Journal of Dentistry 74 (2018) S34–S41

Contents lists available at ScienceDirect

Journal of Dentistry

journal homepage: www.elsevier.com/locate/jdent

Automatic detection and classification of dental fluorosis in vivo using white light and fluorescence imaging

Zhao Liu, Michaela Goodwin, Roger P. Ellwood, Iain A. Pretty*, Michael McGrady

The Dental Health Unit, School of Dentistry, The University of Manchester, M15 6SE, UK

ARTICLE INFO

Keywords:
Fluorosis
Detection
TF index
Automation

ABSTRACT

Objectives: To assess a novel method of automatic fluorosis detection and classification from white light and fluorescent images.

Methods: Dental images from 1,729 children living in two fluoridated and two non-fluoridated UK cities were utilised. A novel detection and classification algorithm was applied to each image and TF scores were obtained using thresholding criteria. These were compared to clinical reference standard images. Comparisons between reference and automated assessments were undertaken to record correct and incorrect classifications and the ability of the system to separate the fluoridated and non-fluoridated populations.

Results: The automated system performed well and was able to differentiate the two populations ($P < 0.0001$) to the same degree as the reference standard. When using the highest score from the clinical assessment the agreement between automated and clinical assessments was 0.56 (Kappa SE = 0.0160, $p < 0.0001$).

Conclusions: Assessment of dental fluorosis is typically undertaken by clinical examiners in epidemiological studies. The training and calibration of such examiners is complex and time consuming and the assessments are subject to bias – frequently because of the examiner's awareness of the water fluoridation status of subjects. The use of remote scoring using photographs has been advocated but still requires trained examiners. This study has shown that image-processing methodologies applied to white light and fluorescent images could automatically score fluorosis and statistically separate fluoridated and non-fluoridated areas. The system requires further refinement to manage confounding factors such as the presence of non-fluoride opacities and tooth stain.

1. Introduction

The detection and diagnosis of enamel fluorosis is fundamental in determining the risks and benefit of the use of fluoride in the prevention of dental caries. In populations with low to moderate exposure to fluoride the clinical appearance of fluorosis presents primarily as diffuse areas of hypomineralisation on the enamel surface. At higher levels of fluoride exposure, the hypomineralisation may be more severe, with post-eruptive pitting and staining appearing on the fluorotic enamel [1].

Traditional methods of assessing fluorosis rely upon the clinical assessment of teeth using either descriptive or aetiological epidemiological indices [2]. A trained clinical examiner relates the interpretation of the clinical presentation to pre-determined criteria for aetiology and severity. Clinical indices involve subjective assessment and raise concerns relating to potential forms of bias. The major form of bias is examiner blinding – a clinical examiner may be aware of the levels of exposure to fluoride in the location where examinations are performed, e.g. the level of fluoride in community drinking water. Examiner reliability (intra-examiner reliability and inter-examiner reliability, where more than one examiner is employed) further compound potential forms of bias. The use of clinical indices to score fluorosis has been criticized in the literature [3–5], particularly in relation to examiner blinding and the subjective nature of the indices used. The York Review and the Medical Research Council report called for the evidence base to be improved, to reduce bias and seek more objective means of measuring fluorosis.

Attempts to address examiner blinding have included transporting participants out of area for clinical examinations, which can be impractical [6,7]. The use of remote scoring of clinical photographs has been successfully used to minimize bias and also to minimize confounding factors such as specular reflection on the tooth surface [8]. Clinical photographs can offer significant advantages over clinical scoring, such as the ability to archive high quality images to be scored by multiple examiners. However, as in the case of direct clinical scoring, remote scoring of images is still prone to the effects of a

* Corresponding author.
E-mail address: Iain.A.Pretty@manchester.ac.uk (I.A. Pretty).

https://doi.org/10.1016/j.jdent.2018.04.021
Received 27 February 2018; Accepted 24 April 2018

subjective index, with variability in examiners' reliability and the effects of thresholding [9,10] i.e. differences between individuals in the application of index criteria, particularly at low levels of fluorosis severity.

The development of Quantitative Light-Induced Fluorescence (QLF) and fluorescent imaging in the field of caries detection has been well documented in the literature [11]. Similarities in the optical behavior of early carious lesions and fluorotic opacities of enamel under fluorescent imaging provided an opportunity to develop the objective assessment and quantification of enamel fluorosis [12]. The major challenge in the development of software algorithms was the difference in appearance between carious and fluorotic opacities (i.e. defined vs. diffuse, respectively).

The earliest work used a software algorithm using MATLAB Version 6.0 (R13, Mathworks, N.Y., USA) to analyse bitmap images obtained with fluorescent imaging [12]. Metrics were derived from QLF analysis relating to area of fluorescence, change in fluorescence relative to sound enamel (Δ F) and an assessment of the extent of fluorescence change, area x Δ F (Δ Q). These were calculated by the use of a "blur" technique whereby each point on the image was replaced by the average fluorescence value of the surrounding pixels. Altering the area of the sample would affect the "blur" effect. Subtracting the blurred image form the original leaving areas considered fluorosis. Analyses of the QLF metrics and comparable clinical scoring could only produce (albeit favourable) associations between the datasets. Statistical tests were problematic owing to the categorical nature of the clinical index (Thylstrup & Fejerskov) and the continuous data produced for the QLF metrics. This could affect the interpretation of the outcomes, which were associations between QLF metrics and clinical scores, not levels of agreement.

