CrossMark

ORIGINAL ARTICLE

In vitro validation of a hand-held optical reflectometer to measure clinically observed erosive tooth wear

Thiago Saads Carvalho[1] · Cristiane Meira Assunção[1,2] · Fabian Jost[1] ·
Walter Bruno Bürgin[1] · Jonas Almeida Rodrigues[2] · Adrian Lussi[1]

Abstract In this study, we analyzed a newly developed optical reflectometer for measuring erosive tooth wear (ETW) in vitro. Three examiners independently assessed the labial surface of 80 deciduous canines and 75 permanent incisors. One examiner performed visual examinations (BEWE), and the other two used the optical pen-size reflectometer to measure surface reflection intensity (SRI) on the same labial surfaces. The examinations were made in duplicate with at least 1 week interval. Intra- and inter-rater agreements were calculated using weighted *kappa* analysis for BEWE, and intraclass correlation coefficients (ICC) as well as Bland-Altman plots for SRI. The teeth were separated into without (BEWE 0) or with (BEWE 1–3) ETW, and SRI cut-off points were calculated. Intra-rater agreement for the visual examination was 0.46 and 0.82 for deciduous and permanent teeth, respectively. Inter-rater and intra-rater agreement for SRI were good (ICC > 0.7; $p < 0.001$). SRI measurements produced high specificity values for deciduous and permanent teeth (≥ 0.74 and ≥ 0.84, respectively), and lower sensitivity values (≥ 0.37 and ≥ 0.64, respectively), but permanent teeth had generally higher SRI values ($p < 0.05$). We observed a significant association between BEWE and SRI ($p < 0.05$). The optical pen-size reflectometer was able to adequately differentiate ETW on permanent teeth, with highly reliable and reproducible measurements, but ETW on deciduous teeth was less

accurately differentiated. The reflectometer is a good candidate for clinical research.

Introduction

Erosive tooth wear (ETW) is the chemical-mechanical process leading to the cumulative loss of hard dental tissue without the involvement of bacteria [1]. Enamel dissolution occurs both at the enamel/acid interface, as well as within a partly demineralized thin softened layer of enamel, in a process called near-surface demineralization [2], leading to loss of minerals, and consequently, loss of tooth substance. So, clinically, ETW is characterized as shiny, silky-glazed, but sometimes dull, excessively smooth tooth surfaces, with the absence of perikymata [3, 4].

Different methods are applied in the laboratory to measure the changes occurring on the enamel surface, but the vast majority of these methods require sample preparation, where the enamel must be ground and polished. This leaves very limited options to measure erosion on native enamel surfaces. ETW in the clinic is basically assessed with a visual examination, and the lesions are then categorized using an erosive wear index, such as the basic erosive wear examination (BEWE) [5].

Recently, an optical device (reflectometer) was newly developed to measure enamel erosion in the laboratory, which showed promising results for monitoring erosion on ground/polished enamel [6–8]. This optical reflectometer was later miniaturized to a pen-size instrument, which could be used on native enamel surfaces [9], and which is small enough to

✉ Thiago Saads Carvalho
  thiagosaads@hotmail.com

1 Department of Preventive, Restorative and Pediatric Dentistry, University of Bern, Freiburgstrasse 7, CH-3010 Bern, Switzerland

2 School of Dentistry, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

be used in the oral cavity. The basic principal of the optical reflectometer is to measure surface reflection intensity (SRI) to quantify the proportion of light from an incident laser beam reflected back from the enamel surface [6, 9]. Smoother enamel surfaces yield greater SRI values, whereas rougher surfaces produce lower SRI values. As this optical pen-size reflectometer is small, it can be used on native enamel surfaces, and as it has showed good results in quantifying early in vitro erosion, it is a potential tool for use in the clinic [9].

Our hypothesis is that SRI values will differ between healthy teeth or teeth with ETW. So, in an attempt to assess this hypothesis, we categorized permanent and deciduous teeth according to the BEWE index, measured their surface reflection intensity (SRI), and verified the association between these variables.

Materials and methods

Enamel specimen selection

In the present study, the sample size (n) and the number of replicate examinations per sample (k) were calculated using the following parameters: for an ideal inter-rater reliability of at least $\rho_0 = 0.9$, but an acceptable reliability of at least $\rho_1 = 0.8$, where $\alpha = 0.05$ and $\beta = 0.2$, we estimated the optimal number of replicate examinations to be $k = 2$ assessments per tooth (each rater should assess each tooth twice), and the minimum sample size was estimated at $n = 46$ teeth (minimum of 46 permanent and 46 deciduous teeth) [10].

