Validity of Two Methods for Assessing Oral Health Status of Populations

Eugenio D. Beltrán, DMD, MPH, MS, DrPH; Dolores M. Malvitz, RDH, MPH, DrPH; Stephen A. Eklund, DDS, MHSA, DrPH

Abstract

Objective: *This investigation assessed two methods for estimating epidemiologic indicators of oral health status among children: (1) a visual-only screening, performed independently by a dental hygienist and a registered nurse; and (2) a parent- or guardian-completed questionnaire. The indicators included dichotomous variables measuring dental caries and treatment needs, presence of sealants, injuries to the anterior teeth, and dental fluorosis. Methods: Following training and calibration, data were collected over an eight-day period in April 1994 among 632 elementary schoolchildren (aged 5 to 12 years) in Monticello, Georgia. Both screening and questionnaire findings were compared pairwise with results from visual-tactile examinations done by a dentist. Validity, represented by sensitivity, specificity, and predictive values, was assessed for screening results from the dental hygienist, the nurse, and the parent-completed questionnaire. Results: Validity was high for screening for caries and treatment needs (>90% for sensitivity, specificity, and predictive values in a sample having 30% to 40% prevalence). Less valid data — mainly an effect of false negatives — were obtained for fluorosis, injuries, and presence of sealants. No significant difference in validity was observed between the nurse and the dental hygienist. One-third of respondents to the questionnaire did not know if their children needed fillings (a proxy for untreated decay) or had received sealants; only knowledge of restorations was comparable to results from screening. Intraexaminer reliability for the two screeners ranged from 85 to 100 for percent agreement and 0.70 to 0.93 for kappa scores. Conclusions: Screening by dental hygienists or nurses can provide valid data for surveillance of dental caries and treatment needs. Training for visual assessment of fluorosis and injuries must be improved to diminish the proportion of false negatives. A parent-completed questionnaire is less effective than visual screening for evaluating oral health status in children. [J Public Health Dent 1997;57(4):206-214]*

Key Words: validity, reliability, visual-tactile examination, screenings, questionnaire, oral health assessment, surveillance.

Surveillance in public health is defined as the ongoing systematic collection, analysis, and interpretation of outcome-specific data for use in the planning, implementation, and evaluation of public health practice (1). Several approaches and sources of data are used in public health surveillance: vital statistics, notifiable disease data, registries, sample surveys, administrative data, and sentinel surveillance data (2). An important element of any surveillance system is the use of the data to advance public health.

No true surveillance systems for oral conditions exist. The prevalence and trends of oral diseases have been monitored through oral health surveys at the national and sometimes state and local levels. These surveys have used some sort of clinical examination and questionnaire. A dental professional generally performs a visual-tactile assessment of the mouth in a sample of the population. Many of these surveys have used the Radike diagnostic criteria and examination protocol (3), modifications of the Radike criteria such as those by the National Institute of Dental Research (4), or those of the World Health Organization (5).

Some differences exist between the diagnostic criteria used in these surveys and those used by clinicians in their practices. In clinical examinations practitioners make complex measurements, stressing the precision of the process, i.e., the detection of the most incipient signs of disease. The greater the complexity, however, the greater the likelihood of unreliable results. Research shows that clinicians apply different criteria (6,7) and differ significantly in their diagnoses and treatment plans (8-11). Oral health surveys have emphasized the reliability of measurements. Experience suggests that visual-tactile examinations can be used to estimate the oral health status of populations, provided the assessment is performed in well-defined samples by trained examiners who use valid indices with known reliability. In oral epidemiology these surveys are considered the standard, even though some underdiagnosis is expected due to the lack of procedures and techniques that are feasible or practical for field examinations (e.g., x-rays, optical transillumination). Still, oral health surveys are complex and demand extensive resources, which might explain why many programs lack data on oral health status (12).

Screenings are defined in the medical model as procedures that can sort out persons who may have a condition from those who may not (13). Those who appear to have the condition are

followed up to obtain a final diagnosis (13). Because screenings are applied to populations, their methods must be simple, inexpensive, and require minimal training for application and interpretation, e.g., the PPD tuberculin test. Screenings must be evaluated against some standard procedure for validity—that is, their sensitivity, specificity, and predictive value. In this evaluation, the proportion of false positives and false negatives are important to patient follow-up and cost (14).

Frankenburg (15) has provided the following 10 criteria for selecting diseases or conditions for screening programs: (1) the condition has to be serious or potentially so, (2) it should be possible to differentiate those with the condition from those at borderline or without it, (3) the prognosis should be improved if the condition is detected and treated early, (4) there should be adequate lead and screening time, (5) the condition should be treatable or controllable, (6) the condition should be prevalent, (7) there should be no harm caused to the individual being screened, (8) there should be a follow-up for those found positive, (9) the procedure should be cost effective, and (10) the program should be acceptable to the public.

