

The Clustering Effects of Surfaces within the Tooth and Teeth within Individuals

Journal of Dental Research
2015, Vol. 94(2) 281–288
© International & American Associations
for Dental Research 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0022034514559408
jdr.sagepub.com

M. Masood¹, Y. Masood², and J.T. Newton³

Abstract

The objectives of this study were 1) to provide an estimate of the value of the intraclass correlation coefficient (ICC) for dental caries data at tooth and surface level, 2) to provide an estimate of the design effect (DE) to be used in the determination of sample size estimates for future dental surveys, and 3) to explore the usefulness of multilevel modeling of cross-sectional survey data by comparing the model estimates derived from multilevel and single-level models. Using data from the United Kingdom Adult Dental Health Survey 2009, the ICC and DE were calculated for surfaces within a tooth, teeth within the individual, and surfaces within the individual. Simple and multilevel logistic regression analysis was performed with the outcome variables carious tooth or surface. ICC estimated that 10% of the variance in surface caries is attributable to the individual level and 30% of the variance in surfaces caries is attributable to variation between teeth within individuals. When comparing multilevel with simple logistic models, β values were 4 to 5 times lower and the standard error 2 to 3 times lower in multilevel models. All the fit indices showed multilevel models were a better fit than simple models. The DE was 1.4 for the clustering of carious surfaces within teeth, 6.0 for carious teeth within an individual, and 38.0 for carious surfaces within the individual. The ICC for dental caries data was 0.21 (95% confidence interval [CI], 0.204–0.220) at the tooth level and 0.30 (95% CI, 0.284–0.305) at the surface level. The DE used for sample size calculation for future dental surveys will vary on the level of clustering, which is important in the analysis—the DE is greatest when exploring the clustering of surfaces within individuals. Failure to consider the effect of clustering on the design and analysis of epidemiological trials leads to an overestimation of the impact of interventions and the importance of risk factors in predicting caries outcome.

Keywords: dental public health, caries, statistics, intra-class correlation coefficient (ICC), design effect (DE), multilevel modeling

Introduction

Epidemiological data are often clustered (hierarchical or nested) in nature; a simple example of clustered data is where students are clustered within a class and classes clustered within schools. This situation is a 3-level data structure, the first level being the students, the second level being the classes, and the third level being the schools. In broader scenarios, clustering can be seen in any setting, for example, in social settings (e.g., individuals → families → neighborhoods) or in geographical settings (e.g., wards → cities → counties → countries). In addition, in longitudinal studies or clinical trials, clustering appears due to aggregates of individuals or repeated measurement of the same subject (Masood, Yusof, et al. 2012a, 2012b; Fleming et al. 2013). However, in dental research, a special type of natural clustering is encountered, where surfaces are clustered within teeth and teeth clustered within individuals. However, frequently in dental research, the surface or tooth is taken as the unit of analysis and the clustering structure of surfaces → teeth → individuals is ignored, and thus each surface or

tooth is treated as an independent observation. This independence of observations treats 5 teeth from 20 patients (i.e., 100 observations) as equivalent to observations of 1 tooth from 100 patients. However, outcomes such as caries experience are likely to be more closely correlated within

¹Centre of Population Oral Health & Clinical Prevention Studies, Faculty of Dentistry, Universiti Teknologi MARA, Shah Alam, Malaysia

²Centre of Studies for Oral Pathology, Faculty of Dentistry, Universiti Teknologi MARA, Shah Alam, Malaysia

³Unit of Dental Public Health and Oral Health Services Research, King's College Dental Institute, King's College London, Denmark Hill, London, UK

A supplemental appendix to this article is published electronically only at <http://jdr.sagepub.com/supplemental>.

Corresponding Author:

M. Masood, Centre of Population Oral Health & Clinical Prevention Studies, Faculty of Dentistry, Universiti Teknologi MARA, Shah Alam, Malaysia.

Email: drmasoodmohd@gmail.com

clusters than between them. For example, the magnitude of caries risk is likely to be more comparable for different teeth within the same individual than would be the case between separate individuals (Hannigan and Lynch 2013).

