Epidemiologic Indices for Measuring the Clinical Manifestations of Dental Fluorosis: Overview and Critique

R.G. Rozier

ADR 1994 8: 39
DOI: 10.1177/08959374940080010901

The online version of this article can be found at:
http://adr.sagepub.com/content/8/1/39
EPIGENETRIC INDICES FOR MEASURING THE CLINICAL MANIFESTATIONS OF DENTAL FLUOROSIS: OVERVIEW AND CRITIQUE

R.G. ROZIER

Department of Health Policy and Administration University of North Carolina Chapel Hill, North Carolina 27599


Abstract—Several indices have been used to describe the clinical appearance of dental fluorosis. The purpose of this paper is to describe and compare the three principal ones in use today: those developed by Dean (1934, 1942), Thylstrup and Fejerskov (1978), and Horowitz et al. (1984). A recent index (Fluorosis Risk Index) developed by Pendrys (1990) is also included in this review. The continued use of Dean’s classification system and derived index (CFI) for more than a half century is testimony to its simplicity and utility. The index has been criticized because the unit of analysis is the person, because criteria are unclear for some categories, or that they lack sensitivity, particularly for severe fluorosis, and because of the way in which data are summarized and reported. The Thylstrup and Fejerskov Index is appealing to clinicians and epidemiologists alike in that it corresponds closely to histological changes that occur in dental fluorosis and to enamel fluoride concentrations, thereby having biological validity. The TSIF described by Horowitz et al. makes a useful contribution because it provides clearer diagnostic criteria and provides for an analysis based on esthetic concerns. The Fluorosis Risk Index appears to be particularly useful for analytical epidemiologic studies, because it is designed to permit a more accurate identification of associations between age-specific exposures to fluoride and the development of dental fluorosis. All three indices in common use today provide useful indices for the study of dental fluorosis. The utility of the Fluorosis Risk Index will be determined as it receives wider use. The selection of one of these indices for use in an epidemiologic study depends in large measure on the purpose of the study. Research needs to continue on the validity of these indices, particularly for mild fluorosis, and on the public’s perception of the cosmetic appearance of teeth with different severity levels of fluorosis.

M any different factors can result in changes in the normal appearance of the enamel (Small and Murray, 1978; Pindborg, 1982). The single most studied causal factor is fluoride, which can result in a range of clinical manifestations referred to collectively as dental fluorosis. This condition can be distinguished from other defects of the enamel based on enamel color, distribution of the condition on the affected tooth or within the mouth, and the extent to which the enamel is left intact.

Enamel changes in populations exposed to differing levels of fluoride were first described by McKay in his initial publication with Black (McKay and Black, 1916). A precise description was provided by McKay (1929) in a subsequent paper, in which he stated, “there are all gradations, from the merest flecking of certain teeth, with white spots scattered about in otherwise normal enamel, to other conditions in which the entire enamel structure is uniformly dead paper white, with or without the presence of the brown discoloration; and still others in which there has been the most pronounced gross destruction of the enamel surfaces by what may be called, for want of a better term, corrosion.”

Several epidemiological indices have been used to describe the clinical appearance of dental fluorosis. The purpose of this paper is to provide an overview of the three principal ones in use today: the classification system developed by Dean (1934), referred to in this paper as Dean’s Index, its subsequent modification, and the Community Fluorosis Index based on this system (Dean, 1942); the Thylstrup and Fejerskov Index (Thylstrup and Fejerskov, 1978; Fejerskov et al., 1988); and the Tooth Surface Index of Fluorosis (Horowitz et al., 1984). The Fluorosis Risk Index, recently described by Pendrys (1990), is also included in this review. Slightly more attention is given to Dean’s Index than to the others, because of its historical significance and widespread use over an extended period, and because it serves as a standard of comparison for all subsequent indices. Previous reviews have been completed on measurement techniques for all enamel opacities, as well as those specific for fluorosis (Small and Murray, 1978; Cutress and Suckling, 1982; Moller, 1982; Horowitz, 1986; Clarkson, 1989). None of these includes a review of the use of these indices, or includes the Fluorosis Risk Index.

While indices other than these four have been used in studies of dental fluorosis, they are not included in this review, because they represent only small refinements in Dean’s approach, or have not received widespread use (Zimmerman, 1954; Forrest and James, 1956; Nevitt et al., 1963; Moller, 1970; Bischoff et al., 1976).
A few indices have been developed to evaluate the cosmetic appeal of teeth based on color alone, the work by Diefenbach et al. (1965) serving as an example. These approaches, used primarily for determining professionals’ perception of the aesthetic appeal of teeth of people drinking optimally fluoridated water, likewise are not included in this review.

**Differential Diagnosis of Dental Fluorosis**

The acceptance of these fluorosis indices rests in large measure on the ability of an examiner to distinguish fluoride-induced changes in the enamel from those that are not fluoride-induced. Several investigators have suggested that some diffuse opacities similar in appearance to dental fluorosis are not caused by fluoride, and have developed descriptive indices for classifying changes in the enamel which require no consideration of etiology (Al-Alousi et al., 1975; Jackson et al., 1975; Ainamo and Cutress, 1982; FDI, 1982; Clarkson and O’Mullane, 1989). This approach, it is argued, alleviates the need for making the sometimes difficult differential diagnosis between fluoride and non-fluoride opacities. Nevertheless, in those studies requiring a clinical determination of fluoride effects on the tooth, one is left with the need to identify fluorosis positively, and thus these descriptive indices are not of direct value.

There is little epidemiologic evidence that opacities characteristic of dental fluorosis can be caused by other factors. Trace elements other than fluoride would seem to be logical choices, yet only two human studies have shown such an association. Strontium (Curzon and Spector, 1977) and zinc (Butler et al., 1985) have been associated with fluorosis-like opacities, but these associations were found to be weak. For example, in a study of dental fluorosis in 16 Texas communities with fluoride levels in the drinking water ranging from 0.2 to 3.3 ppm, water fluoride level accounted for 32% of the variation in dental fluorosis, while various metals accounted for 1% to 6% when considered singly (Segreto et al., 1984).

Nevertheless, care must be taken to make a correct diagnosis. Several criteria are available for use in distinguishing fluorosis in its milder forms from non-fluoride opacities (Zimmerman, 1954; Russell, 1961; Nevitt et al., 1963; Fejerskov et al., 1988; Cutress and Suckling, 1990). While there is no direct evidence, the accuracy of fluorosis diagnosis may be as high as 95% for experienced examiners who give proper attention to examination methods and the use of differential diagnostic criteria (Symposium IV: Fluoride, 1989).

**DEAN’S INDEX**

Early in his research on dental fluorosis, Dean realized the importance of providing a standard classification system for the clinical conditions described by McKay. His collaboration with McKay and his experiences in completing examinations on approximately 2000 subjects in endemic areas of six states resulted in the original classification scale for dental fluorosis, presented in Table 1 (Dean, 1934).

By 1937, Dean alone had examined more than 10,000 subjects in about 185 areas distributed among 16 US states. As could be expected, particularly as observations were extended into non-endemic areas, modifications were made in the classification system (Dean and Elvove, 1937). Dean provided adjustments in the diagnostic criteria for some of the classifications, and the “moderately severe” and the “severe” categories were combined into a single “severe” category providing the six-point measurement scale currently in use (Dean et al., 1939; Dean, 1942). The criteria for these six classifications are given in Table 1.

