Journal of Statistical Research *https://doi.org/10.3329/jsr.v58i1.75408*
2024, Vol. 58, No. 1, pp. 3-31 *ISSN 0256 - 422 X*

ESTIMATING DISEASE PREVALENCE FROM PARTIALLY-SAMPLED CLUSTERS USING THE CONDITIONAL LINEAR FAMILY FOR MULTIVARIATE BERNOULLI DATA

SUSAN L. EDWARDS[⋆,†], JOHN S. PREISSER, BAHJAT QAQISH
Department of Biostatistics, University of North Carloina at Chapel Hill
Gillings School of Global Public Health, Chapel Hill, NC 27599
Email: SusanE@live.unc.edu, John_Preisser@unc.edu, Bahjat_Qaqish@unc.edu

CRISTIANO SUSIN
Adams School of Dentistry, University of North Carloina at Chapel Hill, Chapel Hill, NC 27599
Email: CSusin@email.unc.edu

SUMMARY

In periodontal disease surveillance in human populations, full-mouth clinical examinations to classify the disease status of individuals are the gold standard for estimating periodontitis prevalence. However, conducting full-mouth exams is resource intensive, time consuming, and costly, especially in studies involving thousands of participants. Partial-recording protocols have been utilized in oral health surveys worldwide to gather correlated binary outcomes of periodontal disease on selected teeth in lieu of full-mouth exams. Since the use of partial-recording protocols tends to underestimate disease prevalence, a statistical distributional approach considering the pattern of tooth-level disease in the mouth was proposed to substantially reduce bias for the estimation of periodontitis prevalence. This approach employed multivariate Bernoulli distributions for observation (tooth)-level disease indicators to define formulae for the prevalence of disease (periodontitis) at the cluster (person)-level for various full-mouth case definitions. In turn, prevalence estimators were based on plug-in estimates of parameters from a conditional linear family for binary data gathered under partial recording protocols. Work in this article extended existing prevalence estimators for simple case definitions based on single clinical measures of tooth-level periodontal disease to a definition of severe periodontitis using two measures as defined by the Centers for Disease Control and Prevention and the American Academy of Periodontology, and later adopted by the 2017 World Workshop in Periodontology. Simulations evaluated the finite-sample performance of the proposed estimators and their confidence intervals for three established partial-recording protocols. In general, the prevalence estimators performed well relative to bias and coverage when tooth-level probabilities of disease and within-mouth correlation structures were correctly specified and even when the pattern of tooth-pair correlations was misspecified.

Keywords and phrases: correlated binary data; missing by design; computer simulation; periodontal disease(s)/periodontitis; partial-recording protocol; surveillance

---

⋆ Corresponding author
† Center for Official Statistics, Analytics Practice Area, RTI International, Research Triangle Park, NC 27709

1 Introduction

1.1 Partial-mouth recording protocols and case definitions in the surveillance of periodontitis

The surveillance of health and disease in human populations requires practical methods for ongoing or periodic ascertainment. For example, the National Health and Nutrition Examination Survey (NHANES) has monitored the oral health of the U.S. population since the early 1970s through clinical examinations of the oral cavity. The level of disease is often characterized by prevalence, which is the proportion of individuals in a population that have the disease, according to a specific case definition, at a given time point. Periodontitis is chronic inflammatory disease that leads to the destruction of the supporting tissues around teeth, and at its late stages to tooth loss, impairing function, esthetics, quality of life, and systemic health. The global prevalence of periodontitis is estimated to be approximately 11% and more than 1 billion individuals are affected (Chen et al., 2021). In a full-mouth examination for the surveillance of periodontitis, i.e., gum disease, one or more periodontal measures are recorded on up to 6 sites per tooth from up to 28 teeth in the adult dentition, where third molars are commonly excluded. Thresholds are applied to pocket depth (PD) and/or clinical attachment loss (CAL), measured in full millimeters, at each tooth site, resulting in a set of biologically clustered binary disease indicators. In the surveillance of periodontitis, a case definition is a rule that when applied to the indicators classifies the individual as having periodontitis or not. In this article, we refer to disease status at the site or tooth level (presence vs. absence) as periodontal disease and define periodontitis for person level disease.

In the NHANES, and other surveillance studies, estimating the prevalence of periodontitis in populations has been challenging, in part, because case definitions have varied across studies and over time. While definitions of periodontitis for use in epidemiological surveys often consider multiple clinical factors and continue to undergo development (Tonetti et al., 2018; Holtfreter et al., 2024), most case definitions historically have been based on PD and/or CAL. The simplest definitions have been based on a single cardinal measure, PD or CAL, which often allows direct comparison to results from other studies (Kingman and Albandar, 2002; Susin et al., 2005; Beck et al., 2006; Eke et al., 2010; Alshihayb et al., 2022). Because single measure case definitions were considered inadequate given the different clinical manifestations of periodontal tissue inflammation and destruction (Page and Eke, 2007), definitions of periodontitis with spatial requirements based on thresholds for both PD and CAL were introduced by the Centers for Disease Control and Prevention in conjunction with the American Academy for Periodontology (CDC/AAP) in 2012 (Eke et al., 2012).

In periodontitis surveillance, full-mouth clinical examinations resulting in disease classification of individuals – based on a case definition – are the gold standard to estimate the prevalence of periodontitis. However, conducting full-mouth examinations, which may require up to 30 minutes per person, are costly, resource intensive, increase participant burden, and are often impractical for population research and surveillance, especially in large epidemiological cohort studies involving thousands of participants. To address these barriers, partial recording protocols (PRPs), by which we mean partial-mouth exams, have been utilized in national surveys, including the NHANES from 1988 to 2009, to gather correlated binary outcomes of periodontal disease on a subset of selected

teeth and sites in lieu of time-consuming full-mouth periodontal exams. Specific PRPs, considered in this article, include the Ramfjord (Ramfjord, 1959), the Community Periodontal Index for Treatment Needs (CPITN) (Ainamo, 1982; Chattopadhyay et al., 2008), and random half-mouth (RHM) protocols. Ramfjord and CPITN are fixed site selection methods (FSSMs) where the same set of tooth sites are recorded for each study participant. In contrast, RHM protocols have a two-stage sampling sequence where two mouth quadrants are randomly sampled followed by the selection of fixed tooth sites within all teeth in the selected quadrants. NHANES III from 1988-1994 and the NHANES IV from 1999-2004 both utilized a RHM protocol for periodontal data collection.

Unfortunately, partial-mouth exams underestimate the prevalence of periodontitis when disease from tooth sites not selected in the PRP goes undetected and unconsidered in the application of the case definition. Specifically, assuming no measurement error at the tooth site level, there are false negatives but not false positives at the participant level (Susin et al., 2005; Preisser et al., 2017). The sizeable underestimation of prevalence using the standard disease classification estimator with two different PRPs prompted the NHANES to cease use of partial-mouth exams and implement a full-mouth examination protocol in 2009, albeit with less frequent use over time; NHANES has exclusively used full-mouth periodontal exams ever since. Considering the cost-saving benefit from PRPs, this article develops novel statistical approaches to the estimation of periodontitis prevalence that could ultimately contribute to a return to the regular use of PRPs in large oral epidemiological studies.

1.2 Statistical considerations on estimating periodontitis prevalence

As an alternative to the standard classification prevalence estimator for PRPs, Preisser et al. (2017) proposed formulae for prevalence, defined as the probability of disease for an individual in a homogeneous population. In particular, the formulae are based on an assumed multivariate distribution for the correlated binary disease indicators where each mouth is defined as a cluster and the disease classification at each tooth site is a Bernoulli random variable. While the statistical distribution method (SDM) is a general approach to prevalence estimation in PRPs, it was originally implemented for a single-measure threshold case definition of one or more tooth sites affected. Preisser et al. (2024) extended the SDM to a single-measure threshold case definition of two or more tooth sites affected. In SDM, the data are assumed to follow a multivariate binary distribution in the conditional linear family (CLF) of distributions (Qaqish, 2003b). A particular CLF member is specified by a set of tooth site marginal means (i.e, the probabilities of having disease at each tooth site) and the pairwise correlation matrix of these disease indicators across tooth sites in the full-mouth. In the case of missing teeth, "full-mouth" refers to all of an individual's existing teeth (Preisser et al., 2024).

While the CLF allows for flexible specification of the marginal means and pairwise correlations in the SDM approach, Preisser et al. (2017) defined a prevalence formula for the clustered data that was derived under two working assumptions: (i) the probability of having disease is the same across all sites; and (ii) the correlation of disease status is the same for all pairs of sites within the mouth. Even though these assumptions were unrealistic, using oral examination data from 6,793 participants in the Arteriolosclerosis Risk in Communities study, the new formula yielded periodontitis prevalence estimates from PRPs that were much closer to full-mouth estimates than those based

on disease classification of study participants. Resampling of the cohort showed the proposed estimators gave good precision and accuracy for as few as six tooth sites sampled per individual. On the other hand, Wang and Preisser (2016) conducted simulation studies finding this estimator resulted in approximately 10% bias when the simple assumptions of common disease probability and exchangeable correlation did not hold.

Departing from the overly simple two-parameter CLF model of a constant mean and exchangeable within-mouth correlation model, Shing et al. (2023) introduced flexibly-specified marginal mean and correlation models, which were estimated with generalized estimating equations (GEE), that allowed unequal probabilities of disease across tooth sites and spatial correlation patterns of disease in the mouth. Preisser et al. (2024) employed multiple imputation to impute binary disease status for tooth sites not selected in PRPs via a similarly enriched (i.e., with multiple parameters) CLF model, which, in turn, allowed case classification of individuals based on their imputed full-mouth data. To handle the complexity of the CLF model, Monte Carlo methods were used to estimate the variance of the prevalence estimators. Despite the improvements of SDM relative to the simple estimators of Preisser et al. (2017), the application in Preisser et al. (2024) was limited to single-measure (PD or CAL) threshold case definitions. Furthermore, the evaluations in both articles were limited to random site selection methods (RSSMs), which are PRPs that take a simple random sample of sites from among all 168 sites (excluding third molars) in the adult dentition (Beck et al., 2006). The use of RSSMs are challenging because of the inherent variability introduced in the clinical examination protocols and data recording since different teeth and tooth sites are selected for each participant. To our knowledge, RSSMs have not been used in practice, and they are not considered in this article.

Since the case status of individuals cannot be reliably determined from partial-mouth sampling, which precludes the valid application of the classification approach to prevalence estimation, Section 2 of this article introduces new formulae for disease prevalence as defined by the CDC/AAP severe case definition based on a combination of PD and CAL thresholds. It further includes two single measure case definitions as defined by the marginal components of the CDC/AAP severe case definition for which we assume, as in Preisser et al. (2017), a CLF that implicitly defines a common tooth-level probability of disease (μ) and exchangeable pairwise correlation (ρ) between teeth. Unlike much of the PRP literature including Preisser et al. (2017), the proposed approach applies to tooth-level indicators as the standard tooth-site level case definitions in this article can be reduced to equivalent tooth-level definitions, which simplifies prevalence computations. For the CDC/AAP severe case definition, we define a CLF distribution that has distinct tooth-level means for PD and CAL measures (μ_{PD} and μ_{CAL}, respectively) and a multi-parameter, within-mouth correlation structure. Prevalence estimates are obtained by plugging in GEE estimates (Preisser et al., 2024) of the mean and correlation parameters into the prevalence formulae. Variance estimators are proposed for the prevalence estimators based the delta method (Wang and Preisser, 2016) for the single-threshold case definitions and bootstrap methods for the CDC/AAP severe disease case definition, following the Monte Carlo estimation methods of Preisser et al. (2024). In Section 3, simulation studies evaluate the finite-sample performance of the estimators including bias, efficiency, coverage, and robustness to model misspecification under the Ramfjord, CPITN and RHM PRPs with comparison to full-mouth exams as the gold standard. Finally, Section 4 summarizes results, discusses strengths

and limitations of the SDM method, and identifies future research.

