Original Article

Measurement of White Lesions Surrounding Orthodontic Brackets: Captured Slides Vs Digital Camera Images

Philip Edward Benson, PhD[a]; Anwar Ali Shah, PhD[b]; Derrick Robert Willmot, PhD[c]

Abstract: Digital images provide an efficient means of processing and storing clinical images and are increasingly being used in orthodontic practice. However, for research purposes it must be shown that measurements from digital images are as reliable as those from captured slides, particularly in the area of postorthodontic demineralization. The aim of this in vitro study was to compare the assessment of demineralized white lesions surrounding orthodontic brackets using images produced from captured slides with those from a digital camera. Thirty teeth with orthodontic brackets and a systematic arrangement of artificially induced demineralization were used. Standardized images of the teeth were taken in 35-mm slide format and using a digital camera. The slides were scanned and saved as digital images. All the images were numbered, recoded in a random order, and assessed by a second investigator. The image was examined for the presence or absence of demineralization. This was compared with the actual demineralization pattern, and the positive and negative predictive values were calculated. Twenty images from each group were randomly chosen and duplicated for an assessment of measurement error. The positive and negative predictive values were better from the digital images (0.92 and 0.81) than from the captured slides (0.88 and 0.74). The percentage agreements for the repeat assessments of the same slide were similar (96% for the captured slides and 93% for the digital images). Measurements of enamel demineralization using images from a digital camera are as accurate and reproducible as images captured from a photographic slide. (*Angle Orthod* 2005;75:226–230.)

Key Words: Orthodontics; Image analysis; Demineralization

INTRODUCTION

Clinical photographs are a convenient method of recording the appearance of dental enamel before and after orthodontic treatment. They can be used both as a clinical record and for assessment of iatrogenic damage.[1,2] It has been shown that postorthodontic white spot lesions can be measured reliably from photographic slides converted into digital images.[2] There is an increasing use of direct digital imaging in medicine[3–5] and dentistry[6] because of the advantages of lower cost and less potential error in image processing and storage. Research into the prevention of postorthodontic white spot lesions requires measurements from clinical images, but it is not certain that the resolution of the image from a digital camera is sufficient to allow accurate and reliable measurement of demineralized white spots.

The aim of this study is to compare the assessment and measurement of artificially induced demineralized lesions from a captured 35 mm slide with an image from a digital camera.

MATERIALS AND METHODS

Thirty teeth with orthodontic brackets and a systematic arrangement of artificially induced demineralization were used in this in vitro study. The preparation of the teeth has been described previously.[7]

Production of the images

Standardized photographs of the teeth were taken using a Nikon F301 camera body (Nikon UK Ltd, Kingston upon Thames, UK) with a 90-mm Elicar macro lens (Novoflex,

[a] Senior Lecturer/Honorary Consultant in Orthodontics, Oral Health and Development, School of Clinical Dentistry, University of Sheffield, Sheffield, UK.

[b] Lecturer, Oral Health and Development, School of Clinical Dentistry, University of Sheffield, Sheffield, UK.

[c] Head of Department/Consultant in Orthodontics, Oral Health and Development, School of Clinical Dentistry, University of Sheffield, Sheffield, UK.

Corresponding author: Philip Edward Benson, PhD, Oral Health and Development, School of Clinical Dentistry, University of Sheffield, Claremont Crescent, Sheffield, S Yorks S10 2TA, UK (e-mail: p.benson@sheffield.ac.uk).

Accepted: April 2004. Submitted: March 2004.

Memmingen, Germany) set to a magnification of 1:1. The camera was set to manual with an aperture of f22 and shutter speed of 1/60 of a second. The film used was Kodak Kodachrome 64 ASA (Kodak Limited, Hemel Hempstead, UK) and this was processed by Kodak. A Cokin flash (Cokin SA: Silic 457-94593, Rungis Cedex, France) set to manual ½ was used. The Cokin flash has three tubes. These were arranged with two tubes on either side and one above the tooth to reduce reflections from the flash.[1] A sighting jig was placed in the orthodontic bracket slot to improve the replication of bracket positioning.[7] The slides were scanned using an Epson Expression 1680 Pro scanner (EPSON (UK) Ltd, Hempstead, Herts, UK) to resolution 1600 dpi, 24-bit color (normal), and saved as TIFF images.