Several comments can be made about these criteria. Dean’s description of normal enamel makes it clear that from the beginning, differential diagnostic criteria were being applied to eliminate major developmental disturbances of the enamel which resulted in hypoplasia. It is not clear, however, to what extent efforts were made to eliminate minor opacities, since criteria are not provided.

The unfortunate label of “questionable” used for the second category has created confusion and continues to do so. It is unclear if Dean’s use of this term represented his uncertainty over the effects of fluoride on enamel at low levels of exposure, diagnostic difficulties that he might have experienced, or his belief that this degree of fluorosis was not of enough esthetic concern to merit full consideration in an index. It would appear that the label reflects difficulties he had in distinguishing mild fluoride opacities from non-fluoride ones, and thus his desire to reduce the number of false positives. Initially, upper incisors with thin, irregular, white opaque streaks on the incisal third of the tooth, and premolars with white opacities 2 or 3 mm in extent on the cusps were considered affected at a level between “normal” and “very mild” (Dean et al., 1935). While unwilling to make a diagnosis at the individual level, Dean considered questionable cases a positive sign of fluorosis when found in a community with definite cases. As researchers gained additional experience with the index, some clinical conditions used to describe the questionable category, particularly the snowcapping of posterior teeth, were accepted as definitive indications of fluorosis and added as a criterion for “very mild”.

A third issue related to these criteria is the role that staining plays in the diagnosis of fluorosis. Early work by McKay included separate scores for fluorosed enamel with and without stain, even though his writings from the beginning were clear that he considered stain not to be caused by fluoride, but rather to be an acquired condition. The criteria indicate that staining, while often an accompanying condition, particularly in the more severe forms of fluorosis, was not necessary for a diagnosis. Consequently, stain is largely discounted in application of this index and in studying the biological effects of fluorides on the enamel.

Dean used the area of the tooth affected to describe levels of severity, but in doing so, mixed the use of surfaces and
### TABLE 1

#### DIAGNOSTIC CRITERIA AND WEIGHTING SYSTEM FOR DEAN’S INDEX

<table>
<thead>
<tr>
<th>Classification and Weight</th>
<th>Original Criteria (Dean, 1934)</th>
<th>Modified Criteria (Dean, 1942)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal 0</td>
<td>The enamel presents the usual translucent semi-vitriform type of structure. The surface is smooth and glossy and usually of a pale creamy white color.</td>
<td>The enamel presents the usual translucent semi-vitriform type of structure. The surface is smooth, glossy, and usually of a pale creamy white color.</td>
</tr>
<tr>
<td>Questionable 0.5</td>
<td>...slight aberrations in the translucency of normal enamel, ranging from a few white flecks to occasional white spots, 1 to 2 mm in diameter.</td>
<td>The enamel discloses slight aberrations from the translucency of normal enamel, ranging from a few white flecks to occasional white spots. This classification is utilized in those instances where a definite diagnosis of the mildest form of fluorosis is not warranted and a classification of “normal” not justified.</td>
</tr>
<tr>
<td>Very Mild 1.0</td>
<td>Small, opaque, paper-white areas are scattered irregularly or streaked over the tooth surface. It is principally observed on the labial and buccal surfaces, and involves less than 25% of the tooth surfaces of the particular teeth affected. Small pitted white areas are frequently found on the summits of the cusps. No brown stain is present in the mottled enamel of this classification.</td>
<td>Small, opaque, paper-white areas scattered irregularly over the tooth but not involving as much as approximately 25% of the tooth surface. Frequently included in this classification are teeth showing no more than about 1-2 mm of white opacity at the tips of the summits of the cusps of the bicuspid or second molars.</td>
</tr>
<tr>
<td>Mild 2.0</td>
<td>The white, opaque areas on the surfaces of the teeth involve at least half of the tooth surface. The surfaces of molars, bicuspid, and cuspids subject to attrition show thin white layers worn off and the bluish shades of underlying normal enamel. Faint brown stains are sometimes apparent, generally on the upper incisors.</td>
<td>The white opaque areas in the enamel of the teeth are more extensive but do not involve as much as 50% of the tooth.</td>
</tr>
<tr>
<td>Moderate 3.0</td>
<td>No change is observed in the form of the tooth, but generally all of the tooth surfaces are involved. Surfaces subject to attrition are distinctly marked. Minute pitting is often present, generally on the labial and buccal surfaces. Brown stain is frequently a disfiguring complication. It must be remembered that the incidence of brown stain varies greatly in different endemic areas, and many cases of white opaque mottled enamel, without brown stain, are classified as “moderate” and listed in this category.</td>
<td>All enamel surfaces of the teeth are affected, and surfaces subject to attrition show marked wear. Brown stain is frequently a disfiguring feature.</td>
</tr>
<tr>
<td>Moderately Severe</td>
<td>Macroscopically, a greater depth of enamel appears to be involved. A smoky white appearance is often noted. Pitting is more frequent and generally observed on all the tooth surfaces. Brown stain, if present, is generally deeper in hue and involves more of the affected tooth surfaces.</td>
<td>Includes teeth formerly classified as “moderately severe” and “severe”. All enamel surfaces are affected, and hypoplasia is so marked that the general form of the tooth may be affected. The major diagnostic sign of this classification is the discrete or confluent pitting. Brown stains are widespread, and teeth often present a corroded-like appearance.</td>
</tr>
<tr>
<td>Severe 4.0</td>
<td>The hypoplasia is so marked that the form of the teeth is at times affected, the condition often being manifest in older children as a mild pathologic incisal-occlusal abrasion. The pits are deep and often confluent. Stains are widespread and range from a chocolate brown to almost black in some cases.</td>
<td>Includes teeth formerly classified as “moderately severe” and “severe”. All enamel surfaces are affected, and hypoplasia is so marked that the general form of the tooth may be affected. The major diagnostic sign of this classification is the discrete or confluent pitting. Brown stains are widespread, and teeth often present a corroded-like appearance.</td>
</tr>
</tbody>
</table>
teeth in describing the area affected. For the “very mild” category, the criteria indicate that less than 25% of the tooth surface is affected, for “mild”, less than 1/2 of the tooth, and for “moderate” and “severe”, all surfaces. Thus, it is unclear if criteria are to be applied to the entire visible surface area of a tooth, to selected surfaces, or to some combination of the two, depending on the tooth type being examined. The result is a perception that gaps exist in the criteria.

A final issue is that of loss of enamel. The definition of a “pit” necessary for the “severe” category is not clear in the 1942 diagnostic criteria. Previous descriptions of his classification system identify “small, pitted, white areas...on the summit of cusps” as characteristics of the “very mild” category; “thin, white layers worn off” as characteristic of the “mild” category; and “minute pitting...on the labial and buccal surfaces” as belonging in the “moderate” category (Dean, 1934). In a 1935 publication, discrete pits characteristic of the “moderately severe” category were defined as 1-2 mm in diameter (Dean et al., 1935). Thus, it would seem that a defect in the enamel would need to be at least 1 mm in order to be considered a “pit”.

The focus of Dean’s work was on the community as a whole; it was the community that shared a common risk factor because of a common water supply. An index of fluorosis was needed for the community in order to compare its prevalence in those communities with different levels of fluoride in water supplies, and later to correlate these findings with the caries experience of these communities. The variation of fluorosis within individuals was of less importance than among communities. This study of communities determined his approach to development of an index, as well as its use.

