

A Dynamic Model for Evaluation of the Bias of Influenza Vaccine Effectiveness Estimates from Observational Studies

Kylie E. C. Ainslie, Meng Shi, Michael Haber, and Walter A. Orenstein

1	Web Appendix 1	4
2	Web Appendix 2	8
3	Web Appendix 3	18
4	Web Appendix 4	20
5	Web Appendix 5	21

List of Tables

1	Model Assumptions.	4
2	Input parameter values used in calculations and simulations.	19
3	Bias of Estimates of VE against Symptomatic Influenza and Medically-Attended Influenza in Observational Studies from Monte Carlo Simulations for Different Values of True VE: 5 th , 50 th , and 95 th Quantiles of Bias Under Multiple Sources of Bias	22
4	Alternative X and U probabilities.	23
5	Alternative probabilities of vaccination.	25
6	Median Absolute Value of Bias of Estimated VE against Symptomatic Influenza and Medically-Attended Influenza in Observational Studies from Monte Carlo Simulations for Different Values of Influenza Test Sensitivity and Specificity . . .	28

List of Figures

1	VE estimates from TN studies under source of bias A for different values of true VE	23
2	Bias of VE estimates from PSC, TN, and TCC studies for alternative probabilities of (X, U)	24
3	Bias of VE estimates from PSC, TN, and TCC studies for alternative probabilities of vaccination	26

Web Appendix 1

True Vaccine Effectiveness

1.1 Model Assumptions

Model assumptions are shown in Web Table 1.

Web Table 1: Model Assumptions.

Assumption	Description
(a)	Each person is assigned two binary covariates: health status and health awareness, which are constant over time
(b)	Everyone is either effectively vaccinated prior to the study or remains unvaccinated throughout the study, and vaccination status is determined without error
(c)	Everyone in the study population is susceptible at the beginning of the study
(d)	A person can only have one influenza ARI during the study
(e)	A person can have at most one non-influenza ARI per week
(f)	The probabilities of influenza ARI and non-influenza ARI do not depend on a person's health awareness given his/her health status
(g)	The probability of seeking medical care does not depend on a person's health status given his/her health awareness
(h)	Influenza test specificity and sensitivity do not depend on health status, health awareness, or vaccination status, given a person's influenza infection status
(i)	Every person who seeks medical care is tested for influenza infection
(j)	A person with no ARI does not seek medical care

1.2 True Vaccine Effectiveness Against Symptomatic Influenza

A person is considered a true case of symptomatic influenza if s/he develops an ARI as a result of influenza infection. For a given vaccination status v the probability of symptomatic influenza in week j can be written as follows:

$$\begin{aligned}\mathbb{P}(Y_j = 2|V = v) &= \sum_{x=0}^1 \mathbb{P}(Y_j = 2|V = v, X = x) \times \mathbb{P}(X = x|V = v) \\ &= \sum_{x=0}^1 \gamma_{jvx} \times \mathbb{P}(X = x|V = v)\end{aligned}$$

Because we assume a person can only have one FARI during the study, the probability of symptomatic influenza over the entire study given vaccination status v is

$$\mathbb{P}(\text{person classified as case of symptomatic influenza}|V = v) = \sum_{j=1}^J \sum_{x=0}^1 [\gamma_{jvx} \times \mathbb{P}(X = x|V = v)].$$

where

$$\begin{aligned}\mathbb{P}(X = x|V = v) &= \frac{\mathbb{P}(V = v|X = x)\mathbb{P}(X = x)}{\mathbb{P}(V = v)} \\ &= \frac{\sum_{u=0}^1 \mathbb{P}(V = v|X = x, U = u)\mathbb{P}(X = x, U = u)}{\sum_{x=0}^1 \sum_{u=0}^1 \mathbb{P}(V = v|X = x, U = u)\mathbb{P}(X = x, U = u)}.\end{aligned}$$

Using our model parameters we can rewrite $\mathbb{P}(X = x|V = v)$ as

$$\begin{aligned}\mathbb{P}(X = 0|V = 0) &= \frac{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\ \mathbb{P}(X = 1|V = 0) &= \frac{(1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\ \mathbb{P}(X = 0|V = 1) &= \frac{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\ \mathbb{P}(X = 1|V = 1) &= \frac{\alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}}\end{aligned}$$

When evaluating true VE we assume random vaccination, i.e. vaccination does not depend on health status X or health awareness U ($\alpha_{xu} = \alpha \forall x, u$). The above expressions of $\mathbb{P}(X = x|V =$

v) can be simplified as

$$\begin{aligned}
\mathbb{P}(X = 0|V = 0) &= \frac{(1-\alpha)(\pi_{00}+\pi_{01})}{(1-\alpha)(\pi_{00}+\pi_{01}+\pi_{10}+\pi_{11})} = \pi_{00} + \pi_{01} \\
\mathbb{P}(X = 1|V = 0) &= \frac{(1-\alpha)(\pi_{10}+\pi_{11})}{(1-\alpha)(\pi_{00}+\pi_{01}+\pi_{10}+\pi_{11})} = \pi_{10} + \pi_{11} \\
\mathbb{P}(X = 0|V = 1) &= \frac{\alpha(\pi_{00}+\pi_{01})}{\alpha(\pi_{00}+\pi_{01}+\pi_{10}+\pi_{11})} = \pi_{00} + \pi_{01} \\
\mathbb{P}(X = 1|V = 1) &= \frac{\alpha(\pi_{10}+\pi_{11})}{\alpha(\pi_{00}+\pi_{01}+\pi_{10}+\pi_{11})} = \pi_{10} + \pi_{11}
\end{aligned}$$

Therefore, the true VE against SI (VE_{TSI}) can be written as one minus the ratio of the probabilities of symptomatic influenza in vaccinated and unvaccinated persons:

$$VE_{TSI} = 1 - \frac{\sum_{j=1}^J [\gamma_{j10}(\pi_{00} + \pi_{01}) + \gamma_{j11}(\pi_{10} + \pi_{11})]}{\sum_{j=1}^J [\gamma_{j00}(\pi_{00} + \pi_{01}) + \gamma_{j01}(\pi_{10} + \pi_{11})]}.$$

1.3 True Vaccine Effectiveness Against Medically-Attended Influenza

A person is considered a true case of medically-attended influenza if s/he develops an ARI as a result of influenza infection and seeks medical care for their ARI. Similarly to symptomatic influenza, the probability of medically-attended influenza given vaccination v for a given week j can be written as:

$$\begin{aligned}
&\mathbb{P}(\text{person considered case of medically-attended influenza in week } j|V = v) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \mathbb{P}(Y_j = 2, M_j = 1|V = v, X = x, U = u) \times \mathbb{P}(X = x, U = u|V = v) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \delta_{2vu} \gamma_{jvx} \times \mathbb{P}(X = x, U = u|V = v)
\end{aligned}$$

The probability of medically-attended influenza over the entire study given vaccination status v is

$$\mathbb{P}(\text{person classified as case of medically-attended influenza}|V = v) = \sum_{j=1}^J \left[\sum_{x=0}^1 \sum_{u=0}^1 \delta_{2vu} \gamma_{jvx} \times \mathbb{P}(X = x, U = u|V = v) \right]$$

where

$$\begin{aligned}
\mathbb{P}(X = x, U = u | V = v) &= \frac{\mathbb{P}(V = v | X = x, U = u) \mathbb{P}(X = x, U = u)}{\mathbb{P}(V = v)} \\
&= \frac{\mathbb{P}(V = v | X = x, U = u) \mathbb{P}(X = x, U = u)}{\sum_{x=0}^1 \sum_{u=0}^1 \mathbb{P}(V = v | X = x, U = u) \mathbb{P}(X = x, U = u)}.
\end{aligned}$$