We used 80 deciduous upper canines and 75 permanent upper incisors, which were extracted by dental practitioners in Switzerland. Before the extraction, the patients (and their parents, in case of children) were informed about the use of their teeth for research purposes and their consent was obtained. Because we are using teeth from a pooled bio-bank, the local ethics committee categorizes these kinds of samples as "irreversibly anonymised," and no previous ethical approval was necessary. The selected teeth had different degrees of ETW. They were randomized and each tooth was given a number and maintained in a humid chamber throughout the study.

Basic erosive wear examination

One examiner with over 4 years of clinical experience, blinded to the selection and randomization of the teeth, performed a visual examination of the teeth. In a similar fashion as performed in a clinical exam, each tooth was individually dried, observed under natural light, with no magnifying or optical aids, and the BEWE score for the labial surface of each tooth was recorded. The teeth were not embedded, and the examiner was allowed to handle the teeth during examination. The

BEWE examination was carried out twice, with a 1-week interval between the assessments.

Enamel surface reflection intensity measurement

Two raters (R1 and R2), who were blinded to the BEWE examinations, performed surface reflection intensity (SRI) measurements on the teeth. R1 and R2 had 13 and 12 years of clinical experience, and 1 and 3 years experience with the reflectometer, respectively. For this assessment, each tooth was individually dried, and each rater measured the SRI in the center of the labial surface of each tooth. The SRI measurements were performed twice by each rater, with a 1-week interval between each assessment.

The SRI measurements were carried out using an optical pen-size reflectometer [9]. The reflectometer has a hand-held piece fitted with a laser diode (oeMarket, Cherrybrook, Australia), which directs a laser beam (635 nm) onto the surface of the sample. The reflected light is then captured and measured with a photodiode (FDS100, Thorlabs, Dachau, Germany). The tip of the reflectometer is placed directly onto the enamel surface, and inclined in different angles until the point of highest reflection intensity is registered by a specific software on a notebook connected to the reflectometer [9]. The point of highest reflection intensity is expressed as an SRI value, which is dimensionless. In practical terms, greater SRI values represent greater reflection intensity of the enamel surface, which, in turn, correspond to smoother enamel surfaces (in the clinical situation this corresponds to ETW). In contrast, lower SRI values correspond to rougher enamel surfaces (in the clinical situation this corresponds to sound surfaces).

Further experiments on permanent teeth with ETW

To support part of the discussion, we performed further assessments on the permanent teeth classified as having ETW ($n = 50$). A fourth rater, with no clinical experience, blinded to the BEWE scores and other SRI results, measured SRI on the area of most wear, where ETW was most evident. These results were then compared to those from the other raters (R1 and R2).

Statistical analyses

The intra-rater agreement between the BEWE scores from the first and second BEWE examinations was calculated using weighted *kappa* analysis. To analyze the agreements in SRI, we calculated the intra-class correlation coefficients (ICC) using two-way random models with measures of consistency, and Bland-Altman plots, for raters R1 and R2, respectively. Since each rater performed the SRI measurements twice, we used these two SRI values to calculate the intra-rater

agreements. We later calculated the average of these two measurements of the same rater, and used this average of each rater to calculate the inter-rater agreement.

Furthermore, the BEWE values from the first and second visual examinations were scrutinized. When both values were identical, this value was considered as the BEWE value for the tooth. If the same tooth received discrepant BEWE scores, the lowest score was considered for analysis. We then dichotomized this BEWE score: without (BEWE score 0) or with signs of ETW (BEWE scores 1–3). We used the dichotomized BEWE and the averaged SRI measurements to calculate the Youden J value and, from that, obtained the specific SRI cut-off points for each rater. For these cut-off points, we then calculated the sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV): sensitivity is the fraction of teeth with ETW that were identified as having ETW, and it was calculated as follows: true positive/(true positive + false negative); specificity is the fraction of sound teeth identified as sound, and was calculated as follows: true negative/(true negative + false positive); NPV is the fraction of teeth identified as sound that were in fact sound, and was calculated as follows: true negative/(true negative + false negative); PPV is the fraction of teeth identified as having ETW that had in fact ETW, and was calculated as follows: true positive/(true positive + false positive).