2 Methods

2.1 The measurement of periodontal disease

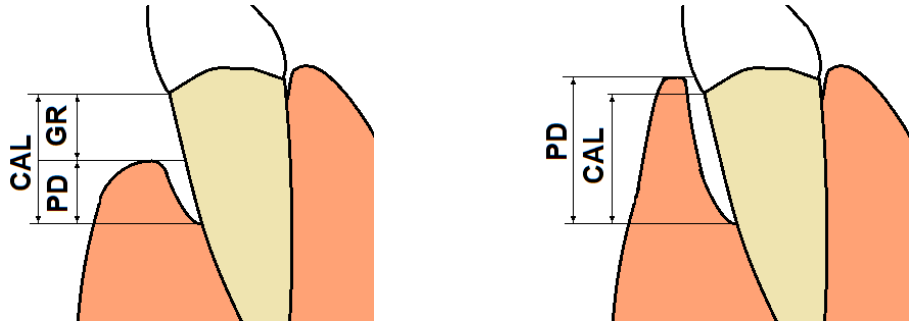
Periodontal diseases (gum disease) are a group of diseases and conditions that affect the soft (gingiva and periodontal attachment) and hard tissues (alveolar bone) surrounding teeth. Gingivitis is the mildest form of periodontal disease and it is characterized by inflammation of the marginal tissues. Periodontitis is a chronic multifactorial inflammatory disease associated with microbial biofilms characterized by destruction of the supporting tissues around the teeth, including periodontal attachment and alveolar bone. If left untreated, periodontitis may lead to tooth mobility, root exposure (gingival recession), halitosis (bad breath), tooth sensitivity, loss of masticatory function, poor esthetics, and ultimately tooth loss. Severe periodontitis has been associated with bacteremia (bacteria in the blood), systemic inflammation, and several systemic diseases and conditions.

Gingival recession (GR) occurs when the gingival margin (gumline) recedes from the border between the smooth enamel of the tooth crown and the tooth root's rough-textured cover; this border is called the cemento-enamel junction (CEJ). In a healthy periodontium, the gingival tissues (gums) are attached to the tooth at the CEJ.

GR is measured as the distance from the gumline to the CEJ; in Figure 1, GR is positive in (a) and negative in (b). PD is the distance from the gumline to the bottom of the pocket formed between the gums and the tooth. CAL measures the total distance between the base of the periodontal pocket and the CEJ (Page and Eke 2007). Since CAL can be calculated as a function of PD and GR (CAL = PD + GR), many surveys only collect two of these three measures. From 2011 to 2014, the NHANES measured GR and PD using a color-coded periodontal probe with measurement rounded to the lowest whole mm; CAL was derived using an algorithm in the data entry program (Dye et al., 2019). When collected, GR, PD, and CAL are measured in whole millimeters (mm) using a probing tool at up to six sites on each assessed tooth. In the NHANES, measurements of PD and GR range from 0 to 9mm and -9 to 9mm, respectively (Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS), 2013). Sites are evenly distributed around the tooth with three sites on the lingual, or tongue, and three sites on the buccal, or cheek, side of the tooth. On each side of the tooth, the sites are distributed as follows: the site closer to the front of the mouth (mesio-), the site closer to the back of the mouth (disto-), and the site halfway between mesio- and disto- sites (mid-). The six sites on each tooth are: mesiolingual, midlingual, distolingual, distobuccal, midbuccal, and mesiobuccal. The four interproximal (IP) sites, sites that are commonly next to another tooth and harder to clean, are mesiolingual, distolingual, mesiobuccal, and distobuccal.

2.2 Case definitions

We consider PD and CAL thresholds in three case definitions of disease at the complete cluster (i.e., mouth) level: the probability of (1) one or more or (2) two or more IP tooth sites with disease



(a) (b)

Figure 1: Periodontal Disease Measurements

not on the same tooth - as defined by a single threshold measure of PD or CAL - and (3) severe disease based on a combination of PD and CAL thresholds as defined below by the CDC/AAP case definition. It is not always possible to express case definitions based on tooth site indicators at the tooth level (e.g., Preisser et al., 2024). However, since disease in the first two case definitions must occur on different teeth, multiple sites on the same tooth with disease only constitute one occurrence of disease for the case definition. This made it possible to rephrase the first two case definitions in terms of tooth-level indicators, which reduces the computational burden of cluster sizes from a maximum of 168 tooth sites (e.g., Preisser et al., 2017, 2024), down to a maximum of 28 teeth. Finally, according to the CDC/AAP, severe periodontitis is defined as a person having at least 2 IP sites, not on the same tooth, with CAL measurements at or above 6mm and at least 1 IP site with PD measurements at or above 5mm (Eke et al., 2012). Because the case definition requires diseased sites for CAL to occur on different teeth (Alshihayb et al., 2022), the CDC/AAP definition of severe periodontitis is equivalent to having two or more teeth with at least one IP site having CAL at or above 6mm and at least one tooth with at least one IP site having PD at or above 5mm. Thus, the CDC/AAP severe case definition is also defined in terms of tooth-level indicators for each measure; with two measures for each tooth, the computational burden is reduced to a maximum of 56 tooth level measurements.

All case definitions for population-based surveillance of periodontitis are based on a full-mouth examination, which in this article and consistent with the CDC/AAP definition, consists of periodontal measurements on the four IP sites of the 28 teeth in the adult dentition, excluding third molars. In practice, a full-mouth exam would include at most 28 teeth (excluding third molars) with 4 IP sites per tooth for a maximum of 112 sites per mouth. Any measurements from non-IP tooth sites are not considered. For simplicity and clarity of exposition, the possibility of individuals with missing teeth is not considered in this article. Extension of the proposed methods to include individuals with less than 28 teeth could follow (Preisser et al., 2024) as discussed in Section 4.

For an individual, 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. Also 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.

The index for an individual, i, is suppressed until it is needed in Section 2.7 on estimation. Then, $\tilde{Y} = \Sigma_{j=1}^{28} Y_j$ is the total number teeth in the individual's mouth with CAL greater or equal to 6mm for at least one IP site, and $\tilde{W} = \Sigma_{j=1}^{28} W_j$ is the total number teeth with PD greater or equal to 5mm for at least one IP site. The prevalence of severe periodontitis is the proportion of individuals in a population of interest that satisfies the CDC/AAP case definition, which is expressed as the probability

$$\pi_{severe} = \Pr(\tilde{W} \geq 1, \tilde{Y} \geq 2).$$

Three full-mouth case definitions of periodontitis prevalence are considered (Table 1). The first two case definitions are marginal population proportions corresponding to a single type of clinical measure from the CDC/AAP severe periodontitis case definition.

In a full mouth exam (without missing teeth), the observed binary data are $\mathbf{W} = (W_1, \ldots, W_{28})^T$ and $\mathbf{Y} = (Y_1, \ldots, Y_{28})^T$ and we can classify each individual as having periodontitis or not. In a PRP, many of the elements of **W** and **Y** are missing by design, and a formula based on an SDM is required to estimate prevalence as the classification method results in underestimation bias. Considering the second case definition, the SDM is given by a model for the mean vector $\boldsymbol{\mu} = \mathbf{E(Y)} = (\mu_1, \ldots, \mu_{28})^T$, where $\mu_j = E(Y_j)$, and its corresponding correlation matrix $R = \text{Corr}(\mathbf{Y})$. Background knowledge on the degree and pattern of periodontitis is used to specify the SDM given by the parameters for $(\boldsymbol{\mu}, R)$. The CLF provides a full distribution for the 2^{28} possible profiles of **Y** in terms of a small or moderately small number of parameters; in the extreme case, only two parameters – common mean and exchangeable correlation, were specified in Preisser et al. (2017). The distribution of **W** can be described similarly.

The CDC/AAP severe case definition required modeling a multivariate binary distribution for 56 tooth-level indicators corresponding to (W_j, Y_j) for each tooth, $j = 1, \ldots, 28$, which required the additional consideration of between-measure correlations, $\text{Corr}(W_j, Y_{j'})$. Evaluating the performance of the SDM approach for estimating periodontitis prevalence with the widely accepted CDC/AAP severe case definition will help to determine whether SDM can produce reliable estimates in large epidemiologic studies.

2.3 Partial-mouth recording protocols (PRPs)

Permanent teeth are classified into four categories: incisors, canines, premolars, and molars. Periodontal disease is more common in premolars and molars than on the incisors and canines. PRPs typically select a subset of teeth and/or tooth sites to estimate full mouth oral health (Figure 2). The Ramfjord method consisted of collecting probing measurements from all six sites on the same six teeth for every participant (Ramfjord, 1959). In 1983, the World Health Organization developed the Community Periodontal Index for Treatment Needs (CPITN) protocol (Ainamo, 1982); while similar to Ramfjord's protocol in that it relies on a fixed set of selected teeth, the CPITN protocol collected probing measurements from all six sites on ten teeth. Random half-mouth (RHM) protocols sample specific sites on all teeth from two randomly chosen quadrants. RHMs have taken several forms: (1) randomly sampling two of the four quadrants without restriction, which provides six possible quadrant pairs in the sample frame; (2) randomly selecting an upper and lower quad-

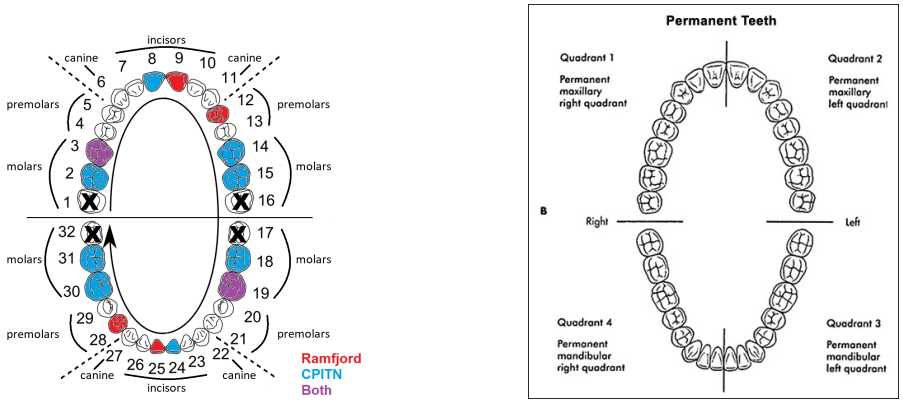
rant giving four possible quadrant pairs, (3) randomly selecting an upper quadrant and taking the contra-lateral (i.e., cross-diagonal) lower quadrant giving two possible quadrant pairs. The latter is equivalent to randomly sampling a lower quadrant and taking its contra-lateral upper quadrant. The exact teeth selected by each of these previously implemented PRPs is detailed below and in Figure 2.

