Diagnostic performance of the universal visual scoring system (UniViSS) on occlusal surfaces

Jan Kühnisch · Katharina Bücher · Volkmar Henschel · Angela Albrecht · Franklin García-Godoy · Ulrich Mansmann · Reinhard Hickel · Roswitha Heinrich-Weltzien

Received: 26 January 2009 / Accepted: 9 February 2010 / Published online: 11 March 2010
© Springer-Verlag 2010

Abstract Aim of this investigation was to evaluate the validity and reproducibility of the universal visual scoring systems (UniViSS) on occlusal surfaces in vitro. The validity study included 65 third molars. Following to reach a UniViSS consensus diagnosis for each surface, all teeth were histologically prepared and evaluated according to the newly developed caries-extension-index (CE-index). The reproducibility study consisted of 149 molars. These teeth were examined by two dentists and four students without any calibration training. In result, the CE-index provided the exact caries depth for each UniViSS criterion, which could further be associated with a distinct preventive or treatment strategy. The mean intra-/inter-examiner weighted Kappa values amounted to 0.685/0.551 (UniViSS/severity) and 0.628/0.542 (UniViSS/discoloration). The cumulative Logit-model underlined that UniViSS/severity criteria were more frequently reproduced in comparison to the UniViSS/discoloration criteria. In conclusion, this study on the diagnostic performance of UniViSS showed encouraging results and give valuable hints for future studies.

Keywords Dental caries · Visual examination · Validity · Reproducibility · Reliability · Caries detection · Caries diagnosis

Introduction

Bearing in mind the high frequency of non-cavitated caries lesions in industrialised nations [1] and the fact that adjunct diagnostic methods, e.g., electrical resistance measurements, fibre-optical transillumination, quantitative light-induced fluorescence and laser fluorescence measurements, do not perform as satisfactorily on such lesions as was hoped [2–6], it became evident to improve visual caries detection and diagnostic methods. Recently introduced methods—criteria by Ekstrand et al. [7, 8] and Nyvad et al. [9], the International Detection and Assessment System [ICDAS, 10] and the Lesion Activity Assessment [11]—included non-cavitated caries lesions, but classify the caries process with only a few criteria. However, due to the fact that the clinical appearance of carious lesions—especially on occlusal surfaces—is complex, a limited set of criteria seems to be of importance for the dental epidemiologist [12, 13] but is unlikely to describe the appearance of
of different findings, both examiners discussed the discordant results and reached an agreement.

**Histological validation** Prior to the histological preparation, a colour photograph of each surface was taken to assist the histological examination later on. After separation of the root from the coronal part, each crown was embedded in cold-polymerising methacrylate (Kallocryl, Speiko, Münster, Germany) and the teeth were catalogued. Each crown was sectioned in buccal-lingual direction into slices of 500 µm thickness with a 200-µm microtome saw (Mikrotrennm MT 1-78-03, Hofer, Switzerland) to find the maximum caries extension of the lesion. In this context, enamel lesions were defined histological as opacity and dentine caries lesions were linked with yellow and/or brown discolorations. To assess the caries extension more precisely, all slices were examined under a light microscope at 16-fold magnification (Stemi SV11 stereomicroscope, Zeiss, Oberkochen, Germany). The slice with the greatest caries extension for each specimen was identified, digitally photographed, and stored for further analyses. Following this, the caries extension was quantified.

**Material and methods**

**Validity study**

Sample size A sample of 65 sound and mostly non-cavitated third molars was selected from a pool of teeth extracted for surgical or orthodontic reasons. Molars with sealants, fillings, cavitations, approximal and/or buccal-lingual caries lesions, and developmental disorders were excluded from this study. After gross debris was removed, teeth were carefully cleaned. In order to prevent bacterial growth, all teeth were stored in separate containers with physiological saline containing 0.02% sodium azide. The used material was part of a report before [14].

Determination of the UniViSS consensus diagnosis All teeth were examined visually using dental magnifying glasses (twofold), the illumination of the dental unit light and compressed air. The (visual) inspection according to UniViSS [14] was carried out under the following principles: For caries lesions detectable in the fissure pattern, the severity (first signs, established lesion, microcavity or dentine exposure) and the discoloration (white, white-brown or brown) were assessed. As the activity assessment—the third step of UniViSS—has to be understood as clinical diagnosis; the present in vitro study design did not include this. Two dentists performed the visual inspection independently (J.K., K.B.). All diagnoses were counterchecked 1 week later to form a consensus diagnosis for each surface. In case of different findings, both examiners discussed the discordant results and reached an agreement.

