Plausibility of Periodontal Disease Estimates from NHANES III

Gary D. Slade, BDSc, DDPH, PhD; James D. Beck, PhD

Abstract

Objectives: *This study investigated possible reasons for observed discrepancies in prevalence estimates and measures of association for periodontal disease between phases (1988–91 and 1991–94) of the third National Health and Nutrition Examination Survey (NHANES III).* ***Methods:*** *NHANES III data on CD-ROM were obtained from the National Center for Health Statistics. Accompanying documentation states that each phase and combined phases constitute national probability samples of the US population. Weighted estimates of prevalence (percent of persons affected) and extent (percent of sites affected) for previously reported thresholds of gingival bleeding (GB), attachment loss (AL), and probing pocket depth (PD) were generated using data from all 15,511 persons aged 13–90 years who received periodontal examinations. Odds ratios for associations between four selected risk indicators and both PD and AL were compared between phases.* ***Results:*** *Phase 2 estimates of GB and PD were as much as 56 percent lower than phase 1 estimates and both were different from combined-phase estimates. However, AL prevalence was consistent between phases. Prevalence differences between phases could be explained in part by examiner variations. Odds ratios for PD differed between phases by as much as one-third, although the direction and precision of associations were not affected, and differences were reduced after controlling for examiner.* ***Conclusions:*** *Combined-phase estimates of GB and PD prevalence and extent differ from previously published estimates derived from Phase 1, apparently because estimates in at least one phase of the NHANES III study are biased. However, associations with selected risk indicators were fairly consistent between phases. [J Public Health Dent 1999;59(2):67-72]*

Key Words: periodontal diseases, epidemiology, statistical methods, NHANES III.

The need for valid epidemiologic data about periodontal disease within the US population is highlighted by the existence of two Healthy People 2000 objectives concerning gingival bleeding and periodontal attachment loss (1) and by new studies linking periodontal disease to systemic health (2,3). The most recent data describing periodontal disease prevalence within the US population come from the third National Health and Nutrition Examination Survey (NHANES III). Public-release data files for phase 1 and 2 of NHANES III are now available on CD-ROM (4). However, in our initial analysis of the data, estimates of periodontal disease for the combined phases differed conspicuously from previously published estimates derived from phase 1 of the same study (5). This paper explores reasons for those differences in more detail by evaluating the plausibility of prevalence estimates and measures of association for periodontal disease within each phase of the NHANES III survey.

Methods

This investigation was a secondary analysis of public-release data from NHANES III, a cross-sectional survey of the health and nutrition status of the US, civilian, noninstitutionalized population aged 2 months and older. The survey methods have been described elsewhere (6). To summarize, a complex, multistage, stratified clustered sample of 39,695 persons was selected; 33,994 of them completed an in-home interview, and 30,818 had a standardized clinical examination conducted at a mobile examination center. The examination included periodontal assessments conducted by six dentists for 15,511 persons aged 13+ years. Periodontal measurements were made at two sites on all teeth for two randomly selected quadrants of the mouth using the NIDR protocol (7). Field work for the survey was conducted from October 1988 to September 1991 (phase 1), and from September 1991 to October 1994 (phase 2).

Survey documentation contained on the public-release CD-ROM states that each phase comprised a national probability sample, and unit record weights were provided that adjust both for differential probability of subject selection and subject nonresponse (4). The documentation further describes how weights and sample design variables included in the data set should be used to generate those estimates. However, the same documentation cautions users about conducting statistical tests of differences between phases. Hence, the focus of this analysis is on epidemiologic and public health interpretations about point estimates of disease.

For this investigation, periodontal disease indices were calculated using thresholds that have been reported previously for the first phase (5). Prevalence measures represent the percentage of persons who have one or more sites with gingival bleeding (GB prevalence), probing pocket depth of 4+ mm (PD prevalence), and attachment loss of 3+ mm (AL prevalence). Corresponding measures of periodontal disease extent represent the mean percentage of sites within the mouth having GB, PD 4+ mm and AL 3+ mm (8). Relative differences between phases were calculated using the difference in estimates between