1. Ramfjord (1959) protocol, which includes two molars (teeth number 3 and 19), two premolars (teeth number 12 and 28), and two incisors (teeth number 9 and 25), for a total of 6 teeth.

2. Community Periodontal Index for Treatment Needs (CPITN) protocol (Ainamo, 1982), which includes eight molars (teeth number 2, 3, 14, 15, 18, 19, 30 and 31) and two incisors (8 and 24), for a total of 10 teeth.

3. Random Half Mouth (RHM; Drury et al., 1996) protocol, which includes 7 teeth from a randomly selected upper quadrant and 7 teeth from a randomly selected lower quadrant (i.e., method 2, above), for a total of 14 teeth. This method collects measurements from four molars, four premolars, two canines, and four incisors per mouth. The teeth numbers observed will vary by participant.

We evaluate the performance of these three PRPs compared to full-mouth exam data. In each of the PRPs, all sites from selected teeth are typically measured. However, under the case definitions in Section 2.2, it would only be necessary to record CAL and PD at the four IP sites on each tooth.



(a) (b)

Figure 2: Fixed Tooth Selection for Ramfjord and CPITN Protocols and Random-Half Mouth Protocol Quadrants

2.4 General prevalence formulae

In this section, general formulae for the prevalence corresponding to the case definitions in Table 1 were derived based on the tooth-level binary variates; computational formulae based on specific

member distributions of the CLF are provided in Section 2.6. In this section, each prevalence expression was factorized as a product of conditional probabilities based upon a standard, but arbitrary, dental numbering system. The binary variates can be reordered in this section giving different (based on tooth identifiers), yet equivalent, factorizations for prevalence. Details can be found in Web-Appendix A.1.

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

$$\pi_{PD} = \Pr(\tilde{W} \geq 1) = 1 - \Pr(\tilde{W} = 0) = 1 - \prod_{j=1}^{28} \big(1 - \zeta_j\big),$$

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

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

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

where

$$\begin{aligned}
\Pr(\tilde{Y} = 0) &= \prod_{j=1}^{28} \big(1 - \eta_j\big), \\
\Pr(\tilde{Y} = 1) &= \sum_{k=1}^{28} \eta_1^{I(k=1)} \left[\eta_k \prod_{j=1}^{k-1} \big(1 - \eta_j\big)\right]^{I(k>1)} \left[\prod_{j=k+1}^{28} \big(1 - \eta_{j|k}\big)\right]^{I(k<28)}, \\
\eta_j &= \Pr\big(Y_j = 1 \big| \sum_{l=1}^{j-1} Y_l = 0\big) \text{ for } j \geq 2, \\
\eta_1 &= \Pr\big(Y_1 = 1\big) = \mu_{Y_1} \text{ and} \\
\eta_{j|k} &= \Pr\left(Y_j = 1 \middle| Y_k = 1, \sum_{l=1; l\neq k}^{j-1} Y_l = 0\right) \text{ for } j \geq 2.
\end{aligned}$$

The case definitions #1 and #2, which are based on single measures (PD or CAL), are general formulae that could be applied to different threshold values. For example, whereas the threshold for CAL of 6mm in Table 1 that is considered in this article defines severe periodontitis, the threshold for CAL of 4mm has been used elsewhere to define moderate periodontitis (Eke et al., 2012).

Case Definition #3: Severe Periodontitis

The CDC/AAP severe periodontitis case definition based on PD and CAL requires the full-mouth vector of length 56 of tooth-level indicators with the first 28 elements corresponding to $W_1, \ldots, W_{28}$ and the last 28 elements corresponding to $Y_1, \ldots, Y_{28}$, such that $\mathbf{U} = (U_1, \ldots, U_{56})^T = (W_1, \ldots, W_{28}, Y_1, \ldots, Y_{28})^T$. The prevalence of severe periodontitis is the joint probability that $\tilde{W} \geq 1$ and $\tilde{Y} \geq 2$ corresponding to the lower right cell in a two by two contingency table defined for $\tilde{W}$ and $\tilde{Y}$

as defined by the thresholds of 1 and 2, respectively:

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

where

$$\begin{aligned}\Pr(\tilde{W} = 0, \tilde{Y} = 0) &= \prod_{j=1}^{56} \big(1 - \theta_j\big), \\ \Pr(\tilde{W} = 0, \tilde{Y} = 1) &= \sum_{k=29}^{56} \theta_k \Big[\prod_{j=1}^{28} \big(1 - \theta_j\big) \Big] \Big[\prod_{j=29}^{k-1} \big(1 - \theta_j\big) \Big]^{I(k>29)} \Big[\prod_{j=k+1}^{56} \big(1 - \theta_{j|k}\big) \Big]^{I(k<56)}, \\ \theta_j &= \Pr\Big(U_j = 1 \Big| \sum_{l=1}^{j-1} U_l = 0\Big), \\ \theta_1 &= \Pr(W_1 = 1) = \mu_{W_1} \text{ and } \theta_{j|k} = \Pr\Big(U_j = 1 \Big| U_k = 1, \sum_{l=1; l\neq k}^{j-1} U_l = 0\Big).\end{aligned}$$

The expression for π_{severe} included terms based upon the probabilities that correspond to the case definitions #1 and #2 via $\Pr(\tilde{Y} \leq 1) = 1 - \pi_{CAL}$ and $\Pr(\tilde{W} = 0) = 1 - \pi_{PD} = \Pi_{j=1}^{28}(1 - \theta_j)$. Also, $\Pr(\tilde{W} = 0, \tilde{Y} = 1) = \Pr(\tilde{W} = 0) \Pr(\tilde{Y} = 1 | \tilde{W} = 0)$ where $\Pr(\tilde{Y} = 1 | \tilde{W} = 0)$ is derived in Web-Appendix A.1.

The formulae in this section are impractical for non-parametric estimation as they involve multinomial distributions with huge numbers of categories corresponding to cells in a 2^{28} contingency table for the first two case definitions and one with 2^{56} possible profiles for the third definition. The next section provides practical computational formulae based on specific expressions for the conditional probabilities $\zeta_j, \eta_j, \eta_{j|k}, \theta_j$ and $\theta_{j|k}$ derived from multivariate binary distributions for the tooth-level measures.

2.5 Statistical distribution model (SDM) method

The SDM approach to estimate periodontitis prevalence models the pattern of disease across all measures in the mouth by assuming a multivariate correlated binary distribution in the CLF. The CLF of distributions was proposed as a method for simulating correlated binary variables with a specified marginal mean vector and pairwise correlation matrix, without the need for explicit specification of the higher order moments of the multivariate binary distribution (Qaqish, 2003b; Preisser and Qaqish, 2014). The CLF has also been used as the basis for maximum likelihood estimation of longitudinal binary data models (e.g., Yang and Chaganty, 2014).

Focusing on CAL, but without loss of generality, background knowledge on periodontitis informs the structures of $\boldsymbol{\mu} = E(\mathbf{Y})$ and $R = \text{Corr}(\mathbf{Y})$, where $\mathbf{Y}$ is the vector of an individual's

Table 1: Full-Mouth Definitions for Periodontitis Cases and Prevalence[1]

(1)	At least 1 tooth affected with maximum PD≥5mm	$\pi_{PD} = \Pr(\tilde{W} \geq 1)$, where $\tilde{W} = \sum_{j=1}^{28} W_j$ & $W_j = \text{I}(\max(\mathbf{PD}_j) \geq 5\text{mm})$
(2)	At least 2 teeth affected with maximum CAL≥6mm	$\pi_{CAL} = \Pr(\tilde{Y} \geq 2)$ where $\tilde{Y} = \sum_{j=1}^{28} Y_j$ & $Y_j = \text{I}(\max(\mathbf{CAL}_j) \geq 6\text{mm})$
(3)	Joint (CDC/AAP Severe Periodontitis)	$\pi_{severe} = \Pr(\tilde{W} \geq 1, \tilde{Y} \geq 2)$

[1]The vectors $\mathbf{PD}_j$ and $\mathbf{CAL}_j$ each contain four IP measurements for the j-th tooth, such that severe periodontitis is defined as ≥2 teeth having at least 1 IP site with CAL≥6mm and ≥1 tooth with at least 1 IP site with PD≥5mm; PD = pocket depth, CAL = clinical attachment loss, IP=interproximal

tooth-level CAL indicators from Section 2.2. The CLF is defined by the following sequence of conditional means (i.e., probabilities, $\lambda_j, j = 1, \ldots, 28$) that fall within the viable range (i.e., [0,1]) with $\lambda_1 = \mu_1$ and, for $j > 2$,

$$\lambda_j(\mathbf{x}_j) = \text{E}(Y_j|\mathbf{X}_j = \mathbf{x}_j) = \Pr(Y_j = 1|\mathbf{X}_j = \mathbf{x}_j) = \mu_j + \mathbf{b}_j^T[\mathbf{x}_j - \text{E}(\mathbf{X}_j)],$$

where

$$\mathbf{X}_j = (Y_1, \ldots, Y_{j-1})^T \text{ and } \mathbf{b}_j = \text{cov}(\mathbf{X}_j)^{-1}\text{cov}(\mathbf{X}_j, Y_j)$$

This family of distributions allows flexibility in defining a distribution of periodontitis across all measurements in a mouth as it allows the specification of unequal means and an arbitrary pairwise correlation structure. The formula above is expressed using the notation for case definition #2, but Y_j could be replaced by W_j for case definition #1 or by U_k, an element from the joint vector of dimension 56 consisting of all W_j and Y_j, for case definition #3, $k = 1, \ldots, 56$.

Moreover, the λ_j-terms, written as $\lambda_j(\mathbf{x}_j)$ to denote their dependency on the values in the ordered sequence of binary variables $\mathbf{x}_j$, become the terms $\zeta_j, \eta_j, \eta_{j|k}, \theta_j$ and $\theta_{j|k}$ according to the particular sequence of 0's and 1's in $\mathbf{x}_j$ used to define each term. An equivalent expression for λ_j is

$$\lambda_j = \mu_j + \sum_{l=1}^{j-1} b_{jl}\big(Y_l - \mu_l\big)$$

where b_{jl} is the l-th element of $\mathbf{b}_j$, which depends upon both marginal means and pairwise correlations. Thus, the model, or distribution, for the correlated, tooth-level binary variates is defined through the specification of the marginal mean vector $\boldsymbol{\mu}$ and pairwise correlation matrix R, which are both parameterized with a relatively small number of parameters. Following Preisser et al. (2017), there are two working assumptions for the single measure case definitions #1 and #2 under which the CLF is equivalent to the Beta-binomial distribution (Qaqish, 2003b) in Web-Appendix A.2:

1. Common mean – probability of disease is the same for all teeth (i.e., μ_{PD} or μ_{CAL}, respectively) and

2. Exchangeable correlation - within-mouth pairwise correlation of disease is constant for all tooth pairs (i.e., ρ_{PD} or ρ_{CAL}, respectively).

