Community Dent Oral Epidemiol 2004; 32 (Suppl. 1): 19–27
Printed in the UK. All rights reserved

COMMUNITY DENTISTRY AND ORAL EPIDEMIOLOGY

A standardized photographic method for evaluating enamel opacities including fluorosis

Judith A. Cochran[1], Clare E. Ketley[2], Leonor Sanches[3], E. Mamai-Homata[4], Anna-Maria Oila[5], Inga B. Árnadóttir[6], Cor van Loveren[7], Helen P. Whelton and Denis M. O'Mullane[1]

[1]Oral Health Services Research Centre, University College Cork, Ireland, [2]Department of Clinical Dental Sciences, University of Liverpool, Liverpool, UK, [3]Departamento de Epidemiologia, Instituto Superior De Ciências da Saúde-Sul, Monte de Caparica, Portugal, [4]Department of Preventive and Community Dentistry, School of Dentistry, University of Athens, Athens, Greece, [5]Institute of Dentistry, University of Oulu, Oulu, Finland, [6]University of Iceland, Faculty of Odontology, Reykjavik, Iceland, [7]Academic Centre for Dentistry Amsterdam, Department of Cariology and Endodontology, Amsterdam, the Netherlands

Cochran JA, Ketley CE, Sanches L, Mamai-Homata E, Oila A-M, Árnadóttir IB, van Loveren C, Whelton HP, O'Mullane DM. A standardized photographic method for evaluating enamel opacities including fluorosis. Community Dent Oral Epidemiol 2004; 32 (Suppl. 1): 19–27. © Blackwell Munksgaard, 2004

Abstract – *Objectives:* The objective of this study was to demonstrate the reproducibility of a standardized photographic technique for recording fluorosis when used by a group of epidemiologists as part of a large multicentred European study. *Methods:* Studies were first carried out to develop the equipment specification and photographic method. The author (JAC) was then trained and calibrated in this method. She was then responsible for the training and calibration of examiners from a further six European study sites. The method involved taking two transparencies of the permanent maxillary central incisors of 8-year-old children, the first after 8 s while the teeth were still wet and the second after 105 s when the teeth had been allowed to dry out naturally. Data were collected at a central location during a training/calibration exercise and subsequently, during the conduct of a large study to measure fluorosis prevalence, at the seven sites. Intra- and interexaminer reproducibility of the photographic method were measured by grading the transparencies produced by all the examiners according to the DDE and TF indices. *Results:* The time period in which the transparencies were taken was to within 4 s among the examiners. Transparencies scored according to the TF index gave a range of Kappa values of 0.45–0.66 for intraexaminer reliability and 0.32–0.55 for interexaminer reliability. When using the DDE index Kappa values ranged from 0.43 to 0.70 for intraexaminer reliability and from 0.34 to 0.69 for interexaminer reliability. *Conclusion:* The photographic method was mostly robust and reproducible when used by epidemiologists from seven European study sites.

Key words: calibration; dental fluorosis; photographic method; standardization

Dr Clare E Ketley, Department of Clinical Dental Sciences, The University of Liverpool School of Dentistry, Liverpool L69 3GN, UK
Tel: +44 151 706 5301
Fax: +44 151 706 5300
e-mail: C.E.Ketley@liverpool.ac.uk

Interest in monitoring levels of fluorosis has increased in recent years, but lack of standardized methods for recording enamel defects combined with the use of a wide variety of, largely subjective, measurement indices, has made comparison between studies very difficult. An obvious solution to this problem would be the development of a standardized system which would provide a permanent record of the condition of the enamel at any given time point. A number of authors (1–11) have reported photographic studies as a means of collecting epidemiological data on enamel opacities. The major benefit of a photographic system is the opportunity for blind scoring of colour transparencies from one or more studies by one or more examiners. However, variation in photographic technique, equipment, lens, lighting system and film quality can all annul the advantages of the photographic method. In addition, such details as lens aperture, film batch and processing are rarely reported.

It has been shown that there is a positive linear relationship between the prevalence of enamel fluorosis recorded using the Thylstrup and Fejerskov (TF) index and drying periods of 15, 45, 75 and up to 105 s (12). None of these reported studies (1–11) gives details of the enamel-drying period before taking photographs of the teeth. Thus the contribution of the drying effect to the prevalence of opacities

observed from colour transparencies in each study compared with the reported prevalence from direct observation of the teeth is unknown. None of these studies gives details of the quality of the transparencies taken, with the exception of one study (11) where it was reported that approximately 2% of transparencies were of unacceptable standard.