Calculation of the CE-index For calculation of the CE-index, first a differentiation was made between histological sound slices (score 0), demineralisations in the enamel (base value=0.x/y) or dentine (base value=1.x/y) and lesions reaching the pulp (base value=2). If an enamel or dentine lesion was present, the corresponding base value has to be combined with the ratio of the caries extension (x) and the overall enamel or enamel/dentine thickness (y) as a second step. This effectively means that the caries extension needs to be measured as the distance between the outer enamel surface or the enamel–dentine junction and the deepest demineralisation point towards the pulp (x), which has to be divided by the enamel or dentine thickness (y). Therefore, the CE-index has to be understood as an index of two components: the base value and the percentage of the x/y ratio (Fig. 2). The CE-index ranges from 0.0 to 2.0.

In detail, the enamel/dentine thickness, as well as the caries extension, was measured with the ImageJ software (National Institutes of Health, USA, downloadable at http://rsb.info.nih.gov) by two blinded dentists (J.K., K.B.). All slices were reassessed 1 week later and a final decision for the caries extension was made for each specimen. In case of different findings, both examiners discussed their discordant measures to reach an agreement.

**Statistical analysis** The data were analysed with Excel 2003 (Microsoft Corporation, Redmond, WA, USA) and SPSS 14.0 (SPSS Inc., Chicago, IL, USA) in order to cross tabulate the findings as well as to calculate mean values (mean), standard deviations (SD), minimum value (min) and the maximum
### Universal Visual Scoring System for pits and fissures (UniViSS occlusal)

#### First step: Lesion Detection & Severity Assessment

<table>
<thead>
<tr>
<th>First visible signs of a caries lesion</th>
<th>Established caries lesion</th>
<th>Microcavity and/or localised enamel breakdown</th>
<th>Dentin exposure</th>
<th>Large cavity</th>
<th>Pulp exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score F</td>
<td>Score E</td>
<td>Score M</td>
<td>Score D</td>
<td>Score L</td>
<td>Score P</td>
</tr>
</tbody>
</table>

#### Second step: Discoloration Assessment

<table>
<thead>
<tr>
<th>Lesion Appearance</th>
<th>Description</th>
<th>Corresponding Visual Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound surface</td>
<td>(Score 0) No cavitations or discolorations are detectable.</td>
<td><img src="example1.png" alt="Example" /></td>
</tr>
<tr>
<td>White (Score 1)</td>
<td>First white spot lesions are unobtrusive and will be detected after prolonged air drying (~5 seconds) in the deepest parts of pits and fissure.</td>
<td><img src="example2.png" alt="Example" /></td>
</tr>
<tr>
<td>White-brown (Score 2)</td>
<td>White-(caramel-)brown lesions will be detected with prolonged air drying (~5 seconds) in the deepest parts of pits and fissure. Brown areas are in most cases visible without air-drying.</td>
<td><img src="example3.png" alt="Example" /></td>
</tr>
<tr>
<td>(Dark) Brown (Score 3)</td>
<td>Small brown spot lesions will be detected even without air drying. This score includes hair-lined brown fissures and/or small brown superficial discolorations. (Lesion ~0.5mm/CPI probe)</td>
<td><img src="example4.png" alt="Example" /></td>
</tr>
<tr>
<td>Greyish translucency (Score 4)</td>
<td>Greyish translucencies are detectable in advanced stage only. In the deepest part of pits and fissures is a greyish translucency as sign of an undermining dentin lesion detectable. Additionally, other discolorations are visible.</td>
<td><img src="example5.png" alt="Example" /></td>
</tr>
</tbody>
</table>

#### Notes
- First white spot lesions are visible even without prolonged air drying. This score indicates hair-lined brown fissures and/or small superficial discolorations. (Lesion ~0.5mm/CPI probe)
- White-(caramel-)brown lesions are detectable without prolonged air drying. Air drying improves the assessment of the surface texture (smooth – rough).
- Greyish translucencies are only detectable on lesion without extensive cavitations.

![Fig. 1](image.png)

Fig. 1 Criteria of the universal visual scoring system for pits and fissures (UniViSS occlusal)
values (max) of the CE-index for each UniViSS score. A CE-index of 0 is associated with sound surfaces and values between 0.01 and 1.0 are corresponding to enamel caries lesions. Dentine caries lesions will be determined by a CE-index between 1.01 and 1.99. A pulpal involvement is related to a CE-index of 2.00. The overall validity of UniViSS can be expressed by the calculation of the sensitivity (SE), the specificity (SP) and the area under the ROC curves (AUC). These results were published elsewhere [14].