Screening protocols for oral cancer or precancerous lesions reported in the literature (16-19) involve the systematic and consistent visual-tactile examination of all anatomical sites of the oral and pharyngeal cavity and neck (20). Few oral conditions, other than oral cancer, have been the subject of screening programs and activities (21-23). For dental caries, screening denotes a fast and simple process by which an examiner visually inspects the oral cavity to detect the presence or absence of specific oral conditions (24), as has been done in "health fairs" (25). Based on Frankenburg's criteria, dental caries, dental fluorosis, and dental injuries screening protocols could be developed as alternatives to visual-tactile examinations.

Few studies have tested the visual and visual-tactile abilities of dental and nondental personnel other than dentists. Only in a few large-scale surveys have dental hygienists been examiners (26-30). These studies reported interexaminer reliability to compare the examinations made by hygienists and dentists. None of these studies evaluated validity. Two stud-

ies from the same research team (29,30) reported measures of validity; however, these measures were on the ability of their research instrument (a multivariable statistical model) to predict future caries patterns. Based mostly on reliability values, these studies concluded that dental hygienists could be used to collect oral health status data. A few studies (25,31) have focused on the use of other personnel—such as dental auxiliaries, dental students, or teachers—for examinations or screenings; however, none of these studies have assessed validity.

Oral health status and behaviors also have been assessed by questionnaires, for example those of the National Health Interview Survey (NHIS) (32) and the Behavioral Risk Factor Surveillance System (BRFSS) (33). These data are self-reported and subject to error, particularly selection and information biases (34,35). Several studies have compared data obtained from clinical examinations with those self-administered questionnaires (36-41), phone interviews (42), and parent-responded questionnaires (43). Despite the different methodologies, populations, oral conditions assessed, and the fact that many studies evaluated reliability instead of validity, researchers concluded that questionnaires appear to be useful tools for obtaining oral health data.

The objective of this investigation was to test the validity of visual screening (protocol 1) and parent- or guardian-completed questionnaire (protocol 2) for assessing the oral health status of schoolchildren, with the goal of using these protocols as tools in a surveillance system for oral conditions. The visual-tactile examination was used as the standard protocol.

Methods

Children in kindergarten through fifth grade (N=838; aged 5 to 12 years) in a rural school in Monticello, Georgia, were invited to participate in this investigation. This school had asked the Georgia Department of Human Resources to evaluate the children's dental status. One week before examination, children were given a package that included a letter of invitation, a consent form, and a multiple-choice questionnaire for their parents. All children who returned questionnaires participated in the examination

(n=632; 75% response rate). Following examination, a form with recommendations regarding the child's need for dental treatment was sent to the parents.

All clinical procedures followed guidelines for infection control in dental settings (44,45). A protocol for injury exposure control consistent with CDC guidelines and the Occupational Safety and Health Administration standards was followed. Institutional review board approval was obtained from both CDC and the University of Michigan.

Measurement Protocols. The standard examination protocol was a visual-tactile examination of all tooth surfaces in each child. All children were examined by a dentist who had been calibrated using NIDR diagnostic criteria for dental caries and presence of sealants (4), Dean's index for dental fluorosis (4), and Bhat's criteria for dental injuries (46). A portable dental chair and light, dental explorers, and plane-surface mirrors were used; no radiographs were exposed. From these examinations, each child was classified into one of the following groups based on dental caries in their primary and permanent teeth: (1) caries free; (2) at least one tooth restored, but no untreated decay; or (3) at least one untreated carious lesion, whether restorations were present or not. In addition, each child was classified as having (yes/no): (1) at least one permanent molar with pit and fissure sealants, (2) at least one anterior tooth with injury or sequelae, and (3) dental fluorosis. Finally, each child was classified into a three-level variable according to urgency of treatment needed (no need, nonurgent, and urgent). Need for urgent treatment was defined as having one or more teeth with extensive tissue destruction, probably involving the pulp, or history of pain or signs of infection. Nonurgent needs included presence of untreated carious lesions or calculus.

Two measurement protocols were tested. Protocol 1 was a visual-only screening of the child's mouth done by a dental hygienist or a registered nurse, each of whom screened about one-half of each class in sequential order. The dental hygienist had not performed oral assessments or provided treatment for more than five years. The nurse had no previous clinical

dental experience. Both screeners received written material on the procedures and diagnostic criteria for each condition to be evaluated.

The protocol allowed the dental hygienist to perform oral screenings after review of this material and without further training. A five-hour training session, which used both slides and patients, familiarized the nurse with dental nomenclature and clinical features of each condition. The protocol required the screener to inspect all four quadrants of the mouth sequentially and to code the results according to five algorithms. Figure 1 depicts the algorithm used to code dental caries (a full set of these algorithms is available upon request). In the particular case of coding for dental caries, some children with restorations were coded as "2" because they also had untreated carious lesions. Screenings were carried out using a portable chair, a flashlight for intraoral illumination, and a tongue blade to facilitate visual access during inspection. Visual screening always preceded the visual-tactile examination for each child. Approximately 5 percent of the children were reexamined by the dentist and the two screeners to estimate intraexaminer reliability. Most replicate assessments were performed 20–25 minutes after the initial dental screenings and 45–60 minutes after the initial visual-tactile examinations.