The position and specification of the flashgun may also affect the image of the tooth surface with regard to specular reflection and lip shadow. For close-up photography the flash may be a point source, ring flash, or linear tubes. The use of a ring flash rather than a linear tube may result in a greater surface area of the tooth being affected by specular reflection. Where the light source is fitted on the lens of the camera, altering the position of the camera angle alters the amount of reflected light, which results in specular reflection. As the angle of light source increases above horizontal the likelihood of resultant lip shadow on the tooth image in the region of the cervical margin increases and the likelihood of specular reflection on the tooth surface decreases. In mild cases of enamel fluorosis, it is usually the incisal edges which are affected (13). A limited degree of lip shadow in the cervical region of the tooth is therefore preferable to specular reflection when grading transparencies of the teeth. When using directional lighting in the 12 o'clock position, a camera flash angle of 45° has been demonstrated to minimize the effects of both specular reflection and lip shadow (14).

Aim

The aim of this study was to demonstrate the level of reproducibility of a photographic method used by a number of examiners to record enamel opacities, including fluorosis, in 8-year-old children as part of a multicentre epidemiological study involving seven European study sites (Project FLINT) (15). A secondary aim was to demonstrate that the colour transparencies taken to represent the subjects' teeth either as 'wet' or 'dry' were taken within defined time restrictions.

Materials and methods

The co-ordinating centre for the study was in Cork, Ireland (site A), and the author had sole responsibility for ensuring that all partners followed the standardized method. Throughout the text, the research groups formed at the seven sites (A–G) may be referred to as partners, epidemiologists, examiners, photographers, or dentists. The dentist (JAC) conducting the training in the photographic method is referred to as the 'Gold Standard'. The sampling procedure for the main study is described elsewhere (16).

Equipment specification and testing

The seven participating dentists each used an identical 35-mm, single-lens, reflex Minolta camera system with the following specification: Minolta 600si classic camera body on manual setting with data back and automatic advance/rewind, Minolta AF 100 mm f/2.8 macro lens magnification ratio 1 : 1, Minolta Macro flash 1200AF, Grid focusing screen, Minolta AC adapter Ac-1000 (115 V type 198–264 V for use in Europe).

A macro lens was necessary to produce a life-size image of the teeth as this is considered to be the conventional magnification ratio to use when photographing two or three teeth (17). The macro flash, although designed with four separate flash tubes, was set to use only one tube in the 12 o'clock position to reduce specular reflection. The flash unit was mounted on the end of the lens to avoid shadows on the image. It was powered from the main electricity supply, allowing an unlimited number of uses and providing a consistent 95% recharge time of approximately 10 s. Each flash unit was tested prior to the study to ensure consistency of light output and recharge time.

Kodachrome ASA 64 colour film was used as it provided optimum archival stability of the colour transparencies (18). All film was purchased from the same production batch and stored in a freezer prior to use to minimize ageing, in line with the manufacturer's recommendation.

Photographic method

Prior to the photography each child was given a new toothbrush and asked to clean their teeth without water or toothpaste. Each child was then positioned leaning against a wall with the Frankfort-maxillary plane parallel to the floor. A cheek retractor was inserted and the child was asked to close the incisors in edge-to-edge contact. Teeth were kept moist using damp cotton wool. When the photographer was ready the research assistant started the timer and at this moment the teeth were allowed to start drying out. Two timed photographs of the permanent maxillary central incisors were taken – one after 8 s while the teeth were still wet and one after 105 s when the teeth had dried out. The research assistant provided a verbal countdown from the timer for the 5 s leading up to the time of each photograph. The actual

time of each exposure was recorded both in writing and automatically on the transparency by means of the data back facility on the camera. The camera was held at a 45° angle during exposure to minimize specular reflection while still aiming to avoid lip shadow.

Dentists and their assistants from each of the seven participating European study sites were trained in this photographic method by the 'Gold Standard' examiner (JAC) in Cork. As most of the participants did not have English as their first language, a video was made in conjunction with the audio-visual department in University College Cork (19), a copy of which was available for each study centre. The 'Gold Standard' photographer/examiner had previously been trained in this method and in the use of the Developmental Defects of Enamel (DDE) and TF indices, following which she demonstrated an extremely high level of both intra- and inter- reproducibility for both indices. She also gained considerable experience in the photographic method through the conduct of preliminary development studies (12, 14).

