Clustering vaccine coverage and estimating the indirect benefits of immunisation - an introduction to methodology

Xiang Li[1], Katy A. M. Gaythorpe[1]

Keywords: Vaccine coverage clustering, Effective vaccine coverage, Indirect vaccine benefit

1. Introduction

Vaccines are one of the most cost-effective health interventions. They have been responsible for substantial reductions in infections and mortality from vaccine preventable diseases (VPDs). In addition to direct protection provided to effectively vaccinated population, an attractive attribute of vaccines is the indirect protection (herd immunity) that un-vaccinated people and un-effectively vaccinated people benefit from. The Vaccine Impact Modelling Consortium (VIMC) aims to deliver an effective, transparent and sustainable approach to generating vaccine impact estimates. The indirect effects of vaccination are included in VIMC dynamic models. However, the proportion of burden averted due to these indirect effects is not often quantified. Understanding the indirect benefits are important for assessing the performance of vaccination programmes and targeting new strategies. This report proposes to establish a useful methodology to facilitate this quantification.

2. Method

The methodology consists of two steps on a cohort by cohort basis. Firstly, it clusters a birth cohort of interest into a group A (effectively vaccinated individuals) and a group B (individuals ineffectively vaccinated or unvaccinated). A core assumption around this

clustering is that group A contains all individuals whose immune system responds to vaccine and builds up protection against certain disease. The second step of the methodology differentiates direct / in-direct vaccine benefit from total vaccine impact. For group A and B, we estimate their probability of survival in a no-vaccination and a with-vaccination scenario. By comparing the group survival in different scenarios, the impact of vaccine conveys on group A is defined as direct vaccine benefit; and any additional vaccine impact due to reduced transmission by immunisation is attributed to group B as indirect vaccine benefit. In addition, the vaccine coverage clustering allows us to assess each immunisation programmes' contribution to the proportion of effectively vaccinated population. This delivers a tool for attributing vaccine impact to specific immunisation activities. This tool will be added into this report at a later stage.

2.1. Vaccine coverage clustering

For birth cohort k at age a, assume they (will) have benefited from $i{+}1$, $i \in \{0, 1, \dots\}$, vaccination activities. Of interest is to cluster the cohort into zero-dose, single-dose and multiple-dose populations. This is approached by step-wise joint Bernoulli distributions.

Denote random variables $V_{(i)}$, $M_{(i)}$ and $Z_{(i+1)}$ from Bernoulli distributions as

$$V_{(i)} = \begin{cases} 1, & \text{if vaccinated in the first } i \text{ doses of vaccination.} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

$$M_{(i)} = \begin{cases} 1, & \text{if vaccinated with multiple doses in the first } i \text{ vaccination activities.} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

$$Z_{(i+1)} = \begin{cases} 1, & \text{if vaccinated in the } (i+1)\text{-th doses of vaccination.} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The probability of each variable taking a value of 1 are $v_{(i)}$, $m_{(i)}$ and $z_{(i+1)}$, i.e. the vaccine coverage, proportion of multiple-dose population and vaccination activity

coverage, respectively. The probability mass function of each random variable can be defined as

$$f(n, v_{(i)}) = v_{(i)}^{n}(1 - v_{(i)})^{1-n}, \tag{4}$$

$$f(n, m_{(i)}) = m_{(i)}^{n}(1 - m_{(i)})^{1-n}, \tag{5}$$

$$f(n, z_{(i+1)}) = z_{(i+1)}^{n}(1 - z_{(i+1)})^{1-n}. \tag{6}$$

Denote the joint distribution of $V_{(i)}$ and $Z_{(i+1)}$ as $F(V_{(i)}, Z_{(i+1)}, \theta_1)$, and the joint distribution of $M_{(i)}$ and $Z_{(i+1)}$ as $F(M_{(i)}, Z_{(i+1)}, \theta_2)$, where θ_1 and θ_2 are the corresponding correlation coefficients. The joint distributions allow estimating vaccine coverage as

$$\begin{aligned} v_{(i+1)} &= Pr(V_{(i+1)} = 1) \\ &= 1 - Pr(V_{(i)} = 0, Z_{(i+1)} = 0), \end{aligned} \tag{7}$$

and multiple-dose proportion as

$$\begin{aligned} m_{(i+1)} &= Pr(M_{(i+1)} = 1) \\ &= Pr(M_{(i)} = 1) + Pr(M_{(i)} = 0 \mid V_{(i)} = 1, Z_{(i+1)} = 1) \\ &= Pr(M_{(i)} = 1) + Pr(V_{(i)} = 1, Z_{(i+1)} = 1) - Pr(M_{(i)} = 1, Z_{(i+1)} = 1) \end{aligned} \tag{8}$$

Hence, single-dose proportion is estimated as

$$s_{(i+1)} = v_{(i+1)} - m_{(i+1)}. \tag{9}$$

2.2. Effective vaccine coverage

Once the proportions of single-dose and multiple dose population are estimated, effective vaccine coverage (e), i.e. the proportion of effectively vaccinated individuals (or the size of group A) can be estimated taking into account vaccine efficacy. In practice, vaccine efficacy may be defined differently. We consider two definitions of vaccine efficacy that VIMC models use - vaccine efficacy as the probability of effective protection (def.1), and

vaccine efficacy as reduction in infection (def.2). We estimate effective vaccine coverage for these two definitions as

$$e = \begin{cases} \lambda_s s + \lambda_m m, & \text{def.1} \\ (1-\gamma(1-\lambda_s))s + (1-\gamma(1-\lambda_m))m, & \text{def.2} \end{cases} \tag{10}$$

where λ represents vaccine efficacy, and γ denotes probability of infection without vaccination or force of infection (FoI). While FoI is metioned here, it is not the focus in this methodology. We use no-vaccination scenario disease burden to approach FoI.

2.3. Vaccine impact decomposition

Denote A^0 and A^1 as survival or non-infection rate of group A in the no-vaccination and with-vaccination scenarios, respectively. Assuming independence between mortality from the focal disease and other causes, i.e. vaccination for a particular disease does not change the probability of mortality from other causes, A is estimated as

$$A^0 = \begin{cases} \frac{\tau P - D}{P}, & \text{if deaths.} \\ 1 - \frac{B^0}{P}, & \text{if cases.} \end{cases} \tag{11}$$

$$A^1 = \begin{cases} \frac{\tau P + B^1}{P}, & \text{if deaths.} \\ 1, & \text{if cases.} \end{cases} \tag{12}$$

where P, D, B^0, B^1 and τ are cohort size, total vaccine impact, disease burden in the no-vaccination scenario, disease burden in the with-vaccination scenario and (average) probability of survival, respectively. Then direct vaccine impact D^A on group A and indirect vaccine benefit D^B on group B are estimated as

$$D^A = e * P * (A^1 - A^0) \tag{13}$$

$$D^B = D - D^A. \tag{14}$$

Table .1: Notations

Group A :	effectively vaccinated population
Group B :	non-effectively vaccinated and un-vaccinated population
z :	vaccination activity coverage
v :	vaccine coverage
s :	proportion of people vaccinated with one dose
m :	proportion of people vaccinated with multiple doses
e :	effective vaccine coverage
θ :	correlation coefficient
λ :	vaccine efficacy
B^0 :	disease burden in a no-vaccination scenario
B^1 :	disease burden in a with-vaccination scenario
D :	vaccine impact
D^A :	direct vaccine impact attributable to group A
D^B :	in-direct vaccine benefit attributable to group B
P :	cohort size
γ :	probability of infection without vaccination
τ :	probability of survival