Community Dental Health (2016) **33,** 145–150
Received 14 October 2015; Accepted 9 December 2015

doi:10.1922/CDH_3873Nor06

Examiner reliability in fluorosis scoring: a comparison of photographic and clinical methods

N.A. Mohd Nor[1, 2], I.G. Chestnutt[1] and B.L. Chadwick[1]

[1]*Department of Applied Clinical Research and Dental Public Health, Cardiff University, UK;* [2]*Department of Community Oral Health and Clinical Prevention, Faculty of Dentistry, University of Malaya, Malaysia*

Objective: To assess examiner reliability when scoring dental fluorosis in Malaysian children using clinical (Dean's Index) and photographic methods. ***Method:*** The upper central incisors of 111 children were examined both clinically and photographically for fluorosis status using Dean's index. Twenty children were re-examined after a two-week interval for intra-examiner reliability by a single examiner. In addition, two independent examiners and the clinical examiner scored 111 photographic images of the same children in a standardized manner. Fluorosis scores were compared individually between examiners for both clinical and photographic scoring. Examiner reliability was assessed using both simple and weighted kappa statistics at tooth level. Sensitivity, specificity, positive-negative predictive values and a Receiver Operating Characteristic (ROC) curve were also calculated to determine the accuracy of the test. ***Results:*** Across the three examiners, the prevalence of fluorosis (Dean's score ≥2) using photographs was lower (ranged from 23% to 26%) than the prevalence recorded by clinical examination (30%). The kappa score for intra-examiner reliability for the duplicate clinical examination was excellent (0.89). Inter-examiner reliability between the photographic method and the clinical examination (gold standard) for each examiner was substantial with weighted kappa values ranging from 0.74 to 0.77. The photographic method indicated higher specificity (99%) than sensitivity (79%) and the area under the ROC curve was also high (0.89) which suggests good accuracy of the diagnostic test. ***Conclusion:*** These results suggest that photographic examination of fluorosis on central incisors can be recorded with good examiner reliability. The recorded fluorosis prevalence was lower using the photographic scores.

Key words: *Dean's index, dental fluorosis, Malaysia, reliability, photographic, clinical*

Introduction

Dental fluorosis is defined as hypomineralization of tooth enamel resulting from exposure to excess levels of fluoride during tooth formation (Dean, 1934). Clinically, dental fluorosis presents as white striations or diffuse parchment-like areas on the enamel. In more severe cases, fluorosed enamel appears pitted and yellowish-brown in colour (Browne *et al*., 2005; Buzalaf and Levy, 2011; Mascarenhas, 2000).

Several epidemiological indices have been developed to describe the clinical appearance of dental fluorosis. No one index has emerged as the agreed standard method, the index of choice to a degree depending on the objective of the study. Dean's Index (Dean, 1934) was the first index reported in the literature. In the intervening years, other indices were developed, the aim being to improve Dean's Index criteria: Thylstrup and Fejerskov Index (TFI; Fejerskov *et al*., 1977), Tooth Surface Index of Fluorosis (TSIF; Horowitz *et al*., 1984) and the Fluorosis Risk Index (Pendrys, 1990). These indices were classified as aetiological indices that specifically measure fluoride induced enamel changes. In contrast, descriptive indices such as the Developmental Defects of Enamel (DDE) index, record enamel defects, based on descriptive criteria, without assuming the aetiology of the defects (FDI, 1982).

Regardless of which index is used in the clinical assessment of fluorosis, the diagnosis can be affected by many factors such as variation in the method of examination, tooth condition (wet or dry), lighting conditions, examiner bias and intra and inter-examiner reliability (Whelton *et al*., 2004). A potential way of overcoming these shortcomings is to use a standardized photographic method for capturing a permanent record of the appearance of the enamel.

There are however, advantages and disadvantages in photographic assessment compared with direct clinical recording of dental fluorosis. The major benefits of photographs are that they capture a permanent record and allow blind scoring. Photography also enables scoring by multiple examiners in multi-site studies and allows repeated assessments of the same images (Cochran *et al*., 2004a; Ellwood *et al*., 1994; Fejerskov *et al*., 1977; Soto-Rojas *et al*., 2008).