The next development in the objective assessment of fluorosis was to test a different software algorithm with QLF. The "blur" technique provided a potential means of objectively quantifying fluorosis, but the ability to discriminate between differing levels of fluorosis in a population was questioned. The "blur" technique was compared to a new convex hull stain algorithm in a population with differing levels of fluoride exposure and presentations of fluorosis severity [13]. As in the earlier study, the QLF imaging was able to demonstrate associations with the clinical scoring. The convex hull algorithm appeared to show stronger associations between the QLF metrics and clinical scores, than the "blur" technique across a range of clinical presentations of fluorosis. The technique utilising QLF and the convex hull algorithm was subsequently tested in an epidemiological setting in a cross-sectional study exploring the effect of social deprivation on dental caries in a fluoridated and non-fluoridated population [14,15]. The results of this study were able to demonstrate associations between the QLF metric Δ Q and clinical scores (Kendall's Tau = 0.332), with the ability to discriminate between fluoridated and non-fluoridated populations (Man Whitney U Test for each of the QLF metrics vs. clinical scores; $p > 0.0001$). A ROC curve of sensitivity vs. specificity demonstrated an excellent level of accuracy (AUC = 0.9164).

Within the data for all the studies using QLF imaging, there were significant outliers, particularly at the lower levels of fluorosis severity. These were explained by the presence of confounding factors, such as extrinsic stain, caries and developmental defects of enamel of non-fluorotic origin. The rationale for this was based upon the effects of each of the confounders on the fluorescent signal obtained from the tooth surface. The consequence of this was there was still a requirement for a diagnosis of fluorosis from the part of a clinician to address potential type I error, or false positives for fluorosis in the presence of confounding factors.

A more recent development has been the use of a dual camera system, which captures a QLF image and a polarised white light image of teeth [16,17]. The dual camera system enabled a white light image of the teeth to be scored remotely using clinical indices, which would be dimensionally identical to the QLF image used for objective analysis. This not only improved the efficiency of image capture, by reducing the examination time for participants and the need for additional equipment and resource to capture separate images for clinical scoring, it also addressed potential issues surrounding the use of photographic images. The dual camera system has the ability of controlling the light settings for the white light images. The use of conventional camera systems limits the ability to control ambient light conditions and the effect of controlling for specular reflection can result in foreshortening of images.

Despite the advantages of remote scoring of clinical images and QLF image analysis of reducing examiner bias and the ability to quantify fluorosis, a number of issues remain. The QLF imaging techniques required the use of a region of interest tool to draw "masks" of the teeth prior to analysis. This process was a manual and time-consuming task, particularly in the case of studies with large numbers of participants. This was a major constraint of the imaging technique and could offset the potential efficiencies of improved image capture and more effective use of resource during the clinical phase of a study.

The aim of this project was to develop an improved method of producing image masks via an automated process by more effectively employing the data obtained from the polarised white and QLF images obtained from the dual-camera system.

A novel software algorithm would then be used to classify the observations from the images to attribute a score for fluorosis based on the criteria described in the Thystrup & Fejerskov index for fluorosis.

2. Materials and methods

2.1. Imaging system

A white light and fluorescence imaging equipment, developed at Dental Health Unit, University of Manchester, applies a cross-polarised light and auto-fluorescence technique for the acquisition of both white light and fluorescence images of teeth. This imaging device comprises a high-resolution 3 CCD colour camera (HV-F31 Hitachi Kokusai Electric Inc.) fitted with a 25 mm focal length lens (TF25DA–8 B Fujinon Corporation), offering an effective field of view of 37.5 mm × 28 mm and a depth of field of 10 mm, which are able to capture images of maxillary incisors and canines.

Two LED ring illuminators are located in front of the camera. The outer ring consists of 60 white LEDs with an emission band from 450 nm to 625 nm (B5-430-JD Roithner Laser- Technik GmbH). The illumination is cross-polarised by means of two linear polarisers, one placed in front of the camera lens (SKR FIL POL-LIN/25,5, Stemmer Imaging, Ltd.) and the other on top of the white LEDs (45668, Edmund Optics). The inner ring, having 60 near-UV LEDs centered at 405 nm (B5-437-CVD, Roithner LaserTechnik GmbH), provides the excitation light for green fluorescence of dental tissues. The long-wavelength tail from the near-UV LEDs is removed using a blue glass filter to circumvent overlapping between excitation and emission spectra. The emission light is filtered by a 515 nm long-pass yellow filter (45069, Edmund Optics).

A custom-built geometry stabilizer, comprising an adjustable head and chin rests, is used to stabilize the participant in a reproducible manner and minimize motion artifacts. The imaging device is mounted onto the stabilizer, allowing the camera to be moved and focused.

The camera and illuminators are controlled by custom-written software. Camera settings, such as gain, shutter speed, image pixel resolution etc., are pre-configured for a study design, which is locked unchangeable throughout the capture of study images. Camera white balance is calibrated before each study. Altering the output of a current controller connected to the LEDs controls the brightness of the white and near-UV light. The light achieves ideal brightness when the average pixel intensity of a standard gray card reached a pre-defined value with a small tolerance of ± 0.05. This ensures images captured at different time points are comparable. During image acquisition, white light and