Additionally, we have also plotted receiver operating characteristic (ROC) curves for both raters (R1 and R2), considering deciduous and permanent teeth, and we also calculated the area under each ROC curve (AUC). All calculations were made separately for deciduous and permanent teeth.

For the statistical analyses, we used SPSS v. 22 and SAS v. 9.4, and experiment wise probability for significance (α) was set at 0.05.

Results

The Bland-Altman plots with confidence intervals (Fig. 1) represent the agreements between the raters (inter-rater), as well as the intra-rater agreements for rater 1 (R1) and rater 2 (R2), according to the two types of teeth (permanent and deciduous). The results show that permanent teeth had higher SRI values, with a higher proportion of permanent teeth in the range of SRI 30 to 50 than deciduous teeth. The intra-rater agreements for rater 2 (R2) were more disperse in both types of teeth, with a larger range of differences than those for rater 1 (R1) and the inter-rater agreements. The inter-rater plots (Fig. 1) also show a slight, but quite distinct, decrease in the difference of SRI values from both raters (R1−R2), as the inter-rater average increases (average of R1 and R2).

The cut-off points, sensitivity, specificity, NPV, and PPV for each rater and for each type of tooth (permanent and deciduous) are displayed on Table 1. We observed generally

high specificity and positive predictive values, especially for permanent teeth (values above 0.8).

Figure 2 shows that there were no significant differences in SRI values between the different BEWE scores in deciduous teeth, whereas, on permanent teeth, we observed significantly lower SRI values on sound teeth (BEWE = 0) in comparison to teeth categorized as having ETW (BEWE ≥ 1).

The intra-rater weighted *kappa* agreement values for the visual examinations (BEWE scores) were 0.82 and 0.46 for permanent and deciduous teeth, respectively. Among the deciduous teeth, 39 were classified as BEWE score 0, 32 teeth had BEWE score 1, and 9 teeth had BEWE score 2. From the permanent teeth, 25 were classified as BEWE score 0, 40 teeth had BEWE score 1, 9 had BEWE score 2, and 1 tooth had BEWE score 3. Regarding the reflectivity assessments, R2 obtained consistently higher SRI values than R1 ($p < 0.001$), and permanent teeth had generally higher SRI values than deciduous teeth ($p < 0.05$).

The intra- and inter-rater reliability for SRI measurements were all above 0.7 ($p < 0.001$), for both deciduous and permanent teeth (Table 2). There was a significant association between the BEWE scores and the SRI values ($p < 0.05$; Table 3). The reflectometer yielded a very small proportion of false positives, and it did not overestimate ETW on permanent teeth (Table 3). In regards to deciduous teeth, the reflectometer underestimated ETW, producing more false negative results (Table 3).

The ROC curves are represented in Fig. 3. The reflectometer presented low overall diagnostic performance for deciduous teeth, where both raters produced quite low AUC values, 0.62 ± 0.06 for R1 ($p = 0.056$) and 0.63 ± 0.06 for R2 ($p = 0.039$). On the other hand, higher AUC values were observed in permanent teeth, representing better overall diagnostic performance from the reflectometer, where AUC values were 0.74 ± 0.06 for R1 ($p = 0.001$) and 0.81 ± 0.05 for R2 ($p < 0.001$).

In regards to the further experiments performed by the fourth rater, we observed no significant differences ($p = 0.191$) between the SRI values measured either at the center of the labial surface or on the area where ETW was most evident (area of most enamel loss). This rater also presented good reproducibility of SRI measurements of ICC = 0.90 ($p < 0.001$).

Discussion

When enamel comes in contact with an acid, the partial mineral dissolution causes a roughening and softening of the surface, leaving it more vulnerable to mechanical forces, which subsequently remove the softened layer, leaving behind a smoother tooth surface and, hence, higher SRI values [8]. The optical reflectometer is a sensitive method to measure