Using our model parameters we can rewrite $\mathbb{P}(X = x, U = u | V = v)$ as

$$\begin{aligned}
\mathbb{P}(X = 0, U = 0 | V = 0) &= \frac{(1-\alpha_{00})\pi_{00}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\
\mathbb{P}(X = 0, U = 1 | V = 0) &= \frac{(1-\alpha_{01})\pi_{01}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\
\mathbb{P}(X = 1, U = 0 | V = 0) &= \frac{(1-\alpha_{10})\pi_{10}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\
\mathbb{P}(X = 1, U = 1 | V = 0) &= \frac{(1-\alpha_{11})\pi_{11}}{(1-\alpha_{00})\pi_{00} + (1-\alpha_{01})\pi_{01} + (1-\alpha_{10})\pi_{10} + (1-\alpha_{11})\pi_{11}} \\
\mathbb{P}(X = 0, U = 0 | V = 1) &= \frac{\alpha_{00}\pi_{00}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
\mathbb{P}(X = 0, U = 1 | V = 1) &= \frac{\alpha_{01}\pi_{01}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
\mathbb{P}(X = 1, U = 0 | V = 1) &= \frac{\alpha_{10}\pi_{10}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
\mathbb{P}(X = 1, U = 1 | V = 1) &= \frac{\alpha_{11}\pi_{11}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}}
\end{aligned}$$

Again, when evaluating the true VE, we assume random vaccination ($\alpha_{xu} = \alpha \forall x, u$), thus the above expressions of $\mathbb{P}(X = x, U = u | V = v)$ can be simplified as

$$\begin{aligned}
\mathbb{P}(X = 0, U = 0 | V = 0) &= \frac{(1-\alpha)\pi_{00}}{(1-\alpha)(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{00} \\
\mathbb{P}(X = 0, U = 1 | V = 0) &= \frac{(1-\alpha)\pi_{01}}{(1-\alpha)(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{01} \\
\mathbb{P}(X = 1, U = 0 | V = 0) &= \frac{(1-\alpha)\pi_{10}}{(1-\alpha)(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{10} \\
\mathbb{P}(X = 1, U = 1 | V = 0) &= \frac{(1-\alpha)\pi_{11}}{(1-\alpha)(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{11} \\
\mathbb{P}(X = 0, U = 0 | V = 1) &= \frac{\alpha\pi_{00}}{\alpha(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{00} \\
\mathbb{P}(X = 0, U = 1 | V = 1) &= \frac{\alpha\pi_{01}}{\alpha(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{01} \\
\mathbb{P}(X = 1, U = 0 | V = 1) &= \frac{\alpha\pi_{10}}{\alpha(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{10} \\
\mathbb{P}(X = 1, U = 1 | V = 1) &= \frac{\alpha\pi_{11}}{\alpha(\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11})} = \pi_{11}
\end{aligned}$$

Therefore, the true VE against medically-attended influenza (VE_{TMAI}) is

$$VE_{TMAI} = 1 - \frac{\sum_{j=1}^J \{\gamma_{j10} [\delta_{210}\pi_{00} + \delta_{211}\pi_{01}] + \gamma_{j11} [\delta_{210}\pi_{10} + \delta_{211}\pi_{11}]\}}{\sum_{j=1}^J \{\gamma_{j00} [\delta_{200}\pi_{00} + \delta_{201}\pi_{01}] + \gamma_{j01} [\delta_{200}\pi_{10} + \delta_{201}\pi_{11}]\}}$$

Web Appendix 2

Estimates of Vaccine Effectiveness

2.1 Passive Surveillance Cohort Study

In a PSC study, a person is considered a case in week j if:

- they have an ARI in week j (i.e., $Y_j > 0$)
- they seek medical care for their ARI in week j (i.e., $M_j = 1$)
- they test positive for influenza infection (i.e., $T_j = 1$)

Let C_j^* be the event that the person was identified as a case in week j ($j = 1, \dots, J$), i.e. $C_j^* = \{T_j = 1, M_j = 1\}$. Then the event that the person is a case is the union of all the events C_j^* , $j = 1, \dots, J$. We can write the event C_j^* as the union of two disjoint events: $A_j^* = \{T_j = 1, M_j = 1, Y_j = 1\}$ and $B_j^* = \{T_j = 1, M_j = 1, Y_j = 2\}$, where $A_1^*, \dots, A_J^*$ are independent and $B_1^*, \dots, B_J^*$ are mutually exclusive. Let $A^* = \bigcup_{j=1}^J A_j^*$, $B^* = \bigcup_{j=1}^J B_j^*$, and $C^* = \bigcup_{j=1}^J C_j^*$. As in an ASC study, we can write the probability of being a case in a PSC study (C^*) as $\mathbb{P}(C^*) = 1 - \mathbb{P}(\bar{A}^*)\mathbb{P}(\bar{B}^*)$. Let C_j^* be the event that the person was identified as a case in week j ($j = 1, \dots, J$), i.e. $C_j^* = \{T_j = 1, M_j = 1\}$. Then the event that the person is a case is the union of all the events C_j^* , $j = 1, \dots, J$. We can write the event C_j^* as the union of two disjoint events: $A_j^* = \{T_j = 1, M_j = 1, Y_j = 1\}$ and $B_j^* = \{T_j = 1, M_j = 1, Y_j = 2\}$, where $A_1^*, \dots, A_J^*$ are independent and $B_1^*, \dots, B_J^*$ are mutually exclusive. Let $A^* = \bigcup_{j=1}^J A_j^*$, $B^* = \bigcup_{j=1}^J B_j^*$, and $C^* = \bigcup_{j=1}^J C_j^*$. As in an ASC study, we can write the probability of being a case in a PSC study (C^*) as $\mathbb{P}(C^*) = \mathbb{P}(A^*) + \mathbb{P}(B^*) - \mathbb{P}(A^* \cap B^*)$.

Using the parameters from our model, we can write the probability of being an PSC case

for a given (v, x, u) as

$$\mathbb{P}(C^*|V = v, X = x, U = u) = 1 - \prod_{j=1}^J (1 - \tau_1 \delta_{1vu} \beta_{jvx}) + \sum_{j=1}^J (\tau_2 \delta_{2vu} \gamma_{jvx}) - \sum_{j=1}^J \sum_{k=1}^J [\tau_1 \delta_{1vu} \beta_{jvx} \times \tau_2 \delta_{2vu} \gamma_{kvx}],$$

where $k \neq j$.

$$\begin{aligned} \mathbb{P}(C^*|V = v) &= \sum_{x=0}^1 \sum_{u=0}^1 \left\{ 1 - \prod_{j=1}^J (1 - \tau_1 \delta_{1vu} \beta_{jvx}) + \sum_{j=1}^J (\tau_2 \delta_{2vu} \gamma_{jvx}) - \sum_{j=1}^J \sum_{k=1}^J [\tau_1 \delta_{1vu} \beta_{jvx} \times \tau_2 \delta_{2vu} \gamma_{kvx}] \right. \\ &\quad \times \left. \mathbb{P}(X = x, U = u|V = v) \right\}, \end{aligned}$$

where $k \neq j$.

Thus, the VE estimate from an PSC study is

$$VE_{PSC} = 1 - \frac{\mathbb{P}(C^*|V = 1)}{\mathbb{P}(C^*|V = 0)}$$

2.2 Test-Negative Study

2.2.1 Test-Negative Case

The probability of being a case in week j is the probability of being diagnosed with a influenza ARI in week j . The probability of having a influenza ARI in week j is the same as the probability of having a influenza ARI in week j and not having a influenza ARI in weeks 1 to $j - 1$. For a given (v, x, u) we can write

$$\mathbb{P}(Y_j = 2|V = v, X = x) = \mathbb{P}(Y_j = 2, \mathbf{Y}_{j-1} < \mathbf{2}|V = v, X = x) = \gamma_{jvx}.$$

