data for 50 animals and 50 humans has been obtained. Figure 6 highlights the structure of data currently collected in the study and some options for further data collection. By using the current data and adding simulated data, it is possible to compare each of the options with regard to their effectiveness at increasing the precision of the vaccine efficacy.

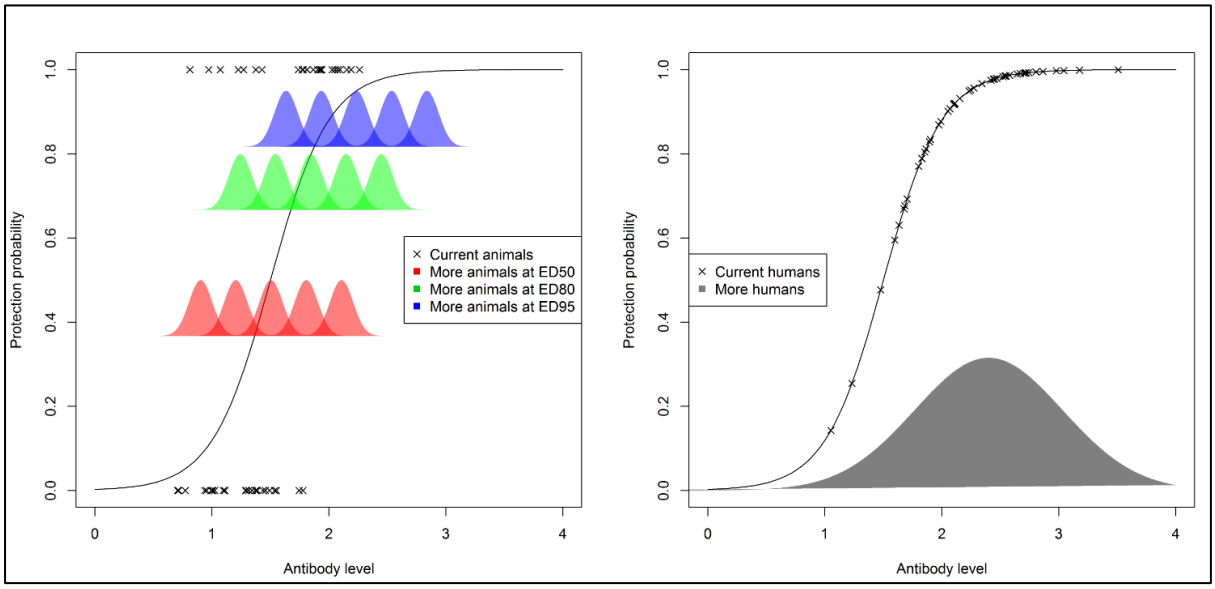

Figure 6: Illustration of available choices from example preliminary study

For this particular case, if all options were of equal cost, then Figure 7 shows that adding animals at the ED95 level would be best. If changing the antibody level of the animals were not possible, then adding 50 humans would be the best option instead of adding 50 animals at the level of ED50.

