VESS

Anirudh Tomer

May 2021

1 Cohort study

The vaccination status of a subject in the study is denoted by v, where $v = 1$ represents that a subject is vaccinated, and $v = 0$ represents unvaccinated subjects. The infection status of a subject is represented by c, where $c = 1$ represents infections (called cases hereforth), and $c = 0$ represents uninfected subjects (called controls hereforth). Let there be $\mathcal{B}$ vaccine brands and $\mathcal{S}$ strains/genetic modifications of viruses. The strain of each case subject is denoted by S and the brand with which a subject is vaccinated is denoted by B. When $B = 0$ it means subject is unvaccinated, and similarly when $S = 0$ it means subject is a control. In total we have $(1 + \mathcal{S}) \times (1 + \mathcal{B})$ types of subjects based on the combinations of vaccinated with brands, unvaccinated, cases with strains, and controls.

1.1 Typically available information Q

Let the available pieces of information be denoted by:

1. The proportion of brand $b \in \{1, \dots, \mathcal{B}\}$ in the vaccinated subjects is given by $Q^1_b = \Pr(B = b \mid v = 1)$.
2. The proportion of strains $s \in \{1, \dots, \mathcal{S}\}$ in unvaccinated cases is given by $Q^2_s = \Pr(S = s \mid c = 1, v = 0)$.
3. Overall vaccine coverage in the source population from which we sample our subjects is given by $Q^3 = \Pr(v = 1)$.

4. Overall attack rate in the unvaccinated subjects of source population is given by $Q^4 = \Pr(c = 1 \mid v = 0)$.

5. The efficacy of brand b against strain s denoted by the $Q^5_{sb} = E_{sb}$ element of the vaccine efficacy matrix E with dimensions $\mathcal{S} \times \mathcal{B}$. Specifically, $E_{sb} = 1 - \dfrac{\mathrm{R}(S = s, c = 1 \mid B = b, v = 1)}{\mathrm{R}(S = s, c = 1 \mid B = 0, v = 0)}$, where,

$$\mathrm{R}(S = s, c = 1 \mid B = b, v = 1) = \frac{\Pr(S = s, c = 1, B = b, v = 1)}{\Pr(S = s, c = 1, B = b, v = 1) + \Pr(S = 0, c = 0, B = b, v = 1)},$$
$$\mathrm{R}(S = s, c = 1 \mid B = 0, v = 0) = \frac{\Pr(S = s, c = 1, B = 0, v = 0)}{\Pr(S = s, c = 1, B = 0, v = 0) + \Pr(S = 0, c = 0, B = 0, v = 0)}$$

1.2 Calculations

As mentioned earlier, we have $(1 + \mathcal{S}) \times (1 + \mathcal{B})$ types of subjects. The probability of being in each such subject category is given by the elements of the matrix F of probabilities shown below.

$$F = \begin{array}{c} \\ \textbf{Control} \\ \textbf{Strain 1} \\ \vdots \\ \textbf{Strain } \mathcal{S} \end{array} \begin{array}{cccc} \textbf{Unvaccinated} & \textbf{Brand 1} & \dots & \textbf{Brand } \mathcal{B} \\ \left[\begin{array}{c} \Pr(S = 0, c = 0, B = 0, v = 0) \\ \Pr(S = 1, c = 1, B = 0, v = 0) \\ \vdots \\ \Pr(S = \mathcal{S}, c = 1, B = 0, s = 0) \end{array}\right. & \begin{array}{c} \Pr(S = 0, c = 0, B = 1, v = 1) \\ \Pr(S = 1, c = 1, B = 1, v = 1) \\ \vdots \\ \Pr(S = \mathcal{S}, c = 1, B = 1, v = 1) \end{array} & \begin{array}{c} \dots \\ \dots \\ \ddots \\ \dots \end{array} & \left.\begin{array}{c} \Pr(S = 0, c = 0, B = \mathcal{B}, v = 1) \\ \Pr(S = 1, c = 1, B = \mathcal{B}, v = 1) \\ \vdots \\ \Pr(S = \mathcal{S}, c = 1, B = \mathcal{B}, v = 1) \end{array}\right] \end{array}$$

