A Web-appendix of 'Estimating Disease Prevalence From Partially-sampled Clusters Using the Conditional Linear Family for Multivariate Bernoulli Data'

A.1 General Prevalence Estimator Derivation

A.1.1 Case Definition #1: At least 1 tooth affected with maximum PD $\geq$ 5mm

Let $W_j = 1$ if the maximum PD measurement across the four IP sites on the j-th tooth is greater or equal to 5mm, and 0 otherwise for $j = 1, \ldots, 28$. Then $\tilde{W} = \sum_{j=1}^{28} W_j$ is the total number teeth with PD greater or equal to 5mm for at least one IP site.

$$\pi_{PD} = \Pr(\tilde{W} \geq 1) = 1 - \Pr(\tilde{W} = 0)$$

$$\begin{aligned}
\Pr(\tilde{W} = 0) = \Pr\left(\sum_{j=1}^{28} W_j = 0\right) &= \Pr(W_1 = 0, W_2 = 0, \ldots, W_{27} = 0, W_{28} = 0) \\
&= \Pr(W_1 = 0)\Pr(W_2 = 0|W_1 = 0)\Pr(W_3 = 0|W_1 = 0, W_2 = 0)\cdots \\
&\quad \Pr(W_{28} = 0|W_1 = 0, \ldots, W_{27} = 0) \\
&= (1 - \Pr(W_1 = 1))\prod_{j=2}^{28}\Big(1 - \Pr(W_j = 1|W_1 = 0, \ldots, W_{j-1} = 0)\Big) \\
&= (1 - \Pr(W_1 = 1))\prod_{j=2}^{28}\left(1 - \Pr\left(W_j = 1|\sum_{k=1}^{j-1} W_k = 0\right)\right) = \prod_{j=1}^{28}(1 - \zeta_j)
\end{aligned}$$

where $\zeta_j = \Pr\left(W_j = 1|\sum_{k=1}^{j-1} W_k = 0\right)$ for $j \geq 2$ and $\zeta_1 = \Pr(W_1 = 1) = \mu_{W_1}$.

A.1.2 Case Definition #2: At least 2 teeth affected with maximum CAL $\geq$6mm

Let $Y_j = 1$ if the maximum CAL measurement across the four IP sites on the j-th tooth is greater or equal to 6mm, and 0 otherwise for $j = 1, \ldots, 28$. Then $\tilde{Y} = \sum_{j=1}^{28} Y_j$ is the total number of teeth in the individual's mouth with CAL greater or equal to 6mm for at least one IP site.

$$\pi_{CAL} = \Pr(\tilde{Y} \geq 2) = 1 - \Pr(\tilde{Y} = 0) - \Pr(\tilde{Y} = 1)$$

By similar logic as in case definition #1, $\Pr(\tilde{Y} = 0) = \prod_{j=1}^{28}(1 - \eta_j)$ where we define $\eta_j = \Pr(Y_j = 1|\sum_{k=1}^{j-1} Y_k = 0)$ for $j \geq 2$ and $\eta_1 = \Pr(Y_1 = 1) = \mu_{Y_1}$.

For the second probability,

$$
\begin{aligned}
\Pr(\tilde{Y}=1) &= \Pr\left(\sum_{j=1}^{28} Y_j = 1\right) \\
&= \sum_{j=1}^{28} \Pr\left(Y_j = 1, \sum_{k\neq j}^{28} Y_k = 0\right) \textit{ sum of all the probabilities that result in } \sum_{j=1}^{28} Y_j = 1 \\
&= \Pr(Y_1=1, Y_2=0,\ldots,Y_{28}=0) + \Pr(Y_1=0,Y_2=1,Y_3=0,\ldots,Y_{28}=0)+\ldots \\
&+ \Pr(Y_1=0,\ldots,Y_{26}=0,Y_{27}=1,Y_{28}=0) + \Pr(Y_1=0,\ldots,Y_{27}=0,Y_{28}=1).
\end{aligned}
$$

Consider 3 cases: (1) first $Y_1 = 1$ (2) middle $Y_k = 1$ (3) last $Y_{28} = 1$

First Term ($k = 1$):

$$
\begin{aligned}
&\Pr(Y_1=1,Y_2=0,\ldots,Y_{28}=0) \\
&\quad = \Pr(Y_1=1)\Pr(Y_2=0|Y_1=0)\Pr(Y_3=0|Y_1=0,Y_2=0)\cdots\Pr(Y_{28}=0|Y_1=0,\ldots,Y_{27}=0) \\
&\quad = \Pr(Y_1=1)\prod_{j=2}^{28}\Pr\left(Y_j=0|Y_1=1, \sum_{l=1;l\neq 1}^{j-1} Y_l = 0\right) = \eta_1 \prod_{j=2}^{28}\left(1-\eta_{j|1}\right)
\end{aligned}
$$

where $\eta_{j|1} = \Pr\left(Y_j = 1\middle|Y_1 = 1, \sum_{l=1;l\neq 1}^{j-1} Y_l = 0\right)$ and $\eta_1 = \Pr(Y_1 = 1) = \mu_{Y_1}$.

Middle Term ($1 < k < 28$):

$$
\begin{aligned}
&\Pr(Y_1=0,\ldots,Y_{k-1}=0,Y_k=1,Y_{k+1}=0,\ldots,Y_{28}=0) = \\
&\quad \Pr(Y_1=0)\Pr(Y_2=0|Y_1=0)\cdots\Pr(Y_{k-1}=0|Y_1=0,\ldots,Y_{k-2}=0)\cdot \\
&\quad \Pr(Y_k=1|Y_1=0,\ldots,Y_{k-1}=0)\Pr(Y_{k+1}=0|Y_1=0,\ldots,Y_{k-1}=0,Y_k=1)\cdots \\
&\quad \Pr(Y_{27}=0|Y_1=0,\ldots,Y_{k-1}=0,Y_k=1,Y_{k+1}=0,\ldots,Y_{26}=0)\cdot \\
&\quad \Pr(Y_{28}=0|Y_1=0,\ldots,Y_{k-1}=0,Y_k=1,Y_{k+1}=0,\ldots,Y_{27}=0) =
\end{aligned}
$$

$$
\begin{aligned}
\Pr(Y_1=0)\left[\prod_{j=2}^{k-1}\Pr\left(Y_j=0|\sum_{l=1}^{j-1}Y_l=0\right)\right]\Pr\left(Y_k=1|\sum_{l=1}^{k-1}Y_l=0\right)\left[\prod_{j=k+1}^{28}\Pr\left(Y_j=0\middle|Y_k=1,\sum_{l=1;l\neq k}^{j-1}Y_l=0\right)\right] \\
= \left[\prod_{j=1}^{k-1}\left(1-\eta_j\right)\right]\eta_k\left[\prod_{j=k+1}^{28}\left(1-\eta_{j|k}\right)\right]
\end{aligned}
$$

where $\eta_j = \Pr\left(Y_j = 1\middle| \sum_{l=1}^{j-1} Y_l = 0\right)$ for $j \geq 2, \eta_1 = \Pr(Y_1 = 1) = \mu_{Y_1}$, and $\eta_{j|k} = \Pr\Big(Y_j = 1\Big|Y_k = 1, \sum_{l=1;l\neq k}^{j-1} Y_l = 0\Big)$ for $j \geq 2$.

