

Contextual effects in the occurrence of periodontal attachment loss and necrotizing gingival lesions among adolescents

Lopez R, Frydenberg M, Baelum V. Contextual effects in the occurrence of periodontal attachment loss and necrotizing gingival lesions among adolescents. *Eur J Oral Sci* 2009; 117: 547–554. © 2009 The Authors. Journal compilation © 2009 Eur J Oral Sci

The aim of this study was to assess and quantify the random effects resulting from clustering in the following individual-level periodontal outcomes: presence of clinical attachment loss of ≥ 1 mm (CAL1), presence of clinical attachment loss of ≥ 3 mm (CAL3), and presence of necrotizing ulcerative gingivitis (NUG); or in the following class-level periodontal outcomes: number of students with CAL1, number of students with CAL3, and number of students with NUG. Mixed-effects logistic regression analysis was used to model these outcomes among 9,162 adolescents in 310 classes in 98 schools spread over 20 communes in the Province of Santiago, Chile, who had been examined for clinical attachment level and NUG, and had completed questionnaires on oral health-related behaviors. The results of all six analyses demonstrated statistically significant random effects, which in all analyses were particularly related to the schools, whereas the class effects were smaller and the commune random effects were almost negligible. The random effects were quantified using the median odds ratio (MOR), and the class-level MOR ranged between 1.05 and 1.51, whereas the school-level MOR values ranged from 2.07 to 2.39. The results of the study demonstrate the potential of the application of multilevel modeling to periodontal epidemiologic data, over and beyond the conventional use of the technique to account for the intrinsic sites–teeth–subject hierarchy in periodontal data.

Rodrigo Lopez¹, Morten Frydenberg², Vibeke Baelum^{3,4}

¹Department of Periodontology, School of Dentistry, Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark;

²Department of Biostatistics, School of Public Health, Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark; ³School of Dentistry, Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark;

⁴Department of Epidemiology, School of Public Health, Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark

Rodrigo Lopez, Department of Periodontology, School of Dentistry, Faculty of Health Sciences, University of Aarhus, Vennelyst Boulevard 9, DK-8000 Aarhus C, Denmark

Telefax: +45-89-424141
E-mail: rlopez@odont.au.dk

Key words: adolescent; cross-sectional studies; periodontitis; cluster analysis; logistic models

Accepted for publication May 2009

The introduction of multilevel modeling to periodontal research (1–3) has provided a new way of exploring old problems (4). Until multilevel modeling became an option it was usually necessary to aggregate the site-based recordings into subject-based summary statistics to fulfill basic requirements for valid statistical inference, but multilevel modeling permits statistical analyses that make explicit use of the natural hierarchy existing in most clinical periodontal data, with sites nested in teeth that are nested in subjects (1, 5–7). Although such accounting for the natural hierarchy has sometimes caused a change in the interpretation of the results (1, 4, 6), multilevel modeling in periodontal research seems so far mainly to have served the purpose of taking into account the site–tooth–subject correlation structure using a population-averaged approach (5, 8–14) in which the heterogeneity is considered a nuisance parameter. Only little attention has been paid to the estimated variance components (1, 13–15), and the heterogeneity of the responses has rarely been a focus of interest.

However, multilevel modeling can fruitfully be extended beyond the subject level to explore, or to account for – as appropriate, the effects of clustering above and beyond the subject level. Two situations seem

particularly suitable for such uses of multilevel analysis: in periodontal epidemiology, the sampling methods used often involve sampling of clusters somewhere in the study subject-identification process (16–18). Even the end-stage selection of study subjects may be based on clusters – for example, when including in the study all students in a school (19, 20), in a class (21), or in a household (22, 23) – and such clustering might well be important for the interpretation of the study results. However, the perhaps most intriguing use of multilevel modeling in periodontal epidemiology is for the identification or assessment of contextual effects. Contextual effects are effects that are not directly captured by subject-level attributes or covariates (24, 25), and could reflect the effects of the norms, values, and beliefs prevailing in the individuals' social context (26), or certain characteristics of the physical environment in which the individual is living. A good example of a contextual effect is the influence of safe walking environments for the walking activity of older adults (27).

Contextual effects have been reported for oral health outcomes, and the residents of disadvantaged neighborhoods are more likely to assess their oral health as only fair or poor, compared with similar persons in more

advantaged neighborhoods, even after adjustment for subject-specific socioeconomic factors (28–30). While the results of BOWER *et al.* (31) did not confirm the presence of contextual effects related to deprivation in hard clinical outcomes, such as the number of sound teeth or the presence of one or more pockets ≥ 4 mm among Scottish adults, the results presented by SANDERS *et al.* (32) have demonstrated considerable neighborhood contextual effects in the perhaps hardest clinical outcome of them all: the number of teeth retained. This study convincingly demonstrated that poor persons benefit from living in affluent neighborhoods, whereas wealthier persons do not lose their oral health advantages when living in poor neighborhoods (32). Several explanations were offered for these findings, including greater availability of dental services in wealthy neighborhoods, greater accessibility of nutritious foods in wealthy neighborhoods, or an undermining effect of area deprivation on the social capital, creating disorder, mistrust, and social exclusion of people living in poor neighborhoods (32).