Intra- and inter-examiner reproducibility study

Sample size and examiners A separate sample of 149 sound and non-cavitated third molars was selected from a pool of teeth extracted for surgical or orthodontic reasons. The inclusion criteria, specimen preparation and storage followed the same principles as mentioned above. For testing the reproducibility the inaugurator of UniViSS (D1) as well as six additional examiners took part in this study. One of these dentists had >6 years of clinical experience and practise (D2) and the other, only some months (D3). Four of the investigators were clinically low-experienced students in their fifth year of studies (S1 to S4).

Diagnostic procedure Prior to the beginning of this study all examiners were introduced to the study protocol and the use of UniViSS. The theoretical and practical training consisted of a 30-min hands-on for all examiners which included basic information about the diagnostic principles of non-tactile visual examination, the usage of the CPI probe as measuring instrument, the help of standardised illumination of each specimen with the dental operation light and the necessity of careful air drying. Extensive and detailed calibration training was not performed. The visual decision process for each occlusal surface according to UniViSS included the following evaluation steps: In case of a detectable caries process (1) the most progressed severity stage (first signs, established lesion, microcavity and dentine exposure) as well as (2) the corresponding discoloration (white, white-brown, brown and greyish) was registered for each specimen. The third (clinical) UniViSS step of an activity assessment was not part of this in vitro study protocol. All examiners were encouraged to form their surface-related diagnosis within a time interval of 30 s. Furthermore, a coloured UniViSS chart with typical examples for each score (Fig. 1) was distributed to all examiners. Each evaluation cycle was repeated after a minimum interval of 2 weeks to safeguard the blindness of each investigator between the measurement cycles. Two investigators collected a third series of measuring data (D1 and S1). Hence, all participating examiners obtained a total of 4,768 diagnoses for the occlusal fissure pattern.

Statistical analysis The data analysis was performed using SAS release 9.1 (SAS Institute Inc., Cary, NC, USA). Weighted Kappa values (wK) were calculated as measure of agreement for categorical data to determine the intra- and inter-examiner reproducibility [16] The reproducibility was assessed as low for wK below 0.40, moderate for wK between 0.41 and 0.60, good for wK between 0.61 and 0.80 and excellent for wK between 0.81 and 1.00 [17]. In addition, a cumulative Logit-model was fitted [18]. The dependence of each visual decision on the occlusal surfaces examined was modelled by a random normal effect $\mu$ ($\mu_i$=149) with standard deviation Sigma. The model uses cumulative Logits to describe the transition probability from one diagnostic level to another ($\alpha_j$). The influence of
the examiners is assumed as equal on all levels, i.e., a common \( \beta \) at all stages \( j \). There are \( J \) diagnostic levels. For all inter-examiner comparison were used the inaugurator of UniViSS as reference examiner (D1); the \( \beta \) values representing existing differences between the examiner D1 and each of the other investigators. The level of significance was set at \( p < 0.05 \). If the \( p \) value was lower than 0.05, then the null hypothesis that the examiner do not differ in their scoring has to be rejected in favour of the alternative hypothesis, which states that differences exist. The following equation describes the used model:

\[
\log \it{it} P(Y_{it} \leq j[u_i]) = \alpha_j + x_{it}^\prime \beta + z_{ui}, \quad j = 1, \ldots, J - 1
\]

The cumulative Logits were calculated as follows:

\[
\log \it{it} P(Y \leq j|x) = \log \frac{P(Y \leq j|x)}{1 - P(Y \leq j|x)}, \quad j = 1, \ldots, J - 1
\]

Results

Validity study Table 1 summarises the results of the CE-index in relation to each of the UniViSS scores. The results showed that for sound occlusal surfaces no demineralisations were detectable. Fissures with ‘First signs’ of a caries process had a mean CE-index of 0.6, which indicates that most lesions were found in the enamel. With respect to the maximum value of 1.3, some exceptions occurred. The UniViSS severity score ‘Established lesion’ showed a heterogeneous distribution of the CE-index in relation to the UniViSS discoloration score. The following findings were observed: (1) The caries extension progresses with increased discoloration scoring. (2) The CE-indices for ‘Established lesions’ with white (0.4) and white-brown discolorations (0.7) indicate an enamel progression only. Contrary to this, ‘Established lesions’ with brown discolorations (1.2) and greyish translucencies (1.5) progressed regularly into dentine. The CE-index of 1.5 and 1.7 for ‘Microcavities’ and ‘Dentine exposure’ indicated that these lesions had histologically progressed into the middle of the dentine on average.