Special cases of "middle term"

$k = 2$: $(1-\eta_1)\eta_2\prod_{j=3}^{28}\left(1-\eta_{j|2}\right)$

$k = 27$: $\prod_{j=1}^{26}\big(1-\eta_j\big)\eta_{27}\big(1-\eta_{28|27}\big)$

Last Term ($k = 28$):

$\Pr(Y_1 = 0, \ldots, Y_{27} = 0, Y_{28} = 1) = \Pr(Y_1 = 0)\prod_{j=2}^{27}\Big[\Pr\Big(Y_j = 0\big|\sum_{l=1}^{j-1} Y_l = 0\Big)\Big]\Pr\Big(Y_{28} = 1\big|\sum_{l=1}^{n-1} Y_l = 0\Big) = \big[\prod_{j=1}^{27}(1-\eta_j)\big]\eta_{28}$.

Bring all the terms together:

$$\begin{aligned}
\Pr(\tilde{Y} = 1) &= \Pr\big(\sum_{j=1}^{28} Y_j = 1\big) = \text{1st Term} + \ldots + \text{Middle Terms} + \ldots + \text{28th Term} \\
&= \eta_1 \prod_{j=2}^{28}\big(1-\eta_{j|1}\big) + \sum_{k=2}^{27}\Big[\prod_{j=1}^{k-1}\big(1-\eta_j\big)\Big]\eta_k\Big[\prod_{j=k+1}^{28}\big(1-\eta_{j|k}\big)\Big] + \Big[\prod_{j=1}^{27}\big(1-\eta_j\big)\Big]\eta_{28} \\
&= \sum_{k=1}^{28} \eta_1^{I(k=1)}\Big[\eta_k\prod_{j=1}^{k-1}\big(1-\eta_j\big)\Big]^{I(k>1)}\Big[\prod_{j=k+1}^{28}\big(1-\eta_{j|k}\big)\Big]^{I(k<28)}
\end{aligned}$$

where (repeating the definitions above)

$$\eta_1 = \Pr(Y_1 = 1) = \mu_{Y_1}, \eta_j = \Pr\Big(Y_j = 1\big|\sum_{l=1}^{j-1} Y_k = 0\Big) \text{ for } j \geq 2, \text{ and}$$

$$\eta_{j|k} = \Pr\Big(Y_j = 1\big|Y_k = 1, \sum_{l=1;l\neq k}^{j-1} Y_l = 0\Big) \text{ for } j \geq 2.$$

A.1.3 Case Definition #3: Severe Periodontitis - $\pi_{severe} = \Pr(\tilde{W} \geq 1, \tilde{Y} \geq 2)$

Consider a 2x2 contingency table defined for $\tilde{W}$ and $\tilde{Y}$ defined by the thresholds of 1 and 2, respectively.

$$\begin{aligned}
1 - \pi_{severe} &= 1 - \Pr(\tilde{W} \geq 1, \tilde{Y} \geq 2) = \Pr(\tilde{W} = 0) + \Pr(\tilde{Y} \leq 1) - \Pr(\tilde{Y} \leq 1, \tilde{W} = 0) \\
&= \Pr(\tilde{W} = 0) + \Pr(\tilde{Y} \leq 1) - \Pr(\tilde{Y} \leq 1|\tilde{W} = 0)\Pr(\tilde{W} = 0) \\
&= \Pr(\tilde{Y} \leq 1) + \Pr(\tilde{W} = 0)\big(1 - \Pr(\tilde{Y} \leq 1|\tilde{W} = 0)\big) \\
&= \Pr(\tilde{Y} = 0) + \Pr(\tilde{Y} = 1) + \Pr(\tilde{W} = 0)\big(1 - \Pr(\tilde{Y} = 0|\tilde{W} = 0) - \Pr(\tilde{Y} = 1|\tilde{W} = 0)\big)
\end{aligned}$$

We will use the dental order for teeth with W_j's ordered first and then Y_j's, such that, $\mathbf{U} = (W_1 \ldots W_{28}, Y_1 \ldots Y_{28})^T$.

From case definition #1, $\Pr(\tilde{W} = 0) = \Pr\big(\sum_{j=1}^{28} W_j = 0\big) = \Pr\big(\sum_{j=1}^{28} U_j = 0\big) = \prod_{j=1}^{28}\big(1-\theta_j\big)$ where $\theta_j = \Pr\Big(U_j = 1\big|\sum_{l=1}^{j-1} U_l = 0\Big)$ and $\theta_1 = \Pr(W_1 = 1) = \mu_{W_1}$.

From case definition #2, $\Pr(\tilde{Y} = 0)$ and $\Pr(\tilde{Y} = 1)$ will be calculated from a vector of

Y values only, analogous to their definitions above, but using U_j terms.

By similar logic as that of case definition #1, $\Pr(\tilde{Y} = 0|\tilde{W} = 0) = \Pr\left(\sum_{j=1}^{28} Y_j = 0|\tilde{W} = 0\right) = \Pr\Big(\sum_{j=29}^{56} U_j = 0\big|\sum_{j=1}^{28} U_j = 0\Big) = \prod_{j=29}^{56}\left(1-\theta_j\right)$, where $\theta_j = \Pr\Big(U_j = 1\big|\sum_{l=1}^{j-1} U_l = 0\Big)$.

Next, $\Pr(\tilde{Y} = 1|\tilde{W} = 0) = \Pr\left(\sum_{j=1}^{28} Y_j = 1|\tilde{W} = 0\right) = \Pr\Big(\sum_{j=29}^{56} U_j = 1\big|\sum_{j=1}^{28} U_j = 0\Big) = \Pr\Big(U_{29} = 1, \sum_{j=30}^{56} U_j = 0\big|\sum_{j=1}^{28} U_j = 0\Big) + \ldots + \Pr\Big(U_{56} = 1, \sum_{j=29}^{55} U_j = 0\big|\sum_{j=1}^{28} U_j = 0\Big)$.

