Estimating Vaccine Efficacy From Data With Recruitments

Claude Lefèvre

Université Libre de Bruxelles
Institut de Statistique et de Recherche Opérationnelle, C.P. 210,
B-1050 Bruxelles, Belgique
(e-mail: clefevre@ulb.ac.be)

Abstract. Vaccine induced protection against infection is often random. A concept of protective vaccine efficacy, depending on the mean relative susceptibility of vaccinated individuals, is considered for a large vaccine trial in which participants are recruited over a period of time. Bounds are derived that make statistical inference possible under weak assumptions about the transmission process, irrespectively of the type of protection induced by the vaccine.

This is a joint work with Niels Becker (The Australian National University, Canberra, Australia) and Sergey Utev (University of Nottingham, Nottingham, United Kingdom).
Keywords: Protective vaccine efficacy, Bounds, Estimation.

1 Introduction

A standard concept of vaccine efficacy is defined as

$$VE_{\mathrm{P}} = 1 - \frac{c_{\mathrm{V}}}{c_{\mathrm{U}}}, \tag{1}$$

c_{U} and c_{V} representing the proportions of cases among unvaccinated and vaccinated individuals, respectively.

As pointed out by, e.g., [Smith and Fine, 1984], vaccine efficacy depends on the type of the protection induced by a vaccine. Two types of vaccine response are usually discussed. A first case is when a vaccinee receives either complete protection or no protection against infection (i.e. the vaccine confers a complete/no (CN in short) protection). A second case is when every vaccinee receives exactly the same partial protection (i.e. the vaccine confers a partial/uniform (PU in short) protection).

Recently, [Becker and Utev, 2002] introduced a class of vaccine responses that includes CN and PU protection as particular cases. Shortly, if at time t, the force of infection acting on an unvaccinated susceptible individual is $\lambda(t)$, then the force of infection acting on a vaccinated susceptible individual is reduced to $A\lambda(t)$, A denoting a discrete random variable with probability distribution

$$\Pr(A = a_j) = p_j, \quad j = 1, \ldots, r, \tag{2}$$

where the possible values a_j are in $[0, 1]$. These authors proposed for this class a concept of protective vaccine efficacy given by

$$VE_{\mathrm{P}} = 1 - \mathrm{E}A. \tag{3}$$

When a vaccine induces CN protection, (3) yields $VE_{\mathrm{P}} = p_1$, i.e. the probability that the vaccinee is completely protected. For PU protection, (3) becomes $VE_{\mathrm{P}} = 1 - a_1$, i.e. the per-contact reduction in the probability of disease transmission.

Estimating VE_{P} from data on the eventual numbers of vaccinated and unvaccinated cases requires to specify assumptions about the type of vaccine response. [Becker and Utev, 2002] showed, however, that for a standard model of epidemics in a large uniformly mixing community, the inequality

$$1 - \frac{c_{\mathrm{V}}}{c_{\mathrm{U}}} \;\le\; VE_{\mathrm{P}} \;\le\; 1 - \frac{\ln(1 - c_{\mathrm{V}})}{\ln(1 - c_{\mathrm{U}})}, \tag{4}$$

holds independently of all types of protection induced. These bounds are estimable from data on the eventual numbers of vaccinated and unvaccinated cases, and seem to be close enough to be used for inference about VE_{P}.

Our purpose in the present paper is to show how to extend the analysis made in [Becker and Utev, 2002] to a more general model (i) based on less restrictive assumptions about the force of infection and (ii) allowing for recruitments of participants over time (which is useful for large field trials and/or for rather rare diseases). As a key result, we will obtain lower and upper bounds that are analogous to (but different fom) those given in (4). Furthermore, we will then prove that if the vaccination coverage remains constant over time, the lower bound can provide a good estimate of the vaccine efficacy.

This is a joint work with Niels Becker (The Australian National University, Canberra, Australia) and Sergey Utev (University of Nottingham, Nottingham, United Kingdom).

2 An epidemic model with vaccination

Denote by A the relative susceptibility of a vaccinated individual, such as defined by (2). Vaccinated individuals for which $A = a_j$ are said to be of type j, and unvaccinated individuals are said to be of type U. In practice, only unvaccinated (U) and vaccinated (V) individuals can be distinguished.

The population sizes are described by a deterministic model (valid for large trials). Let $N_{\mathrm{U}}(t)$ be the number of unvaccinated trial participants recruited by time t, and let $N_{\mathrm{V}}(t)$ be the number of vaccinated trial members recruited by time t. Initially, there are n individuals of whom a fraction u are unvaccinated and a fraction v are vaccinated ($u + v = 1$). In Section 3, the proportion of vaccinated trial participants will be assumed to be always

as large as its initial level v. In particular, the vaccination coverage can then remain constant.

At time 0, the numbers of susceptible trial participants are given by

$$S_{\mathrm{U}}(0) = nu, \quad \text{and} \quad S_j(0) = nvp_j, \quad j = 1, \ldots, r.$$

Let $\lambda(t)$ be the force of infection on an unvaccinated individual. Then, the number of unvaccinated trial members who are susceptible to infection at time t is ruled by the differential equation

$$dS_{\mathrm{U}}(t) = -\lambda(t) S_{\mathrm{U}}(t)\, dt + dN_{\mathrm{U}}(t).$$

For the vaccinated members, these numbers are governed by the differential equations

$$dS_j(t) = -a_j \lambda(t) S_j(t)\, dt + p_j\, dN_{\mathrm{V}}(t), \quad j = 1, \ldots, r.$$