When the within-cluster correlation or intraclass correlation coefficient (ICC) is very small (i.e., observations within the cluster are almost independent), the impact of the clustering on analysis can be ignored. However, where the ICC is high, the impact of clustering is likely to be much greater. Therefore, the degree to which statistical analysis should be modified as a result of clustering can be assessed by examining the estimated within-cluster correlation or ICC of different clusters. The ICC measures the similarity among outcome observations within the same cluster compared with that among observations in different clusters. The information contributed by each cluster in data analysis is inversely proportional to the ICC of the observations (Kerry and Bland 1998; Eldridge et al. 2006). Clustering of data has 2 important implications for the design and analysis of studies: a design effect (DE) whereby the sample size calculation is adjusted to account for clustering and the analytical methods selected to adjust estimates for clustering effects (Fleming et al. 2013).

Higher ICC values necessitate an increase in the required sample size in a clustered survey to maintain power analogous to a nonclustered survey. This increase in sample size can be determined by the DE, which is related to the ICC. The DE must be considered in the planning of oral health studies to ensure adequate statistical power. The DE is the ratio of the sample size required for a clustered design to that required for a design with independent samples to achieve the same power. For example, to obtain equal statistical power, a clustered design with $DE = 2$ requires twice as many observations as a nonclustered design. To calculate the sample size needed for a study, the DE for the anticipated cluster sizes and ICC must be considered. In study designs with more than one level (e.g., 3 levels in surfaces \rightarrow teeth \rightarrow individuals) of clustering, separate estimates of ICC and DE can be calculated for each level of clustering and used sequentially to adjust the required sample size (Litaker et al. 2013).

Clustered dental data are often analyzed by using classic statistical techniques, which are mostly based on the assumption of independent observations. Because observations within clusters violate the assumption of independence, these methods may not be appropriate (Litaker et al. 2013). For example, using classic regression approaches in analyzing clustered data results in biased regression estimates and standard errors (Loc Giang et al. 2011; Mamai-Homata et al. 2012). Therefore, if there is a significant correlation between the observations (i.e., a high ICC), the possibility of a type I error increases (Burnside et al. 2007, 2013). One alternative is to ignore clustering by undertaking a separate analysis for each subject—such an approach reduces the number of

observations considerably in each analysis, leading to an increased risk of type II errors. The most appropriate method of analysis for clustered data is multilevel modeling, which accounts for the correlation between clusters by modeling intercepts and regression coefficients as random (Diez Roux 2002). Caries data naturally fall into a 3-level structure, with the individual participant as the top or level 3 unit, tooth as the level 2 unit, and surface as the level 1 unit. Multilevel models work by splitting the variance in outcome into components for each level of the model, so random effects at tooth and participant levels are estimated in the modeling process. These random effects are assumed to follow a normal distribution with a mean of 0 and a variance that is estimated in the modeling process. Simulation studies have shown that parameter estimates are fairly robust to violations of this assumption (Burnside et al. 2007).

To date, multilevel models have not been used for the analysis of caries data from epidemiological surveys, and thus the potential of multilevel modeling to enhance efficiency and understanding of the risk factors is unknown. Therefore, there are 3 objectives of this study: 1) to provide an estimate of the value of the ICC for dental caries data at the tooth and surface levels, 2) to provide an estimate of the DE to be used in the determination of sample size estimates for future dental surveys, and 3) to explore the usefulness of multilevel modeling of cross-sectional survey data by comparing the model estimates derived from multilevel and single-level models of the same data.

Methods

Data from United Kingdom Adult Dental Health Survey 2009 (ADHS) were analyzed. A description of the ADHS is provided in the Appendix. The outcome variable was caries experience at the surface or tooth level. We recoded each tooth or surface as 0 for sound or 1 for carious or restored. Missing surfaces or teeth were excluded from the analysis as the reason for the missing surface or tooth was not known—assuming that they were missing as a result of caries would overestimate the ICC. This excluded 19,025 of 180,889 teeth (10.5%) and 98,612 of 904,400 surfaces (10.9%).

