The in vitro and in vivo reproducibility of a video-based digital imaging system for tooth colour measurement

Wen Luo⁎, Mojgan Naeeni, Suzanne Platten, Jinfang Wang, Jianing N. Sun, Stephen Westland, Andrew Joiner

Unilever Oral Care, Quarry Road East, Bebington, Wirral, CH63 3JW, UK
Unilever Oral Care, 66 Linxin Road, Changning District, Shanghai, 200335, China
University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK

ARTICLE INFO
Keywords: Colour
Tooth whiteness
Repeatability
Reproducibility
Clinical

ABSTRACT
Objectives: To assess the robustness of a new custom built video-based digital imaging system (VDIS) for measuring tooth colour and whiteness under in vitro and in vivo conditions.
Methods: The VDIS imaging system was developed for tooth colour measurement and evaluated in vitro and in vivo. The in vitro validation used extracted human teeth (HT, n = 14) stored in water and VITA Classical shade guide tabs (SG, n = 16). These were measured by the VDIS at baseline, 5 min, 2 h, 1 week and 2 weeks to evaluate the system repeatability. For in vivo validation, adult volunteers (male/female, n = 34) with two natural, unrestored central incisors had their teeth imaged using the VDIS at baseline, 5 min and 2 h (3 images each) by two different operators to evaluate time and operator effects. Between taking individual images, subjects moved from the imaging-frame to assess the effect of re-positioning on reproducibility. From the in vitro and in vivo images, the average tooth RGB values were obtained, and the CIELAB values and a tooth whiteness index WIO value were calculated. Repeatability and reproducibility of VDIS imaging system was assessed using appropriate repeated measurement analysis techniques and ANOVA.
Results: The measurement variations in vitro were between 1 and 2 units of ΔWIO and the average colour differences were less than 1 ΔE*ab unit. For the in vivo study, analysis of the CIELAB parameters and WIO showed that subject variability accounted for between 82 and 99% of the observed variability in the measurement process. The operator variability was less than 0.5% and the overall measurement error was found to be only 0.3% for WIO. Across assessment times the variability was less than 0.5%.
Conclusions: The dental imaging system V-DIS was shown to be a highly reproducible means for tooth colour and whiteness measurement.
Clinical significance: Digital imaging based techniques give a highly reproducible approach to measuring tooth colour.

1. Introduction
Tooth colour is important to patients and consumers who wish to enhance their smile and also to professionals who want to match tooth colour for aesthetic restorations and whitening procedures [1]. The colour of teeth is influenced by a combined effect of their intrinsic and extrinsic colourations [2,3], and is frequently quantified by visual assessment using a commercial tooth shade guide, or more objectively, by colour measuring instruments [1,4]. There are a number of instruments that have been used for measuring tooth colour in vitro and in vivo, including colorimeters, spectrophotometers, spectroradiometers and digital cameras [1]. Colorimeters and spectrophotometers have been shown to be reliable, have good repeatability and are accurate for colour matching [5,6]. However, since they are contact-measurement devices, measurement errors may occur due to factors such as the curvature of the tooth surface, light loss caused by tooth translucency [7,8], ambient light [9] and fogging of the optical lens during in vivo measurement [10].

Non-contact colour measurement systems, such as spectroradiometers and digital cameras, which use external light sources and do not need to directly attach apertures onto the tooth surface [11,12], may minimise the systematic error due to translucency and surface curvature [13]. From comparison studies between digital imaging and contact-measurement methods, both spectrophotometric and digital
image methods presented sufficient and validated objective evaluation of tooth bleaching efficacy [14]. In another study, it was found that digital camera imaging is reliable in tooth colour quantification, whereas spectrophotometry (colorimetry) gave inaccurate absolute values for tooth colours but gave the same ranking order as the digital-imaging method [13]. A meta-analysis of tooth whitening studies over a 4-year time frame confirmed the suitability of the approach and reliability of digital image analysis for long-term tooth whitening studies [15]. In addition, digital imaging gives further advantages by providing a permanent database of images that can be analysed and re-investigated at a later date; it is relatively quick and simple in terms of training and operation, and does not require a clinician [16].

Dental-imaging systems for tooth colour measurement usually consist of a digital camera and a light source as the two key elements. Commercial single-lens reflex (SLR) cameras [17] and industrial cameras (e.g. 3CCD cameras) [18,19] have been used as the image-acquisition devices. Dual daylight lamps [17,18], halogen lamps with UV fluorescent tubes [13,20,21] and ring light sources [8,19,22] have been used as the illuminant for taking tooth images in vivo. The captured images are usually analysed by converting the camera RGB values (device-dependent colour space) into CIE XYZ or CIELAB values (device-independent colour spaces) for tooth colour measurement. The three-dimensional colour coordinates are transformed into a single scale whiteness index in some tooth whitening studies [23–25], e.g. the tooth whiteness index (WHO) that was proposed based on the CIE whiteness index specifically for quantifying tooth whiteness perception [26].