Digital images of the teeth were taken using a FinePix S1 Pro camera (FUJIFILM Electronic Imaging Ltd, Hemel Hempstead, UK) with a 105 mm/2.8 AF Micro Nikkor lens and Nikon SB29 Speedlight flash. The camera was set to manual with an aperture of f22 and a shutter speed of 1/125 of a second. The image size was set at 1440 × 960 pixels, image quality fine, and ISO sensitivity 400. The flash was set with the tubes to the left and right of the tooth. Images were captured on 32-MB CompactFlash™ card (Lexar Media Inc, Fremont, Calif) as Joint Photographic Experts Group (JPEG) images.

The photographs and digital images of each tooth were repeated after one week. Twenty images of captured slides and 20 digital images were randomly chosen and duplicated to determine the within-assessor method reproducibility. All the images were numbered and recoded in a random order, placed on CD-ROM, and given to a second person (Dr Shah) who had not been involved in the production of the images and was not informed about the nature of the investigation.

Analysis of the images

The four corners of the bracket were examined and a subjective assessment made as to the presence or absence of a demineralized white spot. When it was determined that demineralization was present, an analysis was carried out using the image analysis software Image Pro-Plus (Version 4.5.0.29, Media Cybernetics, Silver Spring, Md), as described previously.[7] Each image was calibrated using a known distance between the tie-wings of the bracket. An area of interest was drawn by hand around the demineralized region to record the area of demineralization (mm^2) and mean gray level. In addition, an area of normal enamel judged to be free of demineralization was measured and the mean gray level recorded.

The luminance intensity proportionality (LI%) was used to standardize the baseline measurement.[8] This defines the white lesion gray level as being a percentage of sound enamel gray level (black being zero and bright white being 255), according to the following formula:

TABLE 1. Sensitivity, specificity, PPV and NPV from the subjective assessment for the presence or absence of demineralization of all the demineralized areas (14, seven, and three days, N = 240 assessments) and for just the 14 and seven-day demineralization (N = 200 assessments)[a]

	14-, seven- and three-d Demineralization		14- and seven-d Demineralization	
	Captured Slides	Digital Images	Captured Slides	Digital Images
Sensitivity	0.47	0.63	0.67	0.73
Specificity	0.96	0.97	0.96	0.97
PPV	0.88	0.92	0.85	0.88
NPV	0.74	0.81	0.89	0.91

[a] NPV indicates negative predictive values; PPV, positive predictive values.

$$LI\% = \left[\left(\frac{\text{Mean grey level of white lesion}}{\text{Mean grey level of sound enamel}}\right) - 1\right] \times 100$$

Statistics

The two methods of recording the teeth were assessed in the following way.

- To assess the validity of the scoring from the photographic and digital images, the subjective evaluations regarding the presence or absence of demineralization by the judge were used. This was compared with the actual demineralization pattern carried out on each tooth. The sensitivity, specificity, positive predictive value, and negative predictive value were calculated[9] for the 240 assessments made with each technique (30 teeth; two images of each tooth; four sides to each bracket).
- The reproducibility of the two methods was compared by calculating the intraclass correlation coefficients[10] for the readings from the 20 repeated images. Sites were used only when there was agreement that demineralization was present on the first and second image. The values for demineralized lesion area (mm^2), mean lesion gray level, mean sound enamel gray level, and LI percent were compared.
- Agreement between the readings from the two methods was assessed using the limits of agreement[11] and a paired *t*-test for systematic error.[12] The measurements of demineralized lesion area (mm^2), mean lesion gray level, mean normal gray level, and LI percent from the two methods were compared. Values from the sites were used only when there was a reading from both the captured slide and the digital image. The readings for the first and second image were averaged for each technique.