Reproducibility study Analyses of all inter-examiner data from the first measurement series showed that 53% (UniViSS/severity) and 61% (UniViSS/discoloration) of all measurements were repeated consistently. Viewing the results from the second measurement series, in 57% (UniViSS/severity) and 66% (UniViSS/discoloration) of all inter-examiner comparisons concordant diagnoses were made. This distinct tendency for higher accordance of the diagnoses obtained by the second measurements was confirmed by the wk values (inter-examiner data) of the first and second series. The inter-examiner wk values rose from 0.520 to 0.576 for the UniViSS/severity criteria and from 0.510 to 0.565 for the UniViSS/discoloration criteria. The intra-examiner reproducibility amounted to 0.685 (UniViSS/severity) and 0.628 (UniViSS/discoloration). The intra- and inter-examiner wk values are summarised in Table 2 for the UniViSS/severity criteria and for the UniViSS/discoloration criteria.

The results of the cumulative Logit-model are shown in Table 3. For UniViSS/severity both dentists and two out of four students did not show a significant difference to the inaugurator of the method (D1); S1 and S4 staged the severity significantly inferior. In case of UniViSS/discoloration a different observation was obtained: Only the experienced dentist (D2) reproduced their findings in concordance to the
inaugurator of the method (D1); the other dentist (D3) and all four students (S1-S4) registered significantly different discoloration scores in comparison to D1.

Discussion

As main result of this study it was shown that the newly developed CE-index provided quantitative information about the exact caries depth of a distinct diagnostic score. Based on these results it will be possible to assign a probable preventive or operative treatment strategy to a certain diagnostic scores (Table 1). While the traditionally used validity parameters do not provide such detailed information, the CE-index will help to determine clearer diagnostic thresholds for intervention strategies. This aspect is of clinical importance as the indication for a restorative treatment no longer depends on the simple fact that a caries lesion penetrated the enamel–dentin junction [19].

In order to investigate the validity of UniViSS, it was possible to evaluate each diagnostic score separately with the CE-index (Table 1). ‘First signs’ of a caries lesion were mainly associated with an enamel caries (CE-index of 0.6); but it should be noted that in few cases with white-brown or brown discolorations the outer third of dentine was reached. Therefore, this criterion seems to be mainly associated with lesions that require a preventive treatment strategy only. ‘Established lesions’ showed a much greater heterogeneity:

### Table 2 Weighted Kappa values for the intra- and inter-examiner reproducibility of the UniViSS/severity and UniViSS/discoloration criteria

<table>
<thead>
<tr>
<th>Reproducibility (wK)</th>
<th>UniViSS/severity</th>
<th>UniViSS/discoloration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intra-examiner</td>
<td>Inter-examiner</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Min</td>
</tr>
<tr>
<td>D1</td>
<td>0.646</td>
<td>0.585</td>
</tr>
<tr>
<td>D2</td>
<td>0.821</td>
<td>0.532</td>
</tr>
<tr>
<td>D3</td>
<td>0.680</td>
<td>0.578</td>
</tr>
<tr>
<td>S1</td>
<td>0.605</td>
<td>0.560</td>
</tr>
<tr>
<td>S2</td>
<td>0.756</td>
<td>0.577</td>
</tr>
<tr>
<td>S3</td>
<td>0.655</td>
<td>0.520</td>
</tr>
<tr>
<td>S4</td>
<td>0.634</td>
<td>0.506</td>
</tr>
<tr>
<td>Mean</td>
<td>0.685</td>
<td>0.551</td>
</tr>
</tbody>
</table>

wK weighted Kappa

Table 3 The cumulative Logit-model represents the comparisons between the reference examiner (D1) and each of the others for the UniViSS severity and discoloration criteria