The photographers underwent a calibration exercise following training. The seven photographers (the 'Gold Standard' and the six photographers from the other sites) photographed the permanent maxillary incisor teeth of nine 8-year-old children in one school in Cork. The transparencies were then mixed in a random fashion with nine sets of seven transparencies of the same subjects taken in a previous study to reduce the likelihood of memorizing grades. The transparencies were viewed on a light box in a darkened room (background light meter reading between 6.5 and 7.0 Electron Volts (EV)) and graded using both the DDE index and the TF index by the 'Gold Standard' examiner (JAC). No information about the subject or photographer was available at the time of grading. The grades of the transparencies of each photographer were compared with the grades for all the transparencies taken by the 'Gold Standard' photographer using Cohen's kappa value of reproducibility (20).

To ensure that skills learnt in Cork were transferable to the site appropriate to the individual examiners, each photographer was provided with two rolls of film to pilot the method. On return of the first set of exposed rolls of film to Cork, assessment and comments were made by the 'Gold Standard' examiner. The second set of rolls of film was then exposed and returned to Cork following assimilation of these comments.

During the conduct of the main study the intra-examiner reproducibility of each photographer was measured by taking a repeat for 15% of the transparencies. Additionally interexaminer reproducibility was measured by the 'Gold Standard' examiner visiting each study site to repeat 15% of the photographs. All the transparencies of the children's teeth, including intra- and interexaminer repeats, were viewed and graded by the 'Gold Standard' as described above using two different indices: the TF index (21, 22), which is an aetiological index for measuring fluorosis, and the DDE index, which is a descriptive index (23–25).

Statistics

To summarize the variability in the data for the range of time periods at which the transparencies were taken, box-and-whisker charts were plotted. To demonstrate that the photographers and the equipment used to take the transparencies could produce reliable data, an indirect approach was adopted for comparing two transparencies of the same teeth. If the grades assigned to the teeth on each transparency were the same then it could be inferred that the transparencies were close to identical and therefore that the photographers and equipment were producing reliable data.

The level of agreement between pairs of grades was determined using Cohen's kappa statistic (20).

Results

Return of the 'pilot' films revealed several problems: cervical margins of the teeth did not appear on film; the interproximal area of the central incisors was not centred; the 1 : 1 ratio was inaccurate; timing was inaccurate or clocks were not synchronized. The problems were addressed in the form of a written report specific to the needs of each photographer and the results obtained from the second pilot rolls of film appeared largely satisfactory. To obtain a measure of the 1 : 1 ratio for the main study a 1-cm scale was printed. Each photographer was then requested to photograph this scale at the beginning of each roll of film. Photographers were also instructed to tape the focusing barrel if there was a problem keeping it in a fixed position. Subsequently a 1 : 1 ratio was consistently achieved.

Calibration of the photographic method

Comparisons between the TF and DDE grades assigned to the teeth viewed from the transparencies taken by each photographer and those taken by the 'Gold Standard' photographer in the calibration

Table 1. Kappa and percentage agreement for TF and DDE grades assigned to the permanent maxillary central incisors viewed from colour transparencies taken by six photographers, and compared with those of the 'Gold Standard' photographer

Country of photographer	No. of observations	TF index % agreement and kappa	DDE Grade % agreement and kappa	DDE Extent % agreement and kappa
B	(18)	78%	100%	56%
		0.65 (Good)	1.00 (Very good)	0.33 (Fair)
C	(18)	78%	94%	89%
		0.65 (Good)	0.73 (Good)	0.79 (Good)
D	(18)	78%	94%	72%
		0.64 (Good)	0.77 (Good)	0.57 (Moderate)
E	(16)	94%	100%	94%
		0.83 (Very good)	1.00 (Very good)	0.88 (Very good)
F	(18)	89%	94%	56%
		0.83 (Very good)	0.77 (Good)	0.25 (Fair)
G	(18)	67%	100%	78%
		0.49 (Moderate)	1.00 (Very good)	0.65 (Good)

exercise are presented in Table 1. Agreements ranged from 'fair' to 'very good'. For example, in the case of the agreement on the TF index between the transparencies taken by the photographer from site C and the author, there was 78% agreement and a kappa value of 0.65. These figures would indicate that the photographic technique used by the site C team was similar to that of the 'Gold Standard'. The poorest kappa for the TF index was 0.49 ('moderate'). This was found when transparencies from site G were compared with the transparencies taken by the 'Gold Standard'. There was mostly good agreement for DDE grades and mostly moderate agreement for extent of DDE defect.