The disadvantages of using photographs are firstly variation in photographic technique between studies such as differences in equipment, lens, lighting system and the quality of the image produced. Secondly, difficulties in capturing teeth images due to lack of accessibility especially for posterior teeth mean that photographs have only been used to record the anterior teeth, mainly incisors and canines. This could result in under reporting of the prevalence of dental fluorosis. In contrast, the greater detail provided by photographs may well result in over reporting prevalence (Cochran *et al*., 2004a; Soto-Rojas *et al*., 2008).

Correspondence to: Nor Azlida Mohd Nor, Department of Applied Clinical Research and Dental Public Health, Cardiff University Dental School College of Biomedical and Life Sciences, Heath Park, Cardiff CF14 4XY, UK. Email: mohdnorn@cardiff.ac.uk

There are a number of studies that have assessed the prevalence of fluorosis using photographs alone (Cochran *et al*., 2004b; Ellwood *et al*., 1994; Tavener *et al*., 2007) and clinical examinations compared with photographs (Stephen *et al.*, 2002). Several studies have compared photographic methods with clinical examination using the (DDE) Index (Nunn *et al*., 1993; Wong *et al*., 2005), the Fluorosis Risk Index (Cruz-Orcutt *et al*., 2012; Soto-Rojas *et al*., 2008) and the TF index (Ellwood *et al*., 1996). Some studies compared several indices against each other using both clinical and photographic methods (Pretty *et al*., 2012; Sabieha and Rock, 1998). However, information is scarce on how Dean's Index compares with photographic methods. Therefore, the present study aims to compare examiner reliability and the relative prevalence and severity scores resulting from clinical and photographic assessment of fluorosis using Dean's index in Malaysian children.

Method

Data from this study were obtained from a larger ongoing Malaysian fluoridation study. The 111 children studied in this exercise were aged 9 and 12 years old and were lifetime residents in a fluoridated community (0.5-0.7mg/L) located in Shah Alam, Selangor. To be included in the study, the children were required to have no medical contraindications to undergoing a clinical dental examination and have had informed written consent provided by their parent or guardian.

Fluorosis was scored by three examiners. Examiner 1 undertook both clinical and photographic assessment, whilst examiners 2 and 3 participated only in the photographic assessment. Examiner 1 received extensive training in the use of Dean's Index as part of the Malaysian National Oral Health Survey. The training for fluorosis assessment involved an online training module (Whelton *et al.*, 2014), theoretical explanation and clinical assessment on clinical subjects. The same online training module was used by the two photographic only examiners (2 and 3). This online training generated kappa scores for intra-examiners reliability and was repeated until each examiner reached very good to excellent kappa values. Prior to conduct of the study, Examiner 1 repeated the online training module and received intra-oral photographic technique training from the Audio-Visual Department, School of Dentistry, Cardiff University.

Clinical examinations were conducted by a trained and calibrated examiner (examiner 1). Clinical recording of fluorosis was conducted under natural light with the subject sitting on a chair in the upright position using a disposable mirror, CPITN probe and gauze for plaque removal (if necessary). Maxillary central incisors were evaluated using Dean's Index in a wet condition (0, normal; 1, questionable; 2, very mild; 3, mild; 4, moderate; 5, severe). If fluorosis was present, diagnosis was based on the condition of the maxillary central incisors. If the two central incisors were not equally affected, the condition of the least affected tooth was recorded. Twenty children were re-examined after a two-week interval to assess intra-examiner agreement.

Immediately after the clinical examination, digital images of the maxillary incisors were taken with a digital SLR camera, Nikon 90D body, sigma 105mm f/2.8 macro lens and sigma macro ring flash E140. The photographic technique used in this study followed the method described by Cochran and colleagues (2004a). A cheek retractor was inserted into the child's mouth and they were instructed to keep their head still and place their teeth edge to edge. If it was not possible to maintain edge to edge incisal contact, the child was instructed to bring their upper and lower central incisors into the same vertical plane as far as possible. The photographs were taken while the teeth were still wet. Children were asked to moisten their teeth before the photograph was taken. If this was not possible damp cotton wool was used to keep the teeth moist. Most of the photographs only involved one exposure per child. However on occasion, where the examiner was not satisfied with the first photograph (such as issues with specular reflection), further exposures were attempted.