Protocol 2 was a 17-item, multiple-choice questionnaire completed by a parent or guardian when permission to participate was granted. The questionnaire asked about conditions present in the child's mouth, demographic characteristics, and family socioeconomic status. Questions about dental fluorosis and need for dental treatment were not included; "don't know" options were available. The

FIGURE 1
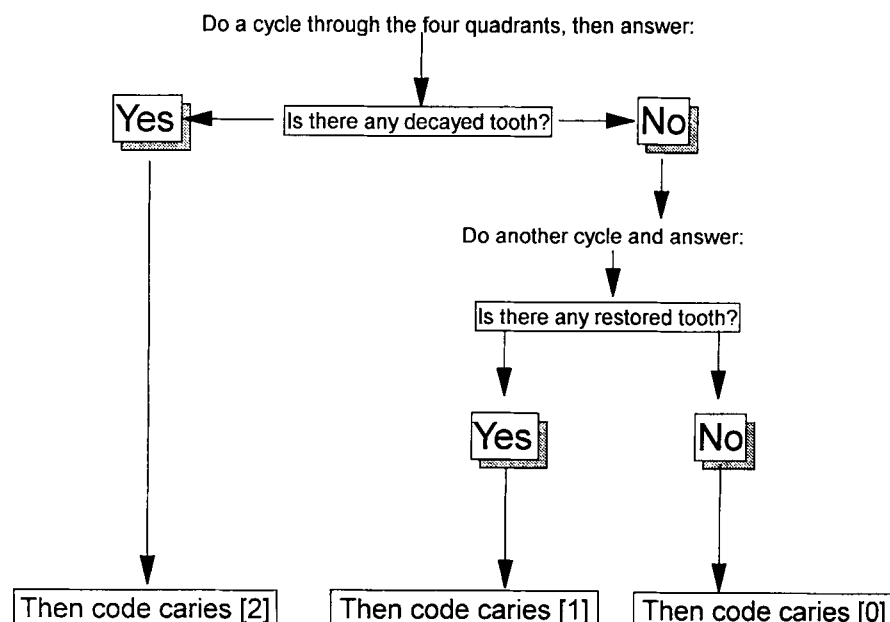### Visual Screening Algorithm for Coding Dental Caries

TABLE 1
Validity of Visual Screening Versus Visual-tactile Examination in Evaluating Presence of Selected Oral Conditions, Both Screeners Combined

Condition	N*	Sensitivity (%)	Specificity (%)	Predictive Value (%)		Prevalence (%)†
				Positive	Negative	
Caries‡						
Restorations present	421	94.0	99.6	99.4	96.2	39.7
Untreated decay	434	93.6	96.9	95.3	95.8	39.9
Caries experience§	632	94.3	96.6	97.5	92.3	58.5
Fluorosis	632	78.6	95.9	90.9	89.7	34.0
Injuries	632	80.0	98.2	85.7	97.3	11.9
Sealants	632	59.1	99.7	92.9	97.0	7.0
Treatment¶						
Nonurgent	554	73.3	95.7	81.7	93.1	20.9
Urgent	482	98.4	99.8	98.4	99.7	12.9
Any treatment needed	632	83.4	95.4	89.0	92.9	30.5

*Total sample size available (including the reference group) to assess validity for each condition.
†As determined by the visual-tactile examination.
‡Caries free was used as the reference.
§Caries experience was defined as having treated or untreated carious lesions.
¶No need for treatment was used as the reference.

TABLE 2
Validity of Visual Screening by Dental Hygienist Versus Visual-tactile Examination in Evaluating Presence of Selected Oral Conditions

| Condition | N* | Sensitivity (%) | Specificity (%) | Predictive Value (%) | | Prevalence (%)† |
				Positive	Negative	
Caries‡						
Restorations present	188	98.0	100.0	100.0	94.1	40.4
Untreated decay	215	94.8	94.1	92.9	95.7	44.7
Caries experience§	309	93.7	94.1	96.2	90.3	61.5
Fluorosis	309	88.2	95.5	88.2	95.5	27.5
Injuries	309	80.6	98.9	90.6	97.5	11.7
Sealants	309	59.1	99.7	92.9	97.0	7.1
Treatment¶						
Nonurgent	266	79.4	91.6	74.6	93.5	23.7
Urgent	221	97.1	99.5	97.1	99.5	15.4
Any treatment needed	309	86.7	91.2	83.5	93.0	34.0

*Total sample size available (including the reference group) to assess validity for each condition.
†As determined by the visual-tactile examination.
‡Caries free was used as the reference.
§Caries experience was defined as having treated or untreated carious lesions.
¶No need for treatment was used as the reference.

questionnaire was reviewed and tested by persons with expertise in questionnaire development and by a convenience sample of parents and schoolteachers.

Data Management and Statistical Methods. Data were entered directly into computer files using customized data entry programs in the Epi Info software (47). Following data cleaning and editing, epidemiologic indicators were calculated using SAS software. Later, the screening and the questionnaire data sets were matched with data from the standard protocol using all eligible records to maximize sample size. Pairwise matching of the screening data with the standard protocol produced 3x3 contingency tables for dental caries and treatment urgency and 2x2 tables for all other conditions. Pairwise matching of questionnaire data with the standard protocol produced 2x2 tables. To assess the validity of each alternative protocol in each condition, a 2x2 table was prepared; for dental caries and treatment urgency in the screening protocol, only the appropriate cells from the 3x3 table were included.