The third case definition introduces two additional pairwise correlations to measure the correlation between disease due to PD and CAL on the same tooth (ρ_{same}) and on different teeth (ρ_{diff}). A CLF distribution is specified for **U**, a vector defined in Section 2.4 that contains both W_j's and Y_j's. A six-parameter model is specified at the tooth-level for indicators of disease based on PD and CAL measurements,

$$\boldsymbol{\mu} \equiv \mathrm{E}(\mathbf{U}) = \begin{bmatrix} \mu_{PD} \\ \vdots \\ \mu_{PD} \\ \mu_{CAL} \\ \vdots \\ \mu_{CAL} \end{bmatrix}_{56,1} \quad \& \; R \equiv \mathrm{Corr}(\mathbf{U}) = \begin{bmatrix} 1 & & \rho_{PD} & \rho_{same} & & \rho_{diff} \\ & \ddots & & & \ddots & \\ \rho_{PD} & & 1 & \rho_{diff} & & \rho_{same} \\ \rho_{same} & & \rho_{diff} & 1 & & \rho_{CAL} \\ & \ddots & & & \ddots & \\ \rho_{diff} & & \rho_{same} & \rho_{CAL} & & 1 \end{bmatrix}_{56,56}$$

2.6 CLF-based prevalence formulae

Prevalence estimation, based on a case definition of periodontitis, used a specified CLF for the pattern of periodontal disease measured by the tooth-level binary variates. The working assumptions for the first two case definitions stated in Section 2.5 result in the following prevalence formulae,

$$\begin{aligned} \pi_{PD} = \Pr(\tilde{W} \geq 1) = 1 - \prod_{j=1}^{28} \left(1 - \zeta_j\right) &= 1 - \prod_{j=1}^{28} \frac{(1-\mu_{PD})(1-\rho_{PD}) + (j-1)\rho_{PD}}{1 + (j-2)\rho_{PD}} \\ &= 1 - \prod_{j=1}^{28} \left(1 - \frac{(1-\rho_{PD})\mu_{PD}}{1 + (j-2)\rho_{PD}}\right) \end{aligned}$$

$$\begin{aligned} \pi_{CAL} = \Pr(\tilde{Y} \geq 2) = 1 &- \prod_{j=1}^{28} \left(1 - \frac{(1-\rho_{CAL})\mu_{CAL}}{1 + (j-2)\rho_{CAL}}\right) \\ &- \frac{28\mu_{CAL}(1-\rho_{CAL})}{1 + 26\rho_{CAL}} \prod_{j=1}^{27} \frac{(1-\mu_{CAL})(1-\rho_{CAL}) + (j-1)\rho_{CAL}}{1 + (j-2)\rho_{CAL}} \end{aligned}$$

For the third case definition, a simplified formula cannot be obtained so we use the general formula for π_{severe} from Section 2.4, while recognizing that the θ_j and $\theta_{j|k}$ terms are specific types of λ_j terms and so have expressions depending upon the six model parameters from Section 2.5.

We note that the CLF provides a full distribution for the 2^{28} possible profiles of **Y** (or **W**) in terms of a small or moderately small number of marginal mean and pairwise correlation parameters. Unlike in Section 2.4, changing the order of the binary indicators may result in a different CLF in the sense that inferred higher order nuisance parameters may change. Even though the ordering of

teeth is arbitrary given their three-dimensional spatial orientation to one another, the impact of the chosen ordering on the value of the prevalence is generally expected to be limited. In general, while the CLF does not include all possible correlated binary distributions, it has been recommended for clustered binary data due to its good coverage of such distributions (Preisser and Qaqish, 2014).

2.7 Estimation of intermediate parameters in the CLF-based prevalence estimator

Prevalence estimates were calculated for each case definition with respect to each PRP using estimates of the mean and pairwise correlations. Means were estimated with the following method of moments equations:

$$\hat{\mu}_{PD} = \frac{\sum_{i=1}^{K} \tilde{W}_i}{\sum_{i=1}^{K} m_i} \text{ and } \hat{\mu}_{CAL} = \frac{\sum_{i=1}^{K} \tilde{Y}_i}{\sum_{i=1}^{K} m_i} \tag{2.1}$$

where $i = 1, \ldots, K$ subjects, $\tilde{W}_i$ is the number of diseased teeth due to PD $\geq$ 5mm for person i, $\tilde{Y}_i$ is the number of diseased teeth due to CAL $\geq$ 6mm for person i, and m_i is the number of teeth selected by the PRP for person i. While m_i is constant for all subjects in our analysis, it will vary based on the PRP being analyzed; for Ramfjord, $m_i = 6$; for CPITN, $m_i = 10$; for RHM, $m_i = 14$; and for full-mouth, $m_i = 28$.

Within-mouth pairwise correlations were estimated with GEE-type estimators (Zeger and Liang, 1986). Define the residuals,

$$r_{PD_{ij}} = \frac{w_{ij} - \hat{\mu}_{PD}}{\sqrt{\hat{\mu}_{PD}(1 - \hat{\mu}_{PD})}} \text{ and } r_{CAL_{ij}} = \frac{y_{ij} - \hat{\mu}_{CAL}}{\sqrt{\hat{\mu}_{CAL}(1 - \hat{\mu}_{CAL})}}$$

Let $s_{ij} = 1$ denote whether the j-th tooth in the i-th mouth (participant) is selected by the PRP. Then

$$\begin{aligned} \hat{\rho}_{PD} &= \frac{\sum_{i=1}^{K} \sum_{j \neq k} s_{ij} r_{PD_{ij}} s_{ik} r_{PD_{ik}}}{\sum_{i=1}^{K} m_i(m_i - 1)}, \hat{\rho}_{CAL} = \frac{\sum_{i=1}^{K} \sum_{j \neq k} s_{ij} r_{CAL_{ij}} s_{ik} r_{CAL_{ik}}}{\sum_{i=1}^{K} m_i(m_i - 1)}, \\ \hat{\rho}_{same} &= \frac{\sum_{i=1}^{K} \sum_{j=1}^{28} s_{ij} r_{PD_{ij}} r_{CAL_{ij}}}{\sum_{i=1}^{K} m_i}, \text{ and } \hat{\rho}_{diff} = \frac{\sum_{i=1}^{K} \sum_{j \neq k} s_{ij} r_{PD_{ij}} s_{ik} r_{CAL_{ik}}}{\sum_{i=1}^{K} m_i(m_i - 1)} \end{aligned} \tag{2.2}$$

For case definition #1, $\hat{\mu} = \hat{\mu}_{PD}$ and $\hat{\rho} = \hat{\rho}_{PD}$; for case definition #2, $\hat{\mu} = \hat{\mu}_{CAL}$ and $\hat{\rho} = \hat{\rho}_{CAL}$. For equal cluster sizes ($m_i = m$) as for the PRPs in this article, estimators from the paired estimating equations of Prentice (1988) without covariates reduces to the formulae for the site-level probabilities $\hat{\mu}$ and pairwise correlations presented in 2.1 and 2.2. Prentice's GEE approach also provides an estimate of the joint covariance matrix of the intermediate parameter estimates that define the prevalence formula in Section 2.6 that, for case definitions #1 and #2, is used to determine the variance of the prevalence estimate via the delta method as described in the next section. Prentice's GEE was implemented in this article with the SAS macro GEECORR (Shing et al., 2021).

When estimating the intermediate parameters for $\hat{\pi}_{severe}$ with GEECORR, it was possible that the estimated pairwise correlation matrix was not positive semi-definite. When this happened, the closest semi-positive definite matrix was found using the Frobenius norm (Higham, 1988); if needed, diagonal values were reset to one to adhere to the structure of a correlation matrix. However, this requirement sometimes caused the final correlation matrix to remain non-positive semi-definite but resulted, nonetheless, in prevalence estimates within the plausible range (i.e., between 0 and 1).

2.8 Variance estimation for the CLF-based prevalence estimator

Common Mean and Exchangeable Correlation

For the two single measure case definitions, asymptotic variance estimates for $\hat{\mu}$ and $\hat{\pi}$ were calculated using the delta method; details are in Web-Appendix A.3. An intercept only model with a logit link for the marginal mean, logit$(\mu) = \beta$, and the identity link for the correlation model was fit using the paired estimating equations approach of Prentice (1988), which provided estimates of Cov$(\beta, \hat{\rho})$.

For case definition #1, let $\pi_{PD} = \Pr(\tilde{W} \geq 1) = 1 - \psi_0$, where $\psi_0 = \Pr(\tilde{W} = 0)$ and define $\psi_0^* = \ln(\psi_0)$. Following Wang and Preisser (2016), and given the logit link, the variance for the prevalence based on case definition #1 is var$(\hat{\pi}_{PD}) = e^{2\psi_0^*} * \text{var}(\hat{\psi}_0^*)$, where the var$(\hat{\psi}_0^*)$ is defined in Web-Appendix A.3.

For case definition #2, let $\pi_{CAL} = \Pr(\tilde{Y} \geq 2) = 1 - \psi_{0,1}$, where $\psi_{0,1} = \psi_0 + \psi_1$, $\psi_0 = \Pr(\tilde{Y} = 0)$ and $\psi_1 = \Pr(\tilde{Y} = 1)$. Consider $\psi_{0,1}^* = \ln(\psi_0 + \psi_1)$. Then the variance for the prevalence in case definition #2 is var$(\hat{\pi}_{CAL}) = e^{2\psi_{0,1}^*} * \text{var}(\hat{\psi}_{0,1}^*)$, where the var$(\hat{\psi}_{0,1}^*)$ is defined in Web-Appendix A.3.

Next, we consider the variance formula for the severe case definition based on the model with two means and four correlations. Since there was no closed form of this variance, bootstrap methods were applied to each simulated replicate to estimate within simulation variances of $\boldsymbol{\mu}$, R, and π_{severe}. The bootstrap method used in this paper has 6 steps.

Step 1. From the PRP data with K individuals (clusters), draw a sample of K individuals with replacement as the r-th replicate.

Step 2. Calculate estimates $\hat{\boldsymbol{\mu}}$ and $\hat{R}$ using method of moments and GEE estimators as described above.

Step 3. Estimate $\hat{\pi}_{severe}$ from a CLF-compatible distribution defined by $\hat{\boldsymbol{\mu}}$ and $\hat{R}$.

Step 4. Repeat steps 1-3 with the r-th replicate 200 times.

Step 5. Calculate the variance of $\hat{\boldsymbol{\mu}}$,$\hat{R}$, and $\hat{\pi}_{severe}$ for the r-th replicate based on results from the 200 bootstrap samples.

Step 6. Repeat steps 1-5 for each replicate.