RESULTS

The sensitivity, specificity, positive predictive value, and negative predictive value for the captured slides and digital images are shown in Table 1. The results for the digital images were better than those for the captured slides be-

TABLE 2. The intraclass correlation coefficients for the repeat readings of demineralized lesion area (mm²), mean lesion gray level, mean sound gray level, and luminance intensity (LI) percent from the repeated images for both captured slide and digital images (N = 18 sites for area, lesion, and LI% and 11 sites for sound)

	Captured Slide	Digital Image
Area	0.96	0.92
Lesion	0.94	0.99
Sound	0.87	0.96
LI%	0.81	0.89

TABLE 3. Limits of agreement comparing the readings of the captured slide and digital image, including demineralized lesion area (mm²), mean lesion gray level, mean sound gray level, and luminance intensity (LI) percent (N = 27 sites for area, lesion, and LI% and 17 for sound)

	Limits of Agreement		
	SD	Lower	Upper
Area	0.64	−1.57	1.00
Lesion	13.77	−18.46	36.61
Sound	12.93	−13.76	37.97
LI%	9.96	−22.79	17.07

TABLE 4. Summary measures of the paired differences between captured slide and digital image with 95% confidence intervals of the differences and results of a paired *t*-test between the readings of demineralized lesion area (mm²), mean lesion gray level, mean sound gray level, and luminance intensity (LI) percent (N = 27 sites for area, lesion, and LI% and 17 for sound)

	Paired Differences		95% CI of the Difference		
	Mean	SD	Lower	Upper	*P*
Area	−0.29	0.64	−0.54	−0.03	.029
Lesion	9.08	13.77	3.63	14.52	.002
Sound	12.11	12.93	5.46	18.76	.001
LI%	−2.86	9.96	−6.80	1.08	.148

cause 63% of sites with demineralization were correctly identified from the digital image compared with only 47% from the captured slide. Most of the unidentified sites were those subjected to the demineralizing environment for three days. Only 20% of the three-day sites with demineralization were correctly diagnosed from the captured slide compared with 50% from the digital image. When these values were excluded from the calculation, the sensitivity improved significantly to 0.67 for the captured slide (Table 1). The specificity or the ability to correctly identify a site when it was not demineralized was good for both the captured slide (0.96) and the digital image (0.97).

Some consider the positive and negative predictive values to be more relevant than the sensitivity and specificity because they examine the probability that the diagnosis is correct.[13] The positive predictive values were good for both the captured slides (0.88) and the digital images (0.92). Thus, when it was determined that demineralization was present, it was indeed present for 88% of sites from the captured slide and 92% of sites from the digital image. The negative predictive values were lower for both the captured slide (0.74) and the digital image (0.81), but both values improved significantly when the sites that were demineralized for only three days were excluded (0.89 and 0.91, respectively).

The agreement for the assessments on the 20 repeated images was very good. In 77 of the 80 assessments (20 images with four assessments equivalent to the four sides of each bracket), or 96%, there was agreement as to the presence or absence of demineralization for the captured slides. The equivalent figure for the digital images was 74 of 80 or 93% agreement. Agreement between the two techniques was less good with 86% agreement as to the presence or absence of demineralization between the captured slide and the digital image.

There were 18 sites with readings from the original image and the repeat image for demineralized area and mean lesion gray scale and 11 sites for the mean gray scale of sound enamel and hence LI%. The discrepancy between the two occurred because there were several teeth with more than one demineralized area measured and only one reading from sound enamel was taken for each tooth. The intraclass correlation coefficients for these repeated readings are shown in Table 2. The results showed excellent reliability except for the LI percent, which showed moderate agreement for the captured slides and substantial agreement for the digital images.[14]

There were 28 sites with readings from the captured slide and digital image for demineralized area and mean lesion gray scale and 19 sites for the mean gray scale of sound enamel and hence LI%. The discrepancy between the two was again because there were more demineralized sites than teeth and only one reading from sound enamel was taken for each tooth. The limits of agreement and paired differences between the captured slides and digital images were large (Tables 3 and 4), suggesting poor agreement between the two techniques. There were systematic differences in the readings for lesion area ($P = .029$), mean lesion gray scale ($P = .002$), and mean sound enamel gray scale ($P = .001$). The measured size of the lesion area was larger (mean 0.29 mm) from the digital image compared with the captured slide. Both the mean lesion gray level (mean difference 9.08) and the mean sound enamel gray levels (mean difference 12.11) were more toward the black end of the gray level spectrum for the digital images. This suggests that these areas were significantly darker when analyzed from the digital image than from the captured slide.