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (Standard error)</th>
<th>p value</th>
<th>Lower confidence limit</th>
<th>Upper confidence limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>UniViSS severity criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β1 (D1–D2)</td>
<td>-0.239 (0.17)</td>
<td>0.1520</td>
<td>-0.5668</td>
<td>0.0889</td>
</tr>
<tr>
<td>β2 (D1–D3)</td>
<td>0.075 (0.16)</td>
<td>0.6468</td>
<td>-0.2468</td>
<td>0.3962</td>
</tr>
<tr>
<td>β3 (D1–S1)</td>
<td>-0.409 (0.15)</td>
<td>0.0058*</td>
<td>-0.6984</td>
<td>-0.1202</td>
</tr>
<tr>
<td>β4 (D1–S2)</td>
<td>-0.054 (0.16)</td>
<td>0.7401</td>
<td>-0.3732</td>
<td>0.2657</td>
</tr>
<tr>
<td>β5 (D1–S3)</td>
<td>0.069 (0.16)</td>
<td>0.6750</td>
<td>-0.2553</td>
<td>0.3931</td>
</tr>
<tr>
<td>β6 (D1–S4)</td>
<td>-1.229 (0.16)</td>
<td>&lt;0.0001*</td>
<td>-1.5527</td>
<td>-0.9051</td>
</tr>
<tr>
<td>UniViSS discoloration criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β1 (D1–D2)</td>
<td>0.151 (0.16)</td>
<td>0.3538</td>
<td>-0.1698</td>
<td>0.4719</td>
</tr>
<tr>
<td>β2 (D1–D3)</td>
<td>-0.355 (0.16)</td>
<td>0.0310*</td>
<td>-0.6769</td>
<td>-0.0329</td>
</tr>
<tr>
<td>β3 (D1–S1)</td>
<td>-0.745 (0.15)</td>
<td>&lt;0.0001*</td>
<td>-1.0339</td>
<td>-0.4558</td>
</tr>
<tr>
<td>β4 (D1–S2)</td>
<td>-1.301 (0.17)</td>
<td>&lt;0.0001*</td>
<td>-1.6275</td>
<td>-0.9745</td>
</tr>
<tr>
<td>β5 (D1–S3)</td>
<td>-0.664 (0.16)</td>
<td>&lt;0.0001*</td>
<td>-0.9881</td>
<td>-0.3406</td>
</tr>
<tr>
<td>β6 (D1–S4)</td>
<td>-1.764 (0.17)</td>
<td>&lt;0.0001*</td>
<td>-2.0942</td>
<td>-1.4343</td>
</tr>
</tbody>
</table>

*p values <0.05 are indicating a significant difference between the tested investigators, which means that the tested comparison revealed an inferior reproducibility
Established white-spot lesions were mainly associated with caries extending into the enamel and possibly needing preventive care. In contrast to this, ‘Established lesions’ with white-brown or brown discoloration showed a more heterogeneous distribution. While the mean CE-index of 0.7 and 1.2 indicated a caries process extending into the enamel or the dentine beneath the enamel–dentine junction, the maximum values of 1.7 for both scores showed that variations could be possible. These exceptions underline the known clinical problem of the difficulty to assess the caries extension on non-cavitated occlusal lesion by visual means correctly. Consequently, occlusal ‘Established lesions’ need clinically a diagnostic ‘safety net’ under inclusion of additional diagnostic methods, e.g., bitewing radiographs and/or laser fluorescence measurements, to detect caries lesions that have progressed far into dentine and need restorative treatment [20–23].

‘Microcavities’ and ‘Dentine cavities’ on occlusal surfaces were always associated with dentine caries regardless of the state of discoloration (Table 1) and should be therefore restored. Based on our results the potential of the quantitative CE-index was shown. Therefore, this index could help to advance the analysis and interpretation of validation data in future diagnostic studies.

When further comparing the established validity parameters (SE, SP and AUC) of UniViSS on occlusal caries lesions [14] with results from other recently published visual caries detection and diagnostic methods [7–11] the registered validity parameters for UniViSS were found to be in the same order of magnitude. When generalizing previously made suggestions that the sum of SE and SP should be at least ~160% before a diagnostic method could be considered as a legitimate candidate for practical use [24, 25], then the potential of UniViSS was illustrated. For both, the overall caries detection level (SE 100.0%, SP 58.3%) as well as for the dentine caries detection level (SE 62.5%, SP 97.6%), the sum of ~160% was reached [14]. Furthermore, the documented AUC for the caries detection level (0.84) and for the dentine caries detection level (0.81) are high and correspond to other studies on meticulous visual inspection methods [26–34].