3 Simulation Study

3.1 Data generation models

The selection of parameter values in CLF distributions for the random generation of the multivariate vector of binary data, **U**, using the method of Qaqish (2003b) was based on the dental literature and statistical considerations. This presented no difficulties for estimating prevalence under case definitions #1 and #2, which relied on two plug-in estimates, $\hat{\mu}$ and $\hat{\rho}$. This owes to the fact that, for a common mean μ, any choice of correlation such that $-1/(n-1) < \rho < 1$, where n is the number of marginal means, corresponds to a CLF distribution, which ensures $0 < \pi < 1$. However, with regards to prevalence estimation for case definition #3, assignment of the six parameters in the model for **U** needed to correspond to a distribution in the CLF. This meant that satisfying positive definiteness for R and the pairwise Fréchet bounds on the correlation parameters were necessary but not sufficient conditions for the existence of a multivariate binary distribution, most particularly one belonging to the CLF. Software provided by Qaqish (2003a) was used to identify whether a combination of the six parameters for the random variable **U** were CLF compatible, i.e., where there exists a multivariate binary distribution with the given parameter values in the CLF. A detailed summary of CLF violations, Fréchet bounds violations, and adjustments to the pairwise correlation matrix is in Web-Appendix A.4.

Considering that Eke et al. (2012) found that roughly 15% of individuals in their convenience sample of adults 35 or older had PD greater than or equal to 5mm, the CLF parameters were chosen such that π_{PD} ranged from 0.15 to 0.19; similarly, values were chosen such that π_{CAL} was approximately 0.11. Furthermore, based on Michalowicz et al. (2013), a data generation model for **U** was considered with a moderately high correlation between PD and CAL measures on the same tooth given by $\rho_{same} = 0.50$. Reported estimates of severe periodontitis range between 5% and 10% (Chen et al., 2021; Eke et al., 2012).

To narrow and simplify the search of CLF distributions, we added non-mandatory "assumption restrictions" of common means ($\mu_{PD} = \mu_{CAL} = \mu$), common measure-specific (marginal) exchangeable correlations ($\rho_{PD} = \rho_{CAL} = \rho$), and, based on biological considerations of proximity and sameness, $\rho_{diff} \leq \rho \leq \rho_{same}$. Having then focused on distributions for **U** with four non-redundant parameters, the top four panels of Figure 3 plots the Fréchet and positive definite bounds on the exchangeable correlation ρ as a function of the common mean μ for fixed ρ_{diff} and ρ_{same}. Likewise, the bottom two panels of Figure 3 plots Fréchet and positive definite bounds on ρ_{diff} as a function of ρ for fixed mean μ and ρ_{same}. All six panels show the range restrictions on correlations imposed by the CLF method. When the within-tooth correlation for PD and CAL was set to $\rho_{same} = 0.50$ and the across tooth correlation for PD and CAL equaled $\rho_{diff} = 0.15$, the exchangeable correlation was restricted to $0.15 \geq \rho \geq 0.16$. Finally, from these four parameters, "true" prevalence values were determined from the CLF formulae in Section 2.4. Thus, data generation model 1 was defined by:

$$\textbf{Model \#1: } \mu_{PD} = \mu_{CAL} = 0.021; \rho_{PD} = \rho_{CAL} = 0.16, \rho_{same} = 0.50, \rho_{diff} = 0.15; \\ \pi_{PD} = 0.19246, \pi_{CAL} = 0.11490, \pi_{severe} = 0.10082$$

Considering that assignment of $\rho_{same} = 0.50$ substantially restricted the possible values the other model parameters could take to ensure CLF compatibility, the second data generation model sets $\rho_{same} = 0.19$. Parameter values were chosen such that $\pi_{severe} = 0.06$. The second data generation model had many CLF compatible distributional options as shown in the right-hand-side panels of Figure 3. When setting $\rho_{same} = 0.19$ and $\rho_{diff} = 0.09$, the exchangeable correlation was restricted to $0.08 \leq \rho \leq 0.18$, so we defined:

$$\begin{aligned}\textbf{Model \#2:}\ & \mu_{PD} = \mu_{CAL} = 0.020; \rho_{PD} = \rho_{CAL} = 0.15, \rho_{same} = 0.19, \rho_{diff} = 0.09; \\ & \pi_{PD} = 0.19072, \pi_{CAL} = 0.11183, \pi_{severe} = 0.06444\end{aligned}$$

A final data generation model was considered to evaluate the performance of the SDM under model misspecification. Model 3 had varying means between PD and CAL measurements and seven unique correlation values to represent the intra-oral distribution of periodontal disease in a mouth; resulting in a nine parameter model.

$$\begin{aligned}\textbf{Model \#3:}\ & \mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{SQ}} = \rho_{CAL_{SQ}} = 0.19, \\ & \rho_{PD_{AQ}} = \rho_{CAL_{AQ}} = 0.17, \rho_{PD_{CQ}} = \rho_{CAL_{CQ}} = 0.14; \\ & \rho_{same} = 0.19, \rho_{diff_{SQ}} = 0.11, \rho_{diff_{AQ}} = 0.10, \rho_{diff_{CQ}} = 0.09; \\ & \pi_{PD} = 0.15520, \pi_{CAL} = 0.11383, \pi_{severe} = 0.06365\end{aligned}$$

The common exchangeable correlation and different measures on different teeth correlations, ρ and ρ_{diff} in models 1 and 2 respectively, were divided into three distinct parameters to reflect quadrant relationships between measurements (Figure 2). Correlations between teeth within the same quadrant ($\rho_{PD_{SQ}} = \rho_{CAL_{SQ}} = \rho_{SQ}; \rho_{diff_{SQ}}$) were assumed to be the largest, followed by those from adjacent quadrants ($\rho_{PD_{AQ}} = \rho_{CAL_{AQ}} = \rho_{AQ}; \rho_{diff_{AQ}}$). The lowest were those from contra-lateral quadrants ($\rho_{PD_{CLQuad}} = \rho_{CAL_{CLQuad}} = \rho_{CQ}; \rho_{diff_{CQ}}$). This logic resulted in the following assumptions:

$$\begin{aligned}\rho_{same} > \rho_{SQ} > \rho_{AQ} > \rho_{CQ}, \text{ and} \\ \rho_{same} > \rho_{diff_{SQ}} > \rho_{diff_{AQ}} > \rho_{diff_{CQ}}.\end{aligned}$$

The corresponding correlation matrix for $\mathbf{U}$, with its blocked elements ordered by measure (PD, then CAL) and, within measure, by quadrant number (1, 2, 3, 4), under the nine-parameter model is $\mathrm{R} \equiv \mathrm{Corr}(\mathbf{U})$ where

$$\mathrm{R} = \begin{bmatrix} R_{SQ_{PD}} & R_{Adj_{PD}} & R_{CL_{PD}} & R_{Adj_{PD}} & R_{SQ_{diff}} & R_{Adj_{diff}} & R_{CL_{diff}} & R_{Adj_{diff}} \\ & R_{SQ_{PD}} & R_{Adj_{PD}} & R_{CL_{PD}} & & R_{SQ_{diff}} & R_{Adj_{diff}} & R_{CL_{diff}} \\ & & R_{SQ_{PD}} & R_{Adj_{PD}} & & & R_{SQ_{diff}} & R_{Adj_{diff}} \\ & & & R_{SQ_{PD}} & & & & R_{SQ_{diff}} \\ & & & & R_{SQ_{CAL}} & R_{Adj_{CAL}} & R_{CL_{CAL}} & R_{Adj_{CAL}} \\ & & & & & R_{SQ_{CAL}} & R_{Adj_{CAL}} & R_{CL_{CAL}} \\ & & & & & & R_{SQ_{CAL}} & R_{Adj_{CAL}} \\ & & & & & & & R_{SQ_{CAL}} \end{bmatrix}_{56,56}$$

and

$$R_{SQ_{PD}} = R_{SQ_{CAL}} = \begin{bmatrix} 1 & & \rho_{SQ} \\ & \ddots & \\ \rho_{SQ} & & 1 \end{bmatrix}_{7,7}, R_{SQ_{diff}} = \begin{bmatrix} \rho_{same} & & \rho_{diff_{SQ}} \\ & \ddots & \\ \rho_{diff_{SQ}} & & \rho_{same} \end{bmatrix}_{7,7},$$

and the rest had a common correlation for all their elements: $R_{Adj_{PD}} = R_{Adj_{CAL}} = \mathbf{1}_{7,7} * \rho_{AQ}$, $R_{CL_{PD}} = R_{CL_{CAL}} = \mathbf{1}_{7,7} * \rho_{CQ}$, $R_{Adj_{diff}} = \mathbf{1}_{7,7} * \rho_{diff_{AQ}}$, and $R_{CL_{diff}} = \mathbf{1}_{7,7} * \rho_{diff_{CQ}}$.

All data were generated for a full-mouth ($n = 28$ teeth; "FULL" in subsequent results) assuming no missing teeth.

3.2 Evaluation methods

To evaluate the performance of the prevalence estimators from Section 2.7 for their estimands from Section 2.6 under various PRPs, full-mouth data with binary indicators of disease based on PD and CAL measurements for each tooth were simulated from a CLF under one of three models with parameter values described in the previous section. In the first simulation experiment, we simulated 1,000 replicate samples of 500, 1000, and 5000 full-mouth clusters ($m = 28$) for Models 1 and 2. Specifically, the vector of binary variates **U** was generated based on the joint model for the severe case definition described in Section 2.5. In a second simulation experiment, 1000 replicates of 5000 clusters under the nine parameter Model 3 were generated to examine the performance of the proposed estimators under misspecification of the correlation structure.

From the generated FULL data for each individual (i.e., cluster), the binary data values for the fixed-teeth in the Ramfjord and CPITN PRPs were saved as data based on tooth number (Figure 2). For RHM, the data consisted of the teeth from a randomly selected upper quadrant and lower quadrant. The prevalence estimate and its variance was estimated for each simulation replicate. For the single measure case definitions #1 and #2, we use the delta method to estimate the variance; for severe periodontitis as defined in case definition #3, the bootstrap method was used as described in Section 2.8.