Send correspondence to Dr. Slade, Department of Dental Ecology, University of North Carolina, CB #7450, Chapel Hill, NC 27599-7450. Web site: http://www.dent.unc.edu/depts/. E-mail: gary_slade@dentistry.unc.edu. Reprints will not be available. Dr. Beck is also with the Department of Dental Ecology, University of North Carolina, Chapel Hill. Based on an oral presentation made at the EpiForum meeting, March 3, 1998, Minneapolis, MN. Manuscript received: 3/12/98; returned to authors for revision: 5/6/98; accepted for publication: 10/29/98.

phases as the numerator, and the larger of the two estimates as the denominator. For persons aged 18+ years, crude odds ratios for each phase and combined phases were generated for four widely reported risk indicators: sex, race, subject-reported current smoking status, and subject-reported history of diabetes (9). To assess levels of confounding, Mantel-Haenszel adjusted odds ratios were calculated after stratification by phase and compared with the crude odds ratios for the combined phases.

For all analyses, weighted data were analyzed using SUDAAN, which calculates prevalence, extent, odds ratios, and corresponding standard errors that are adjusted for the sampling design (10). However, SUDAAN does not calculate adjusted odds ratios from stratified analyses, so SAS was used to generated adjusted odds ratios, and the standard errors were recalculated using design effects for crude odds ratios obtained from SUDAAN.

Results

Demographic and Oral Health Status. Descriptive findings confirmed that subjects in each phase of NHANES III were representative of the US population with respect to age, sex, and race. The prevalence of edentulism was 9.1 percent in phase 1 and 8.3 percent in phase 2, while mean DMFT was 13.0 and 12.4, respectively. Among persons who received a periodontal examination, a mean of 23.8 and 24.2 sites were measured for AL in phase 1 and 2, respectively.

Prevalence and Extent of Periodontal Conditions. Gingival bleeding and pocket depth measures were consistently lower in phase 2 compared with phase 1, with relative differences ranging from 58 percent for PD extent to 25 percent for GB prevalence (Table 1). Corresponding absolute differences in point estimates were substantially greater than the standard errors of those point estimates. For example, PD extent was 1.9 percent lower in phase 2 compared with phase 1, whereas standard errors for those phase-specific estimates were 0.3 or less. PD prevalence was 14.1 percent lower in phase 2 compared with phase 1, whereas corresponding standard errors were 1.9 or less. The 18 percent relative difference in extent of AL was smaller than the relative differences observed for PD and gingival bleeding, although the absolute difference of 1.8 percent in AL extent was much larger than the corresponding standard errors of 0.5. In contrast, phase-specific point estimates of prevalence and extent of recession (REC) and prevalence of AL were virtually identical.

Further analysis of PD prevalence was undertaken to assess any effects of differences between examiners. During phase 1, one dentist (examiner code 4) conducted 34 percent of examinations and recorded a high prevalence of PD: 41.4 percent compared with the overall phase 1 prevalence of 28.2 percent. During phase 2, examiner 4 examined less than 1 percent of subjects. Instead examiner 5, who examined only 4 percent of subjects in phase 1, examined 45 percent of subjects in phase 2 and recorded a PD prevalence of 13.0 percent during phase 2. Nonetheless, reductions persisted even within examiners, most notably for examiner 1, who did the majority of examinations (51 percent in phase 1 and 54 percent in phase 2) and who recorded a 33 percent reduction in prevalence of PD, from 22.1 percent in phase 1 to 14.8 percent in phase 2.

Measures of Association. Odds ratios for PD differed between phases by 12 to 33 percent although the direction was not consistent: for sex and race, odds ratios were larger in phase 2 compared with phase 1, while for smoking and diabetes, odds ratios were smaller in phase 2 (Table 2). These differences in phase-specific odds ratios represent relatively small amounts of effect modification due to phase; furthermore, none of the phase-specific odds ratios altered interpretation about direction or markedly affected the precision (that is, 95% confidence limits) of the associations. Furthermore, there were only trivial differences between the crude phase 1 and 2 estimates and the adjusted phase 1 and 2 estimates, suggesting relatively little confounding of the smoking-PD relationships by phase. Compared with the findings for PD, less variation was found between phases in associations with AL, with differences between phase-specific odds ratios ranging from 6 to 18 percent (Table 2). As noted for PD, crude odds ratios for AL were very similar to adjusted odds ratios for AL, suggesting relatively little confounding by phase.