Number of photographs taken in the main study

Table 2 shows the number of children photographed and the number of children who had repeat photographs in each study site in the main study. In total, 2063 children took part, ranging from 210 in site G to 327 in site A. Between 14% (site F) and 18% (site E) of children at each study site had repeat photographs taken by the site photographer. Time and financial constraints prevented the team from conducting any repeat photographs in site G. Between 10% (site G) and 18% (sites D and F) of the children at each study site had repeat photographs taken by the 'Gold Standard' photographer.

Reproducibility of the time period for which colour transparencies were taken

To be satisfied that all seven photographers taking part in the main study were able to take transparencies at standardized drying periods box-and-whisker plots were drawn and analysed. The mass of data for the original 'wet', intraexaminer and interexaminer repeat 'wet' transparencies for each

Table 2. Number and percentage of children who had original, intra- and interexaminer repeat colour transparencies taken in each country

		Intraexaminer repeats		Interexaminer repeats	
Country	No. of children photographed (no. of transparencies taken)	No. of children photographed (no. of transparencies taken)	%	No. of children photographed (no. of transparencies taken)	%
A	327 (654)	49 (98)	15	NA[b]	–
B	319 (638)	53 (106)	17	50 (100)	16
C	314 (628)	46 (92)	15	50 (100)	16
D	292 (584)	45 (90)	15	52 (104)	18
E	299 (598)	53 (106)	18	49[c] (98)	16
F	302 (604)	42 (84)	14	53 (106)	18
G	210 (420)	0	0	20 (40)	10
Total	2063 (4126)	288 (576)		274 (548)	

[a]Figures in parentheses are the no. of transparencies taken.
[b]NA, not applicable.
[c]One child had no original photograph taken.

Table 3. Reproducibility for TF grades assigned to original and intraexaminer repeat transparencies and original and inter-examiner transparencies for each country

	TF grades assigned to original and intrarepeat transparencies				TF grades assigned to original and inter-repeat transparencies			
Country	*n*	% agreement	kappa	95% CI	*n*	% agreement	kappa	95% CI
A	196	79	0.66 (good)	0.62, 0.70	N/A	N/A	N/A	N/A
B	210	80	0.65 (good)	0.61, 0.69	200	67	0.43 (moderate)	0.38, 0.48
C	180	69	0.47 (moderate)	0.42, 0.52	200	68	0.45 (moderate)	0.40, 0.50
D	178	69	0.45 (moderate)	0.39, 0.50	207	75	0.45 (moderate)	0.39, 0.51
E	212	75	0.54 (moderate)	0.50, 0.59	188	72	0.55 (moderate)	0.51, 0.60
F	164	74	0.60 (good)	0.56, 0.65	210	70	0.53 (moderate)	0.48, 0.57
G	No data	No data	No data	No data	80	66	0.32 (fair)	0.22, 0.42

Kappa values for TF grades assigned to original transparencies and intraexaminer repeat transparencies compared with those values assigned to original transparencies and interexaminer repeat transparencies for each country.

country was in the range of 7–11 s. The vast majority of the original transparencies were judged comparable in terms of being defined as 'wet'.

The mass of original and intraexaminer repeat data for the 'dry' transparencies ranged from 104 s to 107 s, with the exception of intrarepeat transparencies from one centre, which ranged from 100 s to 111 s. Most timings for original and inter-repeat 'dry' data fell between 104 and 110 s. The vast majority of the original transparencies were judged to be comparable in terms of being defined as 'dry'.

Reproducibility of the grades assigned to original, intra- and inter-repeat colour transparencies taken in the study

As it was concluded that the vast majority of the transparencies were comparable in terms of the tooth enamel drying period, all the transparencies were used for grading teeth using the TF and DDE indices. These grades were then used to assess the reproducibility of the photographic method. Both percentage agreement and kappa values are given with 95% confidence intervals for TF grades assigned by the 'Gold Standard' examiner to both original transparencies and intraexaminer repeat transparencies, and original transparencies and interexaminer repeat transparencies (Table 3).

Using the TF index to grade original and intra-repeat transparencies the reproducibility of the photographer was 'good' in three cases and 'moderate' in three cases. The reproducibility of the photographers with respect to the 'Gold Standard' photographer as assessed by comparing grades for original and inter-repeat transparencies on the TF index was 'moderate' for five of the six photographers. The exception was for the photographer from site G whose kappa value was 'fair' and this was reflected in the confidence interval being much wider than that for the other countries. Overall there was a slightly higher level of intraexaminer reproducibility than interexaminer reproducibility (Fig. 1).