For each simulation scenario, the percent relative bias of the prevalence estimators under each PRP was calculated as $100(\bar{\hat{\pi}} - \pi)/\pi$, where $\bar{\hat{\pi}}$ is the mean of the 1000 prevalence estimates. The percent relative bias of the standard errors was calculated as

$$100\big(\overline{\text{se}(\hat{\pi})} - \Sigma_{\hat{\pi},MC}\big)/\Sigma_{\hat{\pi},MC}$$

where the gold standard is the Monte Carlo standard deviation

$$\Sigma_{\hat{\pi},MC} = \sqrt{\sum_{i=1}^{1000}\left(\hat{\pi}_i - \sum_{i=1}^{1000}\hat{\pi}_i/1000\right)^2}/999$$

and i indexes the 1000 replicates. We calculated coverage of the 95% confidence interval (CI) for π as the proportion of replicates for which the CI contained the true parameter value. Finally, relative

(a) Large Within Tooth Correlation (b) Realistic Severe Prevalence

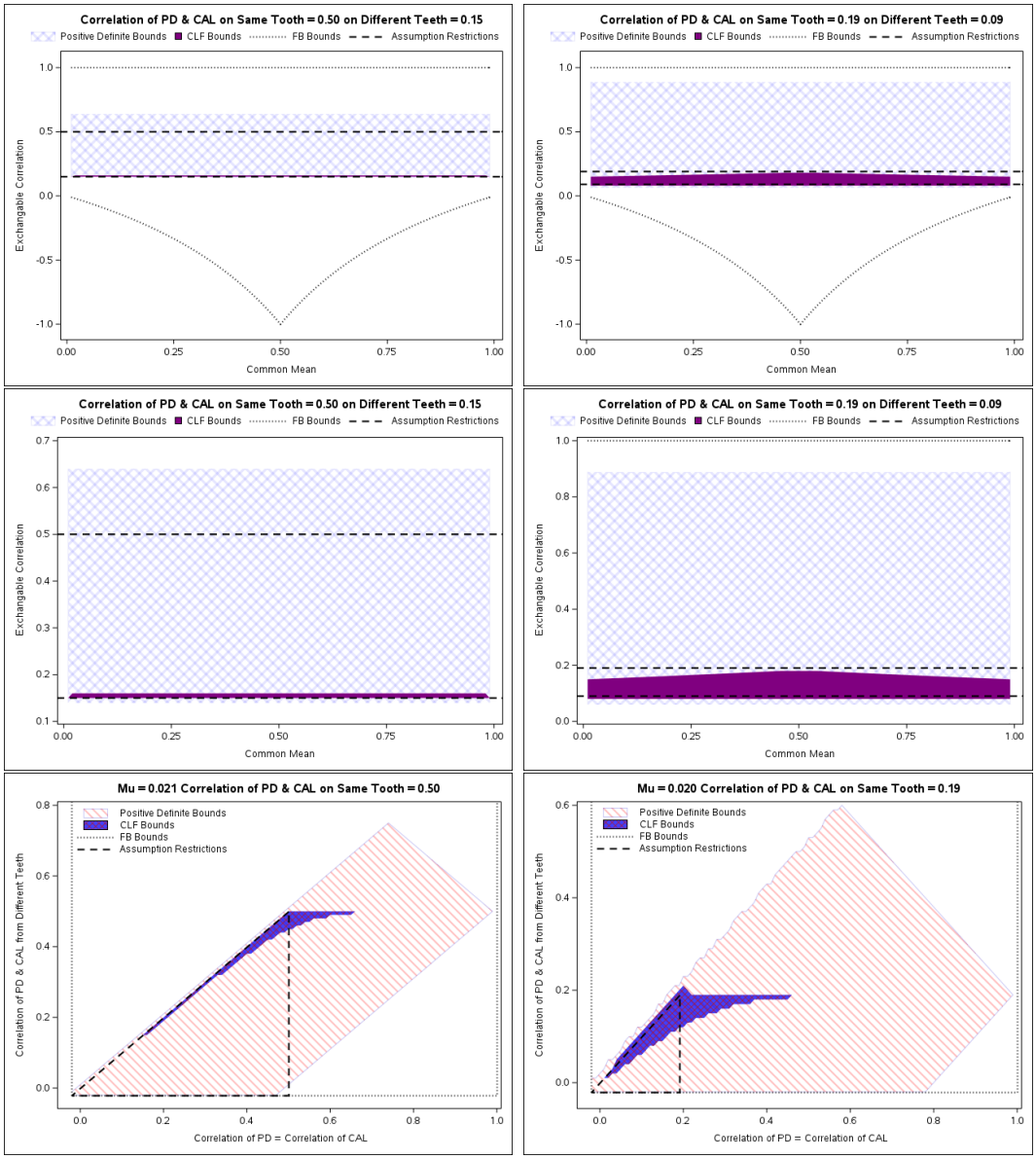


Figure 3: Severe Case Definition Range Restrictions for Model Selection

efficiency was evaluated by taking the ratio of the mean squared error (MSE) of the full-mouth estimate for π and the MSE of the PRP estimate for π,

$$\text{MSE}_{full}/\text{MSE}_{PRP} = \sum_{i=1}^{1000} \left(\hat{\pi}_{full,i} - \pi\right)^2 / \sum_{i=1}^{1000} \left(\hat{\pi}_{PRP,i} - \pi\right)^2,$$

where i indexes the 1000 replicates. All simulations were performed using SAS software Version 9.4. Copyright ©2002-2012 SAS Institute.

3.3 Results when the correlation model is correctly specified

For the two single measure case definitions, the prevalence estimator and its standard error based on a Beta-Binomial distribution, i.e, with a common mean and exchangeable correlation, performed well when the true model was also from a Beta-Binomial distribution (Tables 2 and 3). Prevalence estimates had absolute relative bias below 3%, and closely matched the true prevalence for all sample sizes and PRPs evaluated. Standard error estimates had absolute percent relative bias less than 5%, except in two cases for $K = 500$ individuals and decreased with m_i (the number of teeth selected in the PRP; Table 2). Relative efficiency increased with m_i and was consistent in pattern and magnitude across PRPs for each sample size, K, prevalence estimand, and model (Table 3). Coverage was generally close to the 95% nominal level. For the CDC-AAP severe periodontitis case definition, $\hat{\pi}_{SEV}$ performed well with minimal relative bias, within $\pm 3\%$ of the true values, for both the prevalence and its standard error (Table 4). Average prevalence estimates were similar across all sample sizes and PRPs for each true model, closely matching the true prevalence. As expected, average standard errors decreased with increasing K or m_i. All PRPs and models had coverage rates near the 95% nominal level. Relative efficiency followed a monotonic trend as PRPs with more teeth had larger relative efficiencies. Relative efficiency was about 35% for Ramfjord, 60% for CPITN, and 75% for RHM across all sample sizes. Regardless of the PRP, the severe periodontitis case definition had relative efficiency somewhere between the single measure case definitions #1 and #2. Case definition #1 had the lowest relative efficiency estimates. Case definition #2 had the highest relative efficiency estimates.

3.4 Results when the correlation model is misspecified

When data were generated under Model 3 to examine the impact of correlation structure misspecification, the prevalence estimators $\hat{\pi}_{PD}$, $\hat{\pi}_{CAL}$ and $\hat{\pi}_{SEV}$ and their standard errors generally performed well (Table 5, $K = 5000$). Percent absolute relative bias was below 2% for all PRPs considered except under the RHM protocol, where the underestimation bias for $\hat{\pi}_{PD}$ and $\hat{\pi}_{SEV}$ was 4.0% and 3.7%, respectively. The absolute relative bias of the standard error estimates was under 5.2% for all simulation scenarios. Under the Ramfjord and CPITN protocols, relative efficiency was similar between the single measure case definitions, but noticeably lower for $\hat{\pi}_{SEV}$, the estimate for severe periodontitis. In contrast, relative efficiency under the RHM protocol was 48.4% for $\hat{\pi}_{PD}$ and 53.2% for $\hat{\pi}_{SEV}$ compared to 82.1% for $\hat{\pi}_{CAL}$. In particular, the result that CPITN was more efficient than RHM for estimating $\hat{\pi}_{PD}$, i.e., 68.4% versus 48.4%, was a bit surprising as it selected

four fewer teeth than RHM. Correspondingly, coverage was near the 95% nominal level for all scenarios except for $\hat{\pi}_{PD}$ and $\hat{\pi}_{SEV}$ under RHM where coverage was 85% and 89.2%, respectively, suggesting that RHM may not be a robust PRP method in instances of misspecification of the correlation structure for some prevalence estimators. Intermediate parameters for all models are detailed in Web-Appendix A.4.

4 Discussion

While PRPs have long been-known to reduce the need for resources and costs of periodontal exams in large epidemiological studies and oral health surveys, their use has been questioned in the past decade due to the underestimation bias of periodontitis prevalence estimators owing to the standard approach based on the case classification of study participants using partial examination data. Taking a novel approach, this article extended a statistical distribution method for estimation of periodontitis prevalence under PRPs based on multivariate Bernoulli distributions of tooth-level indicators of periodontal disease, thereby minimizing bias and improving the accuracy of disease estimates. This article's most significant extension of earlier work on SDM (Preisser et al., 2017, 2024) was the development of a formula and a variance estimation procedure for periodontitis prevalence with PRPs for the CDC/AAP case definition of severe periodontitis. In simulation studies, the prevalence estimator was shown to have low bias for three commonly used PRPs. Overall, however, the statistical behavior of the novel prevalence estimator was better under Ramfjord and CPITN than RHM protocols. Therefore, despite sampling more teeth than Ramfjord or CPITN, the conclusions reached by this study recommend against use of RHMs.

This study adds to the evidence that SDM is a valid approach for the estimation of periodontitis prevalence in PRPs. In the simulation experiments, the bias of the prevalence estimator and its standard error were generally under 5% and confidence interval coverage was close to the nominal 95% level. This good performance, even under misspecification of the correlation structure, was somewhat better than reported by Wang and Preisser (2016) in their simulation studies for single measure case definitions, where the bias of standard errors was as high as 10%. Compared to Wang and Preisser, differences in the results observed in this article may be explained by the common within measure tooth-level probabilities in both the data generation and analysis models. In contrast, Wang and Preisser considered a spatial representation of the within measure tooth-level probabilities of disease and the correlation structure in their data generation model. Similarly, using full-mouth periodontal exam data as the gold standard, Preisser et al. (2017, 2024) showed that the "bias" of SDM prevalence estimators in PRPs is generally under 10%, which is much better than the substantial bias often associated with case classification estimators.

Table 2: Prevalence Estimates and their Standard Errors (SE) for All Case Definitions by PRP in Simulation Study with 1,000 Replicates

		# of Individuals					
Data	PRP	500		1,000		5,000	
Generation	(# of teeth)	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	$\mathbf{SE}_{\pi}$
$\boldsymbol{\pi_{PD}}$							
Model 1[1]	RAM (6)	0.1977	0.0322	0.1952	0.0231	0.1932	0.0104
	CPITN (10)	0.1956	0.0260	0.1950	0.0187	0.1929	0.0084
	RHM (14)	0.1960	0.0236	0.1944	0.0168	0.1929	0.0076
	FULL (28)	0.1946	0.0205	0.1940	0.0147	0.1928	0.0066
Model 2[2]	RAM (6)	0.1959	0.0327	0.1940	0.0234	0.1914	0.0106
	CPITN (10)	0.1934	0.0260	0.1932	0.0187	0.1910	0.0084
	RHM (14)	0.1946	0.0236	0.1927	0.0169	0.1908	0.0076
	FULL (28)	0.1932	0.0204	0.1925	0.0146	0.1909	0.0066
$\boldsymbol{\pi_{CAL}}$							
Model 1	RAM (6)	0.1155	0.0176	0.1156	0.0125	0.1149	0.0056
	CPITN (10)	0.1154	0.0152	0.1154	0.0107	0.1151	0.0048
	RHM (14)	0.1153	0.0141	0.1152	0.0100	0.1149	0.0045
	FULL (28)	0.1152	0.0128	0.1153	0.0091	0.1149	0.0041
Model 2	RAM (6)	0.1118	0.0173	0.1127	0.0123	0.1120	0.0055
	CPITN (10)	0.1122	0.0149	0.1122	0.0106	0.1120	0.0047
	RHM (14)	0.1121	0.0139	0.1125	0.0099	0.1119	0.0044
	FULL (28)	0.1121	0.0126	0.1123	0.0089	0.1119	0.0040

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

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

PRP = partial-mouth recording protocol, CLF = conditional linear family, PD = pocket depth,