By similar logic as that of $\Pr(\tilde{Y} = 1)$ from case definition #2 with the $\mathbf{U}_{56,1}$ vector instead of the $\mathbf{Y}_{28,1}$ vector where only terms above 28 must equal 1.

First term: $\Pr\left(U_{29} = 1, \sum_{j=30}^{56} U_j = 0|\sum_{j=1}^{28} U_j = 0\right) = \theta_{29}\prod_{j=30}^{56}\left(1-\theta_{j|29}\right)$

Middle terms ($k = 30, \ldots, 55$):

$\Pr\Big(U_k = 1, \sum_{j=29;j\neq k}^{56} U_j = 0\big|\sum_{j=1}^{28} U_j = 0\Big) = \left[\prod_{j=29}^{k-1}\left(1-\theta_j\right)\right]\theta_k\left[\prod_{j=k+1}^{56}\left(1-\theta_{j|k}\right)\right]$

Last term: $\Pr\Big(U_{56} = 1, \sum_{j=29}^{55} U_j = 0\big|\sum_{j=1}^{28} U_j = 0\Big) = \left[\prod_{j=29}^{55}\left(1-\theta_j\right)\right]\theta_{56}$

Thus $\Pr(\tilde{Y} = 1|\tilde{W} = 0) = \sum_{k=29}^{56}\theta_k\left[\prod_{j=29}^{k-1}\left(1-\theta_j\right)\right]^{I(k>29)}\Big[\prod_{j=k+1}^{56}\left(1-\theta_{j|k}\right)\Big]^{I(k<56)}$,

where $\theta_j = \Pr\Big(U_j = 1\big|\sum_{l=1}^{j-1} U_l = 0\Big)$ and $\theta_{j|k} = \Pr\Big(U_j = 1\big|U_k = 1, \sum_{l=1;l\neq k}^{j-1} U_l = 0\Big)$.

A.2 CLF-based prevalence formulae

Under the working assumptions for the first two case definitions (common mean μ and exchangeable correlation ρ), Qaqish (2003) showed b_{jl} can be further simplified to

$$b_{jl} = \frac{\rho}{1+(j-2)\rho}, l = 1, \ldots, j-1$$

and λ_j can be further simplified to

$$\lambda_j = \frac{(1-\rho)\mu + \rho\sum_{l=1}^{j-1} Y_l}{1+(j-2)\rho}.$$

Applying these equations, the multivariate binary distribution in the CLF simplifies to the beta-binomial distribution (Qaqish, 2003). Using the following parameterization of the beta-binomial distribution for the number of affected teeth

$$\begin{aligned}\Pr(Y = t;\mu,\rho) &= C_t^n\frac{\prod_{j=1}^{t}\left(\mu+(j-1)\tau\right)\prod_{j=1}^{n-t}\left((1-\mu)+(j-1)\tau\right)}{\prod_{j=1}^{n}\left(1+(j-1)\tau\right)}\\ &= C_t^n\frac{\prod_{j=1}^{t}\left(\mu+(1-\rho)+(j-1)\rho\right)\prod_{j=1}^{n-t}\left((1-\mu)(1-\rho)+(j-1)\rho\right)}{\prod_{j=1}^{n}\left(1+(j-2)\rho\right)}\end{aligned}$$

where $n = 28$ is the number of teeth and $\tau = \frac{\rho}{1-\rho}$ for $0 \geq \rho \geq 1$, resulting in the prevalence formulae for π_{PD} and π_{CAL} found in Section 2.6 of the main article.

A.3 Asymptotic Variance Calculations

Following Wang and Preisser (2016), and given the logit link, the variance of $\hat{\mu}$ in either the first or second case definition can be estimated as follows:

$$\mu = \frac{e^{\beta}}{1+e^{\beta}} \text{ and } \mathrm{var}(\hat{\mu}) = \frac{\partial \mu}{\partial \beta}\mathrm{var}(\hat{\beta})\frac{\partial \mu}{\partial \beta} = \mu^2(1-\mu)^2\mathrm{var}(\hat{\beta})$$

For the first two case definitions, determining an expression for prevalence and calculating the variance of $\hat{\pi}$ requires manipulation of the beta-binomial distribution in Appendix A.2.

A.3.1 Case Definition #1: At least 1 tooth affected with maximum PD $\geq$ 5mm

From case definition #1, consider the prevalence formula with $\mu = \mu_{PD}$ and $\rho = \rho_{PD}$. Then

$$\pi_{PD} = \Pr(\tilde{W} \geq 1) = 1 - \Pr(\tilde{W} = 0) = 1 - \psi_0 = 1 - e^{\psi_0^*} = 1 - (1-\mu)\prod_{j=2}^{28}\left(1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right),$$

where $\psi_0^* = \ln(\psi_0)$ and

$$\begin{aligned}
\psi_0 = \Pr(\tilde{W} = 0|\mu,\rho) &= \frac{\prod_{j=1}^{28}(1-\mu+(j-1)\tau)}{\prod_{j=1}^{28}(1+(j-1)\tau)} = \prod_{j=1}^{28}\frac{1-\mu+(j-1)\tau}{1+(j-1)\tau} = \prod_{j=1}^{28}\frac{1-\mu+(j-1)\frac{\rho}{1-\rho}}{1+(j-1)\frac{\rho}{1-\rho}} \\
&= \prod_{j=1}^{28}\frac{\frac{(1-\mu)(1-\rho)+\rho(j-1)}{1-\rho}}{\frac{(1-\rho)+(j-1)\rho}{1-\rho}} = \prod_{j=1}^{28}\frac{(1-\mu)(1-\rho)+(j-1)\rho}{(1-\rho)+(j-1)\rho} = \frac{(1-\mu)(1-\rho)}{1+(1-2)\rho}\prod_{j=2}^{28}\frac{(1-\mu)(1-\rho)+(j-1)\rho}{1+(j-2)\rho} \\
&= (1-\mu)\prod_{j=2}^{28}\frac{(1-\mu)(1-\rho)+(j-1)\rho}{1+(j-2)\rho} = (1-\mu)\prod_{j=2}^{28}\frac{1-\rho-\mu+\mu\rho+(j-1)\rho}{1+(j-2)\rho} \\
&= (1-\mu)\prod_{j=2}^{28}\frac{1-j\rho-2\rho-\mu+\mu\rho}{1+(j-2)\rho} = (1-\mu)\prod_{j=2}^{28}\left(\frac{1-(j-2)\rho}{1+(j-2)\rho} - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right) \\
&= (1-\mu)\prod_{j=2}^{28}\left(1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right)
\end{aligned}$$