All measures of validity — sensitivity, specificity, positive predictive value, and negative predictive value— and the prevalence of the condition in the sample were calculated using

standard procedures (48). When appropriate, standard errors and confidence intervals were estimated using binomial approximation. Data were stratified to control for sociodemographic variables. Data from duplicate examinations and screenings were used to check for intraexaminer reliability tested by percent agreement and kappa statistics (49).

Results

Validity of Visual Screenings and Type of Screener. Summaries of validity results for screening are included in Tables 1–3. In each table two additional comparisons grouped those with any caries experience (treated or untreated) and those needing any treatment. Tables 1–4 display the total number of subjects (N) used in the statistical procedures to assess validity for each condition, including those in the reference group (those not having the condition).

Specificity for both screeners combined (Table 1) was greater than 95 percent across all oral conditions; the 95 percent confidence intervals varied from 94 to 100 percent (data not displayed). The sensitivity for caries was around 94 percent (95% CI=91%,96%). The lowest sensitivity was found in screening for sealants (59%; 95% CI=55%,63%), followed by nonurgent

treatment needed (73%), fluorosis (79%), and injuries (80%). A positive predictive value of 89 percent or higher was found for most oral conditions; exceptions were nonurgent treatment needed (82%) and injuries (86%). Negative predictive values were 90 percent or higher across all conditions, and many values were over 95 percent.

The dental hygienist (Table 2) screened 309 children with more prevalent untreated decay than in the 323 children screened by the nurse; the sample screened by the nurse had a higher prevalence of fluorosis (Table 3). Results obtained by the dental hygienist and nurse were comparable to those obtained by the dentist (confidence intervals overlapped); nevertheless, some minor differences were found. For example, the dental hygienist performed slightly better than the nurse in assessing fluorosis and determining the need for nonurgent and any treatment (Tables 2 and 3). These differences could be explained, in part, by the differences in prevalence in some of these conditions. The overall lower sensitivity for nonurgent treatment needs may be a consequence of the lower prevalence of these conditions in children screened by the nurse. The lack of a tactile component in the assessments made by

TABLE 3
Validity of Visual Screening by Nurse Versus Visual-tactile Examination in Evaluating Presence of Selected Oral Conditions

Condition	N*	Sensitivity (%)	Specificity (%)	Predictive Value (%)		Prevalence (%)†
				Positive	Negative	
Caries‡						
Restorations present	233	96.7	99.3	98.9	97.9	39.1
Untreated decay	219	92.2	99.3	98.6	95.9	35.2
Caries experience§	323	95.0	98.6	98.8	94.0	55.7
Fluorosis	323	72.3	96.4	93.1	83.8	40.3
Injuries	323	79.5	97.5	81.6	97.2	12.1
Sealants	323	59.1	99.7	92.9	97.1	6.8
Treatment¶						
Nonurgent	288	66.0	99.2	94.6	92.8	18.4
Urgent	261	100.0	100.0	100.0	100.0	10.7
Any treatment needed	323	79.6	99.2	97.2	92.8	27.2

*Total sample size available (including the reference group) to assess validity for each condition.
†As determined by the visual-tactile examination.
‡Caries free was used as the reference.
§Caries experience was defined as having treated or untreated carious lesions.
¶No need for treatment was used as the reference.

TABLE 4
Validity of Parent/Guardian-completed Questionnaire Versus Visual-tactile Examination in Evaluating Presence of Selected Oral Conditions

	N	Excluded n (%)	Sensitivity (%)	Specificity (%)	Predictive Value (%)		Prevalence (%)*
					Positive	Negative	
No caries vs restorations present	401	60 (10)	93.3	89.1	84.5	95.1	40.7
No caries vs untreated decay	305	214 (34)	68.8	88.3	80.4	80.3	40.9
Injuries	611	17 (3)	20.0	87.3	16.9	89.4	11.5
Sealants	423	205 (33)	56.7	89.3	28.8	93.4	7.1

*As determined by visual-tactile examination

both screeners may be the cause of the lower sensitivity for sealants.

Validity of Questionnaires. About one-third of respondents did not know or did not indicate whether their children had received sealants or needed restorative treatment. A lower proportion (10%) did not know if their children had restored teeth. To avoid classification bias, these observations were not included in the assessment of the questionnaire's validity (Table 4). The best results were obtained in evaluating the presence of restorations (93% sensitivity and 95% negative predictive value). Although these values

were similar to those obtained by screening (Table 1), specificity and positive predictive value were lower in the questionnaire (89% and 85%, respectively) than in the screening. Measures of validity were lower for the other three conditions (untreated decay, injuries, and presence of sealants), as well.

To evaluate any potential differences among respondents' ability to correctly assess the presence of restorations, responses were stratified by family income (≥$30,000/year vs other), eligibility for free or reduced-price lunch, number of children in the

family (≥3 vs other), and respondent's level of education (>high school vs other). Respondents with greater family income or having children not eligible for free or reduced-price lunch assessed presence of restorations in their children with higher sensitivity (data not shown).