There are various methods available to calculate ICC for a binary outcome, as discussed by Wu et al. (2012) for randomized controlled trials and Fenn et al. (2004) for cross-sectional survey data. The ICC estimation from the random intercept logistic method suggested by Wu et al. has some limitations. In most cases, this method substantially overestimates the ICC values. Therefore, we used the formula suggested by Fenn et al. for cross-sectional survey data.

$$ICC = \frac{\sigma_b^2}{\sigma_x^2},$$

Equation 1 (Donald and Donner 1987; Fenn et al. 2004)

$$ICC = \frac{\sigma_b^2}{\pi(1-\pi)},$$

Equation 2 (Donald and Donner 1987; Fenn et al. 2004) where σ_b^2 is the between-cluster variance of the outcome variable. For binary outcome measures, $\sigma_x^2 = \pi(1-\pi)$, where π is the average cluster-specific proportion.

Multilevel logistic regression was used to produce an estimate of the ICC using a model that contains no explanatory variables, the so-called intercept-only model or null model. This model partitions the variance in the outcome variable into 2 independent components: σ_b^2 and σ_w^2 . Three null models—model N_1 (surfaces within a tooth), model N_2 (teeth within individual), and model N_3 (surfaces within individual)—were used to calculate the 3 ICC values using equations 1 and 2 (Diez Roux 2002). The sampling distribution of the variance estimates in multilevel logistic regression models is, in general, strongly asymmetric. Therefore, the standard error (SE) may be a poor characterization of the distribution, and confidence intervals (CIs) derived from the SE are likely to be unrepresentative of the data (Wu et al. 2012). Given this difficulty, we estimated the 95% CI of the ICC and DE by using “bootstrapping”—a technique for generating a description of the sampling properties of empirical estimators using random sampling with replacement from the original data set (Hox 2010).

The relationship between DE, cluster size, and ICC is represented in the following equation:

$$DE = 1 + (m-1)ICC,$$

Equation 3 (Fenn et al. 2004)

where DE is the design effect, and m is the average number of respondents per cluster, or average cluster size.

First, we performed an analysis for teeth clustered within individuals; the outcome variable was caries at the tooth level and took the value 0 if the tooth did not have caries and 1 if it had caries. The first model, model 0_t , was a simple logistic regression with no multilevel structure. This model was fitted only as a baseline for comparison with later model 1_t . The next model, model 1_t , was the 2-level model, allowing clustering of the teeth within individuals (Gilthorpe et al. 2000).

Second, the analysis of surfaces within teeth and teeth within individuals was performed; the outcome variable was caries at the surface level, which took the value 0 if the surface did not have caries and 1 if it had caries. Model 0_s was a simple logistic regression with no multilevel structure and formed the point of comparison for model 1_s , model 2_s , and model 3_s . Model 1_s and model 2_s were 2-level models, model 1_s incorporated clustering of the surface within teeth, and model 2_s incorporated clustering of the surfaces within individuals. Finally, model 3_s was a 3-level model allowing clustering of the surface within teeth and teeth within individuals (Gilthorpe et al. 2000).

All analyses were performed with R software; for multilevel analysis, the lme4-package of R software was used (Bates et al. 2014). For each model, a set of explanatory variables was included comprising age, sex, educational qualification, brushing, sugar intake, and use of fluoridated toothpaste. See the Appendix for the description of all these variables. All explanatory variables were modeled at the same level as the outcome variable in model 0_t and model 0_s since they were simple logistic regression models. However, in model 1_t , model 1_s , model 2_s , and model 3_s , the explanatory variables were modeled at the individual level (i.e., at level 2 in 2-level models and at level 3 in 3-level models). Regression estimates (β), SE, and level of significance were compared for all the models. Model fit was assessed by examining various fit indices: Akaike information criterion (AIC), Bayesian information criterion (BIC), deviance, -2 log-likelihood, and the chi-square test (Burnside et al. 2007, 2013).

Results

A total of 6,469 dentate adults were examined for dental caries; 9 individuals had only 1 tooth and were therefore excluded from the clustered analyses. The final sample size included 6,460 individuals, 161,855 teeth, and 805,788 surfaces. The mean number of teeth per person was 25.0, with an average of 4.97 surfaces per tooth and 124.5 surfaces per individual. Table 1 gives the descriptive analysis of all variables at the individual, tooth, and surface levels.