Additional stratified analysis of the race-PD association, which had the largest amount of effect modification in Table 2, resulted in less variation between phases after controlling for examiner. For example, in phase 1, race-PD odds ratios were 2.01 (95%

TABLE 1
Prevalence and Extent of Selected Periodontal Conditions in the US Population Aged 13+ Years for Phase 1 and Phase 2 of NHANES III

Index	% (SE) Estimate for Phase 1 (n=7,472)	Phase 2 (n=8,039)	Phases 1 & 2 (n=15,511)
Gingival bleeding prevalence (% of persons)	62.0 (2.7)	46.6 (1.5)	54.1 (2.1)
Gingival bleeding extent (% of sites)	11.7 (0.8)	7.6 (0.4)	9.6 (0.6)
Pockets 4+ mm prevalence (% of persons)	28.2 (1.9)	14.1 (0.9)	21.0 (1.5)
Pockets 4+ mm extent (% of sites)	3.3 (0.3)	1.4 (0.1)	2.4 (0.2)
Recession 3+ mm prevalence (% of persons)	14.6 (0.8)	16.1 (1.1)	15.4 (0.7)
Recession 3+ mm extent (% of sites)	2.8 (0.2)	3.0 (0.3)	2.9 (0.2)
Loss of attachment 3+ mm prevalence (% of persons)	39.0 (1.2)	37.3 (1.4)	38.1 (0.9)
Loss of attachment 3+ mm extent (% of sites)	10.2 (0.5)	8.4 (0.5)	9.2 (0.3)

TABLE 2
Associations Between Selected Risk Indicators and Periodontal Outcomes in US Population Aged 18+ Years for Phase 1 and Phase 2 of NHANES III

Outcome	Risk Indicator	Crude Odds Ratio (95% CI) Phase 1 (*n*=6,643)	Crude Odds Ratio (95% CI) Phase 2 (*n*=7,022)	Crude Odds Ratio (95% CI) Phases 1 and 2 (*n*=13,665)	Adjusted Odds* Ratio (95% CI) Phases 1 and 2 (*n*=13,665)
PD 4+ mm	Sex (0=female, 1=male)	1.46 (1.27, 1.68)	1.94 (1.61, 2.34)	1.60 (1.41, 1.83)	1.62 (1.42, 1.86)
	Race/ethnicity (0=non-Hispanic white, 1=other)	1.72 (1.29, 2.29)	2.59 (2.04, 3.28)	1.92 (1.53, 2.40)	2.05 (1.63, 2.58)
	Current smoker (0=no, 1=yes)	2.00 (1.69, 2.37)	1.77 (1.34, 2.33)	1.96 (1.70, 2.27)	1.91 (1.65, 2.21)
	Diabetes history (0=no, 1=yes)	1.79 (1.25, 2.55)	1.31 (0.92, 1.85)	1.51 (1.19, 1.93)	1.57 (1.22, 2.02)
AL 3+ mm	Sex (0=female, 1=male)	1.26 (1.10, 1.45)	1.51 (1.31, 1.73)	1.38 (1.23, 1.54)	1.38 (1.23, 1.55)
	Race/ethnicity (0=non-Hispanic white, 1=other)	1.08 (0.92, 1.28)	1.15 (0.99, 1.34)	1.11 (0.99, 1.25)	1.12 (0.99, 1.26)
	Current smoker (0=no, 1=yes)	1.64 (1.43, 1.89)	1.50 (1.24, 1.82)	1.58 (1.40, 1.78)	1.57 (1.40, 1.77)
	Diabetes history (0=no, 1=yes)	2.52 (1.75, 3.63)	3.08 (2.03, 4.67)	2.80 (2.20, 3.56)	2.80 (2.20, 3.56)

*Adjusted odds ratios after stratification by phase.

CI=1.53, 2.66) for examiner 1 and 2.41 (95% CI=1.65, 3.52) for examiner 4, resulting in an examiner-stratified odds ratio of 2.15 (95% CI=1.71, 2.70). In phase 2, race-PD odds ratios were 2.92 (95% CI=2.15, 3.96) for examiner 1 and 2.40 (95% CI=1.67, 3.44) for examiner 5, resulting in an examiner-stratified odds ratio of 2.69 (95% CI=2.13, 3.40). This 20 percent difference between phases in examiner-stratified odds ratios was smaller than the 33 percent difference between crude odds ratios observed in Table 2.