Fig. 1 Bland-Altman plots for the inter- and intra-rater agreements in SRI values, for rater 1 (R1) and rater 2 (R2), according to each type of tooth (deciduous and permanent). *R1 − R2* indicates the difference in SRI values between the raters (SRI values of rater 2 subtracted from the SRI values of rater 1); *average of R1 and R2* indicates the average SRI values between raters 1 and 2 ((SRI values of rater 1 + SRI values of rater 2)/2)
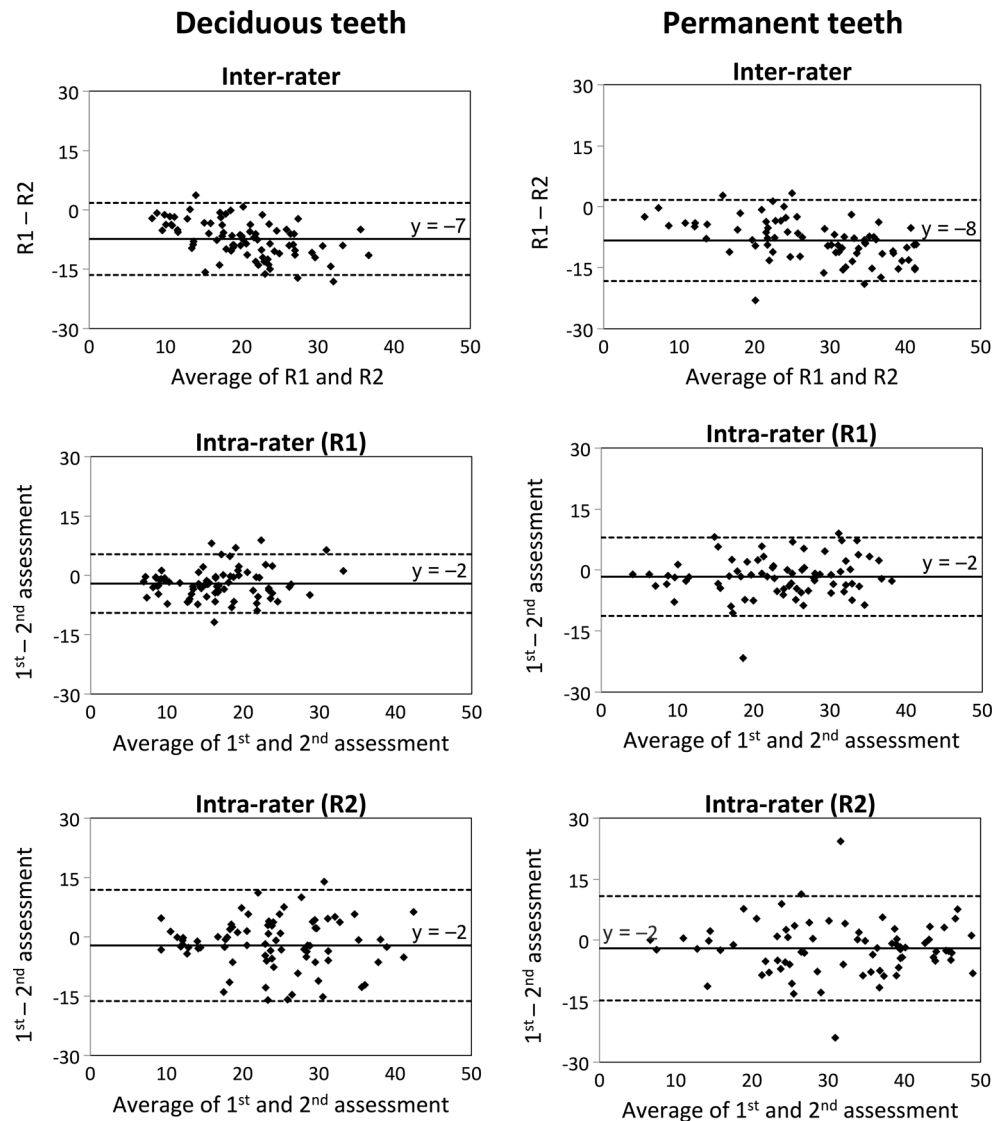
these changes in SRI values [6–9], but until now, there is currently no device to quantify ETW in the clinical setting. The present in vitro study attempted to fill this gap by using the optical reflectometer in an in vitro setting to assess native tooth surfaces containing erosive lesions that had been formed in vivo (before tooth extraction).