Table 2 displays the ICC and DE for teeth within individuals, surfaces within teeth, and surfaces within individuals. All values of the ICC are greater than 0.1, with the highest ICC being that for surfaces clustered within the individual. The most obvious feature of the results is the inverse relation between cluster size and ICC. For larger cluster sizes, even small ICCs might be associated with a substantial DE that should not be ignored in designing studies. The ICC estimate shows that 21% of the variance can be attributed to variation between individuals, 10% of the variance in surface caries is attributable to the individual level, and 30% of the variance in surface caries is attributable to variation between teeth within individuals.

Table 3 shows the values of β , SE(β) for models 0_t and model 1_t , and the ratio of coefficients and SE for model 0_t /model 1_t . Both models show a significant positive relationship of caries with older age, being female, higher educational qualification, and high sugar intake. Significant negative relationships exist between caries and brushing frequency, as well as fluoride usage. The relationship between caries and fluoride usage was not significant in the multilevel analysis for the category of 1,000 to 1,300 parts per million (ppm). If we consider only those results where $P \leq 0.01$, the relationship becomes insignificant for all the fluoride categories. Comparing the multilevel model with

Table 1. Descriptive Analysis of Data at Individual, Tooth, and Tooth Surface Levels.

	Individual (n = 6,469), n (%)	Teeth (n = 161,855), n (%)	Surfaces (n = 805,788), n (%)
Age, y			
16–44	2,837 (43.86)	79,606 (49.18)	396,397 (49.19)
45–64	2,355 (36.4)	57,678 (35.64)	287,114 (35.63)
65 and older	1,277 (19.74)	24,571 (15.18)	122,277 (15.17)
Sex			
Male	2,961 (45.77)	73,795 (45.59)	367,242 (45.58)
Female	3,508 (54.23)	88,060 (54.41)	438,546 (54.42)
Educational qualification			
No	1,499 (23.17)	31,999 (19.77)	159,151 (19.75)
Yes	4,966 (76.77)	129,766 (80.17)	646,192 (80.19)
Brushing frequency			
<Once a day	181 (2.80)	3,897 (2.41)	19,354 (2.40)
Once a day	1,450 (22.41)	34,617 (21.39)	172,365 (21.39)
≥Twice or more a day	4,823 (74.56)	122,994 (75.99)	612,340 (75.99)
Sugar intake			
Low	3,256 (50.33)	81,099 (50.11)	402,109 (49.90)
High	3,213 (49.67)	80,756 (49.89)	403,679 (50.10)
Toothpaste fluoride concentration, ppm			
≤550	996 (15.40)	24,562 (15.18)	122,267 (15.17)
1,000–1,300	1,061 (16.40)	26,591 (16.43)	132,404 (16.43)
1,350–1,500	4,351 (67.26)	109,434 (67.61)	544,819 (67.61)
Cariou tooth			
No	4,543 (70.2)	106,863 (66.02)	—
Yes	1,926 (29.8)	54,992 (33.98)	—
Cariou surface			
No	4,543 (70.2)	—	667,478 (82.84)
Yes	1,926 (29.8)	—	138,310 (17.16)

—, blank.

Table 2. ICC and Design Effect for Various Oral Health–Related Outcome Variables.

	Model N ₁	Model N ₂	Model N ₃
Outcome	Cariou tooth = Yes/No	Cariou surface = Yes/No	Cariou surface = Yes/No
Clustering	Individual	Tooth	Individual
Number of individuals	6,469	—	6,469
Number of teeth	161,855	161,855	—
Number of surfaces	—	805,788	805,788
Average cluster size	25.0	4.97	124.5
Individual-level variance (SD)	0.047 (0.216)	—	0.029 (0.170)
Tooth-level variance (SD)	—	0.015 (0.121)	—
Residual variance (SD)	0.179 (0.423)	0.104 (0.323)	0.104 (0.323)
Average proportion (π)	0.34	0.17	0.11
ICC (95% CI)	0.21 (0.204–0.220)	0.10 (0.063–0.159)	0.30 (0.284–0.305)
DE (95% CI)	6.04 (5.89–6.30)	1.40 (1.24–1.59)	38.0 (36.85–39.11)

CI, confidence interval; DE, design effect; ICC, intraclass correlation coefficient; SD, standard deviation; —, blank.

the simple logistic model, values of β were 4 to 5 times lower and the $SE(\beta)$ 2.3 times lower in multilevel models. All the fit indices suggested that the multilevel models were a better fit than the standard regression models.