We have previously reported the findings of a large periodontal epidemiological study carried out among Chilean adolescents (21). The study participants were identified using cluster sampling, but at the time of reporting the study results, multilevel modeling techniques were not available to us. As some information is available about the cluster-defining units, we decided to undertake a secondary analysis of the data. The purpose of the present study was therefore to model epidemiological data that cluster in school classes, schools, and communes with different socioeconomic profiles, with a view to assess and quantify the variation in three subject-level periodontal-disease outcomes that can be attributed to unmeasured or latent variables at the three data hierarchies.

Material and methods

The data originate in a cross-sectional study carried out (21) for the purpose of identifying cases of early periodontitis among adolescents, 12–21 yr of age, from the Province of Santiago, Chile. The Province of Santiago, which had a

population of about five million at the time of the study, comprised 32 communes, in which 618 schools embrace the school grades corresponding to adolescence. At the time of planning the study it was estimated that 9,200 adolescents would have to be examined in order to identify the appropriate number of cases for an ensuing case-control study (33). The target sample size was obtained by means of a complex two-stage random-cluster sampling procedure, with a random selection of schools comprising the stage-1 sampling and a random selection of classes within schools comprising the stage-2 sampling (21, 34). The target sample size in each school was 100 adolescents, and at least three different classes were included from each school. However, it was not always possible to fulfill the minimum requirement of 100 students per school because there were too few students in the appropriate age range, and in 24 schools fewer than 80 students were available for the present study.

A total of 9,162 adolescents in 310 classes in 98 schools in 20 communes were given a clinical examination, consisting of the direct recording of the clinical attachment level (in mm) in six sites in the first and second molars and in the incisors (21). The gingival tissue was also examined, and necrotizing ulcerative gingivitis was recorded as present if one or more gingival areas displayed signs of such lesions (35). Sociodemographic information on each student was obtained by means of a questionnaire, which among others contained questions on toothbrushing habits, smoking habits, time since last dental visit, diabetic status, parental income, and parental level of education. The data collection was carried out during March–May 2000 by four examiners who worked together two-and-two in the schools. In case of the absence on a particular day of one of the examiners, the principal investigator (RL) acted as a substitute. The inter-examiner and intra-examiner reliabilities of the clinical recordings have been extensively described in a previous publication (36).

Two sets of statistical analyses were performed: one in which the subject was the basic unit of analysis; and one in which the class was the basic analytical unit. In the subject-level analyses, three outcomes were analyzed: presence of clinical attachment loss ≥ 1 mm (CAL1); presence of clinical attachment loss ≥ 3 mm (CAL3); and presence of necrotizing ulcerative gingivitis (NUG). All three subject-based outcome variables were coded as 1 = present/0 = not present. Table 1 shows the distribution of the study subjects and the distribution of the three outcomes across the implicit study hierarchies.

Table 1

Distribution of the study population and the three outcome variables according to the sampling hierarchy

Level	Estimate	Number of students			
		Total	With CAL ≥ 1 mm	With CAL ≥ 3 mm	With NUG
Commune	Median	308	230.5	18.5	26
	IQR	185–671.5	132.5–486	4.5–32	11.5–41
	Range	62–1212	28–920	0.63	(0–86)
School	Median	94.5	67	2	4
	IQR	80–112	46–82	1–7	2–10
	Range	10–152	9–136	0–23	0–27
Class	Median	31	21	1	1
	IQR	23–38	13–28	0–2	0–3
	Range	4–48	1–43	0–12	0–11
Subject	–	9,162	6,342	409	618

CAL, clinical attachment loss; IQR, interquartile range; NUG, necrotizing ulcerative gingivitis.

The selection of covariates for the present analyses was chiefly based on the results of previous analyses (21, 34, 35). Hence, the covariates considered were age (continuous variable centered at 16 yr); gender (1 = boy/0 = girl); toothbrushing less than twice per day (1 = yes/0 = no); smoking (1 = daily smoker/0 = occasional smoker or non-smoker); last dental visit > 1 yr ago (1 = yes/0 = no); being diabetic (1 = yes/0 = no); paternal income ≥ Chilean Pesos (\$) 300,000 (1 = yes/0 = no); and paternal education at least university or technical level (1 = yes/0 = no). For parental income and education the non-responses were recoded as 0. Because the main purposes of the analyses were to identify the variance components corresponding to class, school, and commune, we also included the examiner as a covariate to account for the way in which examiner teams had been distributed among the schools. Table 2 shows the distribution of the covariates across the implicit study hierarchies.