Then $\psi_0^* = \ln(1-\pi_{PD}) = \ln(1-\mu) + \sum_{j=2}^{28}\ln\left(1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right) = \ln(1-\mu) + \sum_{j=2}^{28}\ln\left(\frac{1+(j-2)\rho-(1-\rho)\mu}{1+(j-2)\rho}\right)$. With the logit link, $\mu = e^{\beta}/1+e^{\beta}$. This results in the following partial derivative:

$$\frac{\partial \mu}{\partial \beta} = \frac{e^{\beta}}{1+e^{\beta}} - \frac{e^{2\beta}}{(1+e^{\beta})^2} = \frac{e^{\beta}}{1+e^{\beta}}\left(1 - \frac{e^{\beta}}{1+e^{\beta}}\right) = \mu(1-\mu).$$

Therefore, by the delta method, $\text{var}(\hat{\mu}) = \frac{\partial \mu}{\partial \beta}\text{var}(\hat{\beta})\frac{\partial \mu}{\partial \beta} = \mu^2(1-\mu)^2\text{var}(\hat{\beta})$. An application of the chain rule is useful to calculate $\partial \psi_0^*/\partial \beta$.

$$\frac{\partial \psi_0^*}{\partial \beta} = \frac{\partial \psi_0^*}{\partial \mu} * \frac{\partial \mu}{\partial \beta} = \left(\frac{-1}{1-\mu} + \sum_{j=2}^{28} \left(\frac{-\frac{1-\rho}{1+(j-2)\rho}}{1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}} \right) \right) \mu(1-\mu)$$

where

$$\frac{\partial \psi_0^*}{\partial \mu} = \frac{-1}{1-\mu} + \sum_{j=2}^{28} \left(\frac{-\frac{1-\rho}{1+(j-2)\rho}}{1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}} \right) = \frac{-1}{1-\mu} + \sum_{j=2}^{28} \left(\frac{\rho - 1}{1+(j-2)\rho - (1-\rho)\mu} \right).$$

Calculation of $\partial \psi_0^*/\partial \rho$ is as follows:

$$\begin{aligned}
\frac{\partial \psi_0^*}{\partial \rho} &= \sum_{j=2}^{28} \left(\frac{1+(j-2)\rho}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{(j-2)+\mu}{1+(j-2)\rho} - \frac{(1+(j-2)\rho - (1-\rho)\mu)(j-2)}{(1+(j-2)\rho)^2} \right) \right) \\
&= \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{((j-2)+\mu)(1+(j-2)\rho)}{1+(j-2)\rho} - \frac{(1+(j-2)\rho - (1-\rho)\mu)(j-2)}{1+(j-2)\rho} \right) \right) \\
&= \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{j-2+\mu+(j-2)^2\rho + (j-2)\mu\rho}{1+(j-2)\rho} - \frac{j-2+(j-2)^2\rho - (1-\rho)\mu(j-2)}{1+(j-2)\rho} \right) \right) \\
&= \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{\mu + (j-2)\mu\rho}{1+(j-2)\rho} - \frac{-(1-\rho)\mu(j-2)}{1+(j-2)\rho} \right) \right) \\
&= \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{\mu(1+(j-2)\rho) + (j-2) - (j-2)\rho)}{1+(j-2)\rho} \right) \right) \\
&= \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{\mu(1+(j-2))}{1+(j-2)\rho} \right) \right) = \sum_{j=2}^{28} \left(\frac{1}{1+(j-2)\rho - (1-\rho)\mu} \left(\frac{\mu(j-1)}{1+(j-2)\rho} \right) \right)
\end{aligned}$$

Thus, with an application of the delta method, the variance for the prevalence based on case definition #1 is:

$$\text{var}(\hat{\pi}_{PD}) = \frac{\partial \pi_{PD}}{\partial \psi_0^*}\text{var}(\hat{\psi}_0^*)\frac{\partial \pi_{PD}}{\partial \psi_0^*} = e^{2\psi_0^*}\text{var}(\hat{\psi}_0^*),$$

where

$$\frac{\partial \pi_{PD}}{\partial \psi_0^*} = -e^{\psi_0^*}, \text{var}(\psi_0^*) = \begin{bmatrix} \frac{\partial \psi_0^*}{\partial \beta} & \frac{\partial \psi_0^*}{\partial \rho_{PD}} \end{bmatrix} \Sigma_{\beta,\rho_{PD}} \begin{bmatrix} \frac{\partial \psi_0^*}{\partial \beta} \\ \frac{\partial \psi_0^*}{\partial \rho_{PD}} \end{bmatrix}, \Sigma_{\beta,\rho_{PD}} = \begin{bmatrix} \sigma_\beta^2 & \sigma_{\beta\rho_{PD}} \\ \sigma_{\rho_{PD}\beta} & \sigma_{\rho_{PD}}^2 \end{bmatrix},$$

$$\frac{\partial \psi_0^*}{\partial \rho_{PD}} = \sum_{j=2}^{28} \left[\frac{1}{1+(j-2)\rho_{PD} - (1-\rho_{PD}\mu_{PD}} * \frac{\mu_{PD}(j-1)}{1+(j-2)\rho_{PD}} \right], \text{ and}$$

$$\frac{\partial \psi_0^*}{\partial \beta} = \left[\frac{-1}{1-\mu_{PD}} + \sum_{j=2}^{28} \frac{-(1-\rho_{PD})}{1+(j-2)\rho_{PD} - (1-\rho_{PD})\mu_{PD}} \right] \mu_{PD}(1-\mu_{PD}).$$

