Dentomaxillofacial Radiology (2018) 47, 20170292

birpublications.org/dmfr

RESEARCH ARTICLE

Inter- and intraexaminer reliability of bitewing radiography and near-infrared light transillumination for proximal caries detection and assessment

[1]Friederike Litzenburger, [1]Katrin Heck, [1]Vinay Pitchika, [2]Klaus W Neuhaus, [2]Fabian N Jost, [1]Reinhard Hickel, [3]Anahita Jablonski-Momeni, [4]Alexander Welk, [1]Alexander Lederer and [1]Jan Kühnisch

[1]Department of Conservative Dentistry and Periodontology, University Hospital, LMU Munich, Munich, Germany; [2]Department of Preventive, Restorative, and Pediatric Dentistry, University of Bern, Bern, Switzerland; [3]Department of Pediatric and Community Dentistry, Phillips-University, Marburg, Germany; [4]Department of Restorative Dentistry, Periodontology, Endodontology, Preventive and Pediatric Dentistry, Dental School, University Medicine Greifswald, Greifswald, Germany

Objectives: The purpose of this *in vitro* study was to evaluate the inter- and intraexaminer reliability of digital bitewing (DBW) radiography and near-infrared light transillumination (NIRT) for proximal caries detection and assessment in posterior teeth.
Methods: From a pool of 85 patients, 100 corresponding pairs of DBW and NIRT images (~1/3 healthy, ~1/3 with enamel caries and ~1/3 with dentin caries) were chosen. 12 dentists with different professional status and clinical experience repeated the evaluation in two blinded cycles. Two experienced dentists provided a reference diagnosis after analysing all images independently. Statistical analysis included the calculation of simple (κ) and weighted Kappa (wκ) values as a measure of reliability. Logistic regression with a backward elimination model was used to investigate the influence of the diagnostic method, evaluation cycle, type of tooth, and clinical experience on reliability.
Results: Altogether, inter- and intraexaminer reliability exhibited good to excellent κ and wκ values for DBW radiography (Inter: κ = 0.60/ 0.63; wκ = 0.74/0.76; Intra: κ = 0.64; wκ = 0.77) and NIRT (Inter: κ = 0.74/0.64; wκ = 0.86/0.82; Intra: κ = 0.68; wκ = 0.84). The backward elimination model revealed NIRT to be significantly more reliable than DBW radiography.
Conclusions: This study revealed a good to excellent inter- and intraexaminer reliability for proximal caries detection using DBW and NIRT images. The logistic regression analysis revealed significantly better reliability for NIRT. Additionally, the first evaluation cycle was more reliable according to the reference diagnoses.
Dentomaxillofacial Radiology (2018) **47**, 20170292. doi: 10.1259/dmfr.20170292

Cite this article as: Litzenburger F, Heck K, Pitchika V, Neuhaus KW, Jost FN, Hickel R, et al. Inter- and intraexaminer reliability of bitewing radiography and near-infrared light transillumination for proximal caries detection and assessment. *Dentomaxillofac Radiol* 2018; **47**: 20170292.

Keywords: Dental caries; caries detection and assessment; reliability; digital bitewing radiography; near-infrared transillumination

Introduction

Appropriate treatment of dental caries requires valid and reliable methods to detect proximal caries lesions.[1–3]

Correspondence to: Dr Friederike Litzenburger, E-mail: soechtig@dent.med.uni-muenchen.de

Received 26 July 2017; revised 04 December 2017; accepted 06 December 2017

In addition to visual examination, bitewing radiography is the method of choice to date to detect caries lesions on hidden surfaces.[4–6] Near-infrared light transillumination (NIRT) of posterior teeth is a novel non-ionizing imaging method for the detection of proximal caries