None of the images contained any identifying aspects of the subject's face. A photographic log form enabled the digital images to be linked to a subject identifying code. The digital images were downloaded to a computer for storage and viewing. In those cases where more than one exposure had been taken, the best quality image was selected.

Photographic image scoring took place in Cardiff University, 45 days after clinical examinations in Malaysia. Each photograph was assigned a unique identifying number. The photographs were then mixed randomly for blind fluorosis scoring. All 111 images were included in the assessment and projected onto a white screen using Microsoft PowerPoint in a darkened room. The size of the image projected on screen was approximately 69cm by 38cm. In terms of magnification of the image approximately five times linear magnification from standard photo print size 12.5cm by 7.5cm. All three examiners were seated approximately three metres from the screen and scored the photographs at the same time under identical lighting conditions. Following individual assessment, all examiners re-examined all photographs and discussed scores thoroughly to achieve consensus agreement on the final photographic score. The consensus photographic score was based on the agreement of at least two of the examiners. In the blind scoring protocol, a specific code of 'unable to score image' was also included alongside with Dean's Index code. Any issues with the images such as presence of light reflection or excess camera-flash were noted during the evaluation of each photograph.

This study was reviewed and approved by Cardiff University Dental School Research Ethics Committee (DSREC 14/17a). In addition, permission to conduct the study was obtained from the relevant Ministries in Malaysia namely the Ministry of Health, the Ministry of Education and the State Education Department.

Data were entered analysed using SPSS and STATA software. The tooth-level Dean's score was compared between the same examiner (clinical versus duplicate clinical score; clinical versus photographic score) and different examiners (individual photographic score versus clinical score, Table 1); individual photographic score versus other examiner, Table 2) and individual photographic score versus consensus photographic score, Table 2). The clinical score was used as the gold standard.

Table 1. Inter-examiner agreement of dental fluorosis by clinical examination

	Unweighted data		*Weighted data*	
Clinician Examiners	*Kappa*	*Agreement*	*Kappa*	*Agreement*
Examiner 1 clinical versus Examiner 2 photographs	0.82	93%	0.77	90%
Examiner 1 clinical versus Examiner 3 photographs	0.72	89%	0.74	87%

Table 2. Inter-examiner agreement of dental fluorosis between individual photographic score and consensus photographic score

	Unweighted data		*Weighted data*	
Clinicians	*Kappa*	*Agreement (%)*	*Kappa*	*Agreement (%)*
Examiner 1 versus Examiner 2	0.78	92	0.80	95
Examiner 1 versus Examiner 3	0.72	90	0.85	96
Examiner 2 versus Examiner 3	0.85	95	0.75	89
Examiner 1 versus Consensus	0.83	94	0.91	96
Examiner 2 versus Consensus	0.91	96	0.87	94
Examiner 3 versus Consensus	0.90	96	0.82	92

Note: Consensus photographic score based on the agreement of at least two of the three examiners. Examiner 1 was both a clinical and a photographic examiner, Examiners 2 and 3 were photographic examiners only

Table 3. Weighting matrix used for computing the weighted kappa statistic

	Normal	*Questionable*	*Very mild*	*Mild*	*Moderate*
Normal	1	½	0	0	0
Questionable	½	1	½	0	0
Very mild	0	½	1	½	0
Mild	0	0	½	1	½
Moderate	0	0	0	½	1

A positive diagnosis of fluorosis was based on very mild or greater (Dean's score ≥2) or as no fluorosis (Dean's score 0 or 1). For statistical analysis the data were dichotomised into fluorosis or no fluorosis to simplify analysis. A weighted kappa value was generated for inter-examiner reliability using STATA Software in order to utilize the full range of Dean's Index. The weighted matrix used for computing weighted kappa statistics is shown in Table 3. A weight of 1 was given for exact agreement, a weight of 0.5 was given when examiner disagreed by only one severity level and a weight of 0 was given when examiners disagreed by more than one severity level. Descriptive analysis was used to describe the prevalence of fluorosis. McNemar's Test was used to determine if there were statistically significant differences between the prevalence of fluorosis using clinical and photographic methods (Altman, 1990). Percentage agreement and kappa statistics were used to assess examiner reliability at tooth level. Kappa interpretation was based on the definition by Landis and Koch (1977). Kappa values 0.81 to 1.0 indicate excellent agreement, 0.61 to 0.80 indicate substantial agreement, 0.41 to 0.60 indicate moderate agreement, 0.21 to 0.40 indicate fair and less than 0.20 indicate poor agreement. In addition, sensitivity, specificity, positive-negative predictive values and a Receiver Operating Characteristic (ROC) curve between clinical and consensus photo score were also calculated to determine the accuracy of the diagnostic test.