Table 2

Distribution of the subject-level covariates aggregated according to the three hierarchies in the analysis

Covariate	Level of aggregation			
	Subject n = 9,162	Class n = 310	School n = 98	Commune n = 20
Age (yr)				
Median	0.16	0.41	0.18	0.27
Range	-3.0 to 5.3	-1.8 to 3.7	-0.9 to 2.2	-0.7 to 1.6
Percentage boys				
Median	50.8	50.0	49.7	50.8
Range	-	0-100	0-100	23.9-74.0
Percentage with income ≥ \$300,000 (father)				
Median	39.1	34.6	35.3	32.1
Range	-	0-100	0-95.8	4.4-79.0
Percentage with university/technical education (father)				
Median	30.5	27.8	26.6	22.2
Range	-	0-100	0-90.1	2.4-75.5
Percentage brushing teeth less than twice daily				
Median	29.3	28.1	29.9	29.3
Range	-	0-80	8.2-58.2	12.9-52.4
Percentage last visit to dentist > 1 yr ago				
Median	47.1	46.7	46.7	50.6
Range	-	0-100	12.5-73.3	19.0-70.1
Percentage smoking daily				
Median	24.9	25.0	23.7	24.1
Range	-	0-100	7.5-60.0	16.1-36.3
Percentage diabetics				
Median	0.6	0.0	0.0	0.1
Range	-	0-16.7	0-8	0-2.4
Percentage examined by examiner 1				
Mean	23.9	25.3	25.0	22.4
Range	-	0-100	0-100	0-51.6
Percentage examined by examiner 2				
Mean	22.9	22.1	22.8	21.6
Range	-	0-83.3	0-58.2	0-51.1
Percentage examined by examiner 3				
Mean	24.9	24.6	24.5	26.7
Range	-	0-100	0-80.0	0-54.5
Percentage examined by examiner 4				
Mean	24.2	24.2	24.0	26.1
Range	-	0-100	0-76	0-51.6
Percentage examined by stand-in examiner				
Mean	4.1	3.8	3.7	3.2
Range	-	0-100	0-47.9	0-12.5

For the purpose of the subject-level analyses we fitted random-intercept logistic-regression models of the form:

$$\begin{aligned} \log it[P(Y_{ijk} = 1|x_{ijk}, \zeta_{class}, \zeta_{school}, \zeta_{commune})] \\ = \beta_1 + \beta_2 \cdot x_{2ijk} + \beta_3 \cdot x_{3ijk} + \dots + \beta_p \cdot x_{pijk} \\ + \zeta_{class} + \zeta_{school} + \zeta_{commune} \end{aligned}$$

where x_{ijk} is the vector of subject-level covariates in across individuals i , schools j , and communes k ; and the terms ζ_{class} , ζ_{school} , and $\zeta_{commune}$ denote the random effects at class level, school level, and commune level, respectively.

For the purpose of the class-level analyses the above outcomes were summed over classes, yielding the outcome variables ‘Number of students with CAL1’, ‘Number of students with CAL3’, and ‘Number of students with NUG’. Similarly, covariates were aggregated over classes and the following covariates were considered: mean age (continuous variable centered around 15 yr); number of boys in class; number of students whose father had an income of ≥ \$300,000; number of students whose father had a university or technical education; and number of students examined by each examiner. We elected to reduce substantially the number of covariates for the class-level analyses compared with the subject-level models because of the much smaller number of analytical units available for the class-level analysis (310 vs. 9,162). The excluded covariates were those with the smallest anticipated effect. The models fitted in the class-level analyses were random-intercept logistic-regression models of the form:

$$\begin{aligned} P(Y_{class} = d|x_{classij}, \zeta_{class}, \zeta_{school}, \zeta_{commune}) \\ = \binom{n_{class}}{d} \cdot (P_{class})^d \cdot (1 - P_{class})^{1-d}; \end{aligned}$$

with

$$\begin{aligned} \log it(P_{class}) = \beta_1 + \beta_2 \cdot x_{2ij} + \beta_3 \cdot x_{3ij} + \dots \\ + \beta_p \cdot x_{pij} + \zeta_{class} + \zeta_{school} + \zeta_{commune} \end{aligned}$$

where Y_{class} denotes the number of students in the class with the outcome of interest; $x_{classij}$ is the vector of class-level covariates in school i in commune j ; n_{class} is the number of students in the class; and P_{class} is the probability of a student having the outcome of interest. The terms ζ_{class} , ζ_{school} , and $\zeta_{commune}$ denote the random effects at class level, school level, and commune level, respectively. These random effects follow a normal distribution with a mean value of 0 and a variance that can be estimated. The statistical significance of the random effects was tested using the likelihood-ratio test.

Both the subject-level and the class-level models were fitted using the STATA (37, 38) procedure ‘xtmelogit’ for multilevel mixed-effects logistic regression. This STATA procedure uses maximum-likelihood estimation. For the class-level models, we used binomial response probabilities and n_{class} as the binomial denominator. The fit of the estimated models was evaluated using the postestimation procedure ‘predict’ to obtain the predicted probability of $Y_{ijk} = 1$, which was compared with the observed value of Y at each of the three random-effects levels: class, school, and commune. When the model is correctly specified, the predicted probabilities of $Y_{ijk} = 1$ will approximate the observed proportions with $Y = 1$ at each of the class, school, and commune levels. This was checked graphically by plotting the differences between the observed proportions

and the predicted probabilities. In addition, we evaluated the fit of the estimated models by simulating the responses based on the estimated models, and graphically comparing the simulated responses of each of 20 simulations with the observed responses.