The objectives of this paper is to evaluate the reproducibility of a new custom built video-based digital imaging system (VDIS) for measuring tooth colour and whiteness under in vitro and in vivo conditions.

2. Materials and methods

2.1. System development

A video-based digital imaging system (VDIS) has been developed for measuring tooth colour in vitro and in vivo. The key elements of the hardware are a digital video camera, a polarised and diffused white LED light source and a custom-built system frame. A digital camera (QImaging, Canada) provides high-speed live video during measurements and can capture still images. The camera has a cooling system to help to maintain a constant operating temperature and minimise thermal noise. A ring light (CCS Inc, Japan) is mounted on the camera lens and a diffusion filter (CCS Inc, Japan) is attached to the light source to provide diffused uniform illumination. Two polarising filters (CCS Inc, Japan) are placed, one in front of the camera lens and one in front of the light source, to provide cross-polarisation for excluding specular reflection in the teeth images. The system is connected via a USB connector to a laptop computer (Dell Inc., USA) from where the camera and imaging procedures are operated. A custom-built system frame was made to hold the camera/lighting set with adjustable distance to the teeth of the subject. A subject chin holder and a forehead bar were made to hold the subject’s teeth of the subject. A subject chin holder and a forehead bar were made to hold the subject’s head, and a white ceramic tile (Mt. Baker Research L.L.C., USA) is attached to the chin holder to enable monitoring of the lighting variation. The VDIS system is designed to disassemble into easily transportable pieces, and has been fully engineered to meet the requirements of the European Union Declaration of Conformity for safety under the Laboratory Directives, with fully Conformité Européenne (CE) safety marking.

The image analysis software was written in Matlab (MathWorks Inc., USA). The core algorithm of the software is the camera characterisation model for predicting CIE XYZ values (and CIELAB values) from the camera RGB values. A polynomial regression model was used for this conversion [27]. A Digitizer colourchecker chart (VeriVide, UK) was used as the standard reference to build the model. It contains 240 patches in a 12 cm x 20 cm grid. The colours on this chart give a good coverage of colour space which allows the characterisation model of the camera to suitably predict any colours inside of this range. The CIE XYZ values of each colour patch under D65 illuminant and 2° standard observer provided with the chart were considered as the ‘true’ values. The transform between the camera RGB and XYZ values can be expressed as:

\[ X = AV \]  

where \( X \) represents the XYZ matrix, \( A \) is the transform matrix and \( V \) is the RGB matrix. For different order polynomials, the transform matrix \( A \) is a different size. Consider the 2nd-order polynomial model as an example, the polynomial transform equations are extended as below.

\[
X = a_1 R + a_2 G + a_3 B + a_4 RG + a_5 RB + a_6 GB + a_7 R^2 + a_8 G^2 + a_9 B^2 \\
Y = a_{10} R + a_{11} G + a_{12} B + a_{13} RG + a_{14} RB + a_{15} GB + a_{16} R^2 + a_{17} G^2 + a_{18} B^2 \\
Z = a_{19} R + a_{20} G + a_{21} B + a_{22} RG + a_{23} RB + a_{24} GB + a_{25} R^2 + a_{26} G^2 + a_{27} B^2
\]

(2)

The best-fit regression should minimise the sum of residual square error, then the matrix \( A \) can be calculated by Eq. (3), \( V \) is the transpose of the matrix \( V \).

\[ A = VX'(VV')^{-1} \]  

(3)

In general, different types of camera have different colour rendering characteristics, so that the polynomial transform suitable for one camera may not fit other cameras. Several orders of polynomial should be tested to find the best polynomial transform matrix for a certain camera [27]. In this study, three polynomial transforms (1st-order, 2nd-order and 3rd-order) were tested to find the best mapping between RGB values and CIE XYZ values. Considering the overall colour-rendering ability, the 2nd-order polynomial regression based on the colourchecker chart was chosen for the characterisation model of the V-DIS. Then XYZ values were converted into CIELAB values by Eq. (4), and tooth whiteness index (WHO) values were calculated by Eq. (5) [26].

\[
L^* = 116(Y/Y_n)^{1/3} - 16 \\
a^* = 500[f(X/X_n) - f(Y/Y_n)] \\
b^* = 200[f(Y/Y_n) - f(Z/Z_n)]
\]

(4)

where \( f(Y/Y_n) = (Y/Y_n)^{1/3} \) for \( Y/Y_n > 0.008856 \), otherwise \( f(Y/Y_n) = 7.787(Y/Y_n) + 16/116 \). \( f(X/X_n) \) and \( f(Z/Z_n) \) are similarly defined. \( X_n \), \( Y_n \), \( Z_n \) are the tristimulus values of a perfect white diffuser.

\[ WO = Y + 1075.012(x_n - x) + 145.516(y_n - y) \]  

(5)

where \((x, y)\) and \((x_n, y_n)\) are the chromaticity coordinates of the sample and the reference white respectively.