Table 4 describes the results of the multivariate regression analysis of caries at the surface level, showing a simple logistic regression (model 0_s), a 2-level multilevel model where surfaces were clustered within individuals (model

Table 3. Multivariate Regression Analysis of Carious Teeth as an Outcome Variable, Showing Results of Simple Logistic Regression (Model 0_t), Multilevel Model (Model 1_t) (Level 1 = Tooth; Level 2 = Individual), and Ratio of Regression Coefficients (β_r) and Standard Error (SE_r).

	Model 0 _t		Model 1 _t		Ratio of Model 0 _t /Model 1 _t	
	β	SE	β	SE	β_r	SE _r
Fixed Effect						
Age, y						
16–44			Reference group			
45–64	1.15	0.012***	0.245	0.0054***	4.7	2.2
65 and older	1.38	0.016***	0.294	0.0070***	4.7	2.3
Sex						
Male			Reference group			
Female	0.10	0.011***	0.019	0.005***	5.3	2.2
Education qualification						
No			Reference group			
Yes	0.14	0.014***	0.034	0.0061***	4.1	2.3
Brushing frequency						
<Once a day			Reference group			
Once a day	-0.238	0.040***	-0.056	0.017**	4.3	2.3
≥Twice or more a day	-0.226	0.039***	-0.054	0.017**	4.2	2.3
Sugar intake						
Low			Reference group			
High	0.06	0.011***	0.015	0.0047**	4.0	2.3
Toothpaste fluoride concentration, ppm						
≤550			Reference group			
1,000–1,300	-0.056	0.019**	-0.014	0.0085	4.0	2.2 ^a
1,350–1,500	-0.064	0.015***	-0.015	0.0068*	4.2	2.2 ^b
Random Effect						
Individual	—	—	0.02956	0.171		
Residual	—	—	0.1787	0.422		
Model Fit						
AIC		192,825		189,089.1		
BIC		192,924.9		189,208.9		
Deviance		192,805		189,065.1		
-2 Log-likelihood		-96,402.52		-94,532.56***		

AIC, Akaike information criterion; BIC, Bayesian information criterion; —, blank.

^aSignificantly different at $P \leq 0.05$.

^bSignificantly different at $P \leq 0.01$.

* $P \leq 0.05$. ** $P \leq 0.01$. *** $P \leq 0.001$.

1_s), a 3-level multilevel model where surfaces were clustered within teeth and teeth were clustered within individuals (model 2_s), and the ratio of regression coefficients (β_r) and standard error (SE β_r). All models again showed a significant positive relationship between having carious surfaces and older age, being female, higher educational qualification, and high sugar intake. Similarly, both models found a significant negative relationship between caries at the surface level and brushing frequency and fluoride usage. However, if we consider $P \leq 0.01$, the relationship between caries at the surface level and fluoride intake becomes insignificant.

Discussion

The first aim of this report was to investigate the degree to which caries experience is correlated between teeth in the same mouth and across surfaces on the same tooth. A significant amount of the variability in caries experience at the surface level was explained by being on the same tooth—caries on the surfaces of the same tooth is more highly correlated than surfaces from different teeth. Similarly carious teeth in the same individual are more highly clustered than caries in different individuals. Thus, caries levels in the same individual or the same tooth are more correlated than

Table 4. Multivariate Logistic Regression Analysis of Carious Surfaces as an Outcome Variable, Showing Results of Simple Logistic Regression (Model 0_s), 2-Level Multilevel Model (Model 1_s) (Level 1 = Surfaces; Level 2 = Individual), 3-Level Multilevel Model (Model 2_s) (Level 1 = Surfaces; Level 2 = Tooth; Level 3 = Individual), and Ratio of Regression Coefficients (β_r) and Standard Error (SE_r).