Table 4 shows percentage agreement and kappa values for DDE grades assigned by the 'Gold Standard' examiner to original transparencies and intra-repeat transparencies, and original transparencies and inter-repeat transparencies. The results showed that the reproducibility of the photographers with respect to the 'Gold Standard' photographer, as

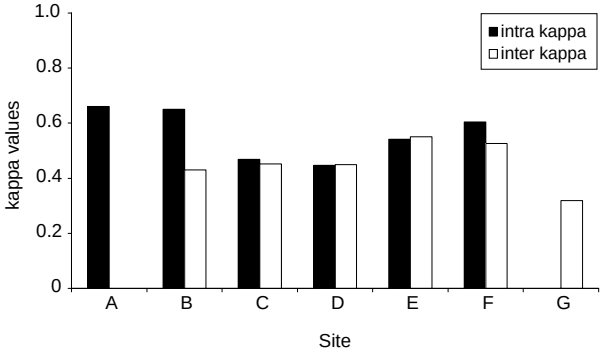

Fig. 1. Kappa values for TF grades assigned to original transparencies and intra-examiner repeat transparencies compared with those values assigned to original transparencies and inter-examiner repeat transparencies for each country.

Table 4. Reproducibility for DDE grades assigned to original and intraexaminer repeat transparencies and original and interexaminer transparencies for each country

	DDE grades assigned to original and intra repeat transparencies				DDE grades assigned to original and inter repeat transparencies			
Country	*n*	% agreement	kappa	95% CI	*n*	% agreement	kappa	95% CI
A	196	84	0.70 (good)	0.67, 0.74	N/A	N/A	N/A	N/A
B	210	83	0.70 (good)	0.67, 0.73	200	72	0.54 (moderate)	0.47, 0.56
C	180	72	0.43 (moderate)	0.37, 0.49	200	80	0.62 (good)	0.58, 0.66
D	178	78	0.59 (moderate)	0.54, 0.64	207	71	0.49 (moderate)	0.44, 0.54
E	212	74	0.58 (moderate)	0.54, 0.63	188	80	0.69 (good)	0.66, 0.73
F	164	80	0.64 (good)	0.60, 0.69	210	83	0.69 (good)	0.65, 0.72
G	No data	No data	No data	No data	80	63	0.34 (fair)	0.25, 0.43

Kappa values for DDE grades assigned to original transparencies and intraexaminer repeat transparencies compared with those values assigned to original transparencies and interexaminer repeat transparencies.

assessed by comparing grades for original and intrarepeat transparencies, was 'good' in three cases and 'moderate' in three cases. The reproducibility of the photographers with respect to the 'Gold Standard' photographer as assessed by comparing grades for original and inter-repeat transparencies was 'good' for three photographers, 'moderate' for two and 'fair' for one.

Comparing the results for agreement on the DDE index (mean kappa value 0.53, Table 4), for agreement on the TF index (mean kappa value 0.43, Table 3), the level of agreement for the DDE index would appear to be slightly higher.

Figure 2 shows the kappa values for intra- and interexaminer results for transparencies, as graded on the DDE index. In sites C, E and F there was a slightly higher level of reproducibility between the individual photographers and the 'Gold Standard' photographer than there was within the individual photographers themselves.

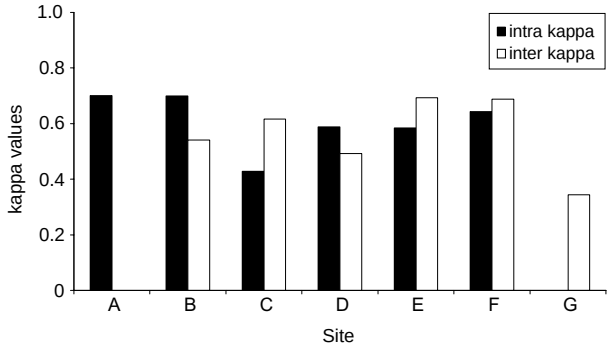

Fig. 2. Kappa values for DDE grades assigned to original transparencies and intra-examiner repeat transparencies compared with those values assigned to original transparencies and inter-examiner repeat transparencies.

Site B was the only case where the kappa value for original and inter-repeat transparencies was lower than that for original and intra-repeat transparencies and was without overlap of confidence intervals using both indices. The confidence intervals for the intra- and inter-repeat data did not overlap for sites B, C and E. The mean value for interexaminer reproducibility for five countries, with the exception of site G, was approximately 0.60 ± 0.11.