Reliability of Measurements. Percent agreement and kappa estimates for each condition were calculated separately for the dentist, the dental hygienist, and the nurse (Table 5). The dentist evaluated caries status and presence of sealants on each surface by using 14 mutually exclusive diagnos-

TABLE 5
Intraexaminer Reliability for Examiner and Screeners

	Caries*		Sealants†		Treatment Urgency‡		Fluorosis§		Injuries¶	
	% Agree.	Kappa	% Agree.	Kappa	% Agree.	Kappa	% Agree.	Kappa	% Agree.	Kappa
Dentist	99	0.97			100	1.0	98	0.92	99	0.92
Hygienist	85	0.77	100	1.00	89	0.77	96	0.92	96	0.70
Nurse	96	0.93	100	1.00	100	1.0	91	0.81	96	0.83

*The dentist evaluated caries status on each surface by using 14 mutually exclusive diagnostic categories; both screeners evaluated caries in the entire mouth by using three possible diagnostic categories.
†The dentist evaluated presence of sealants as part of the diagnosis of caries.
‡Based on three diagnostic categories.
§Based on two diagnostic categories.
¶The dentist evaluated injuries by using four diagnostic categories for each of the upper front teeth; both screeners evaluated injuries in the upper anterior teeth by using two diagnostic categories.

tic categories; therefore, presence of sealants was not analyzed as an independent item for the dentist. The dentist's reliability was evaluated over 14 diagnostic categories for caries and presence of sealants, seven for presence of injuries (46), three for treatment urgency, and two for dental fluorosis. Both the dental hygienist and nurse were evaluated using three-level variables for caries and treatment urgency and two-level variables for presence of sealants, dental fluorosis, and dental injuries. All examiners showed high reliability: agreements were greater than 85%, and kappa statistics were greater than 0.70, a substantial and almost perfect score according to the scale proposed by Landis and Koch (50).

Discussion

Validity of Visual Screening and Type of Screener. Visual screenings can produce data highly comparable to those obtained from visual-tactile examinations. The highest validity was found in the screening for caries and urgent treatment (Table 1) in a sample for which the prevalence of dental caries was comparable to the entire population of Georgia (51) and the United States (4). The protocol for screening did not discriminate between dental caries of the primary and permanent dentitions because its objective was to obtain epidemiologic information on the overall caries experience of the subject. If necessary, the protocol could be modified to measure each dentition separately.

With the exception of negative predictive values for caries experience, predictive values were greater than 95 percent, which indicates that if the visual screening protocol were applied to a sample having a similar caries prevalence, at least 95 out of every 100 positive and negative cases would be true positives and true negatives. These results suggest that the visual screening protocol could provide valid epidemiologic information for dental caries and restorations.

The lower validity of screening for treatment needs in this investigation may indicate uncertainty in translating the diagnosis of untreated decay into the category of nonurgent treatment. A solution would be the use of a computer algorithm incorporated into the data entry program that would automatically assign a code for treatment urgency by taking into account the information entered for caries and responses to yes or no questions about other oral health variables.

Lower validity also was found for other oral conditions (fluorosis, injuries, and presence of sealants). Injuries and presence of sealants were not as prevalent as dental caries in the sample, which might explain some of these results. The low sensitivity for presence of sealants (59%) was probably an effect of the lack of tactile reference. Although fluorosis was as prevalent as caries in the sample, a significant number of false negatives (46 out of 215) were responsible for the lower sensitivity (79%) and negative predictive value (90%).

Some researchers have suggested that validity could be enhanced by improving the diagnostic criteria, retraining, and follow-up of training (16,52). Diagnostic criteria and coding in this investigation were designed to be simple to understand and apply. The validity of screening for injuries and fluorosis could be improved by increasing the length and quality of training and allowing exposure to a wider variety of clinical cases. However, improved training may not affect the validity of screening for dental sealants due to the lack of tactile reference.

Few screening protocols for oral conditions have been proposed and tested; only screening programs for oral mucosal lesions have been tested for validity (16,17,19). Some investigators used findings from these screening protocols to estimate epidemiologic parameters (53-55). Screening programs for other oral conditions, especially dental caries, have been implemented without testing their validity (25). Anecdotal information suggests that screening protocols, such as the one in the Association of State and Territorial Dental Directors' "seven-step model" (24), are being increasingly used, in part because of the financial and logistic difficulties associated with visual-tactile examinations.

The current investigation also focused on potential differences in the validity of measurements obtained by a registered nurse and a dental hygienist. Except for sensitivity for fluorosis and nonurgent treatment, the nurse provided information as valid as that provided by the dental hygienist. In both cases, this finding could be explained in part by differences in prevalence. The hygienist's higher sensitivity and lower specificity values for untreated decay (95% and 94%, respectively) could be explained also by a higher prevalence in her sample

(45%). However, the hygienist had a substantially larger proportion of false positives, with a direct effect on her positive predictive value (93%). Possibly, the hygienist's previous clinical training, focused on evaluating incipient signs of disease, is responsible for the large proportion of false positives, a tendency that could have been aggravated by the absence of a confirmatory tactile reference.