Putting $\Lambda(t) = \int_0^t \lambda(x)\, dx$, the solutions to these equations are respectively given by

$$S_{\mathrm{U}}(t) = \int_{0-}^{t} \exp[\Lambda(x) - \Lambda(t)]\, dN_{\mathrm{U}}(x), \tag{5}$$

and

$$S_j(t) = p_j \int_{0-}^{t} \exp[a_j \Lambda(x) - a_j \Lambda(t)]\, dN_{\mathrm{V}}(x), \quad j = 1, \ldots, r. \tag{6}$$

Let us fix any finite time interval $[0, T]$. The number of unvaccinated trial participants who are cases by time T is

$$C_{\mathrm{U}} = N_{\mathrm{U}}(T) - S_{\mathrm{U}}(T),$$

and the number of vaccinated cases by time T is

$$C_{\mathrm{V}} = N_{\mathrm{V}}(T) - \sum_{j=1}^{r} S_j(T).$$

3 An estimator for the vaccine efficacy

As a first step, we begin by showing how (4) can be generalized to the present framework.

Proposition 1 *Provided that the proportion of vaccinated trial participants remains in the course of time as large as its initial level v, then*

$$1 - \frac{u}{v}\frac{C_{\mathrm{V}}}{C_{\mathrm{U}}} \;\leq\; VE_{\mathrm{P}} \;\leq\; 1 - \frac{\ln[1 - C_{\mathrm{V}}/N_{\mathrm{V}}(T)]}{\ln[1 - C_{\mathrm{U}}/N_{\mathrm{U}}(0)]}. \tag{7}$$

Moreover, the lower bound is attained when the vaccine induces CN protection and the vaccine trial has a non-varying vaccination coverage.

In the proof, a central point is a simple inequality for the expectation of a concave function of a random variable (see, e.g., [Becker and Utev, 2002]): if g is a continuous concave function defined on a finite interval $[c_1, c_2]$, then for any random variable A taking values in $[c_1, c_2]$,

$$g(c_1)\frac{c_2 - \mathrm{E}A}{c_2 - c_1} \;+\; g(c_2)\frac{\mathrm{E}A - c_1}{c_2 - c_1} \;\leq\; \mathrm{E}\, g(A) \;\leq\; g(\mathrm{E}A).$$

Now, let us give some comments on this result. We see that (7) reduces to (4) when recruitment occurs only at time $t = 0$. We also observe that the bounds of (4) still apply when recruitment occurs after time 0, but they ignore data on individuals recruited after time $t = 0$. Obviously, (7) uses data on individuals recruited after time 0, but the upper bound does so only through C_U, C_V and $N_V(T)$. Finally, we indicate that the upper bound in (7) cannot be attained with recruitment after time 0, but it is attained when the vaccine induces PU protection and the only recruitment is at time 0.

As a second step, we are going to derive an approximate estimator for the vaccine efficacy. More precisely, let us assume that all vaccine trial participants are recruited at k different instants during $[0, T]$. Initially, in each group i, $i \in \{1, \ldots, k\}$, there are n_i participants, and an identical vaccination coverage v is applied to each group. In group i, an unvaccinated individual escapes the disease with probability $\pi_i = \exp(-\Lambda_i)$, Λ_i denoting a cumulative force of infection upon this group until time T. A vaccinated individual in group i escapes the disease with probability $\mathrm{E}[(\pi_i)^A]$ where the random variable A has a distribution given by (2).

It is well-known (see, e.g., [Smith and Fine, 1984]) that without recruitment (i.e. when $\pi_i = \pi$), and if the vaccine induces CN protection, the measure (1) constitutes a maximum likelihood estimator of the vaccine efficacy VE_{P}. Hereafter, we will consider the cases, rather frequent in reality, where the different cumulative forces of infection Λ_i are all relatively small. We will then show that the lower bound, $1 - uC_V/vC_U$, derived in (7) provides a good estimator for VE_{P}.

Proposition 2 *Under the condition that* $\max_i(1 - \pi_i) \downarrow 0$*, then*

$$\widehat{VE_{\mathrm{P}}} = 1 - \frac{u}{v}\frac{C_{\mathrm{V}}}{C_{\mathrm{U}}} \tag{8}$$

is asymptotically equivalent to a maximum likelihood estimator of VE_{P}.

In the proof, the starting point is an expression for the global likelihood function L as a function of the unknown parameters $\{\pi_i,\ i = 1, \ldots, k\}$, $\{p_j,\ j = 1, \ldots, r\}$ and $\{a_j,\ j = 1, \ldots, r\}$. To construct L, we will have to introduce the final number of cases among vaccinated and unvaccinated participants in each group.

It is important to underline, however, that the only data needed for this estimator are the final numbers of cases observed at time T.

An asymptotic distribution as $n \to \infty$ can also be derived by using standard statistical arguments. First, a central limit theorem allows us to show that the lower bound type estimator $\widehat{\mathrm{E}A} = uC_V/vC_U$ is approximately normal. Then, using inequalities between integrals of special functions of exponential type, we are able to prove that the asymptotic mean of $\widehat{\mathrm{E}A}$, denoted by a, is given by

$$a \mathrm{sim} \mathrm{E}A + \alpha \quad \text{with} \quad 0 \leq \alpha \leq \varepsilon/8(1-\varepsilon)^2, \tag{9}$$

where $\varepsilon = 1 - \exp[-\Lambda(T)]$ is small by the assumption made before. The variance can also be calculated in a similar way.

References

[Becker and Utev, 2002]N.G. Becker and S. Utev. Protective vaccine efficacy when vaccine response is random. *Biometrical Journal*, pages 29–42, 2002.

[Smith and Fine, 1984]P.G. Smith and P.E.M. Fine. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *International Journal of Epidemiology*, pages 87–93, 1984.