The performance of the reflectometer can be observed in the Bland and Altman diagrams, which are used to show systematic differences in the SRI measurements. The systematic difference lines (solid lines, Fig. 1) represent the average values of the differences between two assessments (R1 − R2 or first − second assessment). In a perfect scenario, it would be preferable to have these lines running through $y = 0$, denoting no systematic differences between the measurements [11, 12]. However, we observed that our systematic deviation lines ranged from $y = -2$ to $y = -8$. This indicates that the raters did not produce the same SRI results for the teeth, especially when the teeth presented higher SRI values (Fig. 1). This may be in connection to the raters' experience with the reflectometer. Raters R1 and R2 had 1 and 3 years of experience with the device, respectively, so the differences in SRI values between them may be a reflection of a learning curve. Moreover,

Table 1 Cut-off points, sensitivity, specificity, negative predictive value (NPV), and positive predictive values (PPV) for each rater and each type of tooth

	Deciduous teeth		Permanent teeth	
	Rater 1	Rater 2	Rater 1	Rater 2
Cut-off point	21	27	25	35
Sensitivity	0.366	0.439	0.640	0.640
Specificity	0.897	0.744	0.840	0.920
NPV	0.574	0.558	0.538	0.561
PPV	0.789	0.643	0.889	0.941

Fig. 2 Box-plots of the SRI values for the different degrees of ETW (BEWE scores), according to the raters (R1 and R2) and to each type of tooth (deciduous and permanent). The *horizontal lines through the boxes* indicate the cut-off point for SRI, *values below the line* indicate absence of ETW, and *values above the line* indicate presence of ETW, according to the reflectometer measurements. According to the clinical criteria, BEWE scores greater than 0 indicate presence of ETW. *p* values below 0.05 indicate significantly different SRI values in the BEWE score group
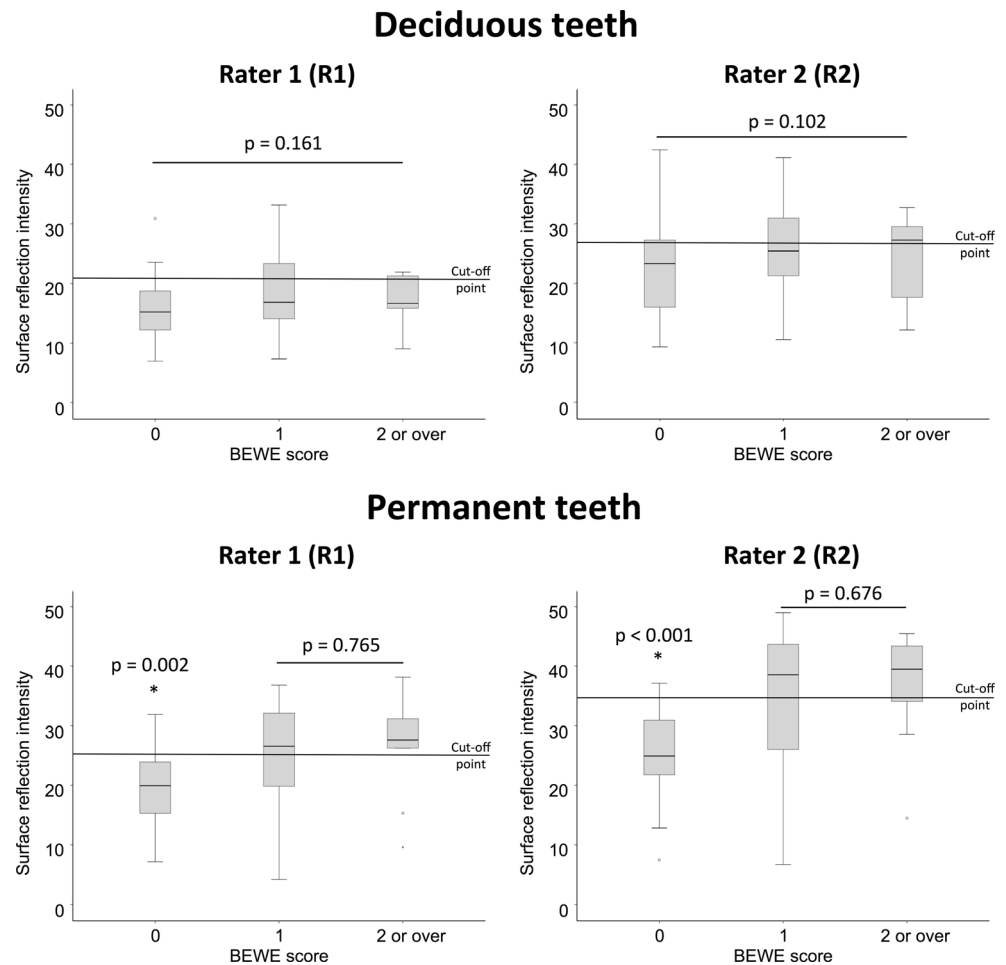
the experience with the device can also explain the different AUC values observed for permanent teeth. We observed that R2 (more experienced rater) produced a slightly better overall diagnostic performance (higher AUC values) than R1 (less experienced rater). However, both raters generated highly reproducible results. This was seen in the inter-rater reproducibility, where strong correlations in SRI measurements were observed for deciduous (ICC = 0.77) and permanent teeth (ICC = 0.86).