CAL = clinical attachment loss, SEV = severity, RAM = Ramfjord protocol, CPITN = Community Periodontal Index

for Treatment Needs protocol, RHM = random half mouth, FULL = full-mouth exam

Table 3: Percent Relative Bias, Percent Relative Efficiency, and Coverage for All Case Definitions by PRP in Simulation Study with 1,000 Replicates

		Number of Individuals											
True	PRP	500				1,000				5,000			
Model	(# of teeth)	**RB**		**RE**	**COV**	**RB**		**RE**	**COV**	**RB**		**RE**	**COV**
		$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π
π_{PD}													
Model 1[1]	RAM (6)	2.70	-9.31	35.77	92.70	1.40	-0.73	39.99	94.70	0.40	-2.84	37.35	93.00
	CPITN (10)	1.62	-4.28	61.52	93.70	1.33	-2.86	58.40	93.60	0.21	2.20	63.64	95.20
	RHM (14)	1.83	-3.74	75.09	94.40	0.98	1.09	78.38	94.70	0.25	0.72	75.54	95.50
	FULL (28)	1.11	-3.78	100.0	93.80	0.80	-0.40	100.0	94.80	0.19	0.91	100.0	95.40
Model 2[2]	RAM (6)	2.74	-4.46	36.29	94.30	1.72	-4.10	37.45	93.40	0.35	-0.30	37.93	94.80
	CPITN (10)	1.40	-5.61	56.77	93.40	1.31	-4.09	58.32	93.80	0.13	1.05	61.15	95.20
	RHM (14)	2.06	-2.90	72.25	94.10	1.04	-1.88	75.47	94.80	0.05	2.12	77.10	96.10
	FULL (28)	1.30	-1.55	100.0	94.60	0.96	-2.07	100.0	95.20	0.09	0.67	100.0	95.20
π_{CAL}													
Model 1	RAM (6)	0.49	0.82	55.44	95.00	0.57	1.69	51.84	95.10	0.00	-0.29	52.04	94.80
	CPITN (10)	0.39	1.97	76.55	94.90	0.45	3.09	71.83	95.50	0.16	-2.93	66.42	93.90
	RHM (14)	0.36	-2.04	81.25	93.60	0.28	0.34	78.60	94.40	-0.01	1.53	83.96	95.90
	FULL (28)	0.22	-1.28	100.0	94.20	0.34	2.93	100.0	95.30	0.00	0.59	100.0	95.40
Model 2	RAM (6)	-0.05	1.21	53.46	94.70	0.76	-0.53	52.56	94.50	0.17	0.95	54.73	95.60
	CPITN (10)	0.31	0.81	71.30	94.20	0.36	-0.29	71.73	94.60	0.18	-2.27	69.36	94.50
	RHM (14)	0.25	1.40	83.19	94.80	0.56	-1.49	80.51	93.80	0.03	1.05	85.85	95.30
	FULL (28)	0.27	0.92	100.0	94.90	0.40	-0.62	100.0	94.70	0.04	-1.09	100.0	94.80

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

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

PRP = partial-mouth recording protocol, CLF = conditional linear family, PD = pocket depth, CAL = clinical attachment loss, SEV = severity, RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol, RHM = random half mouth, FULL = full-mouth exam, RB = percent relative bias, RE = percent relative efficiency, COV = coverage of 95% confidence intervals

Table 4: Prevalence Estimates, their Standard Errors (SE), Percent Relative Bias, Percent Relative Efficiency, and Coverage for CDC/AAP Severe Case Definition (π_{SEV} **- CLF Formula**) by PRP in Simulation Study with 1,000 Replicates

		# of Individuals											
Data	PRP	500				1,000				5,000			
Generation	(# of teeth)	$\hat{\pi}$		$\mathbf{SE}_{\pi}$		$\hat{\pi}$		$\mathbf{SE}_{\pi}$		$\hat{\pi}$		$\mathbf{SE}_{\pi}$	
Model 1[1]	RAM (6)	0.1024		0.0187		0.1028		0.0136		0.1014		0.0063	
	CPITN (10)	0.1023		0.0154		0.1020		0.0109		0.1011		0.0048	
	RHM (14)	0.1021		0.0140		0.1015		0.0098		0.1010		0.0043	
	FULL (28)	0.1013		0.0121		0.1014		0.0085		0.1009		0.0038	
Model 2[2]	RAM (6)	0.0663		0.0175		0.0659		0.0123		0.0647		0.0053	
	CPITN (10)	0.0652		0.0134		0.0653		0.0094		0.0645		0.0042	
	RHM (14)	0.0652		0.0119		0.0652		0.0084		0.0644		0.0037	
	FULL (28)	0.0646		0.0102		0.0649		0.0073		0.0644		0.0033	
True		**RB**		**RE**	**COV**	**RB**		**RE**	**COV**	**RB**		**RE**	**COV**
Model		$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π	$\hat{\pi}$	$\mathbf{SE}_{\pi}$	$\hat{\pi}$	π
Model 1	RAM (6)	1.51	-1.67	40.86	94.40	1.94	0.97	38.25	94.40	0.59	-2.52	34.01	94.20
	CPITN (10)	1.42	-2.23	59.79	94.10	1.19	0.96	59.19	95.00	0.29	-1.33	58.70	94.40
	RHM (14)	1.31	-0.49	75.23	94.90	0.68	3.35	78.11	95.80	0.13	-0.25	75.34	95.10
	FULL (28)	0.49	-1.02	100.0	94.90	0.61	1.62	100.0	95.50	0.03	0.16	100.0	94.40
Model 2	RAM (6)	2.87	-0.01	35.32	93.70	2.20	-0.40	35.31	94.30	0.41	1.54	38.62	96.20
	CPITN (10)	1.18	1.05	62.03	94.20	1.35	-1.25	59.20	94.00	0.11	-2.35	57.28	93.80
	RHM (14)	1.17	-1.03	75.28	92.90	1.13	0.31	76.99	94.60	-0.03	3.37	79.66	95.30
	FULL (28)	0.24	-2.03	100.0	93.60	0.65	-0.95	100.0	94.10	0.01	0.99	100.0	95.00

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

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

PRP = partial-mouth recording protocol, CLF = conditional linear family, PD = pocket depth, CAL = clinical attachment loss, SEV = severity,

RAM = Ramfjord protocol, CPITN = Community Periodontal Index for Treatment Needs protocol, RHM = random half mouth,

FULL = full-mouth exam, RB = percent relative bias, RE = percent relative efficiency, COV = coverage of 95% confidence intervals

This article evaluated multiple sample size combinations to determine the impact of the number of clusters and cluster size (i.e., number of teeth selected in PRPs) on the statistical performance of prevalence estimators. Large epidemiological studies, such as the NHANES surveys (Eke et al., 2018, 2015), the Arteriolosclerosis Risk in Communities Study (Beck et al., 2001; Preisser et al., 2017), and the Hispanic Community Health Study/Study of Latinos (Shing et al., 2023), typically have over 5000 participants, i.e., clusters. Limited funding, international setting, or hard-to-reach populations are a few factors that might cause a study to have fewer than 5000 clusters. With these considerations and under strong parametric assumptions about the distribution of disease within the mouth, simulation results with 500, 1000, and 5000 clusters suggested that the CPITN and RHM protocols were, respectively, about 60% and 75% statistically efficient as full-mouth data collection. This said, if a sample of 5000 clusters can be collected, then the Ramfjord protocol that only samples six teeth per mouth, i.e., cluster, yielded better results, i.e., smaller average standard error, than a sample of 500 or 1000 clusters with full-mouth data. In simulations based on 5000 clusters where the prevalence formula for the CDC/AAP severe case definition was based on a misspecified within-mouth correlation structure, CPITN and RHM protocols were, respectively, 50% and 53% statistically efficient as full-mouth data collection.

One PRP method in our evaluation, RHM, did not perform well under misspecification of the pairwise correlation structure for two of three prevalence estimators. Based on a sample size of $K = 5000$, Ramfjord and, especially, the CPITN protocol were fairly robust to misspecification of the correlation model with minimal relative bias for prevalence and standard error estimates and coverage near the nominal 95% level. Interestingly, RHM, which measures more teeth per individual ($m = 14$) than Ramfjord ($m = 6$) or CPITN ($m = 10$), had mixed results. For case definition #2 it performed well, but for case definitions #1 and #3 (severe disease), RHM had higher relative bias and lower coverage compared to Ramfjord and CPITN. This is due to its quadrant representation. With full-mouth data under a single measure case definition, 52% of the pairwise correlations come from teeth in adjacent quadrants, 26% from teeth in contralateral quadrants, and 22% from teeth in the same quadrant. While these percentages are similar for Ramfjord and CPITN, the RHM PRP oversamples tooth pairs from the same quadrant and thus creates pairwise correlation estimates that are higher than expected with our prevalence analysis model that mistakenly assumes an exchangeable correlation within measure. This result highlights the importance of carefully choosing a PRP method based not only on the case definition but also on assumptions about the underlying distribution of periodontal disease within a cluster, that is, both the intensity and pattern of disease in the mouth.

In regards to the apparent robustness of case definition #2, it produced the best results in terms of relative bias, relative efficiency, and coverage for all data generation models. For case definitions #1 and #2, the full-mouth data can be represented by a 2^{28} contingency table. There is only one way for the contingency table to satisfy case definition #1 (i.e., $1 - \Pr(W = 0)$) - and similarly for the PD component of case definition #3 –, but twenty-nine ways for the contingency table to satisfy case definition #2 because an affected tooth can occur at any one of the 28 teeth or no teeth (i.e., $1 - \Pr(Y = 0) - \Pr(Y = 1)$), which seemingly makes case definition #2 more robust to correlation model misspecification under the RHM protocol.