The interpretation of the fixed effects and the random effects estimated based on the above models is not straightforward (39). Hence, the estimated random effects are conditional on the particular set of covariates considered in the fixed part of the logistic models, just as the fixed-effects estimates, β , have class, school, and commune-specific interpretations. LARSEN *et al.* (39, 40) suggested that the random effects at a given level, say class, should be reported in the form of the median odds ratio (MOR) between the unit at higher propensity and the unit at lower propensity, given the fixed effects. The MOR is always ≥ 1 . If the MOR is 1, there is no variation between clusters; whereas if there is considerable between-cluster variation the MOR will be large. If we consider two students, each with the same given covariate pattern, from two randomly chosen different classes in the same school – hence also in the same commune – we can use the MOR to quantify the class-level variation. Phrased differently, the MOR may be understood to quantify, in odds ratio units, the heterogeneity between classes caused by unmeasured or latent covariates operating at the class level.

In the present study we are not seeking primarily to estimate fixed effects, and it should be borne in mind that it is inappropriate merely to exponentiate the estimated fixed-effect β -values and interpret the results as odds ratios from a conventional logistic regression analysis. Hence, the $\exp(\beta_i)$ value for the above-mentioned three-level model represents the odds ratio for the association between the outcome Y and exposure X_i , conditional on the subjects being from the same class (and thereby also from the same school and the same commune), and having the same values of the remaining covariates.

The schools included in the present study comprise a mix of private and publicly run schools, and the school-based annual tuitions and fees vary from 0 to over 1 million Chilean Pesos (CLP). For the purpose of exploring a possible association with the estimated random effects, we grouped the tuitions and fees in four-fourths – $< \text{CLP}60,000$; $\text{CLP}60,000\text{--}\text{CLP}149,000$; $\text{CLP}150,000\text{--}\text{CLP}760,000$; and $> \text{CLP}760,000$ – and explored the distributional characteristics of the random effects across these groups. Similarly, based on the year 2000 survey of the socioeconomic characteristics of the Chilean provinces (41), we calculated, for each commune in the Province of Santiago, a commune socioeconomic level based on three indicators: the average income; the average number of years of education; and the proportion of unemployed in the workforce. Each of these was scored from 1 (lowest third in socioeconomic position) to 3 (highest third in socioeconomic position), and the scores were summed to form a composite indicator of commune socioeconomic position. This was subsequently used to form three groups of communes, of lower, middle, or higher socioeconomic position.

Results

The results of the multilevel mixed-effects logistic regression of the subject-based outcome, presence of CAL1, clearly indicated notable random effects at both

the class level and the school level (Table 3), indicating considerable heterogeneity in the presence of CAL1 at both class level and school level. Hence, the MOR contrasting two students having the same covariate pattern, but originating from two different classes in the same school, amounted to a MOR = 1.51, whereas two students having the same covariate pattern, but originating from different schools in the same commune, were contrasted by a MOR of 2.39. Both of these random effects were statistically significant, as indicated by the likelihood ratio test, at $P < 0.0001$. The random effect for CAL1 at the commune level was minute and statistically insignificant. Considering the outcomes ‘presence of CAL3’ and ‘presence of NUG’, it was evident that the main random effect occurred at the school level (MOR = 2.25 and MOR = 2.24, respectively). Both class-level and commune-level random effects were statistically insignificant, although the class-level contribution to the median odds ratio contrast was in an order of magnitude of 20–31%, and the commune-level contributions were in an order of magnitude of 14–26% (Table 3).

The estimated fixed effects generally indicated that higher age, brushing teeth less than twice a day, and having had the last visit to the dentist more than a year ago were associated with higher odds for the presence of CAL1, presence of CAL3, and presence of NUG; whereas having a father with a high income and a high level of education was associated with lower odds of the presence of any of the three outcomes. Daily smoking tended to be negatively associated with the outcomes, while being male and reporting being diabetic did not appear to be associated with the outcomes. The examiner had variable, and sometimes considerable, effects on the outcomes (Table 3).

Even though the class-level outcomes ‘number of students with CAL1, CAL3 or NUG’ were modeled using fewer covariates in the fixed-effects part of the models, the results concerning the random effects and their distribution among the class, school, and commune hierarchy essentially confirmed the results obtained in the subject-level analyses. Hence, for the outcome ‘number of students in class with CAL1’, considerable and statistically significant random effects corresponding to the class level (MOR = 1.46) and school level (MOR = 2.12) were found, whereas the commune-level random effect was negligible (Table 4). For the outcomes ‘number of students in class with CAL3 or NUG’, the only significant random effects were observed at the school level (MOR = 2.07 and MOR = 2.10, respectively). The fixed-effects portion of the models indicated a positive association between higher age and a higher proportion of students with the outcome, whereas a higher proportion of students with high paternal income and education were associated with a lower proportion of students with the outcome in question. The examiner effects were variable, but considerably less marked, than the age effect (Table 4).