Discussion

This analysis has demonstrated large differences in population estimates of periodontal disease between phase 1 (1988–91) and phase 2 (1991–94) of NHANES III. Specifically, we found a halving of prevalence and extent of PD, and a reduction of about one-third in prevalence and extent of GB. Furthermore, the 95 percent confidence interval for the combined-phase prevalence estimate of PD prevalence was 18.1 to 23.5 percent, which excludes the previously reported phase 1 estimate of 29.2 percent (5). In contrast, there were relatively subtle differences between phases in measures of association between these periodontal indices and four selected risk indicators, none of which affected "bottom-line" interpretations about the direction or statistical significance of the associations. Furthermore, the phase-differences in odds ratios were reduced after controlling for examiner effects.

At first appearance, it may seem anomalous to find no phase-difference in prevalence of AL, given the large difference in prevalence of PD and equivalence in prevalence of REC. However, this finding can be attributed to site specificity of recession and pocket depth, such that the worst-mouth score for PD does not necessarily occur at the same site as the worst-mouth score for REC. Hence, an individual's worst-mouth score for AL may not represent the sum of their worst-mouth REC and PD. In contrast, the observed 18 percent difference in AL extent between phases is intermediate between the difference of 56 percent in PD extent and essentially no difference in REC extent. For these reasons, extent scores represent a more consistent index for interpreting relationships among REC, PD, and AL.

The observed differences between phases in prevalence and extent are based on eight specific measures of periodontal disease that have arbitrary thresholds to define "disease" status. These indices and thresholds were selected because they have been used widely in oral epidemiology. Although not reported above, smaller absolute differences were observed when other indices were used. For example: prevalence of 6+ sites with gingival bleeding was 16.1 percent in phase 1 and 9.3 percent in phase 2; prevalence of 6+ mm pocket depth was 3.6 percent and 2.0 percent, respectively; and extent of 6+ mm pockets was 0.36 percent and 0.18 percent, respectively. While these represent smaller absolute differences, relative differences between phases remain quite consistent regardless of the threshold and index—namely, an approximate halving of PD and a reduction of about one-third in GB.

Although the observed differences in population estimates for PD and GB appear large, it is appropriate to evaluate whether they are due to chance and whether they are large enough to be of practical relevance. While it is standard practice to rely on *P*-values when judging if differences are due to chance, the NHANES III documentation cautions against conducting hypothesis tests about differences be-

tween phases. Nonetheless, the current results have demonstrated that absolute differences between point estimates of PD and gingival bleeding between phases are substantially greater than standard errors of those point estimates (Table 1). Stated another way, the two-tailed lower 95 percent confidence interval for GB prevalence in phase 1 (62.0–1.96 x 2.7=56.7%) was 10 percentage points higher than the corresponding point estimate for phase 2, while the two-tailed lower 95 percent confidence interval for phase 1 PD prevalence (24.5%) was 10 percentage points higher than the corresponding point estimate for phase 2.

Of course, it remains a matter of interpretation as to whether these "substantial" differences are of practical relevance. In this context, we regard a relevant difference as one that could affect interpretations about the epidemiology of disease, or which may affect public health practice. Precedents from other areas of oral epidemiology certainly suggest that differences of the magnitude observed in this study can be regarded as relevant. For example, based on an observed 32 percent reduction in children's mean DMF values between the 1971–74 NHANES I survey and the 1979–80 US Schoolchildren Survey, Burt concluded, "It was clear that the decline was real" (11). Oral health targets for US adults aged 35–44 years aim to reduce the prevalence of destructive periodontal disease (4+ mm of attachment loss) from a baseline of 24 percent in 1985–86 to no more than 15 percent by the year 2000 (1). And the historically observed twofold difference in mean DMF levels between fluoridated and nonfluoridated communities provided the rationale for an overwhelming endorsement of community water fluoridation as an intervention to reduce caries.

Given that the observed differences in PD and GB are large and epidemiologically substantial, it is appropriate to evaluate whether or not the differences are plausible, that is, reflecting underlying population differences between the two phases of the survey. By design, there is a temporal difference of three years between these phases; however, we reject the possibility that there could have been a halving of PD prevalence within the US population within that time period. Specifically, three years is simply too short a period for a cohort effect of this magnitude to emerge, and there is no evidence of widespread, dramatic increases in adoption of oral preventive practices or periodontal treatment services within this period. For example, surveys of services provided in private dental practice revealed very low levels of periodontal treatment (3.3% of patients seen in 1979), which increased only marginally to 4.1 percent in 1990, the most recent year available (12).