Discussion

The indices used to determine the prevalence of enamel opacities, including fluorosis, are subjective by nature and it is difficult to validate claims that the prevalence of fluorosis is increasing. The development of a reliable and reproducible photographic method for measuring enamel fluorosis has been receiving increasing attention from epidemiologists.

The present study aimed to demonstrate that a standardized photographic method could be robust. The main advantage is that if a method is standardized and shown to be robust, it will be reproducible. A reproducible photographic method enables direct comparisons of the prevalence of enamel opacities to be made for both contemporaneous studies and those conducted at two or more points in time.

During the conduct of the multicentre photographic fluorosis study 2063 children from seven

different European cities were photographed. Approximately 15% of the children photographed in each country had repeat photographs taken by the original photographer and approximately another 15% of the children photographed in each country had repeat photographs taken by the 'Gold Standard' photographer. There were two exceptions: (i) the photographer from site G lost time and film as a result of incorrect exposure of film at the start of the study; (ii) it was not necessary for the 'Gold Standard' photographer to produce her own inter-repeat transparencies. This resulted in 548 transparencies that were used to determine the level of reproducibility for the photographic method. There were also 548 records for exposure time to demonstrate that all the transparencies were taken within defined time restrictions. If the photographic method was robust, it would be possible by definition to teach the method to a number of epidemiologists and obtain reproducible data.

As there had been no previous multicentre studies cited using a photographic method to record enamel opacities it was first necessary to determine to what extent photographic variables would need to be standardized. Lack of set standards for important variables might affect the robustness of the method.

For many, the obvious question is why digital imaging was not used. There are three main advantages of using a digital imaging system. An image can be produced at much greater speed than is the case with conventional photography. Whilst this may be an advantage in some fields of medical photography it would be of no real advantage with respect to monitoring the prevalence of fluorosis. Digital imaging can be used to maintain confidentiality, which may also be an important advantage in some aspects of medical photography more so than may apply to photographing teeth. With images that are formed digitally, it is possible to make measurements of variation in density with accuracy. The disadvantages are that the equipment is much harder to use than the equipment used for conventional photography and thus operators must be very experienced. It is more difficult to obtain the correct exposure, as there is little latitude for error. The cost of a camera alone is approximately eight times more than the equipment used in this study and supporting computer hard- and software costs would be considerable. The use of digital imaging was therefore not feasible. The need for a positioning device to standardize the position of the child relative to the camera was considered but dismissed as it was considered that the minimum of equipment should be used to facilitate the ease with which dental personnel were free to move in and out of schools.

The level of agreement between pairs of grades beyond that which can be expected by chance was determined using Cohen's kappa statistic (20). Guidelines have been given for the interpretation of kappa (26): a value of less than 0.20 is considered 'poor'; 0.21–0.40 is considered 'fair'; 0.41–0.60 is 'moderate'; 0.61–0.80 is 'good'; and 0.81–1.00 is considered to be 'very good'. Landis and Koch considered a kappa value of more than 0.7 to be excellent (27). This approach to measuring the consistency of the photographic technique may give a conservative estimate because there will always be some disagreement because of the inability to grade colour transparencies with absolute consistency. If the colour transparencies are not identical the degree to which this will be reflected will be limited by three factors: the difference between the colour transparencies in terms of observed enamel opacities or enamel fluorosis; the sensitivity of the index; and the reproducibility of the examiner. It cannot be discounted that two colour transparencies taken of the same tooth may differ slightly but not to the extent that the grading is affected.

It is possible to photograph with a high level of reproducibility, but inappropriately. The photographic method requires practice and self-appraisal and, in the initial stages of training, the support of an experienced photographer. Data on the level of agreement for the transparencies from site G graded according to both indices did not reflect as high a level of agreement as was found in other countries. The point has been made that direct comparisons between kappa values should not be made, because the data sets are not the same and differences in prevalence will affect kappa. However, the kappa value between the transparencies taken by the author and the transparencies taken in site G is quite different from the kappa values for the other five countries. This difference is also reflected in the percentage agreements.

Site B was the only country for which there was consistently, across all three indices, no overlap of confidence intervals for intra- and interexaminer kappa values. A notable occurrence of specular reflection on the transparencies from this centre is likely to have been the reason for this. For four out of the five countries, excluding site G that had no intrarepeat data, both intra- and inter-repeat reproducibility of transparencies, as scored by the TF index, were similar. The reproducibility data using the DDE grades for sites C and E were slightly

unexpected, as one would normally expect to find that intra-reproducibility values were higher than inter-reproducibility values.