In summarizing data at the individual level, Dean suggested the “one person, one disease” approach common in medical epidemiology, and sought to classify each person in a manner “that would designate in a general way the degree of involvement found in that person” (Dean, 1934). The worst two teeth are used as a basis for the person-level score. Where two teeth are not affected to the same degree, the convention used in recent years seems to have been to assign a classification based on the less involved tooth of the two exhibiting the highest scores—in other words, to assign a person-level score based on the lowest of the two highest scores (Horowitz et al., 1984; NIDR, 1991). The person-fluorosis score can be used in defining two indices, a qualitative community fluorosis index and a quantitative one.

For the qualitative index, a community is assigned to one of seven descriptive classifications (negative, borderline, slight, medium, rather marked, marked, and very marked) based on a frequency distribution of individual scores within the surveyed community. Although it is possible to relate the prevalence of fluorosis to exposure variables using this index, the ability to demonstrate marked differences in severity among groups is limited, an observation which quickly became apparent to Dean. In a paper read before the Epidemiology Section of the American Public Health Association in 1938, he chose to display previously published data using a quantitative index (Dean and McKay, 1939). This numerical index of clinical severity, referred to as the Community Fluorosis Index, is computed by giving an arbitrary weight to each of the six classifications of fluorosis. The index, with weights ranging from 0 to 4, is defined mathematically as the average weighted score per person surveyed.

The original use of the CFI resulted in the familiar “S” curve, displaying the relationship of the index plotted against the fluoride content of the water supplies of communities with various levels of fluoride (Dean and McKay, 1939). The shape of this relationship probably reflects the lower weight assigned to questionable cases which represent a high percentage of cases at low levels of fluoride exposure, and to a lack of sensitivity of index criteria at higher levels of exposure.

The only reference to examiner reliability in Dean’s many publications was to the independent examinations of 82 continuous residents of Bauxite, Arkansas, done by Dean and McKay. Based on replicate examinations, they concluded that their “agreement with respect to mottled enamel diagnosis was so close that there would be little value in a detailed recording of their individual findings...” (Dean et al., 1938). Detailed accounts for training of the two examiners for the “21-cities” study are given for dental caries, as well as steps to ensure consistency of these caries examinations, yet no reference is made to fluorosis (Dean et al., 1941).

Dean’s index is still in common use today and is the one recommended by the World Health Organization (WHO, 1987). The longevity of the index is testimony to the brilliance and persistence of McKay and Dean when developing the index, to the clinical manifestations of the response of enamel to fluoride, which is both predictable and believed to be in large measure unique, and to the simplicity and utility of the resulting index.

Limitations of Index
Several investigators have pointed out the limitations of Dean’s index (Forrest, 1956; Jackson, 1961; Al-Alousi et al., 1975; Small and Murray, 1978; Thylstrup and Fejerskov, 1978; Horowitz, 1986; Fejerskov et al., 1988; Clarkson, 1989). These limitations can be summarized under four major points. Two points—that the index presupposes the condition and that the diagnostic criteria are unclear, imprecise, or not sensitive enough—have already been discussed in the overview of the index and will not be discussed further.

A third general criticism of Dean’s Index relates to its use of either the person or the community as the unit of measurement. While basing the person-level score on the most severely affected teeth serves the purpose of measuring the maximum dental effects of fluoride exposure, important in determining the safe level of exposure to fluoride from drinking water, it is not useful in those instances where the type of tooth or surface affected is important. Three problems can result from use of this person score. First, aggregation across surfaces or teeth may prevent adequate discrimination among individuals, particularly at high fluoride levels where...
all subjects may have pitting in some teeth. Second, a person-level score may misclassify an individual if esthetic considerations are more important than biological ones, since the most affected teeth tend to be posterior ones, those which have less cosmetic importance than anterior teeth. Finally, the person score does not permit determinations of possible changes in the level of fluoride exposure during tooth development.

Since the CFI is an average, it provides little information about the variation within a population, particularly if measures of variance are not provided, the usual case in reporting CFI values. Further, since it is a weighted average, individual scores may not correlate well with the community score. On average, the response as measured by the CFI may be similar in two communities, when the actual severity of fluorosis differs. A large number of questionable ratings can produce a CFI of the same magnitude as that in a population with fewer questionables, but with some higher scores. A concern pointed out by Fejerskov et al. (1988) is that the index may in fact misrepresent the dose-response relationship between fluoride and fluorosis. The sigmoidal curve referred to earlier may be linear if questionable scores were given equal weight, and if the severe levels were disaggregated.

The final major criticism directed toward Dean’s Index concerns statistical manipulations and reporting, primarily related to use of the CFI. Since the classification system is based on an ordinal scale, and the distribution of scores is usually not normal at some levels of fluoride exposure, the mean is an inappropriate statistic. Further, the weights assigned to each category are arbitrary.

THYLSTRUP AND FEJERSKOV INDEX (TFI)

According to Thylstrup and Fejerskov, they developed their index in order to “refine, modify, and extend the original concepts established by Dean” (Fejerskov et al., 1988). The primary aim was to develop a more sensitive classification system for recording enamel changes found in areas with fluoride in the drinking water at levels above that studied by Dean. The basis for the TFI is appealing to clinicians and epidemiologists alike in that the classification scale corresponds closely to histological changes that occur in dental fluorosis (Thylstrup and Fejerskov, 1978) and to fluoride concentrations found in enamel (Richards et al., 1989).

A 10-point ordinal scale is used to classify enamel changes associated with increasing fluoride exposure (Table 2). As originally proposed, facial and occlusal surfaces were scored with different criteria used at some levels of severity. Most investigations have used only facial surfaces, and by 1988, this had become the recommended procedure. This recommendation is apparently based on the similarity of the different surface scores on the same tooth, and on the difficulty of getting an accurate assessment of fluorosis on occlusal surfaces because of the likelihood that scores will be affected by occlusal wear.

Scores 1-4 reflect increasing involvement of the tooth surface with opaque areas and loss of translucency. The system is based on the premise that all parts of the surface are generally affected, even at the lowest levels of exposure. Thus, thin, white opaque lines running across all parts of the surface, characteristic of a TFI score of “1”, begin to merge to form cloudy areas of increasing size and coverage, until the entire surface exhibits a marked opacity, characteristic of a TFI score of “4”.

In keeping with one of the purposes of the index, Dean’s category of “severe” is divided into 5 categories in order to record the wide variety of clinical changes that can occur in those drinking water with high fluoride levels. Assignment of a score representing one of these categories is based on the proportion of the surface affected. Enamel stains are ignored in assignment of scores throughout the entire scale. The methods used in conducting the examinations distinguish this index from Dean’s Index and from the TSIF. Teeth are to be cleaned and dried before examination.

Replicate examinations were not possible in the field during the original use of this index, but photographs of 125 buccal surfaces were evaluated by the two examiners. Intra-examiner reliability, measured as percent complete agreement, was 76% and 85.6%. Inter-examiner agreement for two different sessions was 72% and 79.2%.

The histological and clinical bases for the criteria used for scoring fluorosis with this index have clarified the way in which fluorosis is distributed over the tooth surface in the case of continuous exposure to constant levels of fluoride during tooth development, as well as the role of enamel loss in fluorosis. The approach used in formulating scoring criteria has had the effect of reducing some of the subjectivity in scoring. Further, cleaning and drying of teeth accentuate the appearance of fluorosis, making diagnoses easier in questionable cases.

Another important aspect of this index is that the authors have provided recommendations for summarizing data and for their presentation. These methods provide both prevalence and severity estimates, and include: (1) percent distributions of scores for all teeth; (2) percent distributions of scores by tooth type; and (3) cumulative percent distributions of subjects by percent of teeth affected at a given TFI score (or greater) per subject.