Also, when $Y_1, \dots, Y_{j-1}$ are given, $M_1, \dots, M_{j-1}$ are independent, thus

$$\begin{aligned}
& \mathbb{P}(\mathbf{M}_{j-1} = \mathbf{0} | \mathbf{Y}_{j-1} < \mathbf{2}, V = v, U = u) \\
&= \mathbb{P}(M_1 = 0 | Y_1 < 2, V = v, U = u) \cdots \mathbb{P}(M_{j-1} = 0 | Y_{j-1} < 2, V = v, U = u) \\
&= \prod_{k=1}^{j-1} \mathbb{P}(M_k = 0 | Y_k < 2, V = v, U = u) \\
&= \prod_{k=1}^{j-1} \frac{\mathbb{P}(M_k = 0, Y_k < 2 | V = v, U = u)}{\mathbb{P}(Y_k < 2 | V = v, X = x)} \\
&= \prod_{k=1}^{j-1} \frac{\mathbb{P}(M_k = 0, Y_k = 0 | V = v, X = x, U = u) + \mathbb{P}(M_k = 0, Y_k = 1 | V = v, X = x, U = u)}{\mathbb{P}(Y_k = 0 | V = v, X = x) + \mathbb{P}(Y_k = 1 | V = v, X = x)} \\
&= \prod_{k=1}^{j-1} \left\{ \frac{\mathbb{P}(M_k = 0 | Y_k = 0, V = v, U = u) \mathbb{P}(Y_k = 0 | V = v, X = x)}{\mathbb{P}(Y_k = 0 | V = v, X = x) + \mathbb{P}(Y_k = 1 | V = v, X = x)} \right. \\
&\quad \left. + \frac{\mathbb{P}(M_k = 0 | Y_k = 1, V = v, U = u) \mathbb{P}(Y_k = 1 | V = v, X = x)}{\mathbb{P}(Y_k = 0 | V = v, X = x) + \mathbb{P}(Y_k = 1 | V = v, X = x)} \right\} \\
&= \prod_{k=1}^{j-1} \frac{1 \times (1 - \beta_{kvx} - \gamma_{kvx}) + (1 - \delta_{1vu})\beta_{kvx}}{(1 - \gamma_{kvx})} \\
&= \prod_{k=1}^{j-1} \frac{1 - \gamma_{kvx} - \delta_{1vu}\beta_{kvx}}{(1 - \gamma_{kvx})}
\end{aligned}$$

Therefore, the probability of being considered a case in week j for a given (v, x, u) is:

$$\begin{aligned}
& \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0}, T_j = 1 | V = v, X = x, U = u) \\
&= \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0}, T_j = 1 | Y_j = 2, V = v, U = u) \mathbb{P}(Y_j = 2 | V = v, X = x) \\
&\quad + \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0}, T_j = 1 | Y_j = 1, V = v, U = u) \mathbb{P}(Y_j = 1 | V = v, X = x) \\
&= \mathbb{P}(T_j = 1 | Y_j = 2) \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | Y_j = 2, V = v, U = u) \mathbb{P}(Y_j = 2 | V = v, X = x) \\
&\quad + \mathbb{P}(T_j = 1 | Y_j = 1) \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | Y_j = 1, V = v, U = u) \mathbb{P}(Y_j = 1 | V = v, X = x) \\
&= \tau_2 \gamma_{jvx} \times \mathbb{P}(M_j = 1 | Y_j = 2, V = v, U = u) \mathbb{P}(\mathbf{M}_{j-1} = \mathbf{0} | \mathbf{Y}_{j-1} < \mathbf{2}, V = v, U = u) \\
&\quad + \tau_1 \beta_{jvx} \times \mathbb{P}(M_j = 1 | Y_j = 1, V = v, U = u) \mathbb{P}(\mathbf{M}_{j-1} = \mathbf{0} | V = v, U = u) \\
&= \tau_2 \delta_{2vu} \gamma_{jvx} \prod_{k=1}^{j-1} \mathbb{P}(M_k = 0 | Y_k < 2, V = v, U = u) + \tau_1 \delta_{1vu} \beta_{jvx} \prod_{k=1}^{j-1} \mathbb{P}(M_k = 0 | V = v, U = u) \\
&= \tau_2 \delta_{2vu} \gamma_{jvx} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{kvx} - \delta_{1vu}\beta_{kvx}}{1 - \gamma_{kvx}} \right) + \tau_2 \delta_{2vu} \gamma_{jvx} \prod_{k=1}^{j-1} (1 - \delta_{1vu}\beta_{kvx} - \delta_{2vu}\gamma_{kvx}).
\end{aligned}$$

The probability of being considered a case in week j for a given v is:

$$\begin{aligned}
& \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0}, T_j = 1 | V = 1) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \left\{ \left[\tau_2 \delta_{21u} \gamma_{j1x} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k1x} - \delta_{11u} \beta_{k1x}}{1 - \gamma_{k1x}} \right) + \tau_1 \delta_{11u} \beta_{j1x} \prod_{k=1}^{j-1} (1 - \delta_{11u} \beta_{k1x} - \delta_{21u} \gamma_{k1x}) \right] \right. \\
&\quad \times \left. \mathbb{P}(X = x, U = u | V = 1) \right\} \\
&= \left[\delta_{210} \gamma_{j10} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k10} - \delta_{110} \beta_{k10}}{1 - \gamma_{k10}} \right) + \tau_1 \delta_{110} \beta_{j10} \prod_{k=1}^{j-1} (1 - \delta_{110} \beta_{k10} - \delta_{210} \gamma_{k10}) \right] \\
&\quad \times \frac{\alpha_{00} \pi_{00}}{\alpha_{00} \pi_{00} + \alpha_{01} \pi_{01} + \alpha_{10} \pi_{10} + \alpha_{11} \pi_{11}} \\
&\quad + \left[\delta_{211} \gamma_{j10} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k10} - \delta_{111} \beta_{k10}}{1 - \gamma_{k10}} \right) + \tau_1 \delta_{111} \beta_{j10} \prod_{k=1}^{j-1} (1 - \delta_{111} \beta_{k10} - \delta_{211} \gamma_{k10}) \right] \\
&\quad \times \frac{\alpha_{01} \pi_{01}}{\alpha_{00} \pi_{00} + \alpha_{01} \pi_{01} + \alpha_{10} \pi_{10} + \alpha_{11} \pi_{11}} \\
&\quad + \left[\delta_{210} \gamma_{j11} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k11} - \delta_{110} \beta_{k11}}{1 - \gamma_{k11}} \right) + \tau_1 \delta_{110} \beta_{j11} \prod_{k=1}^{j-1} (1 - \delta_{110} \beta_{k11} - \delta_{210} \gamma_{k11}) \right] \\
&\quad \times \frac{\alpha_{10} \pi_{10}}{\alpha_{00} \pi_{00} + \alpha_{01} \pi_{01} + \alpha_{10} \pi_{10} + \alpha_{11} \pi_{11}} \\
&\quad + \left[\delta_{211} \gamma_{j11} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k11} - \delta_{111} \beta_{k11}}{1 - \gamma_{k11}} \right) + \tau_1 \delta_{111} \beta_{j11} \prod_{k=1}^{j-1} (1 - \delta_{111} \beta_{k11} - \delta_{211} \gamma_{k11}) \right] \\
&\quad \times \frac{\alpha_{11} \pi_{11}}{\alpha_{00} \pi_{00} + \alpha_{01} \pi_{01} + \alpha_{10} \pi_{10} + \alpha_{11} \pi_{11}}
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0}, T_j = 1 | V = 0) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \left\{ \left[\tau_2 \delta_{20u} \gamma_{j0x} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k0x} - \delta_{10u} \beta_{k0x}}{1 - \gamma_{k0x}} \right) + \tau_1 \delta_{10u} \beta_{j0x} \prod_{k=1}^{j-1} (1 - \delta_{10u} \beta_{k0x} - \delta_{20u} \gamma_{k0x}) \right] \right. \\
&\quad \times \left. \mathbb{P}(X = x, U = u | V = 0) \right\} \\
&= \left[\delta_{200} \gamma_{j00} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k00} - \delta_{100} \beta_{k00}}{1 - \gamma_{k00}} \right) + \tau_1 \delta_{00u} \beta_{j00} \prod_{k=1}^{j-1} (1 - \delta_{100} \beta_{k00} - \delta_{200} \gamma_{k00}) \right] \\
&\quad \times \frac{(1 - \alpha_{00}) \pi_{00}}{(1 - \alpha_{00}) \pi_{00} + (1 - \alpha_{01}) \pi_{01} + (1 - \alpha_{10}) \pi_{10} + (1 - \alpha_{11}) \pi_{11}} \\
&+ \left[\delta_{201} \gamma_{j00} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k00} - \delta_{101} \beta_{k00}}{1 - \gamma_{k00}} \right) + \tau_1 \delta_{101} \beta_{j00} \prod_{k=1}^{j-1} (1 - \delta_{101} \beta_{k00} - \delta_{201} \gamma_{k00}) \right] \\
&\quad \times \frac{(1 - \alpha_{01}) \pi_{01}}{(1 - \alpha_{00}) \pi_{00} + (1 - \alpha_{01}) \pi_{01} + (1 - \alpha_{10}) \pi_{10} + (1 - \alpha_{11}) \pi_{11}} \\
&+ \left[\delta_{200} \gamma_{j01} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k01} - \delta_{100} \beta_{k01}}{1 - \gamma_{k01}} \right) + \tau_1 \delta_{100} \beta_{j01} \prod_{k=1}^{j-1} (1 - \delta_{100} \beta_{k01} - \delta_{200} \gamma_{k01}) \right] \\
&\quad \times \frac{(1 - \alpha_{10}) \pi_{10}}{(1 - \alpha_{00}) \pi_{00} + (1 - \alpha_{01}) \pi_{01} + (1 - \alpha_{10}) \pi_{10} + (1 - \alpha_{11}) \pi_{11}} \\
&+ \left[\delta_{201} \gamma_{j01} \prod_{k=1}^{j-1} \left(\frac{1 - \gamma_{k01} - \delta_{101} \beta_{k01}}{1 - \gamma_{k01}} \right) + \tau_1 \delta_{101} \beta_{j01} \prod_{k=1}^{j-1} (1 - \delta_{101} \beta_{k01} - \delta_{201} \gamma_{k01}) \right] \\
&\quad \times \frac{(1 - \alpha_{11}) \pi_{11}}{(1 - \alpha_{00}) \pi_{00} + (1 - \alpha_{01}) \pi_{01} + (1 - \alpha_{10}) \pi_{10} + (1 - \alpha_{11}) \pi_{11}}
\end{aligned}$$