A.3.2 Case Definition #2: At least 2 teeth affected with maximum CAL $\geq$ 6mm

Let $\mu = \mu_{CAL}$ and $\rho = \rho_{CAL}$. Then

$$\begin{aligned}\pi_{CAL} &= \Pr(\tilde{Y} \geq 2) = 1 - [\Pr(\tilde{Y} = 0) + \Pr(\tilde{Y} = 1)] = 1 - (\psi_0 + \psi_1) \\ &= 1 - (1-\mu)\left(\frac{1 + 26\rho + 27\mu(1-\rho)}{1 + 26\rho}\right)\prod_{j=2}^{27}\left(1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right),\end{aligned}$$

where $\psi_0 = \Pr(\tilde{Y} = 0|\mu, \rho) = (1-\mu)\prod_{j=2}^{28}\left(1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right)$ and

$$\begin{aligned}\psi_1 = \Pr(\tilde{Y} = 1|\mu, \rho) &= C_1^n \frac{\prod_{j=1}^{1}\left(\mu + (j-1)\tau\right)\prod_{j=1}^{n-1}\left(1 - \mu + (j-1)\tau\right)}{\prod_{j=1}^{n}(1 + (j-1)\tau)} = n\frac{\mu\prod_{j=1}^{n-1}(1 - \mu + (j-1)\tau)}{1 + (n-1)\tau\prod_{j=1}^{n-1}(1 + (j-1)\tau)} \\ &= \frac{n\mu}{1 + (n-1)\tau}\prod_{j=1}^{n-1}\frac{1 - \mu + (j-1)\tau}{1 + (j-1)\tau} = \frac{n\mu}{1 + (n-1)\frac{\rho}{1-\rho}}\prod_{j=1}^{n-1}\frac{1 - \mu + (j-1)\frac{\rho}{1-\rho}}{1 + (j-1)\frac{\rho}{1-\rho}} \\ &= \frac{n\mu(1-\rho)}{(1-\rho) + (n-1)\rho}\prod_{j=1}^{n-1}\frac{(1-\mu)(1-\rho) + (j-1)\rho}{1 - \rho + (j-1)\rho} = \frac{n\mu(1-\rho)}{(1-\rho) + (n-1)\rho}(1-\mu)\prod_{j=2}^{n-1}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= \frac{n\mu(1-\rho)(1-\mu)}{(1-\rho) + (n-1)\rho}\prod_{j=2}^{n-1}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right]\end{aligned}$$

Then,

$$\begin{aligned}\psi_0 + \psi_1 &= (1-\mu)\prod_{j=2}^{28}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] + \frac{28\mu(1-\rho)(1-\mu)}{(1-\rho) + 27\rho}\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= (1-\mu)\left(1 - \frac{(1-\rho)\mu}{1 + (28-2)\rho} + \frac{28\mu(1-\rho)}{(1-\rho) + 27\rho}\right)\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= (1-\mu)\left(\frac{(1 + 26\rho - (1-\rho)\mu}{1 + 26\rho} + \frac{28\mu(1-\rho)}{1 - 26\rho}\right)\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= (1-\mu)\left(\frac{(1 + 26\rho - \mu + \rho\mu + 28\mu - 28\rho\mu}{1 + 26\rho}\right)\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= (1-\mu)\left(\frac{(1 + 26\rho + 27\mu - 27\rho\mu}{1 + 26\rho}\right)\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right] \\ &= (1-\mu)\left(\frac{(1 + 26\rho + 27\mu(1-\rho)}{1 + 26\rho}\right)\prod_{j=2}^{27}\left[1 - \frac{(1-\rho)\mu}{1 + (j-2)\rho}\right].\end{aligned}$$

Consider

$$\psi_{0,1} = \ln(\psi_0 + \psi_1) = \ln(1-\mu) + \ln(1+26\rho+27\mu-27\rho\mu) - \ln(1+26\rho) + \sum_{j=2}^{27} \ln\left(1 - \frac{(1-\rho)\mu}{1+(j-2)\rho}\right)$$

Then,

$$\text{var}(\psi^*_{0,1}) = \begin{bmatrix} \frac{\partial\psi^*_{0,1}}{\partial\beta} & \frac{\partial\psi^*_{0,1}}{\partial\rho_{CAL}} \end{bmatrix} \Sigma_{\beta,\rho_{CAL}} \begin{bmatrix} \frac{\partial\psi^*_{0,1}}{\partial\beta} \\ \frac{\partial\psi^*_{0,1}}{\partial\rho_{CAL}} \end{bmatrix},$$

where $\Sigma_{\beta,\rho_{CAL}} = \begin{bmatrix} \sigma^2_\beta & \sigma_{\beta\rho_{CAL}} \\ \sigma_{\rho_{CAL}\beta} & \sigma^2_{\rho_{CAL}} \end{bmatrix}$, $\frac{\partial\psi^*_{0,1}}{\partial\rho_{CAL}} = \sum_{j=2}^{27} \left[\frac{1}{1+(j-2)\rho_{CAL}-(1-\rho_{CAL})\mu_{CAL}} * \frac{\mu_{CAL}(j-1)}{1+(j-2)\rho_{CAL}}\right] + \frac{26-27\mu_{CAL}}{1+26\rho_{CAL}+27\mu_{CAL}(1-\rho_{CAL})} - \frac{26}{1+26\rho_{CAL}}$, $\frac{\partial\psi^*_{0,1}}{\partial\beta} = \frac{\partial\psi^*_{0,1}}{\partial\mu_{CAL}} * \frac{\partial\mu_{CAL}}{\partial\beta} = \left[\frac{-1}{1-\mu_{CAL}} + \frac{27(1-\rho_{CAL})}{1+26\rho_{CAL}+27\mu_{CAL}(1-\rho_{CAL})} + \sum_{j=2}^{27}\left(\frac{\rho_{CAL}-1}{1+(j-2)\rho_{CAL}-(1-\rho_{CAL})\mu_{CAL}}\right)\right]\mu_{CAL}(1-\mu_{CAL})$, and $\frac{\partial\psi^*_{0,1}}{\partial\mu_{CAL}} = \frac{-1}{1-\mu_{CAL}} + \frac{27(1-\rho_{CAL})}{1+26\rho_{CAL}+27\mu_{CAL}(1-\rho_{CAL})} + \sum_{j=2}^{27} \frac{\rho_{CAL}-1}{1+(j-2)\rho_{CAL}-(1-\rho_{CAL})\mu_{CAL}}$.

Thus, the variance for the prevalence in case definition #2 is:

$$\text{var}(\hat{\pi}_{CAL}) = \frac{\partial\pi_{CAL}}{\partial\theta^*_{CAL}} \text{var}(\hat{\theta}^*_{CAL}) \frac{\partial\pi_{CAL}}{\partial\theta^*_{CAL}} = e^{2\psi^*_{0,1}} \text{var}(\hat{\psi}^*_{0,1}),$$

where $\pi_{CAL} = 1 - e^{\psi^*_{0,1}}$ and $\frac{\partial\pi_{CAL}}{\partial\psi^*_{0,1}} = -e^{\psi^*_{0,1}}$.