These strong correlations can be due to the small SRI variation within the same tooth surface. In this regard, we performed further experiments, in which a fourth rater used the

Table 2 Intra-rater and inter-rater reliability values (ICC with 95 % cumulative interval), according to the rater and the type of tooth (deciduous or permanent)

		Deciduous teeth	Permanent teeth
Intra-rater	Rater 1	0.80 (0.71–0.87)	0.84 (0.75–0.89)
	Rater 2	0.72 (0.59–0.81)	0.83 (0.74–0.89)
Inter-rater		0.77 (0.67–0.85)	0.86 (0.78–0.90)

reflectometer on the permanent teeth with ETW. Remarkably, we observed no significant differences between the SRI values measured either at the center of the labial surface or where ETW was most evident. Moreover, there was again a very good reproducibility of SRI measurements. This suggests that ETW occurs on the whole enamel surface, so SRI can be measured on any part of this surface. One must bear in mind, however, that SRI does not measure the amount of enamel lost, nor has been tested in advanced ETW cases where dentine is exposed. So, the present hypothesis is only valid for ETW cases involving enamel.

Considering the results from permanent teeth, we observed that 64 % of teeth with ETW were correctly diagnosed by the reflectometer. The cut-off values for permanent teeth also yielded very high specificity rates, which generated a high proportion of true negative results. This means that the device was able to correctly identify sound surfaces on permanent teeth. Observing the cut-off line in Fig. 2 (for permanent teeth), we see that practically all teeth (≥84 %) with BEWE score 0 (no signs of ETW) had SRI values below the cut-off line, while a large number of permanent teeth with ETW (BEWE 1 or BEWE 2 or over) had SRI values above the

Table 3 Association between the amount of teeth diagnosed with erosive tooth wear (ETW) regarding the BEWE clinical score and the reflectometer (following the cut-off values from Table 1)

			Deciduous teeth				Permanent teeth		
			BEWE		Total		BEWE		Total
			0	≥1			0	≥1	
Rater 1	Presence of ETW (reflectometer)	No	35	26	61	Presence of ETW (reflectometer)	No	21	18	39
		Yes	4	15	19		Yes	4	32	36
	Total		39	41	80	Total		25	50	75
	OR = 4.9 (1.5–19.1); $p = 0.003$					OR = 9.0 (2.8–35.0); $p < 0.001$				
Rater 2	Presence of ETW (reflectometer)	No	29	23	52	Presence of ETW (reflectometer)	No	23	18	41
		Yes	10	18	28		Yes	2	32	34
	Total		39	41	80	Total		25	50	75
	OR = 2.2 (0.9–6.0); $p = 0.047$					OR = 19.6 (4.7–135.8); $p < 0.001$				

cut-off point. In other words, the reflectometer was able to accurately differentiate the sound permanent teeth from those clinically classified with ETW. More specifically, more than 70 % of the teeth classified as having advanced ETW (BEWE scores 2 or over) were correctly identified by the reflectometer, and more than 63 % of the teeth classified as BEWE 1 had