Results

A total of 111 participants were examined clinically and 111 images of these same participants were examined photographically for fluorosis on central maxillary incisors. It was possible to score all 111 images and none were excluded because of poor image quality.

Following re-assessment of 20 children, intra-examiner clinical examination reliability by a single examiner (examiner 1) indicated substantial agreement (89.6%) with a weighted kappa value of 0.89. Intra-examiner agreement between all 111 photographs and corresponding clinical examinations by a single examiner (examiner 1) also indicated substantial agreement with a weighted kappa value of 0.87. Although there was good intra-examiner reliability, Examiner 1 identified a significantly higher prevalence of fluorosis in clinical scores (30%) than photographic scores (23%) (p=0.02).

Table 4 shows the fluorosis prevalence and frequency distribution of Dean's scores for individual examinations of clinical and photographic methods. Most fluorosis cases fell into very mild and mild categories. The prevalence of fluorosis (Dean's score ≥2) using clinical examination was higher than the consensus photographic score (30% vs. 24%, p=0.07).

Table 1 shows inter-examiner reliability between clinical and photographic methods. Inter-examiner reliability between photographic examiners (2 and 3) versus clinical examiner (examiner 1) was found to have substantial agreement using both weighted and simple kappa statistics.

Table 2 shows all examiners demonstrated substantial to excellent inter-examiner reliability for photographic scoring with weighted kappa values ranging from 0.72 to 0.91. There was little difference found between weighted and simple kappa analysis.

Further analysis was carried out using the consensus photographic score versus the clinical examination (gold standard) score. Sensitivity, specificity, positive-negative predictive values and the Receiver Operating Characteristic (ROC) curve were calculated. Table 5 shows that the diagnosis of fluorosis using the photographic method had a higher specificity (99%) than sensitivity (79%); positive predictive value 96%; negative predictive value 92% likelihood ratios (+LR=39;-LR=0.22).

Table 4. Fluorosis prevalence and frequency distribution of Dean's scores for individual examinations for clinical and photographic methods

Method of examination	*Dean's Index score, n (%)*						*Fluorosis**
	0	*1*	*2*	*3*	*4*	*5*	*Prevalence*
	Normal	*Questionable*	*Very mild*	*Mild*	*Moderate*	*Severe*	*n (%)*
Clinical	72 (65)	6 (5)	21 (19)	10 (9)	2 (2)	0	33 (30)
Photographic:							
Examiner 1	73 (66)	12 (11)	15 (14)	9 (8)	2 (2)	0	26 (23)
Examiner 2	77 (69)	5 (5)	21 (19)	6 (5)	2 (2)	0	29 (26)
Examiner 3	75 (68)	11 (10)	10 (9)	14 (13)	1 (1)	0	25 (23)
Consensus photo score	73 (69)	11 (10)	16 (14)	9 (8)	2 (2)	0	27 (24)

Note: *A positive diagnosis of fluorosis is based on a Dean's classification score of very mild or greater. Consensus photographic score based on the agreement of at least two of the three examiners. Examiner 1 was both a clinical and a photographic examiner, Examiner s 2 and 3 are photographic examiner only.

Table 5. Level of agreement in the diagnosis of dental fluorosis between clinical score (gold standard) and consensus photographic score

Photographic	*Clinical Examination*		
Scores	*Fluorosis*	*No fluorosis*	*Totals*
Fluorosis	26 (79%)	1 (1%)	27 (24%)
No fluorosis	7 (21%)	77 (99%)	84 (76%)
Totals	33 (100%)	78 (100%)	111(100%)

Sensitivity 0.79 (79%); specificity 0.99 (99%); accuracy 0.93 (93%); positive predictive value 96%; negative predictive value 92%; likelihood ratio of positive test (+LR) 39; likelihood ratio of the negative test (-LR) 0.22.