A graphic-user-interface (GUI) was developed to allow operators to analyse the captured images, which implemented the camera characterisation model developed specifically for the V-DIS as described above. The procedure of using the post image analysis is: 1) loading an image; 2) selecting a tooth area as the region of interest (ROI); 3) the software calculates the colour values of the ROI and displays the RGB, \( L^*a^*b^* \) and the tooth whiteness index values on the interface, and 4) exporting the data into an excel file.

2.2. In vitro assessment

Precision characterises the degree of mutual agreement or repeatability among a series of individual measurements, values, or results, which means “repeatable, reliable, getting the same measurement each time” [28]. Repeatability is considered as an essential property of VDIS due to the main application of the system in tooth whitening studies to assess colour changes of the same set of teeth over time.

An in vitro validation was conducted to test the repeatability of the
system over time with two sets of samples that represent tooth colour: extracted human teeth (HT, n = 14) and Vita Classical shade guide tabs (SG, n = 16). The two sets of samples were measured by the VDIS at baseline, 5 mins, 2 h, 1 week and 2 weeks. The extracted human teeth, obtained for research purposes, according to Human Tissue Act procedures and with informed consent, were mounted in acrylic resin blocks by embedding the roots into cold-cure acrylic resin (Simplex Rapid, Kemdent, Wiltshire, UK). They were stored in deionised water during the entire period to keep the specimens fully hydrated, since dehydration of enamel can cause obvious colour changes [29]. The Vita guide tabs were measured dry at all time points.

From each collected image, by using the post-image analysis software, the average R, G, B values of the whole tooth specimen were obtained. The mean CIELAB and the WIO values were then calculated.

The colour difference $\Delta E_{ab}$, which has been widely used, was calculated between colours of the baseline and each of the following time points. The CIEDE2000 total colour difference, which has been introduced by CIE for correcting the non-uniformity of the CIELAB colour space for small colour differences, was also calculated [30]. This new formula ($\Delta E_{00}$) has been recommended for tooth colour evaluation since it performed closer to human visual responses than $\Delta E_{ab}$ [31,32].

### 2.3. In vivo validation

Precision analysis was also carried out in vivo in order to determine the viability of the VDIS system using human volunteers and different operators. The protocol, information sheet and informed consent for this human study were reviewed and approved by an independent Unilever Research and Development Ethics Committee. Adult male and female subjects (aged 18–65) from the Wirral area, UK, were invited to participate in this study. All subjects had to be in good general health to be considered suitable. All subjects had an oral examination and were required to have healthy oral soft and hard tissues and two normally aligned natural upper central incisors, free from restorations visible from the labial surface. Images of the central incisors were taken at baseline, 5 mins, 120 mins (3 images each) by two different operators to evaluate time and operator effects. Subjects were requested not to drink or eat between the three measurements (Fig. 1). The order in which the operators imaged each subject was randomised.

To collect an image, the subject was given a sterile plastic cheek retractor and protective eye goggles to wear, then placed their chin on the VDIS chin rest and forehead against the forehead rest. The operator monitored the RGB readings from the white tile in the live video to ensure the lighting intensity was stable, and one digital image of the subject’s teeth was captured after the position of the teeth was kept still in the centre of the video. The first image taken of each subject served as a guide to enable identical positioning of the teeth for subsequent imaging.

Between taking individual images, subjects moved from the imaging-frame to assess the effect of re-positioning on reproducibility. From the images, by using the post image analysis software, the average R, G, B values of the upper central incisors of each image were obtained. The mean CIELAB and the WIO values were then calculated.

### 2.4. Statistical analysis

The in vitro data was investigated using descriptive statistical analysis and paired t-tests to compare consistency and variability of all colour parameters and differences at different time points for teeth and shade guide tabs.

The in vivo data was analysed using ANOVA to calculate variance components and standard deviations. In the in vivo study the factors which could potentially effect VDIS include subjects, operator and time. All analyses were conducted using JMP Pro 11 statistical software (SAS Institute Inc., Cary, NC, USA).