DISCUSSION

This study was designed to compare the validity, reproducibility, and agreement between measurements of artificial demineralization taken from captured slides and images from a digital camera. It has shown that white spots surrounding orthodontic brackets can be assessed as accurately and reproducibly from images of teeth produced using a digital camera as from digital images captured from a clinical slide. There was also some evidence that it was possible to detect demineralization at an earlier stage from the digital images because 50% of areas subjected to three days of demineralization were detected from the digital image compared with only 20% from the captured slide. There was, however, poor agreement between the measurements of mean lesion gray scale, sound enamel, and LI% from the captured photograph and digital camera image.

Similar results for validity and reproducibility have been found in other studies. In ophthalmology, Henricsson et al[3] found that there was good to excellent agreement for grading diabetic retinopathy between color slides and digital color images. There was exact agreement between grades in 82% of cases, which compares with 86% in this study, although they were using a 10-point scale unlike our dichotomous scale. van Leeuwen et al[5] found that stereodigital images were as good as stereo 35-mm color transparencies for the grading of age-related maculopathy (ARM). Their assessments ranged from a two-grade assessment for neurovascular age-related macular degeneration (AMD) with a 99% agreement to an eight-grade stage of ARM assessment with a 63% agreement.

Studies using continuous data have also found good agreement between digital and conventional images. Musadiq et al[4] compared paired fundus fluorescein images from patients acquired with a digital system and photographic film. There was a mean difference of 0.32 mm between measurements of the images, with limits of agreement −0.88 to 0.24 mm, which is comparable with the results of the area measurements in this study.

The reproducibility of the luminance percent was poorer than for the other outcomes. The luminance percent is used to overcome the problems of differences in light reflection between images. The gray level of sound enamel on one image will be different from the gray level of sound enamel on another image, even if they were taken with the same camera and the same technique. By using a percentage of the normal gray level for the white lesion, an attempt is made to resolve the differences in light intensity with the actual levels. This technique is frequently used to resolve the problem of variation in baseline readings between individuals.[8]

The level of agreement between the two techniques was poor, particularly for the measurements of mean lesion gray scale, mean gray scale of sound enamel, and LI%. There was a small increase in the measurement of the lesion area from the digital image. This might be due to the more peripheral areas with reduced demineralization being assessed compared with the captured slide. The mean gray level for the lesion and normal enamel were significantly more toward the dark end of the spectrum than for the captured slides. This probably reflects the differences in set-up, flash, and scanning of the two images. However, it does suggest that, although each technique on its own is reproducible, it would not be satisfactory to combine data from captured slides and digital images in any clinical study because of the lack of agreement between the two techniques.

The quality of a digital image will depend on the type of camera used. There are three levels of digital camera in the market. The fully automatic ‘point and shoot’ cameras are small, light, and convenient but produce lower-resolution images and are less flexible when taking pictures in the mouth. The Prosumer cameras have more advanced features such as through the lens focusing, but if high-quality intraoral images are to be achieved, then the ideal cameras are professional cameras based on the single lens reflex design, such as the FinePix S1 Pro camera used in this study. These cameras allow full flexibility regarding exposure controls and accessories, such as the use of a macro lens for close-up work and multipoint flash units to capture important differences in enamel color. These cameras also produce high-resolution images. The digital images for this study were 1440 pixels wide and 960 pixels high or 1.38 megapixels, which compared favorably with the scanned images, which ranged from 1313 × 1642 pixels to 984 × 1387 pixels (2.16 to 1.36 megapixels).