2.2.2 Test-Negative Control

For a given (v, x, u) the probability of being considered a TN control in week j can be expressed as:

$$\begin{aligned}
& \mathbb{P}(Y_j = 1, M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | V = v, X = x, U = u) \\
&= \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | Y_j = 1, V = v, X = x, U = u) \mathbb{P}(Y_j = 1 | V = v, X = x)
\end{aligned}$$

The second term in the above expression is simply β_{jvx} , while the first term can be further written as:

$$\begin{aligned}
& \mathbb{P}(M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | Y_j = 1, V = v, X = x, U = u) \\
&= \mathbb{P}(M_j = 1 | Y_j = 1, V = v, X = x, U = u) \times \prod_{k=1}^{j-1} \mathbb{P}(M_k = 0 | V = v, X = x, U = u) \\
&= \delta_{1vu} \times \prod_{k=1}^{j-1} \mathbb{P}(M_k = 0 | V = v, X = x, U = u),
\end{aligned}$$

where

$$\begin{aligned}
& \mathbb{P}(M_k = 0 | V = v, X = x, U = u) \\
&= \mathbb{P}(M_k = 0 | Y_k = 0, V = v, X = x, U = u) \mathbb{P}(Y_k = 0 | V = v, X = x) \\
&+ \mathbb{P}(M_k = 0 | Y_k = 1, V = v, X = x, U = u) \mathbb{P}(Y_k = 1 | V = v, X = x) \\
&+ \mathbb{P}(M_k = 0 | Y_k = 2, V = v, X = x, U = u) \mathbb{P}(Y_k = 2 | V = v, X = x) \\
&= 1 \times (1 - \beta_{kvx} - \gamma_{kvx}) + (1 - \delta_{1vu})\beta_{kvx} + (1 - \delta_{2vu})\gamma_{kvx} \\
&= 1 - \delta_{1vu}\beta_{kvx} - \delta_{2vu}\gamma_{kvx}.
\end{aligned}$$

Thus, for a given (v, x, u) the probability of being considered a TN control in week j is

$$\mathbb{P}(Y_j = 1, M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | V = v, X = x, U = u) = \delta_{1vu}\beta_{jvx} \prod_{k=1}^{j-1} (1 - \delta_{1vu}\beta_{kvx} - \delta_{2vu}\gamma_{kvx}).$$

The probability of being considered a TN control in week j for a given v is:

$$\begin{aligned}
& \mathbb{P}(Y_j = 1, M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | V = 1) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \left\{ \delta_{11u}\beta_{j1x} \prod_{k=1}^{j-1} (1 - \delta_{11u}\beta_{k1x} - \delta_{21u}\gamma_{k1x}) \times \mathbb{P}(X = x, U = u | V = 1) \right\}. \\
&= \delta_{110}\beta_{j10} \prod_{k=1}^{j-1} (1 - \delta_{110}\beta_{k10} - \delta_{210}\gamma_{k10}) \times \frac{\alpha_{00}\pi_{00}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
&+ \delta_{111}\beta_{j10} \prod_{k=1}^{j-1} (1 - \delta_{111}\beta_{k10} - \delta_{211}\gamma_{k10}) \times \frac{\alpha_{01}\pi_{01}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
&+ \delta_{110}\beta_{j11} \prod_{k=1}^{j-1} (1 - \delta_{110}\beta_{k11} - \delta_{210}\gamma_{k11}) \times \frac{\alpha_{10}\pi_{10}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\
&+ \delta_{111}\beta_{j11} \prod_{k=1}^{j-1} (1 - \delta_{111}\beta_{k11} - \delta_{211}\gamma_{k11}) \times \frac{\alpha_{11}\pi_{11}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}}
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{P}(Y_j = 1, M_j = 1, \mathbf{M}_{j-1} = \mathbf{0} | V = 0) \\
&= \sum_{x=0}^1 \sum_{u=0}^1 \left\{ \delta_{10u}\beta_{j0x} \prod_{k=1}^{j-1} (1 - \delta_{10u}\beta_{k0x} - \delta_{20u}\gamma_{k0x}) \times \mathbb{P}(X = x, U = u | V = 0) \right\}. \\
&= \delta_{100}\beta_{j00} \prod_{k=1}^{j-1} (1 - \delta_{100}\beta_{k00} - \delta_{200}\gamma_{k00}) \times \frac{(1 - \alpha_{00})\pi_{00}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}} \\
&+ \delta_{101}\beta_{j00} \prod_{k=1}^{j-1} (1 - \delta_{101}\beta_{k00} - \delta_{201}\gamma_{k00}) \times \frac{(1 - \alpha_{01})\pi_{01}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}} \\
&+ \delta_{100}\beta_{j01} \prod_{k=1}^{j-1} (1 - \delta_{100}\beta_{k01} - \delta_{201}\gamma_{k01}) \times \frac{(1 - \alpha_{10})\pi_{10}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}} \\
&+ \delta_{101}\beta_{j01} \prod_{k=1}^{j-1} (1 - \delta_{101}\beta_{k01} - \delta_{201}\gamma_{k01}) \times \frac{(1 - \alpha_{11})\pi_{11}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}}
\end{aligned}$$

2.2.3 Test-Negative-Based Estimate of Vaccine Effectiveness

To obtain the final 2x2 table for a TN study, we collapse across X and U for each week j and then sum over all weeks $1, \dots, J$. Finally, we can obtain the VE estimate from a TN study

using the OR from the final 2x2 table. The VE estimate from a TN study is

$$VE_{TN} = 1 - OR_{TN} = \frac{a \times d_{TN}}{b \times c_{TN}},$$

where

$$\begin{aligned} a &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{21u} \gamma_{j1x} \prod_{k=1}^{j-1} \left[\frac{1 - \gamma_{k1x} - \delta_{11u} \beta_{k1x}}{1 - \gamma_{k1x}} \right] \times \mathbb{P}(X = x, U = u | V = 1) \right] \right\} \\ b &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{20u} \gamma_{j0x} \prod_{k=1}^{j-1} \left[\frac{1 - \gamma_{k0x} - \delta_{10u} \beta_{k0x}}{1 - \gamma_{k0x}} \right] \times \mathbb{P}(X = x, U = u | V = 0) \right] \right\} \\ c_{TN} &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{11u} \beta_{j1x} \prod_{k=1}^{j-1} (1 - \delta_{11u} \beta_{k1x} - \delta_{21u} \gamma_{k1x}) \times \mathbb{P}(X = x, U = u | V = 1) \right] \right\}. \\ d_{TN} &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{10u} \beta_{j0x} \prod_{k=1}^{j-1} (1 - \delta_{10u} \beta_{k0x} - \delta_{20u} \gamma_{k0x}) \times \mathbb{P}(X = x, U = u | V = 0) \right] \right\}. \end{aligned}$$

2.3 Traditional Case-Control Study

2.3.1 Probability of no ARI in weeks $1, \dots, j$

To determine the probability of being considered a TCC control in week j , we need to determine the probability that a person did not have either a influenza ARI or non-influenza ARI prior to, or in, week j . Because Y_{Fj} is independent of Y_{Nj} for all j , we can determine the probability of not having a influenza ARI separately from the probability of not having a non-influenza ARI.