A.4 Intermediate Parameter Tables

Table A.1: Intermediate Parameter Estimates and their Standard Errors from the Beta-Binomial Analysis Model for Case Definitions #1 and #2 (π_{PD} and π_{CAL}) in Simulation Study with 1,000 Replicates and 500 or 1,000 Individuals

		# of Individuals							
Data	PRP	500				1,000			
Generation	(# of teeth)	$\hat{\mu}$	SE_{μ}	$\hat{\rho}$	SE_{ρ}	$\hat{\mu}$	SE_{μ}	$\hat{\rho}$	SE_{ρ}
π_{PD}									
Model 1[1]	RAM (6)	0.0212	0.0035	0.1592	0.0430	0.0211	0.0025	0.1588	0.0318
	CPITN (10)	0.0211	0.0031	0.1579	0.0343	0.0211	0.0022	0.1586	0.0254
	RHM (14)	0.0210	0.0030	0.1567	0.0304	0.0211	0.0021	0.1585	0.0225
	FULL (28)	0.0211	0.0028	0.1582	0.0264	0.0211	0.0020	0.1590	0.0195
Model 2[2]	RAM (6)	0.0202	0.0034	0.1488	0.0423	0.0201	0.0024	0.1479	0.0310
	CPITN (10)	0.0201	0.0030	0.1492	0.0335	0.0200	0.0021	0.1477	0.0245
	RHM (14)	0.0201	0.0029	0.1469	0.0293	0.0200	0.0020	0.1483	0.0219
	FULL (28)	0.0201	0.0026	0.1484	0.0253	0.0200	0.0019	0.1480	0.0186
π_{CAL}									
Model 1	RAM (6)	0.0210	0.0035	0.1567	0.0425	0.0211	0.0025	0.1590	0.0319
	CPITN (10)	0.0210	0.0031	0.1556	0.0337	0.0211	0.0022	0.1587	0.0251
	RHM (14)	0.0210	0.0030	0.1566	0.0301	0.0210	0.0021	0.1586	0.0225
	FULL (28)	0.0210	0.0028	0.1575	0.0261	0.0211	0.0020	0.1591	0.0195
Model 2	RAM (6)	0.0200	0.0034	0.1469	0.0413	0.0201	0.0024	0.1467	0.0305
	CPITN (10)	0.0200	0.0030	0.1468	0.0321	0.0200	0.0021	0.1478	0.0239
	RHM (14)	0.0200	0.0029	0.1485	0.0292	0.0200	0.0020	0.1473	0.0211
	FULL (28)	0.0201	0.0026	0.1486	0.0248	0.0200	0.0019	0.1480	0.0180

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$

data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;, \pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$

data analysis model correctly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,

RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,

RHM = random half mouth, FULL = full-mouth exam

Table A.2: Intermediate Parameter Estimates and their Standard Errors from the Beta-Binomial Data Analysis Model for Case Definitions #1 (π_{PD}) in Simulation Study with 1,000 Replicates and 5,000 Individuals

Data	PRP	# of Individuals=5,000			
Generation	(# of teeth)	$\hat{\mu}$	SE_μ	$\hat{\rho}$	SE_ρ
Model 1[1]	RAM (6)	0.02099	0.00111	0.15915	0.01480
	CPITN (10)	0.02100	0.00100	0.15958	0.01180
	RHM (14)	0.02099	0.00095	0.15928	0.01055
	FULL (28)	0.02099	0.00088	0.15933	0.00909
Model 2[2]	RAM (6)	0.02000	0.00107	0.14939	0.01458
	CPITN (10)	0.02003	0.00096	0.15014	0.01152
	RHM (14)	0.01999	0.00091	0.14983	0.01025
	FULL (28)	0.02001	0.00084	0.15000	0.00876
Model 3[3]	RAM (6)	0.01499	0.00091	0.13397	0.01583
	CPITN (10)	0.01501	0.00081	0.13692	0.01255
	RHM (14)	0.01500	0.00078	0.14589	0.01167
	FULL (28)	0.01501	0.00071	0.13896	0.00958

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$

data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;, \pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$

data analysis model correctly specified

[3] $\mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{samequad}} = \rho_{CAL_{samequad}} = 0.17, \rho_{PD_{adjquad}} = \rho_{CAL_{adjquad}} = 0.14,$

$\rho_{PD_{CLquad}} = \rho_{CAL_{CLquad}} = 0.11; \rho_{same} = 0.19, \rho_{diff_{samequad}} = 0.11, \rho_{diff_{adjquad}} = 0.10, \rho_{diff_{CLquad}} = 0.09;$

$\pi_{PD} = 0.1552, \pi_{CAL} = 0.1138$, and $\pi_{severe} = 0.0637$; weighted estimate of exchangeable correlation,

$\rho_{PD} = \rho_{CAL} = (168 * 0.17 + 392 * 0.14 + 196 * 0.11)/756 = 0.13889$; data analysis model incorrectly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,

RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,

RHM = random half mouth, FULL = full-mouth exam

Table A.3: Intermediate Parameter Estimates and their Standard Errors from the Beta-Binomial Data Analysis Model for Case Definition #2 (π_{CAL}) in Simulation Study with 1,000 Replicates and 5,000 Individuals

Data Generation	PRP (# of teeth)	# of Individuals=5,000 $\hat{\boldsymbol{\mu}}$	$\mathbf{SE}_{\boldsymbol{\mu}}$	$\hat{\boldsymbol{\rho}}$	$\mathbf{SE}_{\boldsymbol{\rho}}$
Model 1[1]	RAM (6)	0.02099	0.00111	0.15970	0.01488
	CPITN (10)	0.02101	0.00100	0.15924	0.01179
	RHM (14)	0.02097	0.00095	0.15915	0.01053
	FULL (28)	0.02098	0.00088	0.15943	0.00909
Model 2[2]	RAM (6)	0.02000	0.00107	0.14899	0.01434
	CPITN (10)	0.02003	0.00096	0.14974	0.01131
	RHM (14)	0.01999	0.00091	0.14948	0.01000
	FULL (28)	0.02000	0.00084	0.14960	0.00852
Model 3[3]	RAM (6)	0.02001	0.00104	0.13397	0.01395
	CPITN (10)	0.02002	0.00093	0.13588	0.01095
	RHM (14)	0.02001	0.00090	0.14506	0.01023
	FULL (28)	0.02001	0.00081	0.13845	0.00847

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$ data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;, \pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$ data analysis model correctly specified