The images from the digital camera were saved as JPEG format, which is described as a ‘‘lossy’’ file format because some of the image information is lost when it is decompressed. The advantage of this format is that compression of the file and consequently smaller file formats are easier to save and transmit electronically. However, the format was designed for compressing images that will be seen by the human eye because it takes advantage of the limitations of the human eye in detecting small color changes less accurately than changes in brightness. This study involved the detection of demineralization with the human eye so this should not have been a problem. For future studies that might involve automatic analysis by a machine, a different file format may be more appropriate.

Another useful property of the JPEG file format is that the amount of information lost can be altered by varying the amount of compression. Wenzel et al[15] showed that JPEG images could be compressed to ratios of 12:1 before accuracy and image quality were significantly affected for caries diagnosis. Eraso et al[16] showed that the diagnostic quality of digital radiographs used to assess periapical lesions was significantly affected when they were saved as JPEG images with a high compression ratio of 48:1 or 64:1. The images for this study were saved with a compression ratio of 8:1, which has not adversely affected diagnostic accuracy.

CONCLUSIONS

Measurements of enamel demineralization using images from a digital camera are as accurate and reproducible as images captured from a photographic slide.

There was poor agreement between the captured slide and digital camera images in the assessment of mean lesion gray scale, the mean gray scale of sound enamel, and LI% measurements of the lesion. Therefore, it would be unwise to combine the results if using the two methods of recording enamel white spots.

REFERENCES

1. Benson PE, Pender N, Higham SM. Enamel demineralization assessed by computerised image analysis of clinical photographs. *J Dent.* 2000;28:319–326.
2. Willmot DR, Benson PE, Pender N, Brook AH. Reproducibility of quantitative measurement of white enamel demineralization by image analysis. *Caries Res.* 2000;34:175–181.
3. Henricsson M, Karlsson C, Ekholm L, Kaikkonen P, Sellman A, Steffert E, Tyrberg M. Color slides or digital photography in diabetes screening—a comparison. *Acta Ophthalmol Scand.* 2000; 78:164–168.
4. Musadiq M, Patsoura E, Hughes S, Yang YC. Measurements of linear dimensions on fundus photographs: comparison between photographic film and digital systems. *Eye.* 2003;17:619–622.
5. van Leeuwen R, Chakravarthy U, Vingerling JR, Brussee C, Hooghart AJ, Mulder PG, de Jong PT. Grading of age-related maculopathy for epidemiological studies: is digital imaging as good as 35 mm film? *Ophthalmology*. 2003;110:1540–1544.
6. Sandler PJ, Murray AM, Bearn D. Digital records in orthodontics. *Dent Update.* 2002;29:18–24.
7. Benson PE, Pender N, Higham SM. Quantifying enamel demineralization from teeth with orthodontic brackets—a comparison of two methods. Part 1: repeatability and agreement. *Eur J Orthod.* 2003;25:149–158.
8. Pocock SJ. *Clinical trials: A practical Approach.* Chichester, NY: Wiley; 1983, 190.
9. Benson PE, Pender N, Higham SM. Quantifying enamel demineralization from teeth with orthodontic brackets—a comparison of two methods. Part 2: validity. *Eur J Orthod.* 2003;25:159–165.
10. Fleiss JL. *Design and Analysis of Clinical Experiments.* New York, NY: Wiley; 1986, 1–32.
11. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135–160.
12. Houston WJ. The analysis of errors in orthodontic measurements. *Am J Orthod.* 1983;83:382–390.
13. Altman DG. *Practical Statistics for Medical Research.* 1st ed. London, UK: Chapman & Hall; 1991, 410–411.
14. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med.* 1987;6:441–448.
15. Wenzel A, Gotfredsen E, Borg E, Grondahl HG. Impact of lossy image compression on accuracy of caries detection in digital images taken with a storage phosphor system. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 1996;81:351–355.
16. Eraso FE, Analoui M, Watson AB, Rebeschini R. Impact of lossy compression on diagnostic accuracy of radiographs for periapical lesions. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2002;93:621–625.