The area under the curve (AUC) was high (0.89) when consensus photographic score was compared to clinical score. The AUC results closer to maximum value of 1 suggest good accuracy of the diagnostic test in the differentiation between fluorosis and non-fluorosis.

Discussion

The present study's analysis focused on examiner reliability on fluorosis assessment and is part of larger on-going research into Malaysian fluoridation. The key findings of this study were fluorosis prevalence was higher using clinical examinations than the photographic method and both intra- and inter-examiner reliability was good for photographic assessment.

It is difficult to compare the data from the current study, with other published studies because of the differences in the clinical examination method, photographic technique and indices used in previous studies. In the following comparisons these limitations should be borne in mind.

Findings from this study can be compared with the one previous study assessing agreement of fluorosis diagnosis between clinical against photographic methods using Dean's Index (Pretty *et al*., 2012). That study compared a specific photographic techniques (traditional digital technique versus polarized white light versus quantitative light fluorescence) using two indices (Dean's and TF Indices). These authors reported a higher fluorosis prevalence (Dean's Index) with all photographic methods than with clinically recorded scores. The difference in findings may have been due their drying of teeth prior to taking the photograph, whereas in the present study the teeth were photographed wet. By drying the teeth, the contrast between normal and abnormal enamel may be enhanced which allows a more detailed examination. In contrast, measuring fluorosis with the teeth wet may obscure some of the subtleties of fluorosis (Cochran *et al*., 2004a; Ellwood *et al*., 1994; Thylstrup and Fejerskov, 1978), but it could be argued more nearly reflects conditions of everyday life. In addition, Pretty and coworkers used two different indices in their study and whether the teeth were dried or not during the clinical examination was not clearly discussed. The differences in clinical examination method used for each index may also account for some of the differences.

When comparing the results of the present study with other studies of different indices for measuring fluorosis, the present findings were similar to some (Martins *et al*., 2009; Soto-Rojas *et al*., 2008). Soto Rojas and colleagues reported a higher prevalence of fluorosis by clinical examination (22%) than by the photographic method (18%) using TSFI criteria. The later study by Martins *et al*. (2009) reported that the prevalence of fluorosis was higher when assessed by clinical examination (49%) compared to a photographic method (37%) among 49 Brazilian children, however the fluorosis criteria used was not made clear. However, other studies reported higher fluorosis prevalence using a photographic method (Cruz-Orcutt *et al.* 2012; Wong *et al*., 2012). Cruz-Orcutt *et al.* (2012) dried teeth with the effect described earlier. While, the difference observed by Wong *et al*. (2012) may be explained by the difference in photographic method. Although the authors examined and photographed teeth in a wet condition, they used conventional photographs and not the digital photographs of the present study.

Overall, all the study examiners demonstrated good intra-examiner and inter-examiner reliability between clinical and photographic scoring. There was not much different found between simple kappa or weighted kappa analysis. The substantial to excellent agreement between diagnosis of fluorosis using the photographic method and clinical assessment is in accordance with other studies (Ellwood *et al*., 1996; Martins *et al*., 2009; Sabieha and Rock, 1998; Soto-Rojas *et al*., 2008; Wong *et al*., 2005).

This could be attributed to the standardized photographic technique employed. For example the use of a ring flash reduced shadows in the photographs. In addition, images were viewed at a standard distance from the screen and scored at the same time by all examiners under similar conditions, which aimed to reduce magnification effects and examiner bias during assessment. The viewing conditions during photographic assessment are considered one of the key factors that affect examiner agreement (Tavener *et al*., 2007). That study reported low examiner reliability (kappa value less than 0.60) among ten examiners in fluorosis assessment of 120 images when assessed remotely using different computer monitors. The resultant different lighting, contrast and brightness affected the viewing conditions. In addition, the used of photographs avoids the inherent limitations with a clinical examination such as uncooperative patient and non-uniform lighting.