Figures 1 and 2 show the distribution of the random effects estimated in the subject-based model for the CAL1 outcomes according to the four levels of

Table 3

Results of fitting the subject-based three-level random-effects models to the following outcomes: presence of CAL1; presence of CAL3; or presence of NUG

Random effects	CAL1		CAL3		NUG	
	σ^2	SE	σ^2	SE	σ^2	SE
Class	0.1861	0.0401	0.0769	0.0686	0.0358	0.0515
School	0.6512	0.1402	0.6483	0.2076	0.6779	0.1637
Commune	0.0053	0.0627	0.2174	0.1803	0.1106	0.1135
MOR for the OR comparing two randomly chosen:			MOR	MOR	MOR	MOR
Students in different classes fixed effects, school, commune			1.51	1.30	1.20	1.20
Students in different schools fixed effects, commune			2.39	2.25	2.24	2.24
Students in different communes fixed effects			2.40	2.52	2.38	2.38
Subject level – fixed effects	β	SE(β)	β	SE(β)	β	SE(β)
Age centered at 16 yr	0.18	0.03	0.18	0.05	0.08	0.04
Boy	-0.07	0.06	0.01	0.12	0.00	0.10
Paternal income \geq \$300,000	-0.10	0.06	-0.35	0.15	-0.15	0.12
Paternal university/technical education	-0.12	0.07	-0.01	0.16	-0.24	0.13
Brushing teeth less than twice daily	0.11	0.06	0.21	0.12	0.16	0.10
Last visit to dentist more than 1 yr ago	0.17	0.05	0.30	0.11	0.27	0.09
Smoking daily	-0.12	0.06	0.05	0.12	-0.22	0.11
Diabetic	0.16	0.33	0.51	0.51	0.55	0.41
Examiner 1	1.45	0.19	-1.61	0.25	0.23	0.25
Examiner 2	0.20	0.18	-2.19	0.27	-1.25	0.28
Examiner 3	-0.09	0.17	-1.82	0.26	0.19	0.26
Examiner 4	0.73	0.18	-1.34	0.25	-0.44	0.27

The random effects for class, school, and commune are described in terms of their variance (σ) and standard errors (SE). The median odds ratio (MOR) describes the contrast that arises from comparing two individuals with the same covariate patterns between classes, between schools, and between communes. The β values denote the logistic regression coefficients for the fixed part of the models fitted.

CAL1, presence of clinical attachment loss of ≥ 1 mm; CAL3, presence of clinical attachment loss of ≥ 3 mm; NUG, necrotizing ulcerative gingivitis.

school-based tuitions and fees (Fig. 1) and the three levels of the composite indicator of commune socioeconomic position (Fig. 2). The school-level random effects were clearly related to the school-based tuitions and fees in a gradient-like manner. The relationship between the random effects and the composite indicator of commune socioeconomic position was less clear, and only the school-level random effect seemed to be associated.

Discussion

The results of the present analysis clearly indicate the presence of contextual effects, predominantly related to the school level. The observation that the commune level did not appear to be a major source of random effects is perhaps not so surprising, taking into account the actual size of the communes. The smallest commune had about 70,000 inhabitants, whereas the largest commune had more than 400,000 inhabitants, and considerable within-commune heterogeneity is to be expected for communes of this size. Unfortunately, whereas we have plenty of data characterizing a vast number of commune-level socioeconomic indicators, we do not have information that characterizes the schools any further than by the information on the requested annual tuitions and fees.

Plausible, but speculative, explanations for the observed school-level random effects include neighborhood effects related to the school catchment area, effects related to the school’s particular philosophical or religious attachment, or effects related to the schools’ main orientation towards technical or academic disciplines, all of which may influence the collective norms, values, and beliefs prevailing in a given school.

Care should always be exercised when interpreting the results of multilevel mixed-effects logistic regression analyses. First of all it should be noted that the estimated random effects shown in Tables 3 and 4 are conditional on the fixed effects and on the adopted hierarchical structure. That is to say, the random effects are random effects, given the particular selection of covariates associated with the fixed effects, represented by the subject-level covariates in Table 3 and the class-level covariates in Table 4. It is thus possible that additional or different subject-level and class-level covariates might have resulted in different random-effects estimates. This observation hinges on the continued discussion of the relative merits of multilevel analysis vs. conventional individual single-level analysis (24). Critics of the use of multilevel analysis models argue that if context or group-level characteristics affect the health of individuals, they must somehow manifest

Table 4

Results of fitting three-level random-effects models to the following class-based outcomes: number of students with CAL1; number of students with CAL3; and number of students with NUG

Random effects	Number of students in class with					
	CAL1		CAL3		NUG	
	σ^2	SE	σ^2	SE	σ^2	SE
Class	0.1576	0.0369	0.0539	0.0642	0.0025	0.0446
School	0.4664	0.1059	0.5522	0.1841	0.6059	0.1497
Commune	0.0298	0.0545	0.0511	0.0964	0.0757	0.0913
MOR for the OR comparing two randomly chosen:			MOR	MOR	MOR	MOR
Classes fixed effects, school, commune			1.46	1.27	1.05	1.05
Classes in different schools fixed effects, commune			2.12	2.07	2.10	2.10
Classes in different communes fixed effects			2.16	2.07	2.20	2.20
Class level – fixed effects	β	SE(β)	β	SE(β)	β	SE(β)
Mean age centered at 16 yr	0.1471	0.0350	0.1846	0.0605	0.0373	0.0454
No. of boys	-0.0003	0.0067	0.0052	0.0090	-0.0052	0.008
No. with paternal university/technical education	-0.0136	0.0153	-0.0037	0.0271	-0.0743	0.0226
No. with paternal income \geq \$300,000	-0.0120	0.0131	-0.0778	0.0246	0.0229	0.0196
Examiner 1	0.0399	0.0144	0.0310	0.0214	0.0047	0.0173
Examiner 2	-0.0267	0.0134	-0.0113	0.0216	-0.0154	0.0176
Examiner 3	-0.0114	0.0140	-0.0005	0.0198	0.0171	0.0166
Examiner 4	-0.0170	0.0128	0.0360	0.0189	-0.0199	0.0151
Stand-in examiner	0.0057	0.0207	0.0787	0.0231	0.0045	0.023