TOOTH SURFACE INDEX OF FLUOROSIS (TSIF)

The Tooth Surface Index of Fluorosis was developed in the early 1980’s by investigators from the National Institute of Dental Research when conducting a survey to assess the prevalence of dental caries and dental fluorosis in Illinois communities having optimal and above-optimal concentrations of natural fluoride in drinking water. In this index, a separate score is given to each facial and lingual surface of anterior teeth and to each buccal, occlusal, and lingual surface of posterior teeth. Tooth surfaces are not dried before the examination, the rationale being primarily an esthetic one. The thought is that teeth should be assessed in...
TABLE 2

DIAGNOSTIC CRITERIA AND SCORING SYSTEM
FOR THE THYLSTRUP AND FEJERSKOV INDEX

<table>
<thead>
<tr>
<th>Score</th>
<th>Original Criteria (Thylstrup and Fejerskov, 1978)</th>
<th>Modified Criteria (Fejerskov et al., 1988)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal translucency of enamel remains after prolonged air-drying.</td>
<td>The normal translucency of the glossy, creamy-white enamel remains after wiping and drying of the surface.</td>
</tr>
<tr>
<td>1</td>
<td>Narrow white lines located corresponding to the perikymata.</td>
<td>Thin white opaque lines are seen running across the tooth surface. The lines correspond to the position of the perikymata. In some cases, a slight “snowcapping” of cusps/incisal edges may also be seen.</td>
</tr>
</tbody>
</table>
| 2     | Smooth surfaces
More pronounced lines of opacity which follow the perikymata. Occasionally confluence of adjacent lines.   | The opaque white lines are more pronounced and frequently merge to form small cloudy areas scattered over the whole surface. “Snowcapping” of incisal edges and cusp tips is common. |
|       | Occlusal surfaces Scattered areas of opacity < 2 mm in diameter and pronounced opacity of cuspal ridges.         | Merging of the white lines occurs, and cloudy areas of opacity occur spread over many parts of the surface. In between the cloudy areas, white lines can also be seen. |
| 3     | Smooth surfaces
Merging and irregular cloudy areas of opacity. Accentuated drawing of perikymata often visible between opacities. | The entire surface exhibits a marked opacity or appears chalky white. Parts of the surface exposed to attrition appear less affected. |
|       | Occlusal surfaces
Confluent areas of marked opacity. Worn areas appear almost normal but usually circumscribed by a rim of opaque enamel. | The entire surface is opaque, and there are round pits (focal loss of the outermost enamel) that are less than 2 mm in diameter. |
| 4     | Smooth surfaces
The entire surface exhibits marked opacity or appears chalky white. Parts of surface exposed to attrition appear less affected. | The small pits may frequently be seen merging in the opaque enamel to form bands that are less than 2 mm in vertical height. In this class are also included surfaces where the cuspal rim of facial enamel has been chipped off, and the vertical dimension of the resulting damage is less than 2 mm. |
|       | Occlusal surfaces
Entire surface exhibits marked opacity. Attrition is often pronounced shortly after eruption.                 |                                                                                                          |
| 5     | Smooth and occlusal surfaces
Entire surface displays marked opacity with focal loss of outermost enamel (pits) < 2 mm in diameter.       |                                                                                                          |
| 6     | Smooth surfaces
Pits are regularly arranged in horizontal bands < 2 mm in vertical extension.                               |                                                                                                          |
|       | Occlusal surfaces
Confluent areas < 3 mm in diameter exhibit loss of enamel. Marked attrition.                               |                                                                                                          |
| 7     | Smooth surfaces
Loss of outermost enamel in irregular areas involving < 1/2 of entire surface.                             | There is a loss of the outermost enamel in irregular areas, and less than half the surface is so involved. The remaining intact enamel is opaque. |
|       | Occlusal surfaces
Changes in the morphology caused by merging pits and marked attrition.                                      |                                                                                                          |
| 8     | Smooth and occlusal surfaces
Loss of outermost enamel involving > 1/2 of surface.                                                          | The loss of the outermost enamel involves more than half the enamel. The remaining intact enamel is opaque. |
| 9     | Smooth and occlusal surfaces
Loss of main part of enamel with change in anatomical appearance of surface. Cervical rim of almost unaffected enamel is often noted. | The loss of the major part of the outer enamel results in a change of the anatomical shape of the surface/tooth. A cervical rim of opaque enamel is often noted. |
Criteria for the 8-point scale used in classifying surfaces are given in Table 3. Four points can be made concerning these criteria. First, as with the TFI, there is no questionable category, and thus a positive score is assigned to the first signs of fluorosis. Second, TSIF scores 1-3 are based on the area of the tooth surface affected, derived by visually coalescing all areas of fluorosis and relating that area to the total visible enamel of a particular surface. The area of a surface affected can provide a useful indication of severity when exposure to fluoride is continuous during the development of the tooth. With interrupted fluoride exposure, however, only a portion of the surface may be affected, and to various degrees depending on the level of exposure. An inaccurate assessment of severity can be provided in those instances where continuous and interrupted exposures result in identical scores based on surface area affected, yet the severity of fluorosis differs.

The third point concerns the score of “4”, given when staining is present in conjunction with any of the three conditions indicative of lower scores. If one’s primary interest is measuring the esthetic consequences of fluorosis, the index is probably an ordinal one. However, if the index is being used solely as an indicator of the biological effects resulting from fluoride exposure, the TSIF must be considered a nominal scale rather than an ordinal one according to the purest definition of scales. Since staining of any portion of the surface is given precedence over the proportion of the surface affected with fluorosis, a score of “4” may be given when an area equivalent to a TSIF score of “1” is present. However, stain rarely occurs before a “mild” level of fluorosis is reached according to Dean’s Index, most likely corresponding to a score of “3” in the TSIF, and should not be of major concern if one wishes to treat the index scale as ordinal.

Fourth, the esthetic orientation of this index is evident once again in the higher degrees of fluorosis, where staining occurring in conjunction with discrete pitting is given a separate score. Finally, a distinction between discrete pitting and more advanced confluent pitting is made, making the index more sensitive than Dean’s to higher degrees of fluorosis.

Results obtained with the use of this index are presented as various frequency distributions of surface scores or of the maximum mouth score, either for all surfaces or for specific surfaces such as facial surfaces of anterior teeth or occlusal surfaces of first molars. With use of the TSIF, examiner reliability may be of more concern than with Dean’s Index or the modified TFI because of the larger number of assessments to be made (72 surfaces vs. the usual maximum of 28 teeth or buccal surfaces in children). The added lingual surfaces are more difficult to visualize than buccal surfaces, which also adds to the concern about examiner reliability. The original publication documented examiner reliability for the two examiners involved in its use. Agreement was assessed for total and maximum scores for all tooth surfaces combined, and for the labial surfaces of the six maxillary anterior teeth. Kappa values were presented only for maximum scores. Intra-examiner agreement ranged from 0.66 to 0.83 for the two examiners and for all teeth combined and for anterior teeth only. Interexaminer agreement as measured by the kappa was 0.35 for all teeth, and 0.54 for anterior teeth, indicating some difficulty in achieving an acceptable level of agreement for all teeth.