Few studies have reported on the quality of the data collected by personnel other than dentists in performing visual or visual-tactile intraoral examinations (26,29). Direct comparison with their results is not possible because of their focus on reliability or use of different methodologies. However, we agree that personnel other than dentists can be trained to perform oral health status assessments. Primary health care workers have been used in Sri Lanka and India to perform screenings to detect oral mucosal lesions (16,56). Observations by nondentists can be highly valid if the criteria are clear and appropriate training is provided. In this investigation we chose a registered nurse because of her availability. Her normal duties include the health assessment of children in Monticello, and we foresee these epidemiologic assessments as part of routine health examinations. What is particularly remarkable from the present findings is that a nurse, who had little formal dental training, was able to use the screening criteria and adequately apply them.

Validity of Questionnaires. Questionnaires completed by parents or guardians were tested for their validity in obtaining status information on a selected group of oral conditions, i.e., presence of restorations, untreated decay, sealants, and injuries to the anterior teeth. The instrument included a "don't know" option in all questions to avoid forcing the parent to "guess" the status of their children. A fairly high proportion of respondents did not know the oral health status of their children. Validity was modest for all four indicators except for restorations (Table 4), suggesting limitations for their use in surveillance.

Many studies have reported on the quality of self-reported information on different aspects of oral health. A series of studies published in the 1970s, 1980s, and early 1990s in Scandinavia (36-41), England and Ireland (43,57),

and the United States (42) investigated the quality of data collected through questionnaires. Data collected in these investigations included the responders' number of teeth and use of dentures, presence of caries and gingival diseases, and personal medical history. These investigations compared the self-appraised oral health status with results from a clinical examination. These studies tested reliability (sometimes labeled as "agreement" or "validity") by using various statistics, including percent agreement, kappa, statistical mean, and Pearson's product moment correlation. Some studies described overreporting, e.g., having more teeth than those present (36,41,43), others underreporting (38,40), or both (37). This inconclusive evidence suggests that the reliability of self- or proxy-reported oral health data is limited.

In this investigation, analysis by sociodemographic variables showed no clear or consistent difference from the overall results. This finding may be explained, in part, by the sociodemographic homogeneity of the children in the sample.

How "Good" is a "Valid" Procedure? A review of publications in the dental literature reveals subjective assessments by investigators on how good they view the quality of their results based on values for validity. For example, for O'Sullivan and Tinanoff (58), a positive predictive value of 87 percent for caries in the maxillary teeth as an indicator of caries risk in the pits and fissures of posterior teeth in the primary dentition one year later was acceptable. Bretz and co-workers (59) concluded that PERIOSCAN™, which showed 91 percent sensitivity and 89 percent specificity, was a valid instrument for detecting bacterial colonization of root surfaces. Pietilä and associates (22) accepted specificity values between 72 percent and 92 percent as valid. Ikeda and associates (17) concluded that a positive predictive value of 71 percent was fairly good. Warnakulasuriya and Pindborg (16) concluded that a positive predictive value of 58 percent was successful. In a recent publication, Brunette (60) presented a sample of sensitivity and specificity values for diagnostic tests used in dentistry taken from the literature. Sensitivity for dental caries was as low as 13 percent and as high as 93 percent. Clearly, a wide range of ac-

ceptance levels exists. In fact, few studies have concluded that the procedures being tested were not valid. This conclusion can mislead potential users who might apply a procedure with borderline validity in a low prevalence sample and reach incorrect conclusions (61).

The problem of defining how "good" a procedure is based on a validity result depends on many aspects, including the nature and quality of the standard used, the prevalence of the condition being assessed (maximizing positive predictive value with minimum effect on negative predictive value), and the proportion of false positives and false negatives the user of the procedure is willing to accept. These arguments reduce the problem from how "valid" a procedure is to "how practical" it is for its intended purpose. In fact, oral epidemiologists trade off validity for reliability in many indices and procedures. Given this circumstance, the screening procedure tested in this investigation appears useful for estimating dental caries in samples having a similar prevalence to this study.

Surveillance systems for oral conditions are so underdeveloped that almost every aspect of these systems requires empirical support. The need for oral health status and treatment needs data is increasing because dental public health programs must support their need for and use of resources. This investigation has tested a visual screening protocol suitable as a surveillance tool for collecting oral health status data. Parent- or guardian completed questionnaires seem to be of limited use mainly because a high proportion of parents did not know their children's oral health status, other than the presence of one or more restorations. Because questionnaires are unlikely to be designed to address only one topic, if necessary, this question on restorations could be included in ongoing structured questionnaires with a wider range of topics.

References

1. Thacker SB, Berkelman RL. Public health surveillance in the United States. Epidemiol Rev 1988;10:164-90.
2. Stroup NE, Zack MM, Wharton M. Sources of routinely collected data for surveillance. In: Teutsch SM, Churchill RE, eds. Principles and practice of public health surveillance. New York: Oxford University Press, 1994:31-85.
3. Radike AW. Criteria for diagnosis of den-

tal caries. Proceedings of the conference on the clinical testing of cariostatic agents. Chicago: American Dental Association, 1968.