Our aim is to calculate every element of F based on the information $Q^1, \dots, Q^5$ listed in Section 2.1. The element F_{00} for unvaccinated controls is given by:

$$\begin{aligned} F_{00} &= \Pr(S = 0, c = 0, B = 0, v = 0) \\ &= \Pr(S = 0, B = 0 \mid c = 0, v = 0)\Pr(c = 0, v = 0) \\ &= \Pr(S = 0 \mid c = 0)\Pr(B = 0 \mid v = 0)\Pr(c = 0, v = 0) \\ &= 1 \times 1 \times \Pr(c = 0 \mid v = 0)\Pr(v = 0) \\ &= \{1 - \Pr(c = 1 \mid v = 0)\}\{1 - \Pr(v = 1)\} \\ &= (1 - Q^4)(1 - Q^3) \end{aligned}$$

The probabilities F_{s0} for unvaccinated cases, where $s \in \{1, \dots, \mathcal{S}\}$ are given by:

$$
\begin{aligned}
F_{s0} &= \Pr(S = s, c = 1, B = 0, v = 0) \\
&= \Pr(S = s, c = 1 \mid B = 0, v = 0)\Pr(B = 0, v = 0) \\
&= \Pr(S = s \mid c = 1, B = 0, v = 0)\Pr(c = 1 \mid B = 0, v = 0)\Pr(B = 0 \mid v = 0)\Pr(v = 0) \\
&= \Pr(S = s \mid c = 1, v = 0)\Pr(c = 1 \mid v = 0)\Pr(v = 0) \\
&= Q_s^2 Q^4 (1 - Q^3)
\end{aligned}
$$

To obtain the probabilities F_{0b} for vaccinated controls, and the probabilities F_{sb} for vaccinated cases, where $s \in \{1, \ldots, \mathcal{S}\}$ and $b \in \{1, \ldots, \mathcal{B}\}$ we use the information given in information Q^5. Specifically,

$$
\begin{aligned}
F_{sb} &= \Pr(S = s, c = 1, B = b, v = 1) \\
&= \Pr(S = 0, c = 0, B = b, v = 1)\left\{\frac{1}{R(S = s, c = 1, B = b, v = 1)} - 1\right\}^{-1} \\
&= F_{0b}\left\{\frac{1}{(1 - E_{sb})R(S = s, c = 1, B = 0, v = 0)} - 1\right\}^{-1} \\
&= F_{0b}\left\{\frac{1}{(1 - Q_{sb}^5)\dfrac{F_{s0}}{F_{s0} + F_{00}}} - 1\right\}^{-1}
\end{aligned}
$$

To solve F_{0b} we use the following equality,

$$\begin{aligned}
F_{0b} + \sum_s F_{sb} &= \Pr(B = b, v = 1) \\
\implies F_{0b} + \sum_s F_{sb} &= \Pr(B = b \mid v = 1)\Pr(v = 1) \\
\implies F_{0b} + \sum_s F_{sb} &= Q_b^1 Q^3 \\
\implies F_{0b} + \sum_s F_{0b}\left\{\frac{1}{(1 - Q_{sb}^5)\dfrac{F_{s0}}{F_{s0} + F_{00}}} - 1\right\}^{-1} &= Q_b^1 Q^3 \\
\implies F_{0b}\left[1 + \sum_s \left\{\frac{1}{(1 - Q_{sb}^5)\dfrac{F_{s0}}{F_{s0} + F_{00}}} - 1\right\}^{-1}\right] &= Q_b^1 Q^3 \\
\implies F_{0b} &= \frac{Q_b^1 Q^3}{\left[1 + \sum_s \left\{\dfrac{1}{(1 - Q_{sb}^5)\dfrac{F_{s0}}{F_{s0} + F_{00}}} - 1\right\}^{-1}\right]}
\end{aligned}$$

2 Case-control study

The vaccination status of a subject in the study is denoted by v, where $v = 1$ represents that a subject is vaccinated, and $v = 0$ represents unvaccinated subjects. The infection status of a subject is represented by c, where $c = 1$ represents infections (called cases hereforth), and $c = 0$ represents uninfected subjects (called controls hereforth). Let there be $\mathcal{B}$ vaccine brands and $\mathcal{S}$ strains/genetic modifications of viruses. The strain of each case subject is denoted by S and the brand with which a subject is vaccinated is denoted by B. When $B = 0$ it means subject is unvaccinated, and similarly when $S = 0$ it means subject is a control. In total we have $(1 + \mathcal{B}) \times \mathcal{S}$ types of cases and $(1 + \mathcal{B}) \times 1$ control subjects. The combinations are based on the combinations of subjects vaccinated with brands, unvaccinated, cases with strains, and controls.