Due to the independence of the Y_{Nj} s, the joint probability of not having a non-influenza ARI in weeks $1, \dots, j$ is simply the product of the marginal probabilities of not having a non-influenza ARI in each week $1, \dots, j$. Thus, for a given (v, x)

$$\begin{aligned} \mathbb{P}(Y_{N1} = 0, \dots, Y_{Nj} = 0 | V = v, X = x) &= \prod_{k=1}^j \mathbb{P}(Y_{Nk} = 0 | V = v, X = x) \\ &= \prod_{k=1}^j (1 - \beta_{kvx}). \end{aligned}$$

2.3.2 Probability of no Influenza ARI

Due to our model assumption that a person can only be infected with influenza once during the season, and thus, only have one influenza ARI, we cannot treat the Y_{Fj} s as independent. To determine the joint probability of not having a influenza ARI in weeks $1, \dots, j$, we will start

with a simple case. For given (v, x) first, consider $j = 2$.

$$\begin{aligned}
\mathbb{P}(Y_{F2} = 0 | Y_{F1} = 0, V = v, X = x) &= 1 - \mathbb{P}(Y_{F2} = 1 | Y_{F1} = 0, V = v, X = x) \\
&= 1 - \frac{\mathbb{P}(Y_{F1} = 0, Y_{F2} = 1 | V = v, X = x)}{\mathbb{P}(Y_{F1} = 0 | V = v, X = x)} \\
&= 1 - \frac{\gamma_{2vx}}{1 - \gamma_{1vx}} \\
&= \frac{1 - \gamma_{1vx} - \gamma_{2vx}}{1 - \gamma_{1vx}}.
\end{aligned}$$

So,

$$\begin{aligned}
\mathbb{P}(Y_{F1} = 0, Y_{F2} = 0 | V = v, X = x) &= \mathbb{P}(Y_{F2} = 0 | Y_{F1} = 0, V = v, X = x) \mathbb{P}(Y_{F1} = 0 | V = v, X = x) \\
&= \frac{1 - \gamma_{1vx} - \gamma_{2vx}}{1 - \gamma_{1vx}} (1 - \gamma_{1vx}) \\
&= 1 - \gamma_{1vx} - \gamma_{2vx}.
\end{aligned}$$

Now, consider $j = 3$.

$$\begin{aligned}
\mathbb{P}(Y_{F3} = 0 | Y_{F1} = 0, Y_{F2} = 0, V = v, X = x) &= 1 - \mathbb{P}(Y_{F3} = 1 | Y_{F1} = 0, Y_{F2} = 0, V = v, X = x) \\
&= 1 - \left[\frac{\mathbb{P}(Y_{F1} = 0, Y_{F2} = 0, Y_{F3} = 1 | V = v, X = x)}{\mathbb{P}(Y_{F1} = 0, Y_{F2} = 0 | V = v, X = x)} \right] \\
&= 1 - \frac{\gamma_{3vx}}{1 - \gamma_{1vx} - \gamma_{2vx}} \\
&= \frac{1 - \gamma_{1vx} - \gamma_{2vx} - \gamma_{3vx}}{1 - \gamma_{1vx} - \gamma_{2vx}}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{P}(Y_{F1} = 0, Y_{F2} = 0, Y_{F3} = 0 | V = v, X = x) &= \mathbb{P}(Y_{F3} = 0 | Y_{F1} = 0, Y_{F2} = 0, V = v, X = x) \\
&\times \mathbb{P}(Y_{F1} = 0, Y_{F2} = 0 | V = v, X = x) \\
&= \frac{1 - \gamma_{1vx} - \gamma_{2vx} - \gamma_{3vx}}{1 - \gamma_{1vx} - \gamma_{2vx}} (1 - \gamma_{1vx} - \gamma_{2vx}) \\
&= 1 - \gamma_{1vx} - \gamma_{2vx} - \gamma_{3vx}.
\end{aligned}$$

For any j , $j = 1, \dots, J$,

$$\mathbb{P}(Y_{Fj} = 0 | Y_{F1} = 0, \dots, Y_{F(j-1)} = 0, V = v, X = x) = \frac{1 - \sum_{k=1}^j \gamma_{kvx}}{1 - \sum_{k=0}^{j-1} \gamma_{kvx}}$$

Thus, the probability that a person does not have an ARI in weeks $1, \dots, j$ is

$$\mathbb{P}(Y_{F1} = 0, \dots, Y_{Fj} = 0 | V = v, X = x) = 1 - \sum_{k=1}^j \gamma_{kvx},$$

where $\sum_{k=1}^J \gamma_{kvx} \leq 1$.

2.3.3 Traditional Case-Control Control

For a person to be considered a TCC control in week j , they did not have an ARI on or prior to week j (i.e., $\mathbf{Y}_j = \mathbf{0}$). Using the results from above, the probability of being a TCC control in week j for a given (v, x) can be written as

$$\begin{aligned} \mathbb{P}(\mathbf{Y}_j = \mathbf{0} | V = v, X = x) &= \mathbb{P}(Y_{F1} = 0, \dots, Y_{Fj} = 0 | V = v, X = x) \\ &\times \mathbb{P}(Y_{N1} = 0, \dots, Y_{Nj} = 0 | V = v, X = x) \\ &= \left[1 - \sum_{k=1}^j \gamma_{kvx} \right] \left[\prod_{k=1}^j (1 - \beta_{kvx}) \right]. \end{aligned}$$

The probability of being a TCC control in week j for a given v is:

$$\begin{aligned} \mathbb{P}(\mathbf{Y}_j = \mathbf{0} | V = 1) &= \sum_{x=0}^1 \left[1 - \sum_{k=1}^j \gamma_{k1x} \right] \left[\prod_{k=1}^j (1 - \beta_{k1x}) \right] \times \mathbb{P}(X = x | V = 1) \\ &= \left[1 - \sum_{k=1}^j \gamma_{k10} \right] \left[\prod_{k=1}^j (1 - \beta_{k10}) \right] \times \frac{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \\ &+ \left[1 - \sum_{k=1}^j \gamma_{k11} \right] \left[\prod_{k=1}^j (1 - \beta_{k11}) \right] \times \frac{\alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}}{\alpha_{00}\pi_{00} + \alpha_{01}\pi_{01} + \alpha_{10}\pi_{10} + \alpha_{11}\pi_{11}} \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}(\mathbf{Y}_j = \mathbf{0} | V = 0) &= \sum_{x=0}^1 \left[1 - \sum_{k=1}^j \gamma_{k0x} \right] \left[\prod_{k=1}^j (1 - \beta_{k0x}) \right] \times \mathbb{P}(X = x | V = 0) \\ &= \left[1 - \sum_{k=1}^j \gamma_{k00} \right] \left[\prod_{k=1}^j (1 - \beta_{k00}) \right] \times \frac{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}} \\ &+ \left[1 - \sum_{k=1}^j \gamma_{k01} \right] \left[\prod_{k=1}^j (1 - \beta_{k01}) \right] \times \frac{(1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}}{(1 - \alpha_{00})\pi_{00} + (1 - \alpha_{01})\pi_{01} + (1 - \alpha_{10})\pi_{10} + (1 - \alpha_{11})\pi_{11}} \end{aligned}$$

2.3.4 Traditional Case-Control-Based Estimate of Vaccine Effectiveness

To obtain the final 2x2 table for a TCC study, we collapse across X and U for each week j and then sum over all weeks $1, \dots, J$. Finally, we can obtain the VE estimate from a TCC study using the OR from the final 2x2 table. The VE estimate from a TCC study is

$$VE_{TCC} = 1 - OR_{TCC} = \frac{a \times d_{TCC}}{b \times c_{TCC}},$$

where

$$\begin{aligned}
a &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{21u} \gamma_{j1x} \prod_{k=1}^{j-1} \left[\frac{1 - \gamma_{k1x} - \delta_{11u} \beta_{k1x}}{1 - \gamma_{k1x}} \right] \times \mathbb{P}(X = x, U = u | V = 0) \right] \right\} \\
b &= \sum_{j=1}^J \left\{ \sum_{x=0}^1 \sum_{u=0}^1 \left[\delta_{20u} \gamma_{j0x} \prod_{k=1}^{j-1} \left[\frac{1 - \gamma_{k0x} - \delta_{10u} \beta_{k0x}}{1 - \gamma_{k0x}} \right] \times \mathbb{P}(X = x, U = u | V = 0) \right] \right\} \\
c_{TCC} &= \left\{ \sum_{x=0}^1 \left[1 - \sum_{k=1}^J \gamma_{k0x} \right] \left[\prod_{k=1}^J (1 - \beta_{k0x}) \right] \times \mathbb{P}(X = x | V = 1) \right\} \\
d_{TCC} &= \left\{ \sum_{x=0}^1 \left[1 - \sum_{k=1}^J \gamma_{k0x} \right] \left[\prod_{k=1}^J (1 - \beta_{k0x}) \right] \times \mathbb{P}(X = x | V = 0) \right\}
\end{aligned}$$

Note, a and b are the same for TN and TCC studies.