The average inter-reproducibility value was 0.48 ± 0.05 for the TF index, 0.6 ± 0.1 for the DDE grade index and 0.48 ± 0.07 for the DDE extent. On average the inter-reproducibility was 'moderate' to 'good' depending on the index. This point demonstrates how important it is for the photographers to be aware of what specular reflection is and how the camera angle may affect this. Specular reflection appeared to be a much more common occurrence than lip shadow. This suggests that photographers are more inclined to hold the camera relative to the teeth at a shallow angle than at a steeper angle. It is recommended that where studies seek to compare the prevalence of enamel opacities from transparencies taken by different photographers the transparencies should first be graded for the prevalence of specular reflection and lip shadow and transparencies failing to meet defined limits should be eliminated from the study (16).

The effect of drying the teeth had, as expected, a significant effect on the prevalence of opacities observed. When transparencies of 'wet' teeth were examined using the TF index, 60% were defined as 'normal' but this dropped to 31% when the teeth were viewed from transparencies of 'dry' teeth. The drying effect had a similar impact on teeth graded using the DDE index. Interestingly drying the teeth had the effect of reducing the number of opacities recorded as demarcated, and the impact was of the order of 53%. This is an important finding as the index may be used on either 'wet' or 'dry' teeth and so it is important that where comparisons are made between studies the examining conditions should be similar with respect to the drying period. In the current study it was found that it was not possible to achieve a level of accuracy of anything less than 4 s in terms of drying period.

Timing may have also been influenced by the photographer's judgement as to whether accuracy of timing was more important than ensuring that the photograph was well composed and focused. No such emphasis was conveyed during training but it became clear after the study that the photographers had different ideas about which was the most important objective. For example, the site C photographer reported that timing was more important than focusing. The 'Gold Standard' photographer believed that the primary objective was to achieve a well-composed and focused picture and that timing was secondary. These beliefs may have been reflected in the results, as there was more spread in the data from site A than for site C for both 'wet' and 'dry' photographs.

The median exposure times for transparencies representing 'wet' and 'dry' teeth were comparable within a few seconds for each country. For the original transparencies of 'wet' teeth from each country, 90% of the time periods at which the transparencies were taken fell between 7 and 12 s; a range of 5 s. Likewise, for the original 'dry' colour transparencies 90% of the data from each country fell between 103 s and 109 s; a range of 6 s. In addition to the latitude for human error in timing, the likelihood that the child may move just at the point of taking the photograph is also responsible for some of the variation in timing. In view of these factors the standardization of the time in seconds at which photographs were taken was considered to be acceptable. It was concluded that subsequent comparisons made between transparencies, for example between original and repeat transparencies, were valid with respect to the similarity of the data for enamel drying periods.

Use of the TF grades assigned to the permanent maxillary incisors, for both the original and repeat transparencies, is an indirect means of establishing the similarity between two transparencies. The basis upon which the comparisons are made depends only upon the observed severity or absence of fluorosis. No judgement is directly passed on whether the composition or exposures of the transparencies are the same. Where there is agreement on the observed severity or absence of fluorosis, it is assumed that factors such as composition and exposure of the transparencies are similar enough for the level of discretion required when grading using either the TF or the DDE index.

As a result of intraexaminer error associated with the TF index, the comparison of TF grades between original and repeat transparencies will tend to underestimate the similarity between the transparencies. It is important to be aware that the 'acceptable' value of kappa depends on the circumstances for which it is being calculated, and the circumstances under and for which it was calculated.

Conclusion

A standardized photographic method was developed and used by seven epidemiologists across Europe. They were able to take colour transparencies of teeth that were both reproducible when measured against transparencies that they had taken themselves and against transparencies taken by the

'Gold Standard' photographer to a moderate or good level as defined by Landis and Koch. The photographic method was mostly shown to be robust, but the success of the project was dependent upon a high standard of training and the dedication of the photographers.

Acknowledgements

Thanks go to Angela Murphy, Eileen MacSweeney and Helen Whelton (Ireland); Liisa Seppa and Erja Komminaho (Finland); Thessaly Athanassouli and Haroula Koletsi-Kounari (Greece); Peter Holbrook and Halla Sigurjóns (Iceland); Annemarie van Leeuwen (the Netherlands); Eduardo Barros Fernandes (Portugal); Roger Ellwood, Mike Lennon and Julia West (UK) and all the participating school children.