Although there was good intra-examiner reliability, examiner 1 demonstrated significantly higher prevalence in clinical score than photographic score. The major difference lay in differentiating between the questionable and very mild categories of fluorosis. The possible explanation may be due to the limitations of Dean's Index that the diagnostic category for the mildest form of fluorosis is unclear, imprecise and lacks sensitivity (Clarkson, 1989; Horowitz 1986). Despite these limitations, Dean's Index has been used extensively because of the simplicity of the index. It also allows historical comparison with previous studies. Other specific fluorosis indices such as TFI and TSIF were found to be more sensitive to detecting the mildest form of fluorosis than Dean's Index (Rozier, 1994). In addition, another possible factor is that drying may have occurred during clinical examination. It is challenging to keep the teeth moist throughout the clinical examination. In the present study, the teeth were re-wetted prior to photography, which may explain the reason for the lower fluorosis score using photographs.

The findings from the present study support results from other studies that the photographic method is a valid and reliable method for assessing fluorosis. Although the study was able to suggest good reliability for photographic assessment, it had some limitations. Firstly, there was potential bias of foreknowledge of the clinical situation by Examiner 1 in photographic scoring. Effort was made to overcome this type of bias, by assessing the photographs 45 days after the clinical examination and the photographs were mixed randomly for blind scoring. Secondly, the distribution of fluorosis distribution within the studied population was based on index teeth scores among selected samples and should not be confused with population prevalence. Lastly, caution should also be taken when interpreting the reliability results, as the variation of prevalence and severity of fluorosis would affect the agreement. For example, the agreement levels will be greater if the sample examined has more people free of fluorosis and the agreement level will be lower when more categories were used in fluorosis classification. The overall distribution of fluorosis score is reported in Table 4. In this study all children examined irrespective of fluorosis status were included in determining the reliability scores.

Conclusion

The results suggest that photographic examination of fluorosis on central incisors can be recorded with good examiner reliability. The reported fluorosis prevalence was lower when using the photographic scoring method.

References

Altman, D.G. (1990): Practical statistics for medical research. London: Chapman & Hall/ CRC.

Browne, D., Whelton, H. and O'Mullane, D. (2005): Fluoride metabolism and fluorosis. *Journal of Dentistry* **33**, 177-186.

Buzalaf, M.A. and Levy, S.M. (2011): Fluoride intake of children: considerations for dental caries and dental fluorosis. *Monograph in Oral Science* **22**, 1-19.

Clarkson, J. (1989): Review of terminology, classifications, and indices of developmental defects of enamel. *Advance Dental Research* **3**, 104-109.

Cochran, J.A., Ketley, C. E., Sanches, L., Mamai-Homata, E., Oila, A.M., Arnadottir, I.B. and O'Mullane, D.M. (2004a): A standardized photographic method for evaluating enamel opacities including fluorosis. *Community Dentistry and Oral Epidemiology* **32**(sup), 19-27.

Cochran, J.A., Ketley, C.E., Arnadottir, I.B., Fernandes, B., Koletsi-Kounari, H., Oila, A.M., and O'Mullane, D.M. (2004b): A comparison of the prevalence of fluorosis in 8-year-old children from seven European study sites using a standardized methodology. *Community Dentistry and Oral Epidemiology* **32** (Sup), 28-33.

Cruz-Orcutt, N., Warren, J.J., Broffitt, B., Levy, S.M. and Weber-Gasparoni, K. (2012): Examiner reliability of fluorosis scoring: a comparison of photographic and clinical examination findings. *Journal of Public Health Dentistry* **72**, 172-175.

Dean, H.T. (1934): Classification of mottled enamel diagnosis. *Journal of the American Dental Association* **21**, 1421-1426.

Ellwood, R., O'Mullane, D., Clarkson, J. and Driscoll, W. (1994): A comparison of information recorded using the Thylstrup Fejerskov index, Tooth Surface Index of Fluorosis and Developmental Defects of Enamel index. *International Dental Journal* **44**, 628-636.

Ellwood, R.P., Cortea, D.F. and O'Mullane, D.M. (1996): A photographic study of developmental defects of enamel in Brazilian school children. *International Dental Journal* **46**, 69-75.