Web Appendix 3

Input Parameters

Web Table 2 shows the parameter input values used in the calculations and simulations. Weekly values of the probabilities of non-influenza ARI (β_s) and influenza ARI (γ_s) are based on the prevalence of illness from surveillance data collected by the Centers for Disease Control and Prevention [1]. Parameters whose input values are a range were varied by 0.1 for each calculation scenario (see section 2.4).

Web Table 2: Input parameter values used in calculations and simulations.

Parameters	Input Values	Parameters	Input Values
π_{11}	0.4	γ_{101}	0.00103726
π_{10}	0.4	γ_{201}	0.0015696
π_{01}	0.1	γ_{301}	0.0018823
π_{11}	0.1	γ_{401}	0.00210545
α_{11}	0.6	γ_{501}	0.00237509
α_{10}	0.3	γ_{601}	0.0030613
α_{01}	0.9	γ_{701}	0.00394527
α_{00}	0.45	γ_{801}	0.00507414
β_{101}	0.020795	γ_{901}	0.006394935
β_{201}	0.02071	γ_{1001}	0.007294342
β_{301}	0.020814	γ_{1101}	0.007909946
β_{401}	0.020773	γ_{1201}	0.008418902
β_{501}	0.024733	γ_{1301}	0.008899444
β_{601}	0.027691	γ_{1401}	0.007516386
β_{701}	0.027538	γ_{1501}	0.005374278
β_{801}	0.027601	γ_{1601}	0.003840503
β_{901}	0.028493	γ_{1701}	0.002726108
β_{1001}	0.034245	γ_{1801}	0.000629997
β_{1101}	0.034375	θ_γ	0.563
β_{1201}	0.03441	ϕ_γ	1.0-2.0
β_{1301}	0.0333	δ_{101}	0.25
β_{1401}	0.027464	θ_{δ_1}	1.0
β_{1501}	0.027515	μ_{δ_1}	0.5-1.0
β_{1601}	0.027607	δ_{201}	0.40
β_{1701}	0.027593	θ_{δ_2}	0.5-1.0
β_{1801}	0.007901	μ_{δ_2}	0.5-1.0
θ_β	0.5-2.0	τ_1	0
ϕ_β	1.0-2.0	τ_2	1

We assume $\mu_{\delta_1} = \mu_{\delta_2}$.

Web Appendix 4

Stochastic Simulations

To validate our calculated results, we used a stochastic simulation program to simulate data from each type of study design. Each simulation consisted of a population of 25,000 people. Each person was assigned a health status, health awareness, and vaccination status. A random sample of 10,000 people were assigned to be the passive surveillance cohort. The remaining people in the population were eligible to be included in one of the case-control studies. The same parameter values used in the calculations were used in the simulation. The parameter value corresponding to the source of bias being assessed was varied by 0.1 while keeping all other parameter values fixed. 1000 simulations were performed for each combination of parameter values. True VE and estimated VE were determined for each simulation. True VE was simulated under random vaccination. VE estimates were obtained from the observed frequencies of cases and controls by vaccination status. The mean true and estimated VE values over all simulations were compared to the calculated values. All simulations were performed in R 3.3.1 [2].

Web Appendix 5

Sensitivity Analyses

5.1 Alternative Values of True Vaccine Effectiveness

To assess whether the value of true VE had an effect on the magnitude of bias, we performed our calculations using different values of true VE (0.2, 0.4, 0.6, and 0.8). Web Table 3 shows the bias of VE estimates under select sources of bias for different values of true VE. Based on our results, bias is dependent on the value of the true VE (particularly for the case-control studies). When the true VE is higher, bias is lower. Much of the biases from A and C in the case-control studies are reduced when the true VE is 0.8; however, if true VE is 0.2, the bias is even larger than our original scenario (true VE=0.437). We illustrate this phenomenon in Web Figure 1 where the magnitude of bias under source of bias A in TN-based VE estimates shrinks as the true VE increases. Our results suggest that more effective vaccines are more robust to sources of bias.

5.2 Alternative Probabilities of (X,U)

We selected two additional sets of input probabilities of X and U to assess the dependency of the amount of bias of VE estimates on our chosen probabilities X and U . Originally, we assumed independence between the probabilities of X and U (i.e., a person's health status did not affect their health awareness). In the first set of alternative probabilities, we selected values such that the probability of having high health awareness ($U = 1$) given a healthy health status ($X = 1$) is greater than the probability of having high health awareness given a frail health status ($X = 0$). In the second set of alternative probabilities, we selected probabilities for the

Web Table 3: Bias of Estimates of VE against Symptomatic Influenza and Medically-Attended Influenza in Observational Studies from Monte Carlo Simulations for Different Values of True VE: 5th, 50th, and 95th Quantiles of Bias Under Multiple Sources of Bias

Source of Bias ¹	Outcome of Interest	True VE	Passive Surveillance Cohort		Test-Negative		Traditional Case-Control	
			Q(5) ²	Q(50)	Q(5)	Q(50)	Q(5)	Q(50)
A	SI and MAI ³	0.20	0.00	0.00	-0.48	0.01	-0.21	0.02
		0.40	0.00	0.00	-0.34	0.01	-0.15	0.02
		0.60	0.00	0.00	-0.24	0.00	-0.09	0.02
		0.80	0.00	0.00	-0.12	0.00	-0.04	0.01
BS	SI and MAI	0.20	-0.08	-0.02	0.00	0.00	-0.10	-0.02
		0.40	-0.06	-0.02	0.00	0.00	-0.06	-0.01
		0.60	-0.04	-0.01	0.00	0.00	-0.03	0.00
		0.80	-0.02	-0.01	0.00	0.00	-0.01	0.00
C	SI	0.20	0.01	0.14	0.02	0.15	0.03	0.15
		0.40	0.01	0.1	0.01	0.11	0.03	0.12
		0.60	0.01	0.07	0.01	0.08	0.03	0.09
		0.80	0.00	0.04	0.01	0.04	0.02	0.05
D	SI and MAI	0.20	-0.16	-0.05	0.00	0.00	-0.13	-0.04
		0.40	-0.11	-0.04	0.00	0.00	-0.08	-0.02
		0.60	-0.08	-0.03	0.00	0.00	-0.05	0.00
		0.80	-0.04	-0.01	0.00	0.00	-0.02	0.00
BS,D,A	SI and MAI	0.20	-0.2	-0.09	-0.46	0.01	-0.41	-0.1
		0.40	-0.14	-0.06	-0.34	-0.01	-0.25	-0.05
		0.60	-0.09	-0.05	-0.25	-0.01	-0.19	-0.03
		0.80	-0.05	-0.02	-0.12	0.00	-0.08	-0.01
BS,D,A,C	SI	0.20	-0.11	0.08	-0.27	0.16	-0.20	0.07
		0.40	-0.09	0.05	-0.21	0.12	-0.17	0.07
		0.60	-0.06	0.04	-0.13	0.08	-0.10	0.05
		0.80	-0.03	0.02	-0.07	0.04	-0.05	0.03
BS,D,A,C	MAI	varies ⁴	-0.16	-0.07	-0.36	0.00	-0.32	-0.07
			-0.13	-0.05	-0.29	0.00	-0.23	-0.05
			-0.08	-0.03	-0.18	0.01	-0.15	-0.02
			-0.04	-0.02	-0.10	0.01	-0.07	-0.01

[1] Sources of bias: A - vaccination affects the probability of non-influenza ARI, BS - Healthy persons have a lower probability of non-influenza ARI and influenza ARI, C - Vaccination lowers the probability of seeking medical care in influenza ARI patients (because of reduced symptom severity), and D - ARI patients with high health awareness have a higher probability of seeking medical care.