The TSIF has two major advantages over Dean’s Index. First, criteria for scoring in the TSIF are clearer, and consequently, subjectivity should be reduced in their application. The other advantage derives from the scoring of surfaces rather than individual teeth. Surface scores allow for
FLUOROSIS RISK INDEX (FRI)

In any study of the association between fluoride exposure and fluorosis, it is important to relate time of exposure closely to the period when enamel is at risk. This linkage is complicated by the large number of enamel surfaces at risk, their development at different times, even separate areas of the same tooth surface, and the long period of time over which they develop. The issue is further complicated by the uncertainty over the relative importance of the secretory and maturation phases of enamel development in the risk of fluorosis. Analytical epidemiologic studies are particularly subject to problems arising from the difficulties involved in relating fluoride exposure to outcomes.

The Fluorosis Risk Index has been designed to address this linkage problem and to respond to the current pattern of fluoride exposure in US children, which can be highly variable during the period when teeth are at risk of fluorosis. The three indices described to this point are based on scoring an individual, an entire tooth, or tooth surface, and do not permit a precise classification of time of exposure in relation to that portion of the enamel that is at risk at that time. The unique feature of the FRI is that each tooth is divided into zones that correspond to the age at which they begin development, and can be related to narrow age-bands of fluoride exposure, such as a 12-month time period.

Enamel surface zones that begin formation during the first year of life (i.e., between birth and the first birthday) and those that begin formation between the third and sixth years of life (i.e., between the second and sixth birthdays) are identified and scored separately. As displayed in the Fig., the buccal surface and the incisal edge/occlusal table of each permanent tooth (excluding third molars) have been divided into four scoring zones: (1) the incisal edge/occlusal table, defined as the enamel surface within one millimeter of the incisal edge of the tooth; (2) the incisal/occlusal third of the buccal surface; (3) the middle third of the buccal surface; and (4) the cervical third of the buccal surface.

The darkly shaded areas in the Fig. designate those portions of the enamel that begin formation during the first year of life and are referred to as Classification I enamel surface zones. The lightly shaded areas designate Classification II enamel surface zones that begin formation during the third through sixth years of life. The clear areas designate unassigned enamel surface zones for which categorization is questionable, based on the available literature on tooth development, or where their development occurred after 5 years.
years of age. Roughly 112 zones are scored, with 10 belonging to Classification I, 48 to Classification II, and the remaining 54 to unassigned zones.

Each zone is scored as either negative for fluorosis, questionable for fluorosis, positive for mild-to-moderate fluorosis, or positive for severe fluorosis, according to the criteria given in Table 4. A surface zone is diagnosed as being positive when 50% or more of the area of a zone being scored is affected—according to the authors, a severity level equivalent to the “mild” category in Dean’s Index. The criterion for this categorization in Dean’s Index, however, is based on the proportion of the entire tooth that is affected, not a surface or portion of a surface. Thus, the two classifications may not be equivalent. Methods to be used are not provided in the publication describing the index (Pendrys, 1990), yet it would appear from its use that teeth are to be dried before examination (Pendrys and Katz, 1989).

Subjects are identified as cases or controls for each of the two surface zones based on the distribution of scores within each. At least two surfaces of a particular zone must have scores of 2 or 3 in order to be considered a case. In the analysis, risk factors are considered separately for each classification.

The single use of this index has been reported by the authors (Pendrys and Katz, 1989). Examiner reliability studies during its initial use produced excellent results. The two examiners achieved interexaminer agreement, as measured by the kappa statistic, of 0.76 and 0.82 for Classification I and Classification II surface zones, respectively. Kappa statistics for intra-examiner agreement varied from 0.83 to 1.00 for the two types of surface classifications. Results from this first study were in agreement with several other studies which have demonstrated an association between use of fluoride supplements and enamel fluorosis. Further, this study provided epidemiological evidence that enamel may be at greater risk for fluorosis during the maturation phase of development than during the secretory phase.

The strengths and weaknesses of the FRI should become more apparent with its use by other investigators. The biological and analytical premises appear sound. While the ability exists to collapse FRI values to yield prevalence data, the index was not designed for this purpose. In most situations, the other three indices discussed in this review are better suited to provide strictly prevalence data. The index is complex, both from a biological perspective and in its application. Acceptable levels of examiner reliability may be difficult to establish due to this complexity. Further, results cannot be compared with those of any of the established indices used in determining prevalence. Therefore, the index cannot serve the dual purpose of providing prevalence estimates as well as estimates of risk.

A primary premise of this index is that scores for the two age-related developmental zones are independent. The differential diagnosis of fluorosis required of this and other fluorosis indices is usually based on an initial overall assessment of the dentition, since the intra-oral distribution of fluorosis is useful in making a diagnosis. Before proceeding with specific tooth calls, the examiner usually makes the initial assessment as to whether the subject is a case or not. This procedure cannot be followed with the FRI because of the likelihood of bias in making determinations for the Classification I and II zones. Approximately two-thirds of the classifications as to a case, control, or questionable in Classifications I and II were in agreement in the 1247 subjects who were assessed in the initial application of the index. Any existing bias in these determinations could weaken or even mask associations between exposure and

---

**Table 4**

**DIAGNOSTIC CRITERIA AND SCORING SYSTEM FOR THE FLUOROSIS RISK INDEX (FRI)**

<table>
<thead>
<tr>
<th>Classification and Score</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative: 0</td>
<td>A surface zone will receive a score of 0 when there is absolutely no indication of fluorosis being present. There must be a complete absence of any white spots or striations, and tooth surface coloration must appear normal.</td>
</tr>
<tr>
<td>Questionable: 1</td>
<td>Any surface zone that is questionable as to whether there is fluorosis present (i.e., white spots, striations, or fluorotic defects cover 50% or less of the surface zone).</td>
</tr>
<tr>
<td>Positive: Mild-to-Moderate: 2</td>
<td>A smooth surface zone will be diagnosed as being positive for enamel fluorosis if greater than 50% of the zone displays parchment-white striations typical of enamel fluorosis. Incisal edges and occlusal tables will be scored as positive for enamel fluorosis if greater than 50% of that surface is marked by the snowcapping typical of enamel fluorosis.</td>
</tr>
<tr>
<td>Positive: Severe: 3</td>
<td>A surface zone will be diagnosed as positive for severe fluorosis if greater than 50% of the zone displays pitting, staining, and deformity, indicative of severe fluorosis.</td>
</tr>
<tr>
<td>Non-fluoride Opacity: 7</td>
<td>Any surface zone that has an opacity that appears to be a non-fluoride opacity.</td>
</tr>
<tr>
<td>Excluded: 9</td>
<td>A surface zone is categorized as excluded (i.e., not adequately visible for a diagnosis to be made) when any of the following conditions exist: incomplete eruption, orthodontic appliances or bands, crowned or restored surfaces, gross plaque, and debris.</td>
</tr>
</tbody>
</table>
TABLE 5

METHODS USED IN STUDIES OF DENTAL FLUOROSIS, BY INDEX, 1980-1992

<table>
<thead>
<tr>
<th>Method</th>
<th>Category</th>
<th>DI (n = 26)</th>
<th>TFI (n = 18)</th>
<th>TSIF (n = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teeth Dried</td>
<td>Gauze/Cotton Rolls</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Compressed Air</td>
<td>-</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>6</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>20</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Teeth Cleaned</td>
<td>Subject brushed</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Plaque removed</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>as needed</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gauze</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Prophylaxis</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>20</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Lighting</td>
<td>Dental</td>
<td>10</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Natural</td>
<td>10</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>6</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Differential</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Diagnostic</td>
<td>Criteria</td>
<td>Yes</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Used</td>
<td>Not reported</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>Examiner</td>
<td>Reliability</td>
<td>Yes</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Determined</td>
<td>Not reported</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Number of Examiners</td>
<td>1</td>
<td>18</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3-4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>1</td>
<td>3</td>
<td>-</td>
</tr>
</tbody>
</table>

outcomes.