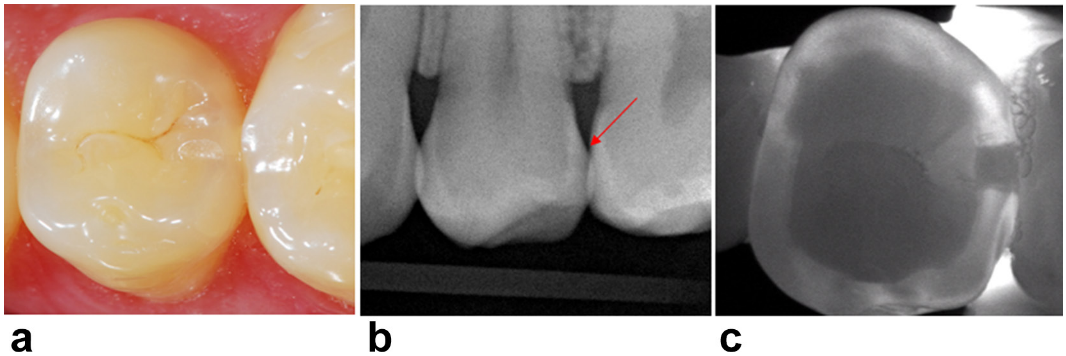

Figure 1 Clinical example of the distal surface of a premolar: (a) no visual signs of a proximal caries lesion; (b) digital bitewing radiograph: radiological translucence involved the enamel–dentin junction and reached the part of the dentin (D3 level); (c) near-infrared light image: a rectangular shading in the enamel with linear contact to the dentino–enamel junction (Score 4). Both, DBW and NIRT image, indicate a dentin caries lesion. DBW, digital bitewing; NIRT, near-infrared light transillumination.

lesions.[7] The technique combines the application of near-infrared light with digital imaging. Optical properties of demineralized dental tissue differ distinctively from those of the surrounding tissue when longwave light is applied. Sound enamel is highly transparent in the near-infrared portion (700 nm^{-1} mm) of the electromagnetic spectrum because the attenuation coefficient is lower than in the visible range (400–700 nm).[8] The altered structure of demineralized enamel with larger numbers of pores and interprismatic water accumulation causes increases in light scattering and absorption in these tissues. Therefore, high-contrast visualization of enamel caries lesions is possible using this technique.

Scattering in dentin is a more complex process, as the dentinal tubules behave as Mie scatterers.[9,10] These scatterers describe the scattering of an electromagnetic plane wave by a homogeneous sphere. The transillumination of dentin does not exhibit a significant decrease in scattering from the ultraviolet to the infrared spectral range that can be attributed to enamel.[11] Therefore, imaging caries lesions at such high contrast is not possible in dentin.

A recently introduced diagnostic device, the DIAGNOcam (KaVo, Biberach, Germany, 2012), uses NIRT to visualize the different dental hard tissues and their condition in posterior teeth. Two near-infrared laser diodes with a wavelength of ~780 nm and a power of 1 mW cm^{-2} transmit light through the alveolar bone and the dental hard tissue. An image of the transilluminated tooth is captured by a charge-coupled device sensor and displayed on a standard monitor (Figure 1). The image displays the tooth from the occlusal surface. High-contrast visualization of enamel caries is possible, whereas dentin caries lesions can only be visualized indirectly.[12]

In previous investigations, digital bitewing (DBW) radiography and NIRT presented equal clinical performance concerning the detection of proximal dentin caries *in vivo*.[12,13] Dentists require practical scoring or classification criteria for the reproducible documentation and monitoring of clinical findings. Therefore, a disease severity scale for proximal caries lesions assessed with NIRT was recommended for this purpose in 2014 (Table 1).[13] To date, only limited information regarding the diagnostic reliability of NIRT compared to DBW radiography has been available.[15] Therefore, the objective of this study was to analyse the inter- and intraexaminer reliability of proximal caries detection using DBW radiography and NIRT, including diagnostic decisions of multiple dentists. Moreover, different influencing variables, such as the involved methods, evaluation cycle, clinical experiences of the investigators and tooth type were included in our logistic regression analysis.