Fig. 3 Receiver operating characteristic (ROC) curves, according to the raters (R1 and R2) and to each type of tooth (deciduous and permanent)
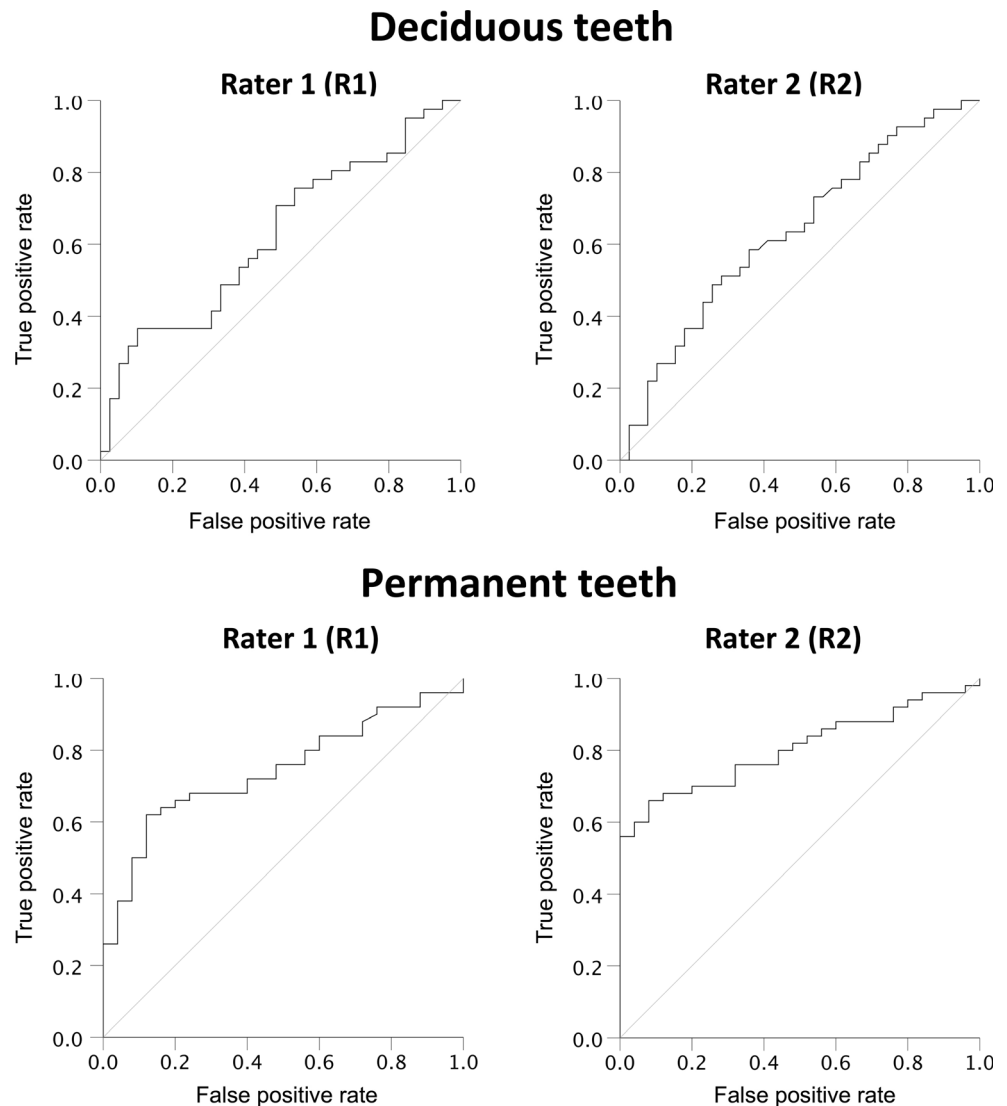
SRI values above the cut-off point. The ROC curves also support these observations, showing that both R1 and R2 raters obtained higher AUC values for permanent teeth (≥0.74) than for deciduous teeth (≤0.63), and that there was a better overall diagnostic performance for permanent teeth.

Considering the results from deciduous teeth, we observed a relatively low proportion (between 37 and 44 %) of the teeth with ETW correctly diagnosed by the reflectometer. This number was considerably lower than those observed for permanent teeth. Observing the cut-off line for deciduous teeth in Fig. 2, we see that the majority of teeth (≥74 %) with no signs of ETW (BEWE score 0) had SRI values below the cut-off line. In contrast, only a small number of deciduous teeth with ETW (BEWE 1 or BEWE 2 or over) lie above the cut-off line. In other words, the reflectometer was not able to accurately discriminate deciduous teeth clinically classified as having ETW. More specifically, only half (56 %) of the deciduous teeth clinically classified as having advanced ETW (BEWE scores 2 or over) were correctly identified by the reflectometer, and less than half of the teeth (47 %) classified as BEWE 1 had SRI values above the cut-off point. This is also clearly observed in the ROC curves, where both raters produced quite low AUC values for deciduous teeth.