Case definition will affect the identification of risk factors. The FRI is conservative in the definition of surface zones and in the severity level chosen for a case. This conservative approach may create difficulties in identifying those factors that carry a small risk, or are widespread in the study population.

There is also a concern about the validity of assessments in Classification I, which are based on only 10 surfaces, the incisal edges of mandibular incisors, maxillary incisors, and occlusal surfaces of first molars. Incisal edges of anterior teeth are more likely to demonstrate clinical signs of fluorosis than many other surfaces, probably because of a smaller amount of dentin, their increased translucency, and the tendency for these surfaces to be drier than others. Further, the occlusal surfaces of first molars are those surfaces most likely to be excluded because of restorations, and presumably dental sealants. Not surprisingly, over 60% of Classification I cases involved the incisal edges of anterior teeth.

USE OF DENTAL FLUOROSIS INDICES

A review of studies published in the English literature since 1980 was undertaken to provide an assessment of the extent to which these indices have been adopted for use in clinical assessments of dental fluorosis, for what purposes, the methods used, and the success investigators have achieved with them. While not intended to be exhaustive, this review does represent virtually all the published studies since 1980.

Dean’s Index

In all, 26 publications reporting use of Dean’s Index were available for review (Rozier and Dudney, 1981; Driscoll et al., 1983, 1986; Segredo et al., 1984; Butler et al., 1985; Mann et al., 1985, 1987, 1990; Schamschula et al., 1985; Brathall et al., 1986; Grobler et al., 1986; Ishii and Suckling, 1986, 1991; Leverett, 1986; Eklund et al., 1987; Brouwer et al., 1988; Bohaty et al., 1989; Evans, 1989; Haikel et al., 1989a, b; Kumar et al., 1989; Songpaisan and Davies, 1989; Evans and Stamm, 1991; Clarkson and O’Mullane, 1992; Lewis et al., 1992; Warnakulasuriya et al., 1992). These studies were conducted in 11 countries (Hong Kong, Hungary, Ireland, Israel, Japan, Morocco, Senegal, South Africa, Sri Lanka, Thailand, and United States) using 78 different study populations with an average of approximately 710 subjects per study. The majority of these are analytical studies examining risk factors for fluorosis, the dental risks and benefits associated with fluoridation, the effects of adjustments in water fluoride levels on fluorosis, and age at which individuals are at risk. A small number of studies report prevalence and trends in fluorosis.

As with all epidemiologic determinations of clinical conditions, estimates of dental fluorosis can be subject to variability due to the different methods used in their clinical determination as well as to examiner differences in these assessments. However, due to the clinical manifestations of the condition and the subjective nature of determinations for fluorosis in its milder forms, estimates may be subject to more variation than many other oral diseases and conditions. Several aspects of fluorosis examination methods—the drying and cleaning of teeth, the type of lighting used, and the use of differential diagnostic criteria to distinguish between fluoride and non-fluoride opacities—for the 51 studies included in the review for this paper are summarized in Table 5, along with the number of studies that provided estimates of examiner reliability, and the number of examiners used in conducting each study.

For those studies using Dean’s Index, if and how teeth were dried and cleaned and the use of differential diagnostic criteria go largely unreported. One can assume that teeth were examined wet in all studies, as specified in the use of this system; however, reporting of this aspect of the methods,
as well as of the others listed in Table 5, would be helpful.

Eleven of the studies using Dean’s Index provided estimates of examiner reliability (Table 6) which generally represent good to excellent agreement. Eight of the studies, however, report only percent agreement scores which may be misleading in the presence of a large number of similar scores. One study found an unacceptable level of examiner agreement during examiner training (kappa = 0.35), but was able to demonstrate 79% examiner agreement during data collection (Songpaisan and Davies, 1989).

The definition of a case differs among the studies, primarily depending on its purpose. Most of the analytic studies use any signs of fluorosis, that is, questionable or higher, for defining a case, and the person as the unit of analysis. However, a few used very mild or greater, and fewer still, moderate or greater for the case definition. For prevalence estimates, very mild or greater was the most common definition of a case, most likely influenced by a

### Table 6

**Examiner Reliability Reported in Studies of Dental Fluorosis, by Index, 1980-1992**

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of Examiners</th>
<th>Unit of Measurement</th>
<th>% Agreement</th>
<th>Kappa Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dean's Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bohaty et al. (1989)</td>
<td>1</td>
<td>NS*</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Clarkson &amp; O’Mullane (1992)</td>
<td>4</td>
<td>Person</td>
<td>&gt;95</td>
<td></td>
</tr>
<tr>
<td>Driscoll et al. (1983)</td>
<td>1</td>
<td>Person</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Eklund et al. (1987)</td>
<td>1</td>
<td>All teeth</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Evans (1989)</td>
<td>2</td>
<td>Index tooth</td>
<td>0.76, 0.83</td>
<td>0.61</td>
</tr>
<tr>
<td>Evans &amp; Stamm (1991)</td>
<td>1</td>
<td>Index &amp; all teeth</td>
<td>0.92, 0.86,</td>
<td>0.68, 0.57+</td>
</tr>
<tr>
<td>Lewis et al. (1992)</td>
<td>1</td>
<td>NS</td>
<td>&gt;95</td>
<td></td>
</tr>
<tr>
<td>Mann et al. (1983)</td>
<td>1</td>
<td>Person</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Mann et al. (1990)</td>
<td>1</td>
<td>Person</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Songpaisan &amp; Davies (1989)</td>
<td>1</td>
<td>Person</td>
<td>79</td>
<td>0.35</td>
</tr>
<tr>
<td>Warnakulasuriya et al. (1992)</td>
<td>1</td>
<td>Person</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td><strong>Thylstrup &amp; Fejerskov Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baelum et al. (1986)</td>
<td>3</td>
<td>Buccal all teeth</td>
<td>65, 66, 66</td>
<td></td>
</tr>
<tr>
<td>Holt et al. (1990)</td>
<td>1</td>
<td>Buccal incisors</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Manji et al. (1986)</td>
<td>1</td>
<td>Buccal all teeth</td>
<td>64-100++</td>
<td>0.67</td>
</tr>
<tr>
<td>Osuji et al. (1988)</td>
<td>1</td>
<td>NS</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Riordan &amp; Banks (1991)</td>
<td>1</td>
<td>Index surface</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Woltgens et al. (1989)</td>
<td>2</td>
<td>Buccal 1/2 teeth</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td><strong>Tooth Surface Index of Fluorosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driscoll et al. (1986)</td>
<td>2</td>
<td>All surfaces</td>
<td>0.48, 0.65</td>
<td></td>
</tr>
<tr>
<td>Heifetz et al. (1988)</td>
<td>2</td>
<td>All surfaces</td>
<td>0.51, 0.64</td>
<td></td>
</tr>
<tr>
<td>Ismail et al. (1990)</td>
<td>3</td>
<td>Person</td>
<td>0.51-0.70+++</td>
<td></td>
</tr>
<tr>
<td>Szpunar &amp; Burt (1988)</td>
<td>1</td>
<td>Person</td>
<td>92</td>
<td>0.85</td>
</tr>
</tbody>
</table>

* Not specified.
+ Same day index tooth, same day all teeth, 1 month apart all teeth, 4 months apart all teeth, respectively.
+++ Agreement by TFH score.
++++ Range for intra- and interexaminer reliability.
desire to be consistent with Dean’s work. More than half of the studies calculated the CFI for use in addressing analytic questions or for presenting prevalence.