2.1 Typically available information Q

Let the available pieces of information be denoted by:

1. The proportion of brand $b \in \{1, \ldots, \mathcal{B}\}$ in the vaccinated subjects is given by $Q^1_b = \Pr(B = b \mid v = 1)$. We also use the assumption that brand proportions in overall population $\Pr(B = b \mid v = 1)$ are almost equal to proportions in controls $\Pr(B = b \mid v = 1, c = 0)$ when the overall attack rate $p(c = 1)$ is low. Specifically,

   $$\Pr(B = b \mid v = 1) = \Pr(B = b \mid v = 1, c = 1)p(c = 1) + \Pr(B = b \mid v = 1, c = 0)\{1 - p(c = 1)\}$$

   Thus, if $p(c = 1)$ is small, then $Q^1_b = \Pr(B = b \mid v = 1) \approx \Pr(B = b \mid v = 1, c = 0)$.

2. The proportion of strains $s \in \{1, \ldots, \mathcal{S}\}$ in unvaccinated cases is given by $Q^2_s = \Pr(S = s \mid c = 1, v = 0)$.

3. Overall vaccine coverage in the source population from which we sample our subjects is given by $Q^3 = \Pr(v = 1)$. We also use the assumption that overall vaccine coverage in the population $p(v = 1)$ is almost equal to the vaccine coverage in controls $p(v = 1 \mid c = 0)$ when the overall attack rate $p(c = 1)$ is low. Specifically,

   $$p(v = 1) = p(v = 1 \mid c = 1)p(c = 1) + p(v = 1 \mid c = 0)\{1 - p(c = 1)\}$$

   Thus, if $p(c = 1)$ is small, then $Q^3 = p(v = 1) \approx p(v = 1 \mid c = 0)$.

4. Number of controls per case that we intend to select $Q^4 = r$.

5. The efficacy of brand b against strain s denoted by the $Q^5_{bs} = E_{bs}$ element of the vaccine efficacy matrix E with dimensions $\mathcal{B} \times \mathcal{S}$. Specifically, $E_{bs} = 1 - \dfrac{\mathrm{O}(B = b, v = 1, S = s \mid c = 1)}{\mathrm{O}(B = b, v = 1, S = 0 \mid c = 0)}$, where,

   $$\mathrm{O}(B = b, v = 1, S = s \mid c = 1) = \frac{\Pr(B = b, v = 1, S = s \mid c = 1)}{\Pr(B = 0, v = 0, S = s \mid c = 1)},$$
   $$\mathrm{O}(B = b, v = 1, S = 0 \mid c = 0) = \frac{\Pr(B = b, v = 1, S = 0 \mid c = 0)}{\Pr(B = 0, v = 0, S = 0, \mid c = 0)}$$

2.2 Calculations

As mentioned earlier, we have $(1+\mathcal{B})\times\mathcal{S}$ cases and $(1+\mathcal{B})\times 1$ controls. The probability of being in each control subject category is given by the elements of the matrix F' of probabilities shown below.

$$F' = \begin{array}{cc} & \textbf{Control} \\ \textbf{Unvaccinated} & \\ \textbf{Brand 1} & \\ \vdots & \\ \textbf{Brand } \mathcal{B} & \end{array} \begin{bmatrix} \Pr(B=0, v=0, S=0 \mid c=0) \\ \Pr(B=1, v=1, S=0 \mid c=0) \\ \vdots \\ \Pr(B=\mathcal{B}, v=1, S=0 \mid c=0) \end{bmatrix}$$

Our aim is to calculate every element of F' based on the information $Q^1, \ldots, Q^5$ listed in Section 2.1. The element F'_{00} for unvaccinated controls is given by:

$$\begin{aligned} F'_{00} &= \Pr(B=0, v=0, S=0 \mid c=0) \\ &= \Pr(B=0 \mid v=0, S=0, c=0)\Pr(v=0 \mid S=0, c=0)\Pr(S=0 \mid c=0) \\ &= 1 \times \Pr(v=0 \mid S=0, c=0) \times 1 \\ &= 1 - \Pr(v=1 \mid c=0) \\ &= 1 - Q^3 \end{aligned}$$