	Model 0 _s		Model 1 _s		Model 2 _s		Ratio of Model 0 _s / Model 1 _s		Ratio of Model 0 _s / Model 2 _s	
	β	SE	β	SE	β	SE	β_r	SE_r	β_r	SE_r
Fixed Effect										
Age, y										
16–44					Reference group					
45–64	1.20	0.007***	0.157	0.004***	0.169	0.004***	7.6	1.8	7.1	1.8
65 and older	1.45	0.009***	0.197	0.005***	0.230	0.005***	7.4	1.4	6.3	1.4
Sex										
Male					Reference group					
Female	0.10	0.006***	0.012	0.004**	0.012	0.004***	8.3	1.5	8.3	1.5
Education qualification										
No					Reference group					
Yes	0.15	0.008***	0.026	0.004***	0.013	0.005**	5.8	2.0	11.5	1.6
Brushing frequency										
<Once a day					Reference group					
Once a day	–0.29	0.021***	–0.052	0.012***	–0.053	0.012***	5.6	1.8	5.5	1.8
≥Twice a day	–0.28	0.021***	–0.049	0.012***	–0.056	0.012***	5.7	1.8	5.0	1.8
Sugar intake										
Low					Reference group					
High	0.09	0.006***	0.014	0.004***	0.016	0.0036***	6.4	1.5	5.6	1.7
Toothpaste fluoride concentration, ppm										
≤550					Reference group					
1,000–1,300	–0.08	0.010***	–0.014	0.006*	–0.013	0.006*	5.7	1.7 ^a	6.2	1.7 ^a
1,350–1,500	–0.07	0.008***	–0.013	0.005**	–0.012	0.005*	5.4	1.6	5.8	1.6 ^a
Random Effect										
Individual	—		0.0182	0.1349	0.0191	0.138	—	—	—	—
Tooth	—		—	—	0.0147	0.121	—	—	—	—
Residual	—		0.1188	0.3447	0.1038	0.322	—	—	—	—
Model Fit										
AIC	688,820		583,111.5		477,469.3		—	—	—	—
BIC	688,935.7		583,250.6		477,619.9		—	—	—	—
Deviance	688,799.8		583,087.5		477,443.3		—	—	—	—
–2 Log-likelihood	–344,399.9		–291,543.8		–238,721.6		—	—	—	—

AIC, Akaike information criterion; BIC, Bayesian information criterion; —, blank.

^aSignificantly different at $P \leq 0.01$.

* $P \leq 0.05$. ** $P \leq 0.01$. *** $P \leq 0.001$.

dental caries measured in different individuals or different teeth. These findings have significant implications for the modeling of dental caries, in that ignoring within-individual or within-tooth correlation could result in biased regression estimates and standard errors. The regression models with and without consideration of the intraclass correlation compared in this study quantify the extent of this bias in estimates of regression statistics (Gunsolley et al. 1994; Chuang et al. 2005).

The effect of clustering on sample size requirements can be substantial, especially for large cluster sizes. Failure to account for clustering typically leads to an underestimation

of the required sample size. If the cluster size is large, even low values of ICC have a considerable impact, for example, in the case of surfaces clustered in an individual. It is also important to note that for dichotomous variables such as caries experience (yes or no), the ICC varies with the prevalence of the outcome, tending to increase with higher prevalence (Litaker et al. 2013). Therefore, the DE values identified in this study should be considered in the light of the prevalence of the carious tooth and carious surfaces in this sample (Table 1), as suggested by Gulliford (2005).

The ICC is a portable parameter that can be compared across studies since it does not depend on the cluster size or

on the numbers of clusters. The DE, on the other hand, is affected by the sample design and is strongly dependent on cluster size (Fenn et al. 2004). Therefore, not only DE but also ICC should be considered for accurate sample size calculation for future studies (Litaker et al. 2013).

This study has shown a substantial difference between the estimated power of multilevel models and classic regression analysis. Traditional regression analysis is known to overestimate the beneficial effects of interventions where clustering is present (Shaffer et al. 2013; Masood and Reidpath 2014). In this analysis, we used fluoridation and frequency of brushing as examples of preventive measures: simple regression analysis ascribes a nearly 6 times greater beneficial effect to such interventions than 2- or 3-level multilevel analysis. It was interesting to see that if we require more precise results with a P value of 0.01, then the preventive effect of fluoride becomes insignificant in multilevel models. Similarly, traditional analysis also overestimates the effect of risk factors; in this analysis, the effect of sugar consumption on dental caries was nearly 6 times higher in the classic logistic regression than in the 2- and 3-level multilevel analysis (Chuang et al. 2001; Masood, Masood, et al. 2012). Although this was not an objective for this study, it is important to discuss the unexpected positive relationship between educational qualification and caries. There could be 2 possible reasons for this relationship: 1) the educational qualification was coded as “yes” or “no,” which might not capture the education effect well, and 2) the outcome was whether or not the tooth or surface was sound or carious/restored. Education may be related to restoration via use of dental services.