**TFI**

The TFI has received widespread use for both descriptive and analytical studies, with 18 studies having been reported during the time period included in the review (Holm and Andersson, 1982; Wenzel and Thystrup, 1982; Larsen et al., 1985a,b, 1986, 1987, 1988, 1989; Baelum et al., 1986; Manji et al., 1986a,b,c; Osuji et al., 1988; Wolgens et al., 1989; Holt et al., 1990; Opinya et al., 1991; Riordan and Banks, 1991; Mabelya et al., 1992). The adoption of this index by investigators is tempered somewhat, however, by the fact that the list of authors for 10 of the 18 studies included one of the original developers of the index. The 18 studies included more than 30 different populations, with an average sample size of approximately 275 subjects per study, and were conducted in nine different countries (Australia, Canada, Denmark, Greenland, Kenya, The Netherlands, Sweden, Tanzania, United Kingdom). Only a few of these studies are descriptive, with most investigating the effect of either a single risk factor—such as fluoride tablets, fluoride gels, and altitude—or multiple risk factors. A few also investigate the age at which teeth are at risk of fluorosis.

A relatively high number of studies do not report if teeth were dried or cleaned (Table 5). Again, if the protocol suggested by the index developers was followed, we can assume that teeth were both dried and cleaned. Since much of the research in fluorosis is currently taking place in both developed and developing countries where access to dental equipment may vary, it is important to indicate in what way teeth were dried, cleaned, and illuminated.

Examiner reliability studies were included in seven of the studies using the TFI, with kappa statistics produced for three and percent agreement scores for four (Table 6). Kappa statistics indicated good to excellent agreement. Percent agreement scores ranged from 65% to 100%.

The majority of studies using the TFI do not provide a single classification for the individual. Rather, prevalence and severity estimates are generally reported at the individual or tooth level according to one or more of the methods suggested by the authors of the index, typically a frequency distribution of scores (Fejerskov et al., 1988). In most of these studies, a limited number of risk factors with few confounders were being studied for their association with fluorosis, which permitted comparisons of groups using the various frequency distributions of TFI scores. With the investigation of multiple risk factors, the use of an individual score becomes more important. Two studies of multiple risk factors and using multivariate techniques have been done, both of which used a TFI of greater than 0 as the case definition (Osuji et al., 1988; Riordan and Banks, 1991).

**TSIF**

Seven studies were available for assessment of the use of the TSIF (Horowitz et al., 1984; Driscoll et al., 1986; Heifetz et al., 1988; Szpunar and Burt, 1988; Bagramian et al., 1989; Woolfolk et al., 1989; Ismail et al., 1990; Williams and Zwemer, 1990). These studies were conducted in the USA and Canada, and included an average of slightly over 600 subjects per study, all children and adolescents. The purpose of these studies included the investigation of risk factors for dental fluorosis, risks and benefits of fluoridation and other fluoride exposures, the affect of dental fluorosis severity on caries susceptibility, and trends in fluorosis.

The methods used in conducting the examinations are generally not reported, yet one would have to assume, at least with regard to drying of teeth, that they were examined wet as specified in the TSIF (Table 5). The use of Russell’s (1961) criteria for differentiating fluoride and non-fluoride opacities appears to be accepted by investigators as a part of the TSIF system.

The four studies using the TSIF and reporting estimates for examiner reliability are displayed in Table 6. All four indicate acceptable levels of examiner agreement, but display a wide variation, with values for the kappa statistic ranging from 0.48 to 0.85.

The TSIF, like the TFI, eliminates the “questionable” category used in Dean’s Index, and thus is based on the premise that any positive score is fluorosis. The assignment of a fluorosis classification to an individual based on the potentially large number of surface scores provides opportunity for variation in case definition, particularly in the absence of recommendations in this regard. The case definition is not uniform in the five analytical studies choosing to use the subject as the unit of analysis, and thus requiring a decision on criteria for individual classification. Subjects were assigned a classification based on 1 or more teeth with evidence of fluorosis (Ismail et al., 1990), at least 2 bilateral teeth with evidence of fluorosis (Woolfolk et al., 1989), and the highest score (Williams and Zwemer, 1990). Case definitions for 2 of these 5 studies were not clear (Szpunar and Burt, 1988; Bagramian et al., 1989).

Several conclusions can be made concerning the use of these three indices. First, all three have been and continue to be used with success by a number of investigators for several different purposes, types of studies, and with different populations. Second, important methods used in conducting the clinical examinations go unreported, and thus it is not known to what extent comparisons of prevalence data for a particular index may be hampered by a lack of comparability. This omission seems particularly important in regard to drying and cleaning of teeth, since the potential effect on fluorosis estimates can be large. Third, examiner reliability is within an acceptable range for those studies that provided evidence of this aspect of the examinations. However, only about half of studies provided such estimates, and fewer still provide the more appropriate conservative estimate derived from use of statistical methods that would account for chance agreements. Fourth, the all-important definition of a case differs among studies. As a result, comparisons of prevalence estimates among studies using the same index are further compromised, and identification of risk factors which might
TABLE 7
PREVALENCE OF DENTAL FLUOROSIS
DERIVED FROM STUDIES COMPARING DI, TFI, AND TSIF

<table>
<thead>
<tr>
<th>Study</th>
<th>Fluoride (ppm) in Drinking Water</th>
<th>Percent Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DI</td>
</tr>
<tr>
<td>Thylstrup &amp; Fejerskov (1978)</td>
<td>3.5, 6.0, 21.0</td>
<td>100</td>
</tr>
<tr>
<td>Wenzel &amp; Thylstrup (1982)</td>
<td>&lt;0.2</td>
<td>3</td>
</tr>
<tr>
<td>Granath et al. (1985)</td>
<td>0.2</td>
<td>20</td>
</tr>
<tr>
<td>Burger et al. (1987)+</td>
<td>1.6</td>
<td>15*</td>
</tr>
<tr>
<td>Driscoll et al. (1986)</td>
<td>Optimal</td>
<td>44</td>
</tr>
<tr>
<td>Horowitz et al. (1984)</td>
<td>2x Optimal</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>3x Optimal</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>4x Optimal</td>
<td>87</td>
</tr>
<tr>
<td>Cleaton-Jones &amp; Hargreaves (1990)</td>
<td>1.56</td>
<td>67</td>
</tr>
</tbody>
</table>

+ Primary teeth, not dried for either index.
** Percent of teeth affected.
*** Maximum TSIF score.
*** Percent of surfaces affected.

be common across several studies made difficult. Finally, complete reporting of prevalence and severity of dental fluorosis in each study, regardless of study type, would facilitate research in dental fluorosis. For example, the presentation of data derived from use of Dean’s Index should, at a minimum, provide prevalence estimates as the percent with very mild or greater, and severity estimates as a percent distribution of all individual scores (person or tooth) and the CFI. These three methods would standardize reporting and provide sufficient information for historical purposes.

STUDIES COMPARING INDICES

Six studies have compared Dean’s Index, the TFI, and the TSIF in various combinations of head-to-head comparisons (Thylstrup and Fejerskov, 1978; Wenzel and Thylstrup, 1982; Horowitz et al., 1984; Granath et al., 1985; Driscoll et al., 1986; Burger et al., 1987; Cleaton-Jones and Hargreaves, 1990). Prevalence estimates provided by these studies are given in Table 7. In each case, the estimates for Dean’s Index are based on the presence of questionable levels of fluorosis or greater.