4. National Institute of Dental Research. Oral health of United States children. Washington, DC: US Public Health Service, 1989; pub no (NIH) 89-2247.

5. World Health Organization. Oral health surveys basic methods. 3rd ed. Geneva: WHO, 1987.

6. Riordan PJ, Espelid I, Tveit AB. Radiographic interpretation and treatment decisions among dental therapists and dentists in Western Australia. Community Dent Oral Epidemiol 1991;19:268-71.

7. Knutsson K, Brehmer B, Lysell L, Rohlin M. General dental practitioners'evaluation of the need for extraction of asymptomatic mandibular third molars. Community Dent Oral Epidemiol 1992;20: 347-50.

8. Rytömaa I, Järvinen V, Järvinen J. Variation in caries recording and restorative treatment plan among university teachers. Community Dent Oral Epidemiol 1979;7:335-9.

9. Mileman P, Purdell-Lewis D, van der Weele LT. Variation in treatment decisions and radiographic caries diagnosis among university teachers. Community Dent Oral Epidemiol 1982;10:329-34.

10. Bader JD, Shugars DA. Understanding dentists' restorative treatment decisions. J Public Health Dent 1992;52:102-10.

11. Bader JD, Shugars DA. Agreement among dentists' recommendations for restorative treatment. J Dent Res 1993;72: 891-6.

12. Centers for Disease Control and Prevention. Core public health functions and state efforts to improve oral health—United States, 1993. MMWR Morb Mortal Wkly Rep 1994;43:201, 207-9.

13. Commission on Chronic Illness. Chronic illness in the United States. Vol. 1. Prevention of chronic illness. Cambridge, MA: Harvard University Press, 1957:45.

14. Morrison AS. Screening in chronic disease. New York: Oxford University Press, 1985.

15. Frankenburg WK. Principles in selecting diseases for screening—criteria in screening test selection. In: Frankenburg WK, Camp BW, eds. Pediatric screening tests. Springfield, IL: Charles C. Thomas, 1975:9-37.

16. Warnakulasuriya S, Pindborg JJ. Reliability of oral precancer screening by primary health care workers in Sri Lanka. Community Dent Health 1990;7:73-9.

17. Ikeda N, Ishii T, Iida S, Kawai T. Epidemiological study of oral leukoplakia based on mass screening for oral mucosal diseases in a selected Japanese population. Community Dent Oral Epidemiol 1991;19:160-3.

18. Bánóczy J, Rigó O. Prevalence study of oral precancerous lesions within a complex screening system in Hungary. Community Dent Oral Epidemiol 1991;19: 265-7.

19. Downer MC, Evans AW, Hughes Hallet CM, Jullien JA, Speight PM, Zakrzewska JM. Evaluation of screening for oral cancer and precancer in a company headquarters. Community Dent Oral Epidemiol 1995;23:84-8.

20. Pindborg JJ. Screening for oral cancer. In: Prorok PC, Miller AB, eds. Screening for cancer. Technical Report Series, 78. Geneva: International Union Against Cancer, 1985.

21. Bruck TL. Dental screening. In: Frankerburg WK, Camp BW, eds. Pediatric screening tests. Springfield, IL: Charles C. Thomas, 1975:270-84.

22. Pietilä T, Pietilä I, Väätäjä P. Early screening for orthodontic treatment. Differences in assessments made by a consultant orthodontist and three public health dentists. Community Dent Oral Epidemiol 1992;20:208-13.

23. Gerstner GE, Clark GT, Goulet J-P. Validity of a brief questionnaire in screening asymptomatic subjects from subjects with tension-type headaches or temporomandibular disorders. Community Dent Oral Epidemiol 1994;22:235-42.

24. Association of State and Territorial Dental Directors, Maternal and Child Health Bureau. Assessing oral health needs. ASTDD seven-step model. Columbus, OH: ASTDD, 1995.

25. Heine CS, Jong A, Casamassimo PS, Osterbrock N, Call RL. Oral health status and behaviors of the elderly and other adults: results of a health screening. Geriatr Dent 1983;3:217-21.

26. Hughes JT, Rozier RG, Ramsey DL. Natural history of dental diseases in North Carolina 1976-1977. Durham, NC: Carolina Academic Press, 1982.

27. Maine Department of Human Services. Survey of the oral health of Maine's fifth-grade public schoolchildren. Augusta, ME: Maine Department of Human Services, 1985.

28. South Carolina Department of Health and Environmental Control. The South Carolina Dental Health and Pediatric Blood Pressure Survey 1982-1983. Columbia, SC: South Carolina Department of Health and Environmental Control.

29. Mauriello SM, Bader JD, Disney JA, Graves RC. Examiner agreement between hygienists and dentists for caries prevalence examinations. J Public Health Dent 1990;50:32-7.

30. Disney JA, Abernathy JR, Graves RC, Mauriello SM, Bohannan HM, Zack DD. Comparative effectiveness of visual/tactile and simplified screening examinations in caries risk assessment. Community Dent Oral Epidemiol 1992;20:326-32.