Plausible differences between phases could exist if each phase were representative of different population subgroups. However, this explanation can be refuted for three reasons. First, the NHANES III documentation states that the survey "was designed so that the survey's first three years, 1988–91, its last three years, 1991–94, and the entire six years were national probability samples" (4). Second, the sample weights correct both for unequal probability of subject selection (inherent in the sampling design) and for differential nonresponse. These corrections are borne out in the current analysis, which revealed no significant demographic differences between phases. Third, it is difficult to envisage circumstances under which phase 1 subjects could be fundamentally different in their levels of gingival bleeding and pocket depth, but essentially equivalent with respect to other oral health status measures including edentulism, DMFT, periodontal recession, and number of periodontal sites measured.

For these reasons, we conclude that there must be substantial bias in at least one of the three survey estimates (that is, phase 1, phase 2, or combined phases) of GB, PD and probably AL extent. In this context, we use the epidemiologic definition for bias of an estimator—namely, "the difference between the expected value of an estimator of a parameter and the true value of this parameter" (13). This definition necessarily assumes that (say) periodontal pocket depth conforms with the epidemiologic construct of a parameter—namely, "a measurable characteristic of a population" (13)—and consequently has a "true" population value. Philosophically, it could be argued that periodontal disease is not a population parameter because the biology of the condition is so complex as to be unmeasurable in principle, at least using existing technology. For example, evidence shows that the process of manual probing can alter the periodontal pocket by penetration of the junctional epithelium (14); hence, our arguably crude methods of measurement may be incapable of quantifying the true depth of a pocket. Furthermore, the presence of a periodontal pocket does not necessarily signify active disease. It is for these reasons that some researchers have concluded that it is "unlikely that any of the current methods [of clinical and radiographic periodontal assessment] can lead to an adequate definition of periodontal disease activity" (15). While this conclusion essentially expresses a concern about validity of pocket depth measurements, there are additional, well-documented difficulties in obtaining adequate reliability of these measures (16).

Despite these concerns about reliability and validity, good reasons exist to continue to advocate the use of pocket depth as a population parameter in epidemiologic studies, in the same tradition that manifestational criteria are used in epidemiologic studies to classify disease—albeit at a crude level, when there is incomplete evidence about its pathology. For example, the manifestational criteria used to define cases of consumption in the 18th century yielded cases of disease that would not have met the criteria for diagnosis of tuberculosis that were adopted following identification of the tubercle bacillus a century later; nevertheless, epidemiologic studies provided valuable information about secular changes in the disease (17).

In the case of periodontal disease, there are additional, pragmatic reasons to view pocket depth as a meaningful population parameter, primarily because it is currently used in the real world of clinical practice. For example, the American Academy of Periodontology advocates screening patients in general dental practice using the periodontal screening and recording (PSR) system, which includes thresholds of pocket depth assessed by manual probing (18). This position is not to deny that assessment of pocket depth is prone to measurement error, and possibly measurement bias, particularly in epidemiologic studies using manual probes. However, meas-

urement errors and biases are methodologic phenomena affecting individual studies and their existence does not nullify the concept that periodontal pocket depth is a population parameter useful for describing occurrence of periodontal disease.

Our conclusion that bias must exist in at least one of the survey estimates does not clarify which of the estimates is least biased. Unfortunately, without further investigation, this question cannot be answered convincingly. Possibly systematic differences among examiners contributed to phase-specific differences. Although subjects were not assigned at random to examiners, large interexaminer differences in prevalence may partly explain variations in prevalence and extent between phases. However, it also should be noted that all examiners had acceptable levels of interexaminer reliability when compared with the "gold standard" dentist (19) and that there was a substantial reduction in PD prevalence between phases for the one dentist who did the majority of examinations.

Sources of bias may be examined through further analyses of data about interexaminer reliability, or of survey results from specific examiners, or of specific places and times that examinations were conducted. Such investigations may permit guidelines to be generated to advise users about appropriate use of these data on PD and GB. However, not all of those data items are available in the public-release CD ROM, and in any event it seems unlikely that definitive evidence will be found to justify use of a specific subset of the data, or that a simple statistical correction could be devised to reduce bias. Hence, unless such guidelines become available, it would seem necessary to limit the descriptive analyses of the parameters studied here to prevalence of AL and prevalence and extent of recession. Given that prevalence data from the first phase have been published, there clearly is a need to be cautious in drawing any conclusions about those estimates because it is not known whether the phase 1 estimate is the least biased.