### 3. Results

#### 3.1. In vitro assessment

Mean and standard errors of the colour parameters (CIELAB values, the whiteness index WIO and the two colour differences, $\Delta E_{ab}$ and $\Delta E_{00}$) were calculated for each time period and the results are shown in Table 1. Similar mean values over time with very low variability were obtained for all values. The mean $\Delta E_{ab}$ values between baseline and different time points for the human teeth was 0.62 or lower and for the shade guide tabs was 0.77 or lower, the corresponding $\Delta E_{00}$ values were 0.52 or lower and 0.71 or lower. For each of the colour values, the pairwise comparisons did not show any significant differences ($p \geq 0.05$) between time points for both groups of human teeth and shade guide tabs.

#### 3.2. In vivo assessment

The study included 33 adult volunteers. The mean and standard errors for the CIELAB and WIO values at different time points and for different operators are shown in Table 2. The CIELAB and a tooth whiteness index WIO were measured three times by two operators at three different time points. Individual source of variation in this data...
Table 1
Mean and standard errors of the in vitro evaluation of the V-DIS.

<table>
<thead>
<tr>
<th>Indices</th>
<th>Group</th>
<th>Baseline(S.E.)</th>
<th>5 min(S.E.)</th>
<th>2 h(S.E.)</th>
<th>1 Week(S.E.)</th>
<th>2 Weeks(S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L*</td>
<td>HT</td>
<td>65.69(0.78)</td>
<td>65.45(0.81)</td>
<td>65.95(0.79)</td>
<td>66.13(0.80)</td>
<td>65.18(0.82)</td>
</tr>
<tr>
<td>SG</td>
<td>67.90(0.90)</td>
<td>67.66(0.86)</td>
<td>67.61(0.85)</td>
<td>68.15(0.92)</td>
<td>67.20(0.83)</td>
<td></td>
</tr>
<tr>
<td>a*</td>
<td>HT</td>
<td>5.11(0.29)</td>
<td>5.39(0.30)</td>
<td>5.33(0.30)</td>
<td>5.28(0.34)</td>
<td>5.35(0.33)</td>
</tr>
<tr>
<td>SG</td>
<td>3.53(0.25)</td>
<td>4.04(0.23)</td>
<td>4.04(0.22)</td>
<td>3.99(0.24)</td>
<td>4.13(0.25)</td>
<td></td>
</tr>
<tr>
<td>b*</td>
<td>HT</td>
<td>21.32(0.63)</td>
<td>21.28(0.65)</td>
<td>21.09(0.65)</td>
<td>21.16(0.65)</td>
<td>21.06(0.63)</td>
</tr>
<tr>
<td>SG</td>
<td>18.57(0.71)</td>
<td>18.30(0.73)</td>
<td>18.29(0.72)</td>
<td>18.19(0.70)</td>
<td>18.02(0.69)</td>
<td></td>
</tr>
<tr>
<td>WIO</td>
<td>HT</td>
<td>−37.14(2.92)</td>
<td>−38.04(3.01)</td>
<td>−36.32(2.91)</td>
<td>−36.09(2.94)</td>
<td>−37.99(3.00)</td>
</tr>
<tr>
<td>SG</td>
<td>−21.73(3.66)</td>
<td>−22.44(3.60)</td>
<td>−22.49(3.55)</td>
<td>−20.95(3.42)</td>
<td>−22.70(3.72)</td>
<td></td>
</tr>
<tr>
<td>ΔE*ab</td>
<td>HT</td>
<td>0.54(0.07)</td>
<td>0.47(0.06)</td>
<td>0.50(0.06)</td>
<td>0.62(0.06)</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>0.73(0.08)</td>
<td>0.77(0.08)</td>
<td>0.64(0.07)</td>
<td>0.55(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔE&lt;sub&gt;00&lt;/sub&gt;</td>
<td>HT</td>
<td>0.34(0.06)</td>
<td>0.27(0.04)</td>
<td>0.42(0.06)</td>
<td>0.52(0.05)</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>0.65(0.07)</td>
<td>0.66(0.07)</td>
<td>0.63(0.07)</td>
<td>0.71(0.06)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Mean colour indices for anterior teeth measured in vivo with the VDIS.