[3] $\mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{samequad}} = \rho_{CAL_{samequad}} = 0.17, \rho_{PD_{adjquad}} = \rho_{CAL_{adjquad}} = 0.14,$ $\rho_{PD_{CLquad}} = \rho_{CAL_{CLquad}} = 0.11; \rho_{same} = 0.19, \rho_{diff_{samequad}} = 0.11, \rho_{diff_{adjquad}} = 0.10, \rho_{diff_{CLquad}} = 0.09;$ $\pi_{PD} = 0.1552, \pi_{CAL} = 0.1138$, and $\pi_{severe} = 0.0637$; weighted estimate of exchangeable correlation, $\rho_{PD} = \rho_{CAL} = (168 * 0.17 + 392 * 0.14 + 196 * 0.11)/756 = 0.13889$; data analysis model incorrectly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,
RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,
RHM = random half mouth, FULL = full-mouth exam

Table A.4: Intermediate Mean Parameter Estimates and their Standard Errors for π_{SEV} (Case Definition #3) in Simulation Study with 1,000 Replicates and 500 or 1,000 Individuals

		# of Individuals							
Data	PRP	500				1,000			
Generation	(# of teeth)	$\hat{\mu}_{PD}$	$SE_{\mu_{PD}}$	$\hat{\mu}_{CAL}$	$SE_{\mu_{CAL}}$	$\hat{\mu}_{PD}$	$SE_{\mu_{PD}}$	$\hat{\mu}_{CAL}$	$SE_{\mu_{CAL}}$
Model 1[1]	RAM (6)	0.0212	0.0035	0.0210	0.0035	0.0211	0.0025	0.0211	0.0025
	CPITN (10)	0.0211	0.0032	0.0210	0.0031	0.0211	0.0022	0.0211	0.0022
	RHM (14)	0.0210	0.0030	0.0210	0.0030	0.0211	0.0021	0.0210	0.0021
	FULL (28)	0.0211	0.0028	0.0210	0.0028	0.0211	0.0020	0.0211	0.0020
Model 2[2]	RAM (6)	0.0202	0.0034	0.0200	0.0033	0.0201	0.0024	0.0201	0.0024
	CPITN (10)	0.0201	0.0030	0.0200	0.0030	0.0200	0.0021	0.0200	0.0021
	RHM (14)	0.0201	0.0028	0.0200	0.0029	0.0200	0.0020	0.0200	0.0020
	FULL (28)	0.0201	0.0026	0.0201	0.0026	0.0200	0.0019	0.0200	0.0019

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$
data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;, \pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$
data analysis model correctly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,
RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,
RHM = random half mouth, FULL = full-mouth exam

Table A.5: Intermediate Mean Parameter Estimates and their Standard Errors for π_{SEV} (Case Definition #3) in Simulation Study with 1,000 Replicates and 5,000 Individuals

Data Generation	PRP (# of teeth)	# of Individuals=5,000			
		$\hat{\boldsymbol{\mu}}_{\boldsymbol{PD}}$	$\mathbf{SE}_{\boldsymbol{\mu_{PD}}}$	$\hat{\boldsymbol{\mu}}_{\boldsymbol{CAL}}$	$\mathbf{SE}_{\boldsymbol{\mu_{CAL}}}$
Model 1[1]	RAM (6)	0.02099	0.00111	0.02099	0.00111
	CPITN (10)	0.02100	0.00100	0.02101	0.00100
	RHM (14)	0.02099	0.00095	0.02097	0.00095
	FULL (28)	0.02099	0.00088	0.02098	0.00088
Model 2[2]	RAM (6)	0.02000	0.00107	0.02000	0.00107
	CPITN (10)	0.02003	0.00096	0.02003	0.00096
	RHM (14)	0.01999	0.00091	0.01999	0.00091
	FULL (28)	0.02001	0.00084	0.02000	0.00084
Model 3[3]	RAM (6)	0.01499	0.00090	0.020006	0.00104
	CPITN (10)	0.01501	0.00081	0.020024	0.00093
	RHM (14)	0.01500	0.00078	0.020010	0.00090
	FULL (28)	0.01501	0.00071	0.020010	0.00081

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$ data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;$, $\pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$ data analysis model correctly specified

[3] $\mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{samequad}} = \rho_{CAL_{samequad}} = 0.17, \rho_{PD_{adjquad}} = \rho_{CAL_{adjquad}} = 0.14,$ $\rho_{PD_{CLquad}} = \rho_{CAL_{CLquad}} = 0.11; \rho_{same} = 0.19, \rho_{diff_{samequad}} = 0.11, \rho_{diff_{adjquad}} = 0.10, \rho_{diff_{CLquad}} = 0.09;$ $\pi_{PD} = 0.1552, \pi_{CAL} = 0.1138$, and $\pi_{severe} = 0.0637$; weighted estimate of exchangeable correlation, $\rho_{PD} = \rho_{CAL} = (168 * 0.17 + 392 * 0.14 + 196 * 0.11)/756 = 0.13889$; data analysis model incorrectly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,
RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,
RHM = random half mouth, FULL = full-mouth exam

Table A.6: Intermediate Pairwise Correlation Parameter Estimates for π_{SEV} in Simulation Study with 1,000 Replicates with 500 and 1,000 Individuals

		# of Individuals							
Data	PRP	500				1,000			
Generation	(# of teeth)	$\hat{\rho}_{PD}$	$\hat{\rho}_{CAL}$	$\hat{\rho}_{same}$	$\hat{\rho}_{diff}$	$\hat{\rho}_{PD}$	$\hat{\rho}_{CAL}$	$\hat{\rho}_{same}$	$\hat{\rho}_{diff}$
Model 1[1]	RAM (6)	0.1592	0.1567	0.4974	0.1482	0.1588	0.1590	0.5009	0.1495
	CPITN (10)	0.1579	0.1556	0.4986	0.1471	0.1586	0.1587	0.5016	0.1487
	RHM (14)	0.1567	0.1566	0.4974	0.1470	0.1585	0.1586	0.4996	0.1485
	FULL (28)	0.1582	0.1575	0.4980	0.1479	0.1590	0.1591	0.5003	0.1491
Model 2[2]	RAM (6)	0.1488	0.1469	0.1878	0.0885	0.1479	0.1467	0.1872	0.0886
	CPITN (10)	0.1492	0.1468	0.1885	0.0890	0.1477	0.1478	0.1881	0.0891
	RHM (14)	0.1469	0.1485	0.1888	0.0883	0.1483	0.1473	0.1882	0.0893
	FULL (28)	0.1484	0.1486	0.1882	0.0885	0.1480	0.1480	0.1882	0.0889

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$

data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;, \pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$

data analysis model correctly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,

RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,

RHM = random half mouth, FULL = full-mouth exam

Table A.7: Intermediate Pairwise Correlation Parameter Estimates for π_{SEV} in Simulation Study with 1,000 Replicates with 5,000 Individuals