This recommendation may seem unusually harsh because it is only by virtue of the unique two-phase design of this study that any evidence of bias was discerned. Potentially, other epidemiologic studies may show large differences in prevalence estimates between (say) subjects seen in the first half of the study compared with those seen during the second half. However, most other surveys select a single, representative sample and then fieldwork progresses (say) from county to county, or from school to school. Under those circumstances, observed differences between initial and later examinations could be due entirely to underlying differences between subgroups of the sampled population. In contrast, the NHANES III study had the unusual feature that each phase was designed (both in terms of sampling and unit record weights) to be independently representative of the US population in each phase. Of course, if the only source of measurement error in this or any other survey was random error (for example, due to large levels of intra- or interexaminer variation), then we would expect no differences in prevalence estimates between equally representative subsamples of the survey. However, differences were observed in this instance; therefore, it is necessary to conclude that bias exists.

This analysis of specific phases and combined phases was undertaken after considering guidelines for users of the NHANES III data that are, at first appearance, unambiguous: "unbiased national estimates of health and nutrition characteristics can be independently produced for each phase as well as for both phases combined" (4). However, such statements could be interpreted to contrast with other statements in the guidelines, for example, "Analysts are encouraged to use all six years of survey results" (4). Furthermore, it may seem contradictory that "although point estimates can be produced separately for each phase, no test is available to test whether those estimates are significantly different from each other" (4), since statistical tests usually are appropriate for independent samples. While this analysis has avoided presenting any formal hypothesis tests, the findings concerning PD and gingival bleeding appear to represent a situation where *P*-values are unnecessary to formulate epidemiologic inferences. However, if for any reason the use of phase-specific estimates (such as the ones presented here) is invalid, perhaps because of undocumented problems in the survey design or implementation, it would appear essential to describe promptly and unambiguously those reasons, so that users of the public data set can make informed decisions about analytic strategies. However, it should be emphasized that there is no evidence that such problems occurred; hence, it appears valid to make phase-specific estimates. Indeed, they are essential for those aspects of data collection that were limited to one phase (for example, serum homocysteine).

It should be emphasized that these findings of substantial bias were not simply a consequence of a frivolous "fishing expedition" to exploit the special two-phase design of this study. As noted in the introduction, our original motive to investigate phase-specific effects in periodontal disease was the conspicuous discrepancy between the previously reported findings (5) and our preliminary analysis of the combined-phase data. Clearly, we would be unjustified in drawing conclusions about bias if we had set out simply to compare (say) 20 oral epidemiologic indices because such an exercise could be expected to produce a statistically significant ($P<.05$) difference in at least one of those comparisons.

While we cannot generalize from this study to previous surveys, it is conceivable that biases have occurred in previous population estimates of periodontal disease, possibly even of the twofold magnitude observed here. If that were the case, it would emphasize a general need for caution when interpreting results from such studies. And whether or not previous surveys have produced biased estimates, there is a particular need to be circumspect when attempting to interpret trends in periodontal disease by comparing results from the current survey with estimates from previous studies.

Compared with the differences observed in prevalence, there was less variation between phases in associations between selected risk indicators and periodontal disease, and the odds ratios were not consistently lower in one phase (Table 2). Differences in odds ratios for PD were greater than 20 percent for three risk indicators (sex, race, and diabetes history), enough to constitute some minimal amount of effect modification due to phase. However, a difference of 18 percent also was observed between phase-specific odds ratios for the diabetes—AL association, indicating that

some level of effect modification can occur, even for an outcome such as AL that did not differ in prevalence between phases. Furthermore, additional analysis of the race-PD association (selected because it had the greatest degree of effect modification due to phase) indicated that the variation could be reduced through adjustment for examiner effects within each phase. For the associations that did not differ substantially between phases, there was little evidence of confounding. Hence, it seems reasonable to use these NHANES III data for examination of associations, even with indices of PD and GB for which prevalence estimates are substantially biased.