There may be a number of reasons why an investigator or health care provider would be interested in the multilevel analysis of caries data, especially for preventive agents or treatments at different levels. Multilevel modeling offers the advantage of allowing greater understanding of the patterns of caries development within the mouth since it allows estimates to be made of the relative variance at individual, tooth, and surface levels (Burnside et al. 2007). Therefore, investigators who are interested in exploring the effect of their intervention in more detail may wish to consider the use of multilevel modeling (Burnside et al. 2013). The ability to quantify relative treatment effect sizes at the tooth level may be important when dental caries presents more commonly on different tooth types, for example, the relative benefit of fluoride varnish on molars and incisors. Investigators may also need to measure the comparative benefit at the surface level, or level 1 (e.g., being able to determine whether a preferential benefit is found on fissures compared with smooth surfaces).

It should be noted that although the results are based on a large sample size, the data set comprises a sample of the population of UK adults with a particular disease level, and the results may vary in populations with higher or lower

levels of caries. Younger age groups will certainly show different patterns due to the different teeth present, deciduous teeth in very young children, and a mixture of deciduous and permanent teeth in older children. The finding of differing probabilities of caries according to differing tooth types, with molars most susceptible, is well established (Reidpath et al. 2014). The current work indicates that the advantages of multilevel modeling in dental caries may lie in a greater understanding of the data structure and within-mouth patterns of caries development, rather than a reduction in required sample sizes (Hannigan and Lynch 2013). In addition to the hierarchy reported here, further levels of clustering exist in this data set, such as individual → neighborhood → district → county. Data on these clusters were not available in the data set owing to confidentiality issues. Another limitation was the inability to include survey design features in this analysis. Multilevel analysis for logistic regression models incorporating survey design features is not currently available (Thomas Lumley, developer of the survey package in R-project, personal communication [StataCorp 2013]).

Conclusion

This study has provided estimates of the ICC for dental caries data: 0.21 (95% CI, 0.204–0.220) at the tooth level and 0.30 (95% CI, 0.284–0.305) at the surface level. The DE used for sample size calculation for future dental surveys will vary on the level of clustering, which is important in the analysis—the DE is greatest when exploring the clustering of surfaces within individuals. In such instances, clustering will have a considerable effect on the required sample size. Failure to consider the effect of clustering on the design and analysis of epidemiological trials leads to an overestimation of the impact of interventions and the importance of risk factors in predicting caries outcome.

Author Contributions

M. Masood, contributed to conception, design, data acquisition, analysis, and interpretation, drafted and critically revised manuscript; Y. Masood, contributed to design, data analysis, and interpretation, drafted and critically revised manuscript; T.J. Newton, contributed to data interpretation, critically revised manuscript. All authors gave final approval and agree to be accountable for all aspects of the work.

Acknowledgments

The authors received no financial support and declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

References

Bates D, Mächler M, Bolker B, Walker S. 2014. Fitting linear mixed-effects models using lme4. *J Stat Software*. In press.