Data	PRP	# of Individuals=5,000			
Generation	(# of teeth)	$\hat{\rho}_{PD}$	$\hat{\rho}_{CAL}$	$\hat{\rho}_{same}$	$\hat{\rho}_{diff}$
Model 1[1]	RAM (6)	0.15915	0.15970	0.50059	0.14956
	CPITN (10)	0.15958	0.15924	0.50024	0.14950
	RHM (14)	0.15928	0.15915	0.49973	0.14927
	FULL (28)	0.15933	0.15943	0.49966	0.14938
Model 2[2]	RAM (6)	0.14939	0.14899	0.19010	0.08958
	CPITN (10)	0.15014	0.14974	0.18965	0.08986
	RHM (14)	0.14983	0.14948	0.18966	0.08975
	FULL (28)	0.15000	0.14960	0.18984	0.08986
Model 3[3]	RAM (6)	0.13397	0.13397	0.19051	0.09830
	CPITN (10)	0.13692	0.13588	0.18999	0.09882
	RHM (14)	0.14589	0.14506	0.19026	0.10195
	FULL (28)	0.13896	0.13845	0.19012	0.09973

[1] $\mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \pi_{PD} = 0.1925, \pi_{CAL} = 0.1149, \pi_{severe} = 0.1008$
data analysis model correctly specified

[2] $\mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09;$, $\pi_{PD} = 0.1907, \pi_{CAL} = 0.1118, \pi_{severe} = 0.0644$
data analysis model correctly specified

[3] $\mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{samequad}} = \rho_{CAL_{samequad}} = 0.17, \rho_{PD_{adjquad}} = \rho_{CAL_{adjquad}} = 0.14,$
$\rho_{PD_{CLquad}} = \rho_{CAL_{CLquad}} = 0.11; \rho_{same} = 0.19, \rho_{diff_{samequad}} = 0.11, \rho_{diff_{adjquad}} = 0.10, \rho_{diff_{CLquad}} = 0.09;$
$\pi_{PD} = 0.1552, \pi_{CAL} = 0.1138$, and $\pi_{severe} = 0.0637$;
weighted estimate of exchangeable correlation, $\rho_{PD} = \rho_{CAL} = (168 * 0.17 + 392 * 0.14 + 196 * 0.11)/756 = 0.13889$;
weighted estimate of common ρ_{diff} correlation, $\rho_{diff} = 0.09963$; data analysis model incorrectly specified

PRP = partial-mouth recording protocol, PD = pocket depth, CAL = clinical attachment loss,
RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol,
RHM = random half mouth, FULL = full-mouth exam

A.5 Analysis of Distributions from each Replicate and PRP

Intermediate estimates of the marginal means and pair-wise correlation matrix for each replicate and PRP could result in a distribution not from the CLF. GEE estimates of the marginal means and pair-wise correlation matrix resulted in a CLF distribution at most 62% of the time and at minimum 16% of the time. From the restrictions depicted in Figure 3, estimates from the first data generation model were more likely to not result in a CLF distribution. This happened for 16.1% of the models fit with the six selected teeth using RAM with 500 individuals. At best, estimates based on the first data generation model resulted in a CLF distribution 56% of the time when using full mouth data and a sample size of 1000. The second data generation model produced estimates that resulted in a CLF distribution between 46% and 62% of the time. The best results occurred with more teeth and larger sample sizes. These results are shown in Table A.8.

When this occurred, prevalence estimates were evaluated from the CLF formula. A sensitivity analysis was conducted using the bootstrap samples to compare estimates from non-CLF distributions and CLF distributions to the true values. It found at most a percent bias of 15% observed for the pair-wise correlation for CAL. For the most part, the severity prevalence and model parameters had percent bias estimates less than 5%.

It was rare for the GEE estimates to produce a pair-wise correlation matrix that violated the Fréchet bounds. This only occurred in the bootstrap samples. When it did occur, the data was removed from the analysis. Smaller sample sizes, smaller cluster sizes, and the number of available CLF distributions appear to affect the GEE estimates producing negative probabilities. When this occurred a Frobenius norm adjustment was applied to the pair-wise correlation matrix. In the simulations, this happened a total of 14 times with 12 of those times occurring with the RAM PRP for the first data generation model; all of these violations were observed with a sample size of 500. This occurred a bit more often in the bootstrap samples but was largely observed for the RAM PRP and the first data generation model with sample sizes of 500 or 1000.

Table A.8: Model Fitting Results from 1,000 Simulations and 200,000 Bootstrap Samples

# of Individuals	500				1000				5000			
Data	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
Generation Model	Sim	Boot	Sim	Boot	Sim	Boot	Sim	Boot	Sim	Boot	Sim	Boot
Non-CLF Distribution												
RAM	75.6%	83.9%	47.0%	54.3%	66.4%	75.3%	43.0%	43.4%	46.6%	50.0%	45.2%	44.5%
CPITN	60.2%	70.8%	41.7%	43.6%	50.4%	59.2%	41.4%	41.5%	46.5%	46.1%	43.0%	43.4%
RHM	52.9%	60.5%	3.6%	43.2%	46.5%	51.0%	39.6%	40.0%	48.2%	46.1%	40.4%	42.4%
FULL	45.9%	47.2%	44.8%	42.7%	43.8%	44.4%	41.4%	40.9%	48.6%	46.9%	38.1%	41.3%
Fréchet Bounds Violations												
RAM	0%	0.1%	0%	0.1%	0%	0%	0%	0%	0%	0%	0%	0%
CPITN	0%	0%	0%	1.5%	0%	0%	0%	0%	0%	0%	0%	0%
RHM	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
FULL	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Adjustments to Pair-wise Correlation Matrix												
RAM	12	12158	1	1010	0	2757	0	12	0	0	0	0
CPITN	1	727	0	2	0	12	0	0	0	0	0	0
RHM	0	2	0	0	0	0	0	0	0	0	0	0
FULL	0	0	0	0	0	0	0	0	0	0	0	0

CLF = conditional linear family, RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol, RHM = random half mouth, FULL = full mouth exam

References

Qaqish, B. F. (2003), "A family of multivariate binary distributions for simulating correlated binary variables with specified marginal means and correlations," *Biometrika*, 90, 455–463.

Wang, R. and Preisser, J. S. (2016), "Prevalence estimation at the cluster level for correlated binary data using random partial-cluster sampling," *The University of North Carolina at Chapel Hill Department of Biostatistics Technical Report Series*, Working Paper 46.