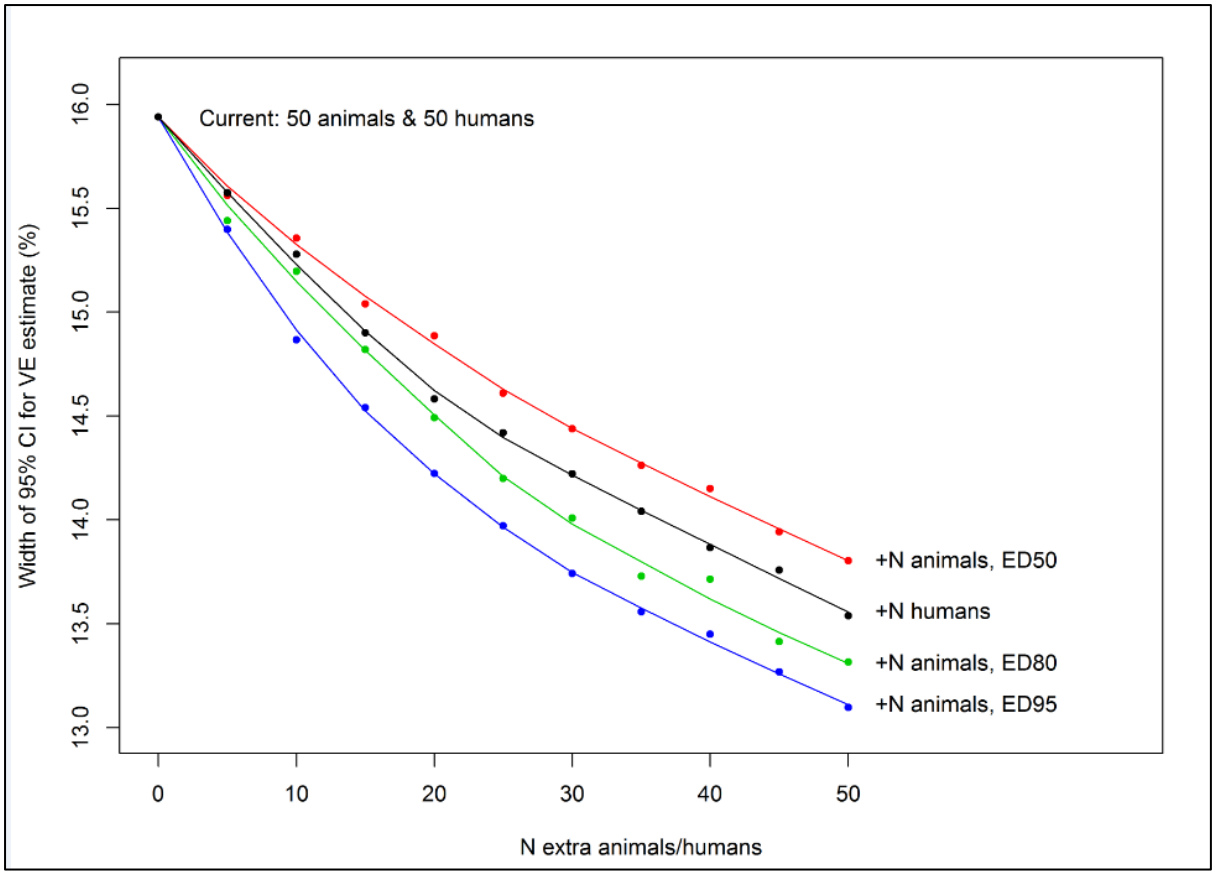

Figure 7: Comparison of choices for example preliminary study

4. Discussion

Estimating the efficacy of a vaccine by applying the protection achieved at a given antibody level in animals to a population of human subjects is an acceptable approach in certain circumstances, as defined by the FDA's Animal Rule. Data from animal studies are used to develop models relating antibody level to protection probability, and then data from a study or studies in human subjects vaccinated with the proposed vaccine regimen are used in combination with the relevant animal models to predict protection in humans, and hence estimate vaccine efficacy.

As both the antibody level induced by the vaccine in animals at the doses tested, the antibody level-protection relationship and the antibody levels seen in human subjects are unknown *a priori*, it is usually the case that standard designs are used for these studies, and there is no explicit *a priori* control over the precision of the VE estimate.

However, this paper shows that simulation can be used after initial data is obtained to provide some guidance on how to design further experiments. Such further experiments may be required so the resulting estimates of efficacy are adequately precise to support conclusions about the vaccine and the proposed regimen, or are adequately powered to detect differences between candidate

regimens. Careful design of experiments can save time, money and animals in line with 3R recommendations.

References

Bayes T. and Price R. (1763). An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London, 53*(0), 370-418. doi:10.1098/rstl.1763.0053

FDA. (2014). Guidance for Industry: Product Development Under the Animal Rule.

Fleming T. (2005). Surrogate Endpoints and FDA's Accelerated Approval Process Health Affairs. *Health Affairs, 24*, 67-78. doi:10.1377/hlthaff.24.1.67

Kohberger RC et al. (2008). Prediction of pertussis vaccine efficacy using correlates of protection model. *Vaccine, 26*, 3516-3521. doi:10.1016/j.vaccine.2008.04.016

NIAID. (2010). Utilization of serologic assays to support efficacy of vaccines in nonclinical and clinical trials: Meeting at the crossroads. Conference Report. *Vaccine, 28*, 4539-4547. doi:10.1016/j.vaccine.2010.04.094