References

1. Ellwood RP, O'Mullane DM. Dental enamel opacities in three groups with varying levels of fluoride in their drinking water. Caries Res 1995;29:137–42.
2. Ellwood RP, Côrtes DF, O'Mullane DM. A photographic study of developmental defects of enamel in Brazilian school children. Int Dent J 1996a; 46:69–75.
3. Ellwood RP, Hawew RM, Worthington HV, Blinkhorn AS. Developmental enamel defects and extrinsic tooth stain in Libyan school children. Community Dent Oral Epidemiol 1996b;24:419–20.
4. Holt RD, Morris CE, Winter GB, Downer MC. Enamel opacities and dental caries in children who used a low fluoride toothpaste between 2 and 5 years of age. Int Dent J 1994;44:331–41.
5. Ishi T, Suckling G. The severity of dental fluorosis in children exposed to water with a high fluoride content for various periods of time. J Dent Res 1991; 70:952–6.
6. Levine RS, Beal JF, Fleming CM. A photographically recorded assessment of enamel hypoplasia in fluoridated and non-fluoridated areas in England. Br Dent J 1989;166:249–52.
7. Nunn JH, Ekanayake L, Rugg-Gunn AJ, Saparamadu KDG. Assessment of enamel opacities in children in Sri Lanka and England using a photographic method. Community Dent Health 1993;10:175–88.
8. Nunn JH, Murray JJ, Reynolds P, Tabari D, Breckon J. The prevalence of developmental defects of enamel in 15–16-year-old children residing in three districts (natural fluoride, adjusted fluoride, low fluoride) in the north east of England. Community Dent Health 1992;9:235–47.
9. Sabieha AM, Rock WP. A comparison of clinical and photographic scoring using the TF and modified DDE Indices. Community Dent Health 1998;15:82–7.
10. Stephen KW, McCall DR, Gilmour WH. Incisor enamel mottling prevalence in child cohorts which had or had not taken fluoride supplements from 0 to 12 years of age. Proc Finn Dent Soc 1991;87:595–605.
11. Fleming CM, Beal JF, Levine RS. Photographic technique for the recording of children's teeth for signs of enamel mottling. J Audiovisual Med Med 1989;12:16–18.
12. Cochran JA. A standardised photographic method for recording enamel opacities including enamel fluorosis. PhD Thesis. University of Cork 1999.
13. Richards A, Kragstrup J, Josephsen K, Fejerskov O. Dental fluorosis developed in post-secretory enamel. J Dent Res 1986;65:1406–9.
14. Cochran JA, O'Mullane DM. Photographing enamel opacities. A question of standardisation. J Dent Res 1997;76:429 (abstract 3328).
15. O'Mullane DM, Cochran JA, Whelton HP. Fluoride ingestion from toothpaste: background to European Union-funded multicentre project. Community Dent Oral Epidemiol 2004;32 (Suppl. 1):5–8.
16. Cochran JA, Ketley CE, Arnadóttir IB, Fernandes B, Koletsi-Kounari H, Oila A-M, et al. A comparison of the prevalence of fluorosis in 8-year-old children from seven European study sites using a standardized methodology. Community Dent Oral Epidemiol 2004; 32 (Suppl. 1):28–33.
17. Goodlin RM. The Complete Guide to Dental Photography. Montreal: Michael Publishing Co.; 1987.
18. Kodak. Professional and Amateur Kodak Colour Films: the Difference Between Them. Kodak Publication No E-6 (H). Hemel Hempstead: Kodak Ltd.
19. Cochran JA. Epidemiological Measurement of Fluorosis. Cork: Audio-Visual Department. University College; 1997.
20. Cohen J. A coefficient of agreement for nominal scales. Educational Psychol Measurement 1960;20: 37–46.
21. Thylstrup A, Fejerskov O. Clinical appearance of dental fluorosis in permanent teeth in relation to histological changes. Community Dent Oral Epidemiol 1978;6:315–28.
22. Fejerskov O, Baelum V, Manji F, Moller IJ. Dental Fluorosis;a Handbook for Health Workers. Copenhagen: Munksgaard; 1988.
23. FDI Commission on Oral Health Research and Epidemiology. Technical Report Number 15. An epidemiological index of developmental defects of dental enamel (DDE Index). Int Dent J 1982;32: 159–67.
24. FDI. An epidemiological index of Developmental Defects of Enamel (DDE Index). Int Dent J 1992;42: 411–26.
25. Clarkson J, O'Mullane D. A modified DDE Index for use in epidemiological studies of enamel defects. J Dent Res 1989;68:445–50.
26. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall; 1991, p. 416–18.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33: 159–74.