Why you need a multinomial roll for infection status

Arseniy Khvorov

11/12/2019

1 How it's supposed to work

F — infected with flu
L — infected with non-flu
V — vaccinated
U — unvaccinated
e — vaccine effectiveness

$$
\begin{gathered}
P(F) = f \\
P(L) = l \\
P(V) = v \\
P(U) = 1 - v = u \\
P(F, V) = vf(1 - e) \\
P(F, U) = uf \\
P(L, V) = vl \\
P(L, U) = ul \\
OR = \frac{P(F, V)P(L, U)}{P(F, U)P(L, V)} = \frac{vf(1 - e)ul}{ufvl} = 1 - e
\end{gathered}
$$

2 How it works with a multinomial roll

$$
\begin{gathered}
P(F|V) = f(1 - e) \\
P(F|U) = f \\
P(L|V) = l \\
P(L|U) = l \\
P(F, V) = P(F|V)P(V) = vf(1 - e) \\
P(F, U) = P(F|U)P(U) = uf \\
P(L, V) = P(L|V)P(V) = vl \\
P(L, U) = P(L|U)P(U) = ul \\
OR = \frac{P(F, V)P(L, U)}{P(F, U)P(L, V)} = \frac{vf(1 - e)ul}{ufvl} = 1 - e
\end{gathered}
$$

It works exaclty how it's supposed to.

3 How it works with a sequential roll

If the first roll works out flu infection and the second roll (non-flu infection) only applies to those not infected with flu, then

$$\begin{aligned} P(F|V) &= f(1-e) \\ P(F|U) &= f \\ P(L|V) &= l(1-f(1-e)) \\ P(L|U) &= l(1-f) \\ P(F,V) = P(F|V)P(V) &= vf(1-e) \\ P(F,U) = P(F|U)P(U) &= uf \\ P(L,V) = P(L|V)P(V) &= vl(1-f(1-e)) \\ P(L,U) = P(L|U)P(U) &= ul(1-f) \\ OR = \frac{P(F,V)P(L,U)}{P(F,U)P(L,V)} = \frac{vf(1-e)ul(1-f)}{ufvl(1-f(1-e))} &= (1-e)\frac{1-f}{1-f(1-e)} \end{aligned}$$

The OR is biased.

Note that the "central assumption" is that $P(L|V) = P(L|U)$ and with a sequential roll

$$\begin{aligned} P(L|V) &= l(1-f(1-e)) \\ P(L|U) &= l(1-f) \end{aligned}$$

the assumption is violated (unless $e = 0$), in fact, vaccination increases the probability of non-flu (since $f(1-e) \leq f$).

Also note that the total proportion infected with non-flu is

$$\begin{aligned} P(L,V) + P(L,U) &= vl(1-f(1-e)) + ul(1-f) \\ &= l(v - vf(1-e) + (1-v)(1-f)) \\ &= l(v - vf(1-e) + 1 - f - v + vf) \\ &= l(vfe + 1 - f) \end{aligned}$$

is not l and the expected number of people infected with non-flu is not Nl where N is the total population size.