In general, the reflectometer results for permanent teeth were more promising than those for deciduous teeth. This is also mirrored on the intra-rater agreements for BEWE scores, where we observed a high intra-rater agreement for permanent teeth, and lower agreement for deciduous teeth. The differences between permanent and deciduous teeth can be twofold. Firstly, the lower overall diagnostic performance of the reflectometer for deciduous teeth is analogous to the difficulty in visually detecting ETW on deciduous teeth [13]. Secondly, the curvature of the surface of the teeth could influence the SRI results, where flatter and smoother surfaces provide higher reflectivity, thus higher SRI values. In other words, the lower SRI values observed in deciduous teeth is, in parts, the outcome of their greater curvature as opposed to the permanent teeth, which have comparably flatter labial surfaces.

Although the present in vitro study used extracted teeth, our results provide a good perspective as to how the reflectometer will work in the clinical environment, but two issues arise. On the one hand, differentiating very early ETW lesions can be challenging, and since there is still no gold standard for the clinical evaluation of ETW, diagnosis is dependable on the visual examination of the dentist. On the other hand, the reflectometer does not measure the amount of tooth surface already lost in the ETW process, so the use of the device in the clinical environment will not substitute clinical diagnosis. In this regard, it is important to bear in mind that more severe ETW can be easily judged in the clinic; ergo, the reflectometer would rather serve as an aid in the decision-making process when diagnosing initial ETW lesions on enamel. Also, the reflectometer could provide valuable numerical data for ETW in permanent enamel. So the device is a promising tool for in vitro, in situ, and clinical research purposes. Still, additional studies are necessary to verify the possible use of the reflectometer in the clinic. It is worth mentioning that we performed the present study focusing on ETW involving enamel. Further experiments should also verify the performance of the reflectometer on advanced ETW with exposed dentine.

In conclusion, the reflectometer was able to adequately differentiate ETW on permanent teeth, and it generated highly reliable and reproducible measurements for native tooth surfaces.

Compliance with ethical standards

Ethical standards   This article did not require the approval of an ethical committee, as detailed in the "Materials and methods" section.

Informed consent   For this type of study, formal written consent is not required.

References

1. Carvalho TS, Colon P, Ganss C, Huysmans MC, Lussi A, Schlueter N, Schmalz G, Shellis RP, Tveit AB, Wiegand A (2015) Consensus report of the European Federation of Conservative Dentistry: erosive tooth wear—diagnosis and management. Clin Oral Investig 19: 1557–1561

2. Shellis RP, Barbour ME, Jesani A, Lussi A (2013) Effects of buffering properties and undissociated acid concentration on dissolution of dental enamel in relation to pH and acid type. Caries Res 47:601–611

3. Lussi A, Schlueter N, Rakhmatullina E, Ganss C (2011) Dental erosion—an overview with emphasis on chemical and histopathological aspects. Caries Res 45(Suppl 1):2–12

4. Ganss C, Lussi A (2014) Diagnosis of erosive tooth wear. Monogr Oral Sci 25:22–31

5. Bartlett D, Ganss C, Lussi A (2008) Basic erosive wear examination (BEWE): a new scoring system for scientific and clinical needs. Clin Oral Investig 12(Suppl 1):S65–S68

6. Rakhmatullina E, Bossen A, Höschele C, Wang X, Beyeler B, Meier C, Lussi A (2011) Application of the specular and diffuse reflection analysis for in vitro diagnostics of dental erosion: correlation with enamel softening, roughness, and calcium release. J Biomed Opt 16:107002

7. Brevik SC, Lussi A, Rakhmatullina E (2013) A new optical detection method to assess the erosion inhibition by in vitro salivary pellicle layer. J Dent 41:428–435

8. Lussi A, Bossen A, Höschele C, Beyeler B, Megert B, Meier C, Rakhmatullina E (2012) Effects of enamel abrasion, salivary pellicle, and measurement angle on the optical assessment of dental erosion. J Biomed Opt 17:97009–1

9. Rakhmatullina E, Bossen A, Bachofner KK, Meier C, Lussi A (2013) Optical pen-size reflectometer for monitoring of early dental erosion in native and polished enamels. J Biomed Opt 18:117009

10. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. Stat Med 17:101–110

11. Huysmans MC, Kühnisch J, ten Bosch JJ (2005) Reproducibility of electrical caries measurements: a technical problem? Caries Res 39:403–410

12. Rodrigues JA, Hug I, Diniz MB, Lussi A (2008) Performance of fluorescence methods, radiographic examination and ICDAS II on occlusal surfaces in vitro. Caries Res 42:297–304

13. Carvalho TS, Lussi A, Jaeggi T, Gambon DL (2014) Erosive tooth wear in children. Monogr Oral Sci 25:262–278