F.D.I Commission on Oral Health, Research and Epidemiology (1982): An epidemiological index of developmental defects of dental enamel (D.D.E Index). *International Dental Journal* **32**, 159-167.

Fejerskov, O., Thylstrup, A. and Larsen, M.J. (1977): Clinical and structural features and possible pathogenic mechanisms of dental fluorosis. *Scandinavian Journal of Dental Research* **85**, 510-534.

Horowitz, H.S., Driscoll W.S., Meyers R.J., Heifetz S.B. and Kingman A. (1984): A new method for assessing the prevalence of dental fluorosis--the Tooth Surface Index of Fluorosis. *Journal of the American Dental Association* **109**, 37-41.

Horowitz, H.S. (1986): Indexes for measuring dental fluorosis. *Journal of Public Health Dentistry* **46**, 179-183.

Landis, J.R. and Koch, G.G. (1977): The measurement of observer agreement for categorical data. *Biometrics* **33**, 159-174.

Martins, C.C., Chalub, L., Lima-Arsati, Y.B., Pordeus, I. A. and Paiva, S.M. (2009): Agreement in the diagnosis of dental fluorosis in central incisors performed by a standardized photographic method and clinical examination. *Cadernos de Saude Publica* **25**, 1017-1024.

Mascarenhas, A.K. (2000): Risk factors for dental fluorosis: a review of the recent literature. *Pediatric Dentistry* **22**, 269-277.

Nunn, J. H., Ekanayake, L., Rugg-Gunn, A.J. and Saparamadu, K. D. (1993): Assessment of enamel opacities in children in Sri Lanka and England using a photographic method. *Community Dental Health* **10,** 175-188.

Pendrys D.G. (1990): The fluorosis risk index: a method for investigating risk factors. *Journal of Public Health Dentistry*, **50**, 291-298.

Pretty, I.A., McGrady, M., Zakian, C., Ellwood, R.P., Taylor, A., Sharif, M.O. and Dye, B.A. (2012): Quantitative light fluorescence (QLF) and polarized white light (PWL) assessments of dental fluorosis in an epidemiological setting. *BMC Public Health* **12**, 366.

Rozier, R.G. (1994): Epidemiologic indices for measuring the clinical manifestations of dental fluorosis: overview and critique. *Advances in Dental Research* **8**, 39-55.

Sabieha, A.M. and Rock, W.P. (1998): A comparison of clinical and photographic scoring using the TF and modified DDE indices. *Community Dental Health* **15**, 82-87.

Soto-Rojas, A.E., Martinez-Mier, E.A., Urena-Cirett, J., Jackson, R.D. and Stookey, G. K. (2008): Development of a standardisation device for photographic assessment of dental fluorosis in field studies. *Oral Health & Preventive Dentistry* **6**, 29-36.

Stephen, K.W., Macpherson, L.M., Gilmour, W.H., Stuart, R.A. and Merrett, M.C. (2002): A blind caries and fluorosis prevalence study of school-children in naturally fluoridated and nonfluoridated townships of Morayshire, Scotland. *Community Dentistry and Oral Epidemiology* **30**, 70-79.

Tavener, J., Davies, R.M. and Ellwood, R.P. (2007): Agreement amongst examiners assessing dental fluorosis from digital photographs using the TF index. *Community Dental Health* **24**, 21-25.

Thylstrup, A. and Fejerskov, O. (1978): Clinical appearance of dental fluorosis in permanent teeth in relation to histologic changes. *Community Dentistry and Oral Epidemiology* **6**, 315-328.

Whelton, H.P., Ketley, C.E., McSweeney, F. and O'Mullane, D.M. (2004): A review of fluorosis in the European Union: prevalence, risk factors and aesthetic issues. *Community Dentistry and Oral Epidemiology* **32**(Sup), 9-18.

Whelton, H., Browne, D., Felicia, P. and Whelton, J. (2014): *E-training for Dean's Index Version 2.* Cork, Ireland: Oral Health Services Research Centre, University College Cork. http://www.fluorosisindex.com

Wong, H.M., McGrath, C., Lo, E.C. and King, N.M. (2005): Photographs as a means of assessing developmental defects of enamel. *Community Dentistry and Oral Epidemiology* **33**, 438-446.