The random effects for class, school, and commune are described in terms of their variance (σ^2) and standard errors (SE). The median odds ratio (MOR) describes the contrast that arises from comparing two classes with the same covariate patterns between classes, between schools, and between communes. The β values denote the logistic regression coefficients for the fixed part of the models fitted.

CAL1, presence of clinical attachment loss of ≥ 1 mm; CAL3, presence of clinical attachment loss of ≥ 3 mm; NUG, necrotizing ulcerative gingivitis.

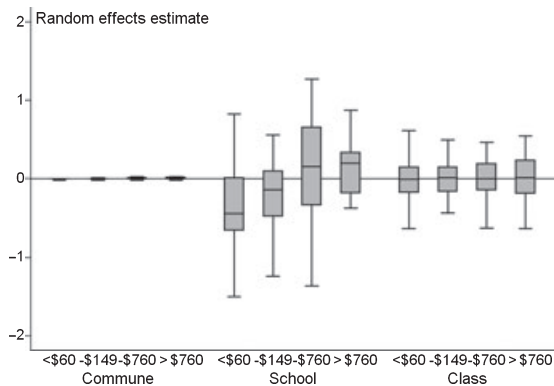


Fig. 1. Box-plots showing the distribution of the random-effects estimates for the subject-based logistic models estimated for the outcome ‘Having CAL ≥ 1 mm’ according to the quartile levels of the contextual variable ‘School-based tuitions and fees’ (see the text for definitions). CAL, clinical attachment loss.

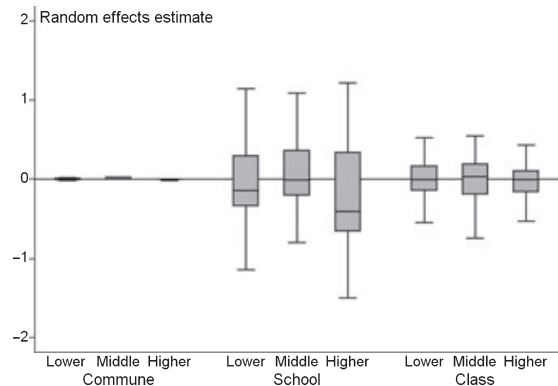


Fig. 2. Box-plots showing the distribution of the random-effects estimates for the subject-based logistic models estimated for the outcome ‘Having CAL ≥ 1 mm’ according to the levels of the contextual variable ‘Commune socioeconomic level’ (see the text for definitions). CAL, clinical attachment loss.

themselves in people’s bodies and operate via individual-level processes. This clearly implies that if enough individual-level covariates are measured, the group-level or contextual effects would disappear. However, this is not to say that we should strive to find the causes of the outcome only at the individual level. Returning to the results of the study by SANDERS *et al.* (32), demon-

strating that living in an affluent neighborhood confers a reduction in the excess risk of tooth loss among poor people, one might conceivably be able to introduce individual-level variables measuring, for example, the dental attendance patterns and the intake of various nutritious foods into the regression equations and thereby effectively eliminate the evidence for the

contextual neighborhood effects. However, if such an approach is taken, it is not a scientifically motivated decision, but a rather political/philosophical decision fuelled by a strong belief in people's ability to make rational and informed choices based on their personal assessment of the available evidence on all pros and cons. However, this 'victim-blaming' approach (42) has limited scientific credibility, and it is common sense that a person's dental-attendance pattern is likely to be influenced by the availability and accessibility of dentists sufficiently nearby, just as a person's intake of junk food clearly depends on the availability in the local area of shops or restaurants selling such foods.

Another aspect that may affect group-level random effects is related to confounding. Suppose an unknown individual-level variable exists that is related to the disease outcome and does not represent an intermediary step in the causal pathway linking a group-level characteristic to the disease. If such is the case, group-level effects will clearly be confounded by this unknown confounder. This is no different from the problem of unmeasured confounding in conventional single-level analysis. If, by contrast, the above hypothetical individual-level variable is in fact an intermediary step in the causal pathway linking a group-level characteristic to the disease, group-level effects will be attenuated if this variable is included in the analysis as a potential confounder. As is well known from conventional single-level analysis, the issue of confounding is not a statistical issue but a problem that can only be resolved through better insight into the determinants of the disease outcome and their mode of action.