Table 1 Caries diagnostic criteria for the assessment of digital bitewing radiographs[14] and near-infrared transillumination images[13]

	Digital bitewing radiography		*Near-infrared transillumination*
Score	*Description*	*Score*	*Description*
0	No caries visible	0	No caries visible
D1	Caries in the outer half of enamel	1	First visible signs of caries within the enamel
D2	Caries in the inner half of enamel	2	Sign of established caries with the enamel
D3	Caries in the outer half of dentin	3	Caries with a point contact to the dentino–enamel junction
D4	Caries in the inner half of dentin	4	Caries with a linear contact with the dentino-enamel junction
		5	Caries visible within dentin

The null hypothesis of this study was that no difference exists in the inter and intraexaminer reliability between DBW radiography and NIRT.

Methods and materials

This *in vitro* reliability study used DBW and NIRT images from patients who were included in a previous clinical diagnostic study.[12] In this trial, 85 subjects (38 males and 47 females, mean age 25.0 years) with full permanent dentition and an overall healthy status were included. Each participant provided written informed consent. The local Ethics Committee approved this study (Project No. 013–12).

Image selection and reference diagnoses

In total, 211 image pairs consisting of DBW radiographs and their corresponding NIRT images were initially preselected according to the following criteria. All images were required to exhibit optimal contrast, brightness and sharpness. Each image exhibited at least one proximal surface with an adjacent tooth that was sound or showed varying degrees of proximal caries lesions. Furthermore, the relevant interproximal space of the selected images was depicted without significant overlapping effects. Surfaces with restorations, secondary caries, residual caries, hypomineralizations and orthodontic appliances were excluded. Before radiological assessment, each patient was asked whether DBW radiographs had been taken within the preceding 4 months. If these current images were available in sufficient quality, they were evaluated. In the case of present justifying indication, new DBW radiographs were made using an intraoral X-ray dental machine with a 203 mm tube (Heliodent DS, Sirona, Bensheim, Germany) including an X-ray field limitation (30 × 40 mm) with a charge-coupled device sensor (Intraoral II, sensor size 30.7 × 40.7 mm, Sirona, Bensheim, Germany). The exposure time was 0.06 s at a cathode voltage of 60 kV and an amperage of 7 mA. A sensor holding device (XPP-DS Digital Sensor Holders for Sirona, Dentsply Rinn, Elgin, IL) was used at all times.

Two investigators (JK and FL) provided a reference diagnosis for all of initially pre-selected pairs of DBW and NIRT images according to the given criteria in Table 1.[13, 14] Both analysed all images independently in two evaluation cycles under optimal conditions with calibrated monitors in a darkened room. Then, both investigators compared their results to determine a reference diagnosis for each surface. If the investigators reached different conclusions, they reassessed the images, discussed their points, and came to a consensus score. Finally, 100 image pairs with a well-balanced distribution of different lesion stages (34 sound, 33 enamel caries, and 33 dentin caries) were included to avoid any under- and overrepresentation of one caries stages.

All selected DBW and NIRT images were assigned to a unique identification number. Furthermore, a mark up was embedded on each image to highlight the relevant site, aiming to avoid misclassifications. To reduce the recognizability and memorability during the study period, each sequence of images was randomly changed between the first and second cycle of examinations for DBW and NILT images, respectively. The randomization list was only known to the principle investigators (JK and FL), who did not participate in the subsequent reliability study. All four image sequences were exported to separate PDF documents for the convenience of the participating dentists.

Reliability study

A group of 12 dentists with different professional status and clinical experience participated in this investigation. Four dentists had less than 2 years of clinical experience, four dentists had between 2 and 5 years of clinical experience, and four of had >10 years of clinical experience. Most of the investigators ($N = 8$) were employed at different universities, whereas the remaining four investigators worked in private practices. All received basic half-a-day theoretical and practical training sessions prior to participation. The dentists were introduced to the NIRT method. Furthermore, the study design and scoring criteria (