The probabilities F_{b0} for vaccinated controls, where $b \in \{1, \ldots, \mathcal{B}\}$ are given by:

$$\begin{aligned} F'_{b0} &= \Pr(B=b, v=1, S=0 \mid c=0) \\ &= \Pr(B=b \mid v=1, S=0, c=0)\Pr(v=1 \mid S=0, c=0)\Pr(S=0 \mid c=0) \\ &= \Pr(B=b \mid v=1, c=0)\Pr(v=1 \mid c=0) \\ &= Q^1_b Q^3 \end{aligned}$$

The probability of being in each case subject category is given by the elements of the matrix F of probabilities shown below:

$$F = \begin{array}{c} \\ \textbf{Unvaccinated} \\ \textbf{Brand 1} \\ \vdots \\ \textbf{Brand } \mathcal{B} \end{array} \begin{array}{ccc} \textbf{Strain 1} & \ldots & \textbf{Strain } \mathcal{S} \\ \end{array} \begin{bmatrix} \Pr(B=0, v=0, S=1 \mid c=1) & \ldots & \Pr(B=0, v=0, S=\mathcal{S} \mid c=1) \\ \Pr(B=1, v=1, S=1 \mid c=1) & \ldots & \Pr(B=1, v=1, S=\mathcal{S} \mid c=1) \\ \vdots & \ddots & \vdots \\ \Pr(B=\mathcal{B}, v=1, S=1 \mid c=1) & \ldots & \Pr(B=\mathcal{B}, v=1, S=\mathcal{S} \mid c=1) \end{bmatrix}$$

Our aim is to calculate every element of F based on the information $Q^1, \ldots, Q^5$ listed in Section 2.1. The probabilities for vaccinated cases are given by:

$$
\begin{aligned}
F_{bs} &= \Pr(B = b, v = 1, S = s \mid c = 1)\\
&= \Pr(B = 0, v = 0, S = s \mid c = 1) O(B = b, v = 1, S = s \mid c = 1)\\
&= F_{0s}(1 - E_{bs}) O(B = b, v = 1, S = 0 \mid c = 0)\\
&= F_{0s}(1 - E_{bs}) \frac{\Pr(B = b, v = 1, S = 0 \mid c = 0)}{\Pr(B = 0, v = 0, S = 0 \mid c = 0)}\\
&= F_{0s}(1 - E_{bs}) \frac{F'_{b0}}{F'_{00}}\\
&= F_{0s}(1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3}
\end{aligned}
$$

To solve for F_{0s} we use the following equality,

$$
\begin{aligned}
F_{0s} &= \Pr(B = 0, v = 0, S = s \mid c = 1)\\
&= \Pr(S = s \mid c = 1, B = 0, v = 0) \Pr(B = 0, v = 0 \mid c = 1)\\
&= \Pr(S = s \mid c = 1, v = 0) \Pr(B = 0 \mid v = 0, c = 1) \Pr(v = 0 \mid c = 1)\\
&= Q^2_s \times 1 \times \Pr(v = 0 \mid c = 1)\\
&= Q^2_s \Pr(v = 0 \mid c = 1)
\end{aligned}
$$

To solve for $\Pr(v = 0 \mid c = 1)$ we use the following equality:

$$
\begin{aligned}
\sum_s F_{0s} + \sum_b \sum_s F_{bs} &= 1\\
\implies \sum_s F_{0s} + \sum_b \sum_s F_{0s}(1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3} &= 1\\
\implies \sum_s F_{0s} \Big\{ 1 + \sum_b (1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3} \Big\} &= 1\\
\implies \sum_s Q^2_s \Pr(v = 0 \mid c = 1) \Big\{ 1 + \sum_b (1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3} \Big\} &= 1\\
\implies \Pr(v = 0 \mid c = 1) \left[\sum_s Q^2_s \Big\{ 1 + \sum_b (1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3} \Big\} \right] &= 1\\
\implies \Pr(v = 0 \mid c = 1) &= \frac{1}{\left[\sum_s Q^2_s \Big\{ 1 + \sum_b (1 - E_{bs}) \frac{Q^1_b Q^3}{1 - Q^3} \Big\} \right]}
\end{aligned}
$$