31. Bovonsantijid Y. Factors associated with the reliability of oral examination by teachers in Phuket primary school oral disease surveillance program (dissertation). Bangkok, Thailand: Mahidol University, 1990.

32. Wagener DK, Nourjah P, Horowitz AM. Trends in childhood use of dental care products containing fluoride: United States, 1983-89. Advance data from Vital and Health Statistics; no 219. Hyattsville, MD: National Center for Health Statistics, 1992.

33. Cleveland J, Holm K, Malvitz D, Hines B. Recent dental exams among adults in Washington State—Behavioral Risk Factor Surveillance System, 1993-1994 [Abstract]. J Dent Res 1996;75(Spec Iss):231.

34. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research. Principles and quantitative methods. New York: Van Nostrand Reinhold, 1986.

35. Rothman KJ. Modern epidemiology. Boston, MA: Little, Brown and Company, 1986.

36. Helöe LA. Comparison of dental health data obtained from questionnaires, interviews and clinical examinations. Scand J Dent Res 1972;80:495-9.

37. Widström E, Nilsson B. Dental health and perceived treatement needs of Finnish immigrants in Sweden. Scand J Soc Med 1984;12:129-36.

38. Könönen M, Lipasti J, Murtomaa H. Comparison of dental information obtained from self-examination and clinical examination. Community Dent Oral Epidemiol 1986;14:258-60.

39. Tervonen T, Knuuttila M. Awareness of dental disorders and discrepancy between "objective" and "subjective" dental treatment needs. Community Dent Oral Epidemiol 1988;16:345-8.

40. Lahti S, Tuutti H, Honkala E. Comparison of numbers of remaining teeth from questionnaires and clinical examination. Proc Fin Dent Soc 1989;85:217-23.

41. Palmqvist S, Söderfeldt B, Arnbjerg D. Self-assessment of dental conditions: validity of a questionnaire. Community Dent Oral Epidemiol 1991;19:249-51.

42. Douglass CW, Berlin J, Tennstedt S. The validity of self-reported oral health status in the elderly. J Public Health Dent 1991;51:220-2.

43. Beal JF, Dickson S. Parental awareness of the dental needs of 5-year-old children in the West Midlands, England. Community Dent Oral Epidemiol 1974;2:91-4.

44. Centers for Disease Control and Prevention. Recommended infection-control practices in dentistry, 1993. MMWR Morb Mortal Wkly Rep 1993;42(RR-8).

45. Summers CJ, Gooch BF, Marianos DW, Malvitz DM, Bond WW. Practical infection control in oral health surveys and screenings. J Am Dent Assoc 1994;125: 1213-17.

46. Kaste LM, Gift HC, Bhat M, Swango PA. Prevalence of incisor trauma in persons 6 to 50 years of age: United States, 1988-1991. J Dent Res 1996;75(Spec Iss):696-705.

47. Dean AG, Dean JA, Burton AH, Dicker RC. Epi Info, Version 6. Atlanta: Centers for Disease Control and Prevention, 1994.

48. Gordis L. Epidemiology. Philadelphia: WB Saunders, 1996.

49. Fleiss JL. Statistical methods for rates and proportions. New York: John Wiley & Sons, 1981.

50. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

51. Alderman EJ, White SL, Johnson WT. Georgia dental disease prevalence survey. Georgia Dent Assoc Action 1994:17-19.

52. Camp BW. Principles of screening. In: Frankerburg WK, Camp BW, eds. Pediatric screening tests. Springfield, IL: Charles C. Thomas, 1975:5-8.

53. Gupta PC, Mehta FS, Daftary DK, et al. Incidence rates of oral cancer and natural history of oral precancerous lesions in a

10-year follow-up study of Indian villagers. Community Dent Oral Epidemiol 1980;8:283-333.

54. Bouquot JE, Gorlin RJ. Leukoplakia, lichen planus, and other oral keratoses in 23,616 white Americans over the age of 35 years. Oral Surg 1986;61:373-81.

55. Axéll T. Occurrence of leukoplakia and some other oral white lesions among 20,333 adult Swedish people. Community Dent Oral Epidemiol 1987;15:46-51.

56. Mehta FS, Gupta PC, Bhonsle RB, Murti PR, Daftary DK, Pindborg JJ. Detection of oral cancer using basic health workers in an area of high oral cancer incidence in India. Cancer Detect Prev 1986;9:218-25.

57. Fenlon MR, McCartan BE. Validity of a patient self-completed health questionnaire in a primary care dental practice. Community Dent Oral Epidemiol 1992; 20:130-2.

58. O'Sullivan DM, Tinanoff N. Maxillary anterior caries associated with increased caries risk in other primary teeth. J Dent Res 1993;72:1577-80.

59. Bretz WA, Eklund SA, Radicchi R, et al.

The use of rapid enzymatic assay in the field for the detection of infections associated with adult periodontitis. J Public Health Dent 1993;53:235-40.

60. Brunette DM. Critical thinking. Understanding and evaluating dental research. Chicago, IL: Quintessence, 1996.

61. Shapiro MF, Lehman AF. The diagnosis of depression in different clinical settings. An analysis of the literature on the dexamethasone suppression test. J Nerv Ment Dis 1983;171:714-20.