At first appearance, this conclusion may seem unusually lenient, or at least irreconcilable with the earlier recommendation to avoid descriptive analyses of the PD and GB data. In principle, however, the greatest threat to validity of associations arises when there is differential misclassification of exposure and disease (20). However, for many exposures of relevance to periodontal disease including the ones studied here, it seems unlikely that the factors contributing to biases in measurement of disease (presumably examiner differences, among other things) would necessarily be greater or lesser among subjects who were exposed versus nonexposed. In other words, the misclassification is likely to be nondifferential, which has the known effect of biasing the bivariate estimate of effect toward the null. Furthermore, at least one of the probable contributors to disease misclassification in this study (namely, examiner) can be used in the analysis to evaluate the extent of effect modification, and, as occurred for the race-PD association evaluated here, adjustment may reduce the amount of bias in odds ratios. In any event, it would appear prudent for investigators to evaluate carefully phase-specific associations, seeking evidence of confounding or effect modification and attempting to correct for such effects when possible. Such an evaluation would at least help to clarify the amount of variation in odds ratios to provide some general indication of the amount of imprecision inherent in these data.

References

1. Department of Health and Human Services. Healthy people 2000. National health promotion and disease prevention objectives. Boston, MA: Jones and Bartlett, 1992.
2. Beck JD, Garcia RG, Heiss G, Vokonas P, Offenbacher S. Periodontal disease and cardiovascular disease. J Periodontol 1996;67(Suppl):1123-37.
3. Offenbacher S, Katz V, Fertik G, et al. Periodontal disease as a possible risk factor for preterm low birth weight. J Periodontol 1996;67:1103-13.
4. National Center for Health Statistics. Third National Health and Nutrition Examination Survey, 1988-1994, NHANES III Examination data file (CD-ROM). Public use data file doc no 76200. Hyattsville, MD: National Center for Health Statistics and Centers for Disease Control and Prevention, 1996. [Available from National Technical Information Service (NTIS), Springfield, VA.]
5. Brown LJ, Brunelle JA, Kingman A. Periodontal status in the United States, 1998-91: prevalence, extent, and demographic variation. J Dent Res 1996;75(Spec Iss): 672-83.
6. National Center for Health Statistics. Plan and operation of the Third National Health and Nutrition Examination Survey, 1988-94 (Vital Health Stat 1994; series 1, no 32).
7. Miller AJ, Brunelle JA, Carlos JP, Brown LJ, Löe H. Oral health of United States adults. The national survey of oral health in US employed adults and seniors, 1985-86. Bethesda, MD: National Institute of Dental Research, 1987. NIH pub no 87-2868.
8. Carlos JP, Wolfe MD, Kingman A. The extent and severity index: a simple method for use in epidemiologic studies of periodontal disease. J Clin Periodontol 1986;13:500-5.
9. Genco RJ. Current view of risk factors for periodontal diseases. J Periodontol 1996; 67(Suppl):1041-9.
10. Shah BV, Barnwell BG, Bieler. SUDAAN user's manual. Release 7.0. Research Triangle Park, NC: Research Triangle Institute, 1996.
11. Burt BA. Dental Caries. In: Burt BA, Eklund SA, eds. Dentistry, dental practice and the community, 4th ed. Philadelphia, PA: WB Saunders, 1992:90-112.
12. American Dental Association. The 1990 survey of dental services rendered. Chicago: ADA, 1994.
13. Last JM. A dictionary of epidemiology. New York: Oxford University Press, 1988.
14. Listgarten MA. Periodontal probing. What does it mean? J Clin Periodontol 1980;7:165-76.
15. Claffey N. Gold standard—clinical and radiographical assessment of disease activity. In: Lang NP, Karring T, eds. Proceedings of the 1st European Workshop on Periodontology. Chicago: Quintessence, 1994:42-53.
16. Goodson JM. Selection of suitable indicators of periodontitis. In: Bader JD, ed. Risk assessment in dentistry. Chapel Hill, NC: University of North Carolina Dental Ecology, 1990:69-74.
17. McMahon B, Pugh TF. Epidemiology principles and methods. Boston, MA: Little, Brown and Co., 1970:48-55.
18. Nasi JH. Background to, and implementation of, the Periodontal Screening and Recording (PSR) procedure in the USA. Int Dent J 1994;44:585-8.
19. Kingman A, Brunelle JA, Winn DM, Drury TF, Kleinman DV. Evaluation of NHANES III examiners' ability to assess oral health conditions (Abstract). J Dent Res 1998;77:115.
20. Rothman KJ. Modern epidemiology. Boston, MA: Little, Brown and Co., 1996: 84-9.