Just as care must be exercised in the interpretation of the random effects, there is no simple interpretation of the fixed effects because these are contingent on the random effects. In fact, the application of a multilevel model is another way of stating that the lower-level covariate estimates, in the present study represented by the individual-level (Table 3) or class-level (Table 4) covariate estimates, are themselves random variables (39) rather than fixed parameters. LARSEN *et al.* (39) and LARSEN & MERLO (40) have suggested the use of the interval odds ratio (IOR) to overcome the problem that the random effects are unobserved. The α -level IOR may be interpreted as showing that the odds ratio for a given fixed effect (e.g. gender) for two randomly chosen subjects is contained in the IOR with a probability of α . The IOR is a prediction interval (not to be confused with a confidence interval), which reflects the influence of the random effects on the odds ratio for a given fixed effect. In the present study, we did not calculate such intervals because we were primarily interested in assessing the random effects.

In conclusion, the results of the present study have demonstrated the feasibility of the use of multilevel modeling approaches to the assessment of contextual effects in periodontal epidemiological data. Our results suggest that multilevel modeling has an important role to play, over and beyond its more traditional use to account for the intrinsic sites–teeth–subject hierarchy, and may fruitfully be used to explore contextual effects

that may explain the variation in the periodontal outcomes.

Acknowledgements – Apart from the support of the authors' institution, this study was partially supported by a grant from the Danish Medical Research Council (Copenhagen, Denmark), which played no role in the study design, collection, analysis, and interpretation of data; or in the writing of the manuscript; or in the decision to submit the manuscript for publication.

References

1. STERNE JAC, JOHNSON NW, WILTON JMA, JOYSTON-BECHAL S, SMALES FC. Variance components analysis of data from periodontal research. *J Periodontol Res* 1988; **23**: 148–153.
2. STERNE JAC, CURTIS MA, GILLET IR, GRIFFITHS GS, MAIDEN MFJ, WILTON JMA, JOHNSON NW. Statistical models for data from periodontal research. *J Clin Periodontol* 1990; **17**: 129–137.
3. GILTHORPE MS, MADDICK IH, PETRIE A. Introduction to multilevel modelling in dental research. *Community Dent Health* 2000; **17**: 222–226.
4. GILTHORPE MS, GRIFFITHS GS, MADDICK IH, ZAMZURI AT. An application of multilevel modelling to longitudinal periodontal research data. *Community Dent Health* 2001; **18**: 79–86.
5. GILTHORPE MS, GRIFFITHS GS, MADDICK IH, ZAMZURI AT. The application of multilevel modelling to periodontal research data. *Community Dent Health* 2000; **17**: 227–235.
6. GILTHORPE MS, ZAMZURI AT, GRIFFITHS GS, MADDICK IH, EATON KA, JOHNSON NW. Unification of the "burst" and "linear" theories of periodontal disease progression: a multilevel manifestation of the same phenomenon. *J Dent Res* 2003; **82**: 200–205.
7. MÜLLER H-P. Dealing with hierarchical data in periodontal research. *Clin Oral Invest* 2009; [Epub ahead of print].
8. MATULIENE G, PIETURSSON BE, SALVI GE, SCHMIDLIN K, BRÄGGER U, ZWALEN M, LANG NP. Influence of residual pockets on progression of periodontitis and tooth loss: results after 11 years of maintenance. *J Clin Periodontol* 2008; **35**: 685–695.
9. TOMASI C, LEYLAND AH, WENNSTRÖM JL. Factors influencing the outcome of non-surgical periodontal treatment: a multilevel approach. *J Clin Periodontol* 2007; **34**: 682–690.
10. PHILLIPS C, PREISSER JS, WHITE RJ, BLAKEY GH, HAUG RH. Prediction of periodontal pathology around third molars using linear mixed effects modeling. *Community Dent Health* 2008; **25**: 89–97.
11. MÜLLER HP. Multilevel modeling of gingival bleeding on probing in young adult carriers of non-JP-2-like strains of *Aggregatibacter actinomycetemcomitans*. *Clin Oral Invest* 2009; **13**: 171–178.
12. DEMMER RT, BEHLE JH, WOLF DL, HANDFIELD M, KEBSCHULL M, CELENTI R, PAVLIDIS P, PAPAPANOU PN. Transcriptomes in healthy and diseased gingival tissues. *J Periodontol* 2008; **79**: 2112–2124.
13. TU Y-K, GILTHORPE MS, GRIFFITHS GS, MADDICK IH, EATON KA, JOHNSON NW. The application of multilevel modeling in the analysis of longitudinal periodontal data – part II: changes in disease levels over time. *J Periodontol* 2004; **75**: 137–145.
14. TU Y-K, GILTHORPE MS, GRIFFITHS GS, MADDICK IH, EATON KA, JOHNSON NW. The application of multilevel modeling in the analysis of longitudinal periodontal data – part I: absolute levels of disease. *J Periodontol* 2004; **75**: 127–136.
15. D'AUTO F, READY D, PARKAR M, TONETTI MS. Relative contribution of patient-, tooth-, and site-associated variability on the clinical outcomes of subgingival debridement. I. Probing depths. *J Periodontol* 2005; **76**: 398–405.
16. HOLTFRETER B, SCHWAHN C, BIFFAR R, KOCHER T. Epidemiology of periodontal diseases in the study of health in Pomerania. *J Clin Periodontol* 2009; **36**: 114–123.