Besides evaluation of diagnostic accuracy, caries activity assessment gained more attention throughout the last years. With respect to the definition of ‘caries activity’, which designates such lesions as active which are showing an ongoing mineral loss due to the metabolic activity of the biofilm [35], the need for longitudinal clinical trials is obvious. For this basic reason, we deliberately excluded an activity assessment from this in vitro study even though it is part of UniViSS [14]. Nevertheless, Braga et al. [11] published a first attempt to quantify caries activity in vitro. The authors included the parameters ‘ICDAS score’, ‘plaque stagnation area’ and ‘surface texture’ into their lesion activity assessment system. While plaque stagnation areas and roughness can be justified aetologically, there is currently no unanimous consensus as to when a plaque stagnation area, plaque presence and/or surface roughness should be rated as such. Further, it has also been taken into account that during the investigation the operator is constantly aware of the visual diagnosis and hence biased when it comes to the decision about activity [1]. All of these uncertainties will be completed by the fact that there are no quantitative reference standards published so far to determine the caries activity itself [36] as well as the validity of the criteria ‘plaque stagnation area’ and ‘roughness’. Therefore, more research to improve the caries activity assessment under inclusion of objective and quantifiable diagnostic criteria is needed. Within the second part of this investigation the reproducibility of UniViSS was comprehensively analysed. According to the wK results the all-over reproducibility can be assessed as good to moderate (Table 2). While good wK values were obtained mostly for the intra-examiner reproducibility, the inter-examiner reproducibility proved to be on a moderate level. Nevertheless, for several inter-examiner comparisons good wK values were found, too (Table 2). The registered wK values show almost the same order of magnitude as the Kappa values that were obtained for the visual criteria by Ekstrand et al. [7, 8] on occlusal surfaces [7, 30, 31, 33, 34].

With respect to the almost identical criteria sequences of ICDAS II and UniViSS/severity comparisons of the reproducibility data between both systems are feasible. The wK values for the intra- and inter-examiner reproducibility for ICDAS II amounted to 0.62 to 0.83 [28] and 0.88 to 0.90 [1] and indicate more favourable results than those for UniViSS from the present study. Since this can be explained by a more extensive calibration training in both ICDAS studies [1, 28], which was explicitly not part of the present study design, it can be assumed that an extensive calibration training should improve the reproducibility for UniViSS substantially.

To our knowledge, this was the first study that used a cumulative Logit-model to assess the reproducibility of a caries diagnostic method. In case of our investigation the developed cumulative Logit-model served to evaluate measurements in comparison to the decisions of the inaugurator of UniViSS (D1). As shown in Table 3, statistically non-significant differences were found for the majority of comparisons for UniViSS/severity. In contrast, the results for the UniViSS/discholoration criteria showed a more heterogeneous pattern (Table 3). Here, only one dentist (D2) was able to obtain similar decisions, which can be explained by longer clinical experience. These findings indicate that clinically and scientifically inexperienced dentists have a higher need for training especially for the
assessment of the UniViSS/discoloration criteria. The main keys for successful calibration training seem to impart theoretical knowledge, to discuss visual decisions on dental photographs and extracted teeth and to perform a clinical training. In case of clinically inexperienced dentists, the training should be intensified and, perhaps, repeated to eliminate possible subjective errors. Under these circumstances it can be hypothesised that the registered reproducibility values could be substantially improved when extensive calibration training is performed.

In conclusion of the present study, the overall validity and reproducibility of UniViSS has to be assessed as encouraging. Based on our findings we could further state that the detailed approach to categorise (non-)cavitated caries lesions with UniViSS provides additional information especially on ‘Established lesions’ for the clinician. With respect to the UniViSS/discoloration score and the CE-index, it was shown that each discoloration score on ‘Established lesions’ was associated with a different caries depth which would further result in different preventive/treatment strategies. As such detailed and clinically important information are not part of recently published visual caries detection and diagnostic systems [7–11], this aspect has to be understood as the unique feature of UniViSS. Nevertheless, further investigations are needed to validate UniViSS with a larger sample size before final treatment recommendations could be drawn. With respect to the results of the reproducibility study, it has to be concluded to adapt the calibration training to the clinical experience of the dentist(s). As the UniViSS/discoloration score seems to be more difficult to reproduce, more emphasis should be paid on this finding in the future. The used cumulative Logit-model enabled statistical comparisons between all examiners and enhanced the informative value of the reproducibility study. Therefore, such models could be used more frequently to investigate intra- and inter-examiner reproducibility in future.

Conflict of interest The authors declare that they have no conflict of interest.

References