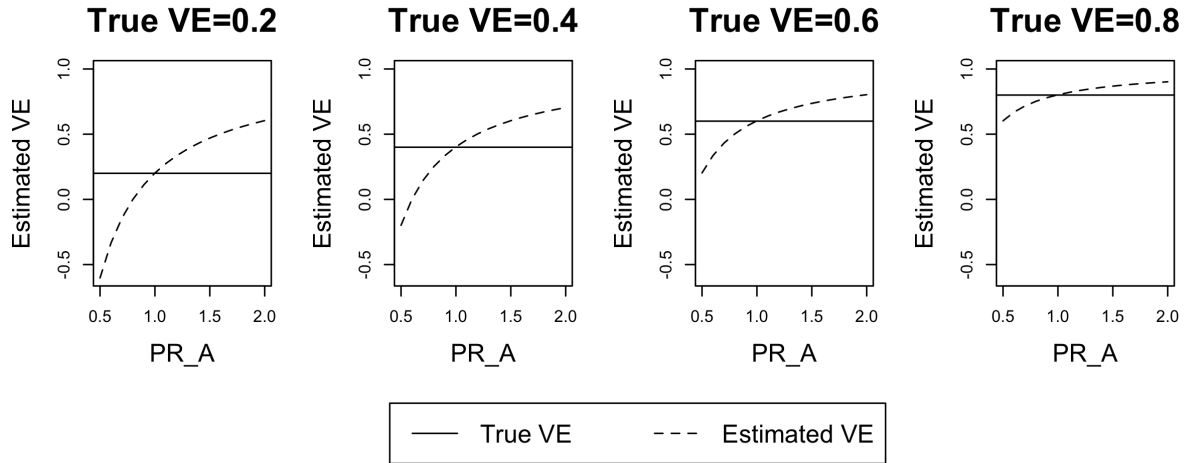
[2] Q(5)=5th quantile, Q(50)=50th quantile (median), and Q(95)=95th quantile.

[3] SI = symptomatic influenza, MAI = medically-attended influenza.

[4] True VE against MAI varies based on the value of the probability ratio when bias C was present.

The definitions of the parameters can be found in Table 3.

Web Figure 1: VE estimates from TN studies under source of bias A for different values of true VE



PR_A=probability ratio under source of bias A (vaccination affects the probability of non-influenza ARI).

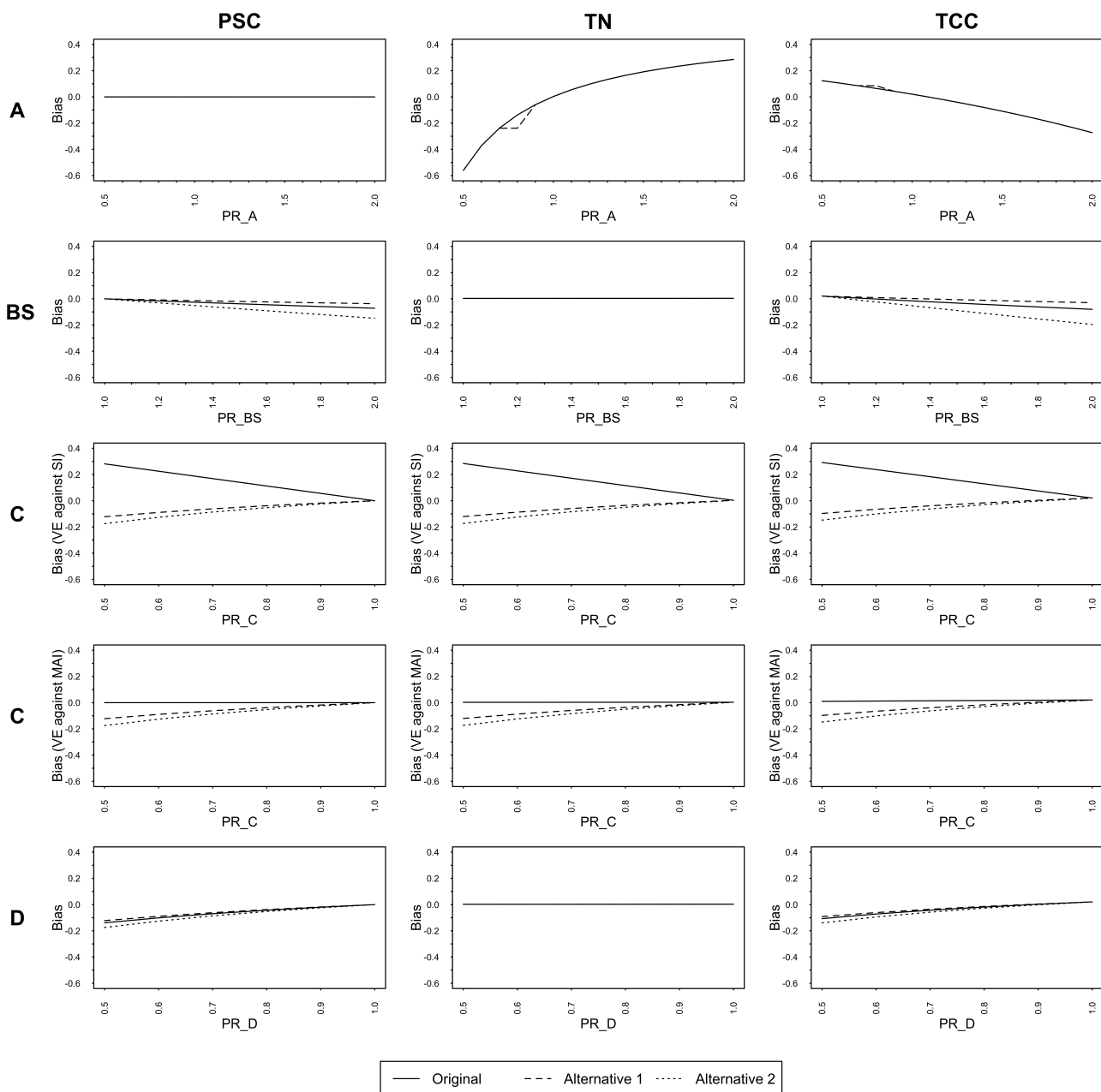
opposite scenario (i.e., the probability of having high health awareness given a healthy health status is *less* than the probability of having high health awareness given a frail health status). Web Table 4 shows the input probabilities for the original scenario and the two alternative scenarios.

Web Table 4: Alternative X and U probabilities.

Scenario	Parameters	Input Values
Original	π_{11}	0.40
	π_{10}	0.40
	π_{01}	0.10
	π_{00}	0.10
$\mathbb{P}(U = 1 X = 1) > \mathbb{P}(U = 1 X = 0)$	π_{11}	0.44
	π_{10}	0.36
	π_{01}	0.06
	π_{00}	0.14
$\mathbb{P}(U = 1 X = 1) < \mathbb{P}(U = 1 X = 0)$	π_{11}	0.32
	π_{10}	0.48
	π_{01}	0.18
	π_{00}	0.02

For each set of alternate input probabilities, bias was calculated in the presence of each source of bias for every study design. When A and C were present, bias was the same for both sets of alternative probabilities when compared to the original probabilities for all study designs. When BS was present, bias using both sets of alternative probabilities differed from the original calculations for all study designs. When D was present, bias was the same as the

Web Figure 2: Bias of VE estimates from PSC, TN, and TCC studies for alternative probabilities of (X, U)



Each row of plots corresponds to a single source of bias and each column of plots corresponds to a single study design. Under sources of bias A, BS, and D, the true VE against symptomatic influenza is equal to the true VE against MAI thus, the bias is the same for both outcomes of interest. Under source of bias C the true VE against MAI differs from the true VE against symptomatic influenza when the probability ratio is not equal to 1. If only a solid line is visible the bias is the same among the different scenarios.

original calculations for TN studies, but differed under both sets of alternative probabilities for PSC and TCC studies (Figure 2).