Table 5: Results from Misspecified Analysis Model[1] for All Case Definitions by PRP in Simulation Study with 1,000 Replicates and 5,000 Individuals per Replicate

	Estimates		**RB**		**RE**	**COV**
PRP	$\hat{\pi}$	$\mathbf{SE}_{\hat{\pi}}$	$\hat{\pi}$	$\mathbf{SE}_{\hat{\pi}}$	$\hat{\pi}$	π
π_{PD}						
RAM (6)	0.1559	0.0101	0.46	-0.58	41.06	94.40
CPITN (10)	0.1542	0.0079	-0.66	1.29	68.43	94.60
RHM (14)	0.1489	0.0069	-4.03	-0.17	48.44	85.00
FULL (28)	0.1529	0.0060	-1.48	-1.29	100.00	92.90
π_{CAL}						
RAM (6)	0.1156	0.0056	1.52	1.26	45.75	93.90
CPITN (10)	0.1152	0.0048	1.21	5.11	68.38	95.60
RHM (14)	0.1130	0.0044	-0.74	3.58	82.07	94.40
FULL (28)	0.1145	0.0039	0.62	2.54	100.00	95.80
π_{SEV}**- CLF Formula**						
RAM (6)	0.0658	0.0058	3.31	5.19	28.21	95.50
CPITN (10)	0.0645	0.0043	1.36	-2.89	49.71	93.60
RHM (14)	0.0613	0.0036	-3.68	0.01	53.15	89.20
FULL (28)	0.0637	0.0032	0.02	-0.04	100.00	95.00

[1] $\mu_{PD} = 0.015, \mu_{CAL} = 0.020; \rho_{PD_{samequad}} = \rho_{CAL_{samequad}} = 0.17, \rho_{PD_{adjquad}} = \rho_{CAL_{adjquad}} = 0.14,$
$\rho_{PD_{CLquad}} = \rho_{CAL_{CLquad}} = 0.11; \rho_{same} = 0.19, \rho_{diff_{samequad}} = 0.11, \rho_{diff_{adjquad}} = 0.10,$
$\rho_{diff_{CLquad}} = 0.09; \pi_{PD} = 0.1552, \pi_{CAL} = 0.1138, \pi_{severe} = 0.0637$

PRP = partial-mouth recording protocol, CLF = conditional linear family, PD = pocket depth,

CAL = clinical attachment loss, SEV = severity, RAM = Ramfjord protocol,

RHM = random half mouth, FULL = full-mouth exam, CPITN = Community Periodontal Index for

Treatment Needs protocol, RB = percent relative bias, RE = percent relative efficiency, COV = coverage of 95% CIs

While our work indicates the SDM method can produce accurate prevalence estimates when the PRP is carefully chosen with regard to knowledge of the structure of periodontal disease within a mouth, further evaluation of the proposed prevalence estimators under model misspecification is needed. Additional evaluation of the proposed estimators under model misspecification that account for varying within measure (PD and CAL) tooth-level probabilities of disease in addition to other complex structures for the pairwise correlations would further support the SDM approach as a viable method of estimating periodontitis prevalence in surveillance studies using PRPs. Another type of model misspecification arises when the means and correlations are correctly specified, but the data are generated from a distribution for correlated binary data that is not from the CLF, such as the multivariate probit (Emrich and Piedmonte, 1991), which could also be explored. In addition to extending the data generation and analysis models, since not all studies can benefit from large cluster sizes, future analysis should extend the evaluation of model misspecification to smaller sample sizes of 500 and 1000 clusters.

The aforementioned discussion suggests that new prevalence estimators defined under various model specifications to account for even more complex mean and pairwise correlation structures than considered in this article, such as considered in Shing et al. (2023), should be explored to solidify the dynamic application of the SDM approach. Another limitation of the general prevalence formulae in Section 2.4 is that they assume there are no missing data within a cluster. That is, every individual has all 28 teeth, which is very unlikely to occur in adult populations, especially in developing countries/regions. In the convenience sample of adults over 35 years of age from Eke et al. (2012), only 32.4% of the sample had no tooth loss, with a sample average of 3.5 missing teeth. Moreover, tooth loss is associated with demographics, socioeconomic, and behavioral factors, which would affect estimates of disease for different population subgroups. Extending the proposed CDC-AAP severe periodontitis prevalence estimate to include individuals with missing teeth is needed to fully establish its use with PRPs as an alternative to full-mouth data collection. A possible adaptation could follow Preisser et al. (2024) who proposed conditioning estimands on the number of missing teeth, and then averaging them across the population distribution of the number of missing teeth in the mouth.

A strength of SDM is that each prevalence formula in Section 2.6 may apply to any underlying distribution of correlated tooth-level binary data from the conditional linear family Qaqish (2003b), but this method requires that the prevalence formula be uniquely determined for each case definition. As evidenced in Section 2.4, the complexity of the formula increases from case definitions #1 to #2 to #3. While a prevalence formula for the "moderate or severe" CDC/AAP case definition should be possible (Eke et al., 2012), the use of multiple imputation, albeit more computationally intensive, may be more feasible for prevalence estimation than SDM for some complex case definitions (Preisser et al., 2024). Another adaptation of the SDM approach would be to use weighted estimates of the mean and correlation parameters of the working CLF with the aim of lessening bias of the prevalence estimator under model misspecification, which, for example, may have led to better performance of RHM in this study.

In summary, when cluster sizes (e.g., the number of teeth in the oral cavity) are large in research involving biological clustering of disease indicators within humans, sampling methods may

be used to reduce cost and respondent burden resulting in partially sampled clusters. This article demonstrated that PRPs are a viable alternative to full-mouth data collection when analyzed by assuming an underlying statistical distribution to calculate periodontitis prevalence under a commonly accepted case definition for severe periodontitis.

References

Ainamo, J. (1982), "Development of the World Health Organization community periodontal index of treatment needs (CPITN)," *International Dental Journal*, 32, 281–291.

Alshihayb, T. S., Sharma, P., Dietrich, T., and Heaton, B. (2022), "Exploring periodontitis misclassification mechanisms under partial-mouth protocols," *Journal of Clinical Periodontology*, 49, 448–457.

Beck, J. D., Caplan, D. J., Preisser, J. S., and Moss, K. (2006), "Reducing the bias of probing depth and attachment level estimates using random partial-mouth recording," *Community Dentistry and Oral Epidemiology*, 34, 1–10.

Beck, J. D., Elter, J. R., Heiss, G., Couper, D., Mauriello, S. M., and Offenbacher, S. (2001), "Relationship of periodontal disease to carotid artery intima-media wall thickness: the atherosclerosis risk in communities (ARIC) study," *Arteriosclerosis, Thrombosis, and Vascular Biology*, 21, 1816–1822.

Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS) (2013), "National Health and Nutrition Examination Survey: Oral health examiners manual," `https://wwwn.cdc.gov/nchs/data/nhanes/2013-2014/manuals/Oral_Health_Examiners.pdf`.

Chattopadhyay, A., Arevalo, O., and Sohn, W. (2008), "Understanding measurement of dental diseases and research participation in practice set-up," *Dental Clinics of North America*, 52, 367–386.

Chen, M. X., Zhong, Y. J., Dong, Q. Q., Wong, H. M., and Wen, Y. F. (2021), "Global, regional, and national burden of severe periodontitis, 1990–2019: An analysis of the Global Burden of Disease Study 2019," *Journal of Clinical Periodontology*, 48, 1165–1188.

Drury, T. F., Winn, D. M., Snowden, C. B., Kingman, A., Kleinman, D. V., and Lewis, B. (1996), "An overview of the oral health component of the 1988–1991 National Health and Nutrition Examination Survey (NHANES III-Phase 1)," *Journal of Dental Research*, 75, 620–630.

Dye, B. A., Afful, J., Thornton-Evans, G., and Iafolla, T. (2019), "Overview and quality assurance for the oral health component of the National Health and Nutrition Examination Survey (NHANES), 2011–2014," *BMC Oral Health*, 19, 1–11.

Eke, P., Thornton-Evans, G., Wei, L., Borgnakke, W., and Dye, B. (2010), "Accuracy of NHANES periodontal examination protocols," *Journal of Dental Research*, 89, 1208–1213.

Eke, P. I., Dye, B. A., Wei, L., Slade, G. D., Thornton-Evans, G. O., Borgnakke, W. S., Taylor, G. W., Page, R. C., Beck, J. D., and Genco, R. J. (2015), "Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012," *Journal of Periodontology*, 86, 611–622.

Eke, P. I., Page, R. C., Wei, L., Thornton-Evans, G., and Genco, R. J. (2012), "Update of the case definitions for population-based surveillance of periodontitis," *Journal of Periodontology*, 83, 1449–1454.

Eke, P. I., Thornton-Evans, G. O., Wei, L., Borgnakke, W. S., Dye, B. A., and Genco, R. J. (2018), "Periodontitis in US adults: national health and nutrition examination survey 2009-2014," *The Journal of the American Dental Association*, 149, 576–588.

Emrich, L. J. and Piedmonte, M. R. (1991), "A method for generating high-dimensional multivariate binary variates," *The American Statistician*, 45, 302–304.

Higham, N. J. (1988), "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra and Its Applications*, 103, 103–118.

Holtfreter, B., Kuhr, K., Borof, K., Tonetti, M. S., Sanz, M., Kornman, K., Jepsen, S., Aarabi, G., Völzke, H., Kocher, T., et al. (2024), "ACES: A new framework for the application of the 2018 periodontal status classification scheme to epidemiological survey data," *Journal of Clinical Periodontology*.

Kingman, A. and Albandar, J. M. (2002), "Methodological aspects of epidemiological studies of periodontal diseases." *Periodontology 2000*, 29, 11–30.

Michalowicz, B. S., Hodges, J. S., and Pihlstrom, B. L. (2013), "Is change in probing depth a reliable predictor of change in clinical attachment loss?" *The Journal of the American Dental Association*, 144, 171–178.

Page, R. C. and Eke, P. I. (2007), "Case definitions for use in population-based surveillance of periodontitis," *Journal of Periodontology*, 78, 1387–1399.

Preisser, J., Shing, T., Qaqish, B., Divaris, K., and Beck, J. (2024), "Multiple imputation for partial recording periodontal examination protocols," *JDR Clinical & Translational Research*, 9, 52–60.

Preisser, J. S., Marks, S. J., Sanders, A. E., Akinkugbe, A. A., and Beck, J. D. (2017), "A new way to estimate disease prevalence from random partial-mouth samples," *Journal of Clinical Periodontology*, 44, 283–289.

Preisser, J. S. and Qaqish, B. F. (2014), "A comparison of methods for simulating correlated binary variables with specified marginal means and correlations," *Journal of Statistical Computation and Simulation*, 84, 2441–2452.

Prentice, R. L. (1988), "Correlated binary regression with covariates specific to each binary observation," *Biometrics*, 1033–1048.

Qaqish, B. F. (2003a), "Bahjat F. Qaqish – Software," `http://www.bios.unc.edu/$\sim$qaqish/software.htm`, retrieved 2021.

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

Ramfjord, S. P. (1959), "Indices for prevalence and incidence of periodontal disease," *The Journal of Periodontology*, 30, 51–59.

Shing, T. L., Preisser, J. S., Sotres-Alvarez, D., Divaris, K., and Beck, J. D. (2023), "Patterns of site-level periodontal disease and within-mouth correlation among older adults in the Hispanic Community Health Study/Study of Latinos," *Community Dentistry and Oral Epidemiology*, 51, 927–935.

Shing, T. L., Preisser, J. S., and Zink, R. C. (2021), "GEECORR: A SAS macro for regression models of correlated binary responses and within-cluster correlation using generalized estimating equations," *Computer Methods and Programs in Biomedicine*, 208, 106276.

Susin, C., Kingman, A., and Albandar, J. M. (2005), "Effect of partial recording protocols on estimates of prevalence of periodontal disease," *Journal of Periodontology*, 76, 262–267.

Tonetti, M. S., Greenwell, H., and Kornman, K. S. (2018), "Staging and grading of periodontitis: Framework and proposal of a new classification and case definition," *Journal of Periodontology*, 89, S159–S172.

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.

Yang, W. and Chaganty, N. R. (2014), "A contrasting study of likelihood methods for the analysis of longitudinal binary data," *Communications in Statistics-Theory and Methods*, 43, 3027–3046.

Zeger, S. L. and Liang, K.-Y. (1986), "Longitudinal data analysis for discrete and continuous outcomes," *Biometrics*, 42, 121–130.

Received: January 30, 2024

Accepted: June 6, 2024