- Burnside G, Pine CM, Williamson PR. 2007. The application of multilevel modelling to dental caries data. *Stat Med*. 26(22):4139–4149.
- Burnside G, Pine CM, Williamson PR. 2013. Statistical power of multilevel modelling in dental caries clinical trials: a simulation study. *Caries Res*. 48(1):13–18.
- Chuang SK, Cai T, Douglass CW, Wei LJ, Dodson TB. 2005. Frailty approach for the analysis of clustered failure time observations in dental research. *J Dent Res*. 84(1):54–58.
- Chuang SK, Tian L, Wei LJ, Dodson TB. 2001. Kaplan-Meier analysis of dental implant survival: a strategy for estimating survival with clustered observations. *J Dent Res*. 80(11):2016–2020.
- Diez Roux AV. 2002. A glossary for multilevel analysis. *J Epidemiol Commun Health*. 56(8):588–594.
- Donald A, Donner A. 1987. Adjustments to the Mantel-Haenszel chi-square statistic and odds ratio variance estimator when the data are clustered. *Stat Med*. 6(4):491–499. Erratum in: *Stat Med*. 1997;16(24):2927–2928.
- Eldridge SM, Ashby D, Kerry S. 2006. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. *Int J Epidemiol*. 35(5):1292–1300.
- Fenn B, Morris SS, Frost C. 2004. Do childhood growth indicators in developing countries cluster? Implications for intervention strategies. *Public Health Nutr*. 7(7):829–834.
- Fleming PS, Koletsi D, Polychronopoulou A, Eliades T, Pandis N. 2013. Are clustering effects accounted for in statistical analysis in leading dental specialty journals? *J Dent*. 41(3):265–270.
- Gilthorpe MS, Maddick IH, Petrie A. 2000. Introduction to multilevel modelling in dental research. *Commun Dent Health*. 17(4):222–226.
- Gulliford MC, Adams G, Ukoumunne OC, Latinovic R, Chinn S, Campbell MJ. 2005. Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data. *J Clin Epidemiol*. 58(3):246–251.
- Gunsolley JC, Williams DA, Schenkein HA. 1994. Variance component modeling of attachment level measurements. *J Clin Periodontol*. 21(4):289–295.
- Hannigan A, Lynch CD. 2013. Statistical methodology in oral and dental research: pitfalls and recommendations. *J Dent*. 41(5):385–392.
- Hox JJ. 2010. *Multilevel analysis: techniques and applications*. 2nd ed. London (UK): Routledge.
- Kerry SM, Bland JM. 1998. Sample size in cluster randomisation. *BMJ*. 316(7142):1455.
- Litaker MS, Gordan VV, Rindal DB, Fellows JL, Gilbert GH. 2013. Cluster effects in a national dental PBRN restorative study. *J Dent Res*. 92(9):782–787.
- Loc Giang Do, Spencer AJ, Roberts-Thomson KF, Hai Dinh Trinh, Thuy Thanh Nguyen. 2011. Oral health status of Vietnamese children: findings from the National Oral Health Survey of Vietnam 1999. *Asia Pac J Public Health*. 23(2):217–227.
- Mamai-Homata E, Topitsoglou V, Oulis C, Margaritis V, Polychronopoulou A. 2012. Risk indicators of coronal and root caries in Greek middle aged adults and senior citizens. *BMC Public Health*. 12:484.
- Masood M, Masood Y, Newton T. 2012. Impact of national income and inequality on sugar and caries relationship. *Caries Res*. 46(6):581–588.
- Masood M, Reidpath DD. 2014. Multi-country health surveys: are the analyses misleading? *Curr Med Res Opin*. 30(5):857–863.
- Masood M, Yusof N, Hassan MI, Jaafar N. 2012a. Assessment of dental caries predictors in 6-year-old school children: results from 5-year retrospective cohort study. *BMC Public Health*. 12:989.
- Masood M, Yusof N, Hassan MI, Jaafar N. 2012b. Longitudinal study of dental caries increment in Malaysian school children: a 5-year cohort study. *Asia Pac J Public Health*. 26:260–267.
- Reidpath DD, Masood M, Allotey P. 2014. How much energy is locked in the USA? Alternative metrics for characterising the magnitude of overweight and obesity derived from BRFSS 2010 data. *Int J Public Health*. 59(3):503–507.
- Shaffer JR, Feingold E, Wang X, Weeks DE, Weyant RJ, Crout R, McNeil DW, Marazita ML. 2013. Clustering tooth surfaces into biologically informative caries outcomes. *J Dent Res*. 92(1):32–37.
- StataCorp LP. 2013. *STATA multilevel mixed-effects reference manual 13*. College Station (TX): StataCorp LP.
- Wu S, Crespi CM, Wong WK. 2012. Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials. *Contemp Clin Trials*. 33(5):869–880.