The first four studies compare results obtained with Dean’s Index and the TFI. Little difference was found in prevalence estimates derived from the use of the two indices. The one study that found a difference led these investigators to suggested modifications in scoring for both indices in order to facilitate comparisons of results (Granath et al., 1985). They propose that scores of “1” in the TFI be scored as normal, and the “questionable” category of Dean be considered as definite fluorosis. In terms of severity, Thylstrup and Fejerskov (1978) found that the TFI was able to identify a difference in fluorosis effect between the two high-fluoride communities, while Dean’s Index was unable to do so. In the three studies done in populations exposed to fluoride levels of 2 times optimal or less, marginal differences were evident in severity scores. The TFI was somewhat more sensitive, a finding which led Granath et al. (1985) to conclude that it was more appropriate for use in clinical trials with limited sample sizes. All four investigators preferred the TFI over Dean’s Index because it appeared to be more appropriate for studying the biological effects of fluorides, or because it was found to be easier to use in the field.

Dean’s Index and TSIF estimates for fluorosis in the same population are provided by the two publications reporting results of studies in Illinois (Horowitz et al., 1984; Driscoll et al., 1986). Higher prevalence estimates were found with the TSIF, and it was further able to identify significant differences between those exposed to 2- and 3-times-optimal fluoride levels, whereas Dean’s Index was unable to do so. Reasons for differences displayed in this Table may be due to the manner in which a person was classified as a case, as well as other differences in the two classification systems.

The final study has used all three indices in the same population. Comparisons among all indices are difficult, since prevalence estimates are based on the person for Dean’s Index, and on surfaces for the other two indices. Results for the TSIF and the TFI are directly comparable, however. For the three categories of normal, mild-to-moderate fluorosis, and severe fluorosis, the two indices were in agreement 92% of the time. When the two were not in agreement, the TFI produced higher scores. In 98 instances, fluorosis was diagnosed as mild-to-moderate with the TFI but not with the TSIF, and in only two instances was the opposite found. Further, the TFI identified 11 scores of severe fluorosis...
which were classified as mild-to-moderate with the TSIF. In no instances was the reverse the case. These investigators concluded that all three indices were used without difficulty in the field, but examination time was greater with the TFI because of the necessity of drying the teeth.

CONCLUSIONS

The four dental fluorosis indices reviewed in this paper provide suitable measurement techniques for the broad array of study questions which have been investigated to date. A prime consideration in choosing among these four indices is the purpose of the study in which the index is to be used. Dean’s Index, the TFI, and the TSIF all can be used for prevalence studies, particularly at fluoride exposure levels below 5 ppm. For general screening in communities to determine the number of people affected, and for historical epidemiological studies in which trends are to be established, Dean’s is often the index of choice because of its simplicity and the availability of previous data. If the “questionable” category in Dean’s Index is considered a positive score for fluorosis and all teeth are scored, prevalence estimates are probably not underestimated to any great extent compared with results obtained with the TSIF and TFI. The primary factors contributing to differences are the drying of teeth with the TFI, and the separate scoring of buccal, lingual, and occlusal surfaces with the TSIF. The TFI and TSIF are a better choice for prevalence studies in which the level of fluoride exposure is above 5 ppm, since they provide more sensitive measures of enamel loss than does Dean’s Index. Of the two, the TFI provides more categories for severe fluorosis and would seem more appropriate for use in populations exposed to very high levels of fluoride found in certain parts of the world.

The TFI seems more appropriate than Dean’s Index or the TSIF for use in clinical trials or analytical epidemiologic studies, primarily because teeth are dried and fluorosis can be identified in its milder forms. The resulting increased sensitivity provides statistical and practical advantages from the possible detection of effects with smaller samples. This feature of the TFI is a particular advantage when potential fluoride effects are small, or when the exposure may be widespread, as is the case with fluoride from toothpaste in developed countries. This choice must consider, however, the case definition, which is often based on public health issues in addition to biological ones. The TSIF provides an esthetic basis for the case definition and could be used in those instances where the desire is to identify risk factors for fluorosis as it appears in its natural state, i.e., when teeth are not cleaned and dried.

For use in analytical epidemiologic studies, the FRI offers an important conceptual advance in fluorosis indices. This index may provide statistical advantages in those populations with interrupted fluoride exposure, since the exposure is more precisely tied to the stage of development when the enamel might be at risk. The index may also be more suitable for those studies investigating the phase of enamel development that is at greatest risk for fluorosis. With more widespread use, the advantages and disadvantages of this index will become more apparent.

Other scientific considerations in choosing an index are examiner reliability and index validity. Dean’s Index, the TFI, and the TSIF have all demonstrated acceptable levels of reliability when used by several investigators in many different studies. More comprehensive documentation is needed in this area, however, in that only slightly over 40% of the studies using these indices since 1980 reported examiner reliability. Researchers should routinely report examiner reliability when using any of these indices, since it will vary among or within studies under various examination conditions or study populations.

Surprisingly little work has been done on the validity of any of these indices. The severity of fluorosis can be affected at any time by the teeth available in the mouth for observation, and over time by remineralization, abrasion, or treatment. Further, quantitative estimates of the false-positive rate for enamel opacities occurring in populations with low levels of fluoride exposure have not been provided. Both factors affect the validity of the measurements, and thus can have an affect on estimates of the relationship between fluoride exposure and enamel effects.

Beyond the purpose of the study and the scientific considerations discussed, practical factors should be considered when selecting an index for use in prevalence surveys. Dean’s Index, the TFI, and the TSIF all appear to be easy to learn and use in the field. The TFI and TSIF require more examination time, simply because of the need to clean and dry the teeth when using the TFI, and because of the increased number of calls made with the TSIF. As with any epidemiologic index, experience with its use is important in providing useful results. In those studies to be undertaken by inexperienced investigators, access to investigators with experience in their use will often be an important determining factor in selection of one of these indices.

These indices are designed primarily for studying the biological effects of fluoride. None adequately addresses the public health significance of dental fluorosis. Objectionable fluorosis has usually been defined as moderate to severe fluorosis. Perhaps this is not for us alone to decide. In the final analysis, a condition that is not considered normal, and that can be prevented or treated with reasonable efforts, may be judged by society at large as a problem. No studies have investigated the public’s perceptions of the different levels of severity of fluorosis. An esthetic index incorporating considerations of the public’s view would provide a useful adjunct to the study of the clinical aspects of fluorosis.

REFERENCES

Al-Alousi W, Jackson D, Compton G, Jenkins OC (1975). Enamel mottling in a fluoride and in a non-fluoride
Dean HT, McKay FS (1939). Production of mottled enamel halted by a change in common water supply. Am J Publ Health 29:590-596.
Dean HT, McKay FS, Elvove E (1938). Mottled enamel survey of Bauxite, Ark., ten years after a change in the common water supply. Publ Hlth Rep 53:1736-1748.
Dean HT, Jay P, Arnold FA, Elvove E (1941). Domestic water and dental caries. II. A study of 2,832 white children aged 12-14 years, of eight suburban Chicago communities, including L. acidophilus studies of 1,761 children. Publ Hlth Rep 56:761-792.
Grobler SR, van Wyk CW, Kotze D (1986). Relationship between enamel fluoride levels, degree of fluorosis and...


McKay FS (1929). The establishment of a definite relation between enamel that is defective in its structure, as mottled enamel, and the liability to decay. *Dent Cosmos* 71:747-755.