5.3 Alternative Probabilities of Vaccination

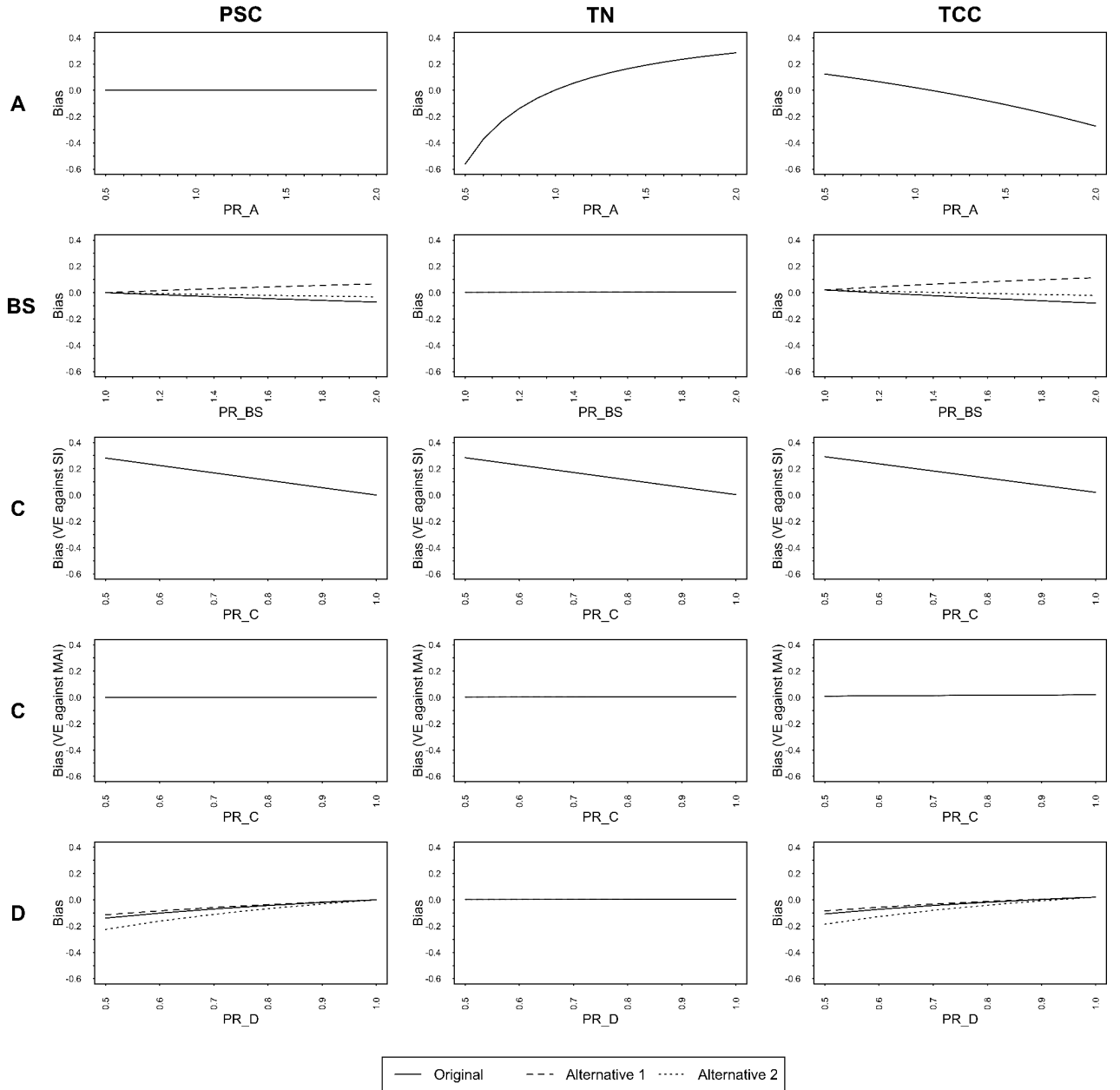
We selected two additional sets of input probabilities of vaccination to assess the dependency of the amount of bias of VE estimates on our chosen probabilities of vaccination. Originally, we selected values such that the probability of being vaccinated was multiplicatively higher in individuals with frail health status ($X = 0$), high health awareness ($U = 1$), or both ($X = 0, U = 1$). In the first set of alternative probabilities we selected values such that the probability of being vaccinated was lower in people with frail health status ($X = 0$), low health awareness ($U = 0$), or both ($X = 0, U = 0$). For the second set of alternative probabilities, we selected values such that persons with frail health status and high health awareness had a much higher probability of vaccination as we expect this group to have the highest probability of vaccination. Web Table 5 shows the alternative input probabilities.

Web Table 5: Alternative probabilities of vaccination.

Scenario	Parameters	Input Values
Original	α_{11}	0.60
	α_{10}	0.30
	α_{01}	0.90
	α_{00}	0.45
	α_{11}	0.60
Alternative 1	α_{10}	0.30
	α_{01}	0.30
	α_{00}	0.15
	α_{11}	0.60
	α_{10}	0.20
Alternative 2	α_{01}	0.95
	α_{00}	0.05

For each set of alternate input probabilities, bias was calculated in the presence of each source of bias for every study design. Bias was the same for each set of alternative input probabilities compared to the original input probabilities for all study designs when A and C were present. Bias of estimates of VE from the cohort studies was the same in the alternative scenarios compared to the original scenario. When BS was present, bias differed in the alternative scenarios compared to the original scenario for all study designs, except for bias in the TN study for alternative scenario 1. When D was present, bias was the same in the alternative scenarios compared to the original scenario for TN studies, but differed for all three alternative scenarios for PSC and TCC studies (Figure 3).

Web Figure 3: Bias of VE estimates from PSC, TN, and TCC studies for alternative probabilities of vaccination



Each row of plots corresponds to a single source of bias and each column of plots corresponds to a single study design. Under sources of bias A, BS, and D, the true VE against symptomatic influenza is equal to the true VE against MAI thus, the bias is the same for both outcomes of interest. Under source of bias C the true VE against MAI differs from the true VE against symptomatic influenza when the probability ratio is not equal to 1. If only a solid line is visible the bias is the same among the different scenarios.

5.4 Sensitivity and Specificity

We varied the values of influenza test sensitivity and specificity to assess how deviations from our assumption of 100% sensitivity and specificity impacted influenza VE estimates. We assessed

the median absolute value of bias for four combinations of sensitivity and specificity: (0.9, 0.9), (0.9, 1.0), (1.0, 0.9), and (1.0, 1.0). Our results suggest that VE estimates are robust to imperfect test sensitivity, but may be more biased when the test has less than perfect specificity (Web Table 6). However, when source of bias C is present and the outcome of interest is symptomatic influenza, estimates of VE from PSC, TN, and TCC have a smaller absolute value of bias when the specificity is 0.9. Additionally, when bias A is present estimates of VE from TN studies have similar median absolute value of bias regardless of the value of test sensitivity and specificity.

Web Table 6: Median Absolute Value of Bias of Estimated VE against Symptomatic Influenza and Medically-Attended Influenza in Observational Studies from Monte Carlo Simulations for Different Values of Influenza Test Sensitivity and Specificity

Source of Bias ¹	Outcome of Interest	Passive Surveillance Cohort		Test-Negative		Traditional Case-Control	
		Sensitivity	Specificity	Specificity	Specificity	Specificity	Specificity
A	SI and MAI ²	0.9	1.0	0.9	1.0	0.9	1.0
		0.9	0.13	0.14	0.11	0.11	0.05
BS	SI and MAI	1.0	0.12	0.12	0.12	0.10	0.05
		0.9	0.15	0.13	0.00	0.13	0.02
C	SI	1.0	0.14	0.11	0.00	0.12	0.01
		0.9	0.06	0.10	0.10	0.05	0.12
	MAI	1.0	0.05	0.10	0.11	0.04	0.12
		0.9	0.16	0.16	0.00	0.14	0.02
D	SI and MAI	1.0	0.15	0.14	0.00	0.12	0.02
		0.9	0.18	0.13	0.00	0.15	0.02
		1.0	0.16	0.12	0.00	0.13	0.02

[1] Sources of bias: A - vaccination affects the probability of non-influenza ARI, BS - Healthy persons have a lower probability of non-influenza ARI and influenza ARI, C - Vaccination lowers the probability of seeking medical care in influenza ARI patients (because of reduced symptom severity), and D - ARI patients with high health awareness have a higher probability of seeking medical care.

[2] SI = symptomatic influenza, MAI = medically-attended influenza.

References

- [1] Centers for Disease Control and Prevention. Influenza (flu), 2017.
- [2] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.