17. ALBANDAR JM, BRUNELLE JA, KINGMAN A. Destructive periodontal disease in adults 30 years of age and older in the United States, 1988–1994. *J Periodontol* 1999; **70**: 13–29.
18. HUGOSON A, SJÖDIN B, NÖRDERYD O. Trends over 30 years, 1973–2003, in the prevalence and severity of periodontal disease. *J Clin Periodontol* 2008; **35**: 405–414.
19. ALBANDAR JM, MURANGA MB, RAMS TE. Prevalence of aggressive periodontitis in school attendees in Uganda. *J Clin Periodontol* 2002; **29**: 823–831.
20. WANG Q-T, WU Z-F, WU Y-F, SHU R, PAN Y-P, XIA J-L. Epidemiology and preventive direction of periodontology in China. *J Clin Periodontol* 2007; **34**: 946–951.
21. LOPEZ R, FERNANDEZ O, JARA G, BAEUM V. Epidemiology of clinical attachment loss in adolescents. *J Periodontol* 2001; **72**: 1666–1674.
22. BAEUM V, FEJERSKOV O, MANJI F. Periodontal diseases in adult Kenyans. *J Clin Periodontol* 1988; **15**: 445–452.
23. SUSIN C, ALBANDAR JM. Aggressive periodontitis in an urban population in Southern Brazil. *J Periodontol* 2005; **76**: 468–475.
24. DIEZ-ROUX AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health* 1998; **88**: 216–222.
25. KRIEGER N. Historical roots of social epidemiology: socioeconomic gradients in health and contextual analysis. *Int J Epidemiol* 2001; **30**: 899–903.
26. MERLO J, CHAIX B, YANG M, LYNCH J, RÅSTAM L. A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomena. *J Epidemiol Community Health* 2005; **59**: 443–449.
27. LI F, FISHER J, BROWNSON RC. A multilevel analysis of change in neighborhood walking activity in older adults. *J Aging Phys Act* 2005; **13**: 145–159.
28. TURRELL G, SANDERS AE, SLADE GD, SPENCER AJ, MARCENES W. The independent contribution of neighborhood disadvantage and individual-level socioeconomic position to self-reported oral health: a multilevel analysis. *Community Dent Oral Epidemiol* 2007; **35**: 195–206.
29. BORRELL LN, TAYLOR GW, BORGNACKE WS, WOOLFOLK MW, NYQUIST LV. Perception of general and oral health in White and African American adults: assessing the effect of neighborhood socioeconomic conditions. *Community Dent Oral Epidemiol* 2004; **32**: 363–373.
30. JAMIESON LM, THOMSON WM. Adult oral health inequalities described using area-based and household-based socioeconomic status measures. *J Public Health Dent* 2006; **66**: 104–109.
31. BOWER E, GULLIFORD M, STEELE J, NEWTON T. Area deprivation and oral health in Scottish adults: a multilevel study. *Community Dent Oral Epidemiol* 2007; **35**: 118–129.
32. SANDERS AE, TURRELL G, SLADE GD. Affluent neighborhoods reduce excess risk of tooth loss among the poor. *J Dent Res* 2008; **87**: 969–973.
33. LOPEZ R, FRYDENBERG M, BAEUM V. Non-participation and adjustment for bias in case–control studies of periodontitis. *Eur J Oral Sci* 2008; **116**: 405–411.
34. LOPEZ R, FERNANDEZ O, BAEUM V. Social gradients in periodontal diseases among adolescents. *Community Dent Oral Epidemiol* 2006; **34**: 184–196.
35. LOPEZ R, FERNANDEZ O, JARA G, BAEUM V. Epidemiology of necrotizing ulcerative gingival lesions in adolescents. *J Periodontol* 2002; **37**: 439–444.
36. LOPEZ R, RETAMALES C, CONTRERAS C, MONTES JL, MARIN A, VAETH M, BAEUM V. Reliability of clinical attachment level recordings: effects on prevalence, extent, and severity estimates. *J Periodontol* 2003; **74**: 512–520.
37. STATA CORP. *Stata statistical software: Release 10*. College Station, TX: StataCorp, 2007.
38. RABE-HESKETH S, SKRONDAL A. *Multilevel and longitudinal modeling using STATA*, 2nd edn. College Station, TX: Stata Press, 2008.
39. LARSEN K, PETERSEN JH, BUDTZ-JØRGENSEN E, ENDAHL L. Interpreting parameters in the logistic regression model with random effects. *Biometrics* 2000; **56**: 909–914.
40. LARSEN K, MERLO J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol* 2005; **161**: 81–88.
41. *Resultados de la encuesta de caracterización socioeconómica nacional (CASEN 2000)*. Santiago, Chile: Ministerio de Planificación y Cooperación División Social, 2002; Available at: <http://www.mideplan.cl>.
42. WATT RG. From victim blaming to upstream action: tackling the social determinants of oral health inequalities. *Community Dent Oral Epidemiol* 2007; **35**: 1–11.