<table>
<thead>
<tr>
<th>Time Operator</th>
<th>L*</th>
<th>a*</th>
<th>b*</th>
<th>WIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1</td>
<td>70.14(0.54)</td>
<td>9.25(0.15)</td>
<td>18.82(0.31)</td>
<td>−28.07(1.57)</td>
</tr>
<tr>
<td></td>
<td>[61.95–78.30]</td>
<td>[7.12–11.11]</td>
<td>[15.8–21.89]</td>
<td>[−47.97–−14.91]</td>
</tr>
<tr>
<td>0 2</td>
<td>70.11(0.52)</td>
<td>9.22(0.14)</td>
<td>18.64(0.29)</td>
<td>−27.63(1.62)</td>
</tr>
<tr>
<td></td>
<td>[63.01–77.23]</td>
<td>[7.48–10.82]</td>
<td>[15.05–18.62]</td>
<td>[−46.94–−3.72]</td>
</tr>
<tr>
<td>5 1</td>
<td>70.63(0.52)</td>
<td>9.41(0.14)</td>
<td>18.81(0.3)</td>
<td>−27.29(1.65)</td>
</tr>
<tr>
<td></td>
<td>[64.38–75.84]</td>
<td>[7.75–11.08]</td>
<td>[15.05–22.88]</td>
<td>[−45.98–−5.13]</td>
</tr>
<tr>
<td>5 2</td>
<td>70.11(0.57)</td>
<td>9.26(0.15)</td>
<td>18.60(0.32)</td>
<td>−27.53(1.61)</td>
</tr>
<tr>
<td></td>
<td>[64.42–75.84]</td>
<td>[7.51–11.08]</td>
<td>[15.11–22.88]</td>
<td>[−45.79–−5.13]</td>
</tr>
<tr>
<td>120 1</td>
<td>68.98(0.56)</td>
<td>9.41(0.13)</td>
<td>18.70(0.3)</td>
<td>−28.41(1.62)</td>
</tr>
<tr>
<td></td>
<td>[62.97–76.33]</td>
<td>[8.15–11.18]</td>
<td>[15.14–22.01]</td>
<td>[−46.84–−3.74]</td>
</tr>
<tr>
<td>120 2</td>
<td>69.86(0.52)</td>
<td>9.38(0.13)</td>
<td>18.61(0.3)</td>
<td>−28.36(1.54)</td>
</tr>
<tr>
<td></td>
<td>[62.56–75.80]</td>
<td>[7.9–14.97]</td>
<td>[11.17–21.77]</td>
<td>[−46.6–−4.42]</td>
</tr>
</tbody>
</table>

Table 3
Variance component analysis of anterior teeth measured in vivo with the VDIS.

<table>
<thead>
<tr>
<th>Component</th>
<th>L*</th>
<th>a*</th>
<th>b*</th>
<th>WIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operator</td>
<td>0.006</td>
<td>0.062</td>
<td>0.001</td>
<td>0.155</td>
</tr>
<tr>
<td>Time</td>
<td>0.018</td>
<td>0.180</td>
<td>0.005</td>
<td>0.792</td>
</tr>
<tr>
<td>Operator*Time</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Subject</td>
<td>8.404</td>
<td>82.400</td>
<td>0.588</td>
<td>85.400</td>
</tr>
<tr>
<td>Subject*Operator</td>
<td>0.089</td>
<td>0.869</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Time*Subject</td>
<td>0.183</td>
<td>1.800</td>
<td>0.015</td>
<td>2.100</td>
</tr>
<tr>
<td>Operator<em>Time</em>Subject</td>
<td>0.493</td>
<td>4.800</td>
<td>0.026</td>
<td>3.800</td>
</tr>
<tr>
<td>Within</td>
<td>1.010</td>
<td>9.900</td>
<td>0.0529</td>
<td>7.700</td>
</tr>
<tr>
<td>Total</td>
<td>10.204</td>
<td>100.000</td>
<td>0.688</td>
<td>100.000</td>
</tr>
</tbody>
</table>
variability will clearly be due to the natural variation in tooth colours expected between different people. Indeed, there is reported to be a large range of tooth colours, in terms of L*a*b* values, even from study populations from the same country [1,7]. The within variability of the three repeats for subjects at all time points was found to contribute between 0.3-9.9% of the observed variability. If the V-DIS system is used in the future to measure potential product effects the inter subject variability will be accounted for in power and sample size calculations together with the design of experiment.

The operator variability was less than 0.4% indicating that operator error is negligible. This low value will be due to a number of factors including the operators being fully trained in the system and a number of control measures in place such as repositioning of subjects’ head and white tile calibration. The variability over time was less than 0.8%. This provides strong evidence of the ability of the measurement system to produce the same tooth colour values over time.

5. Conclusion

The dental imaging system VDIS was shown to be a reproducible and highly robust means of measuring tooth colour over time.

Conflict of interest statement

Wen Luo, Mojgan Naeni, Suzanne Platten, Jinfang Wang, Jianing Sun and Andrew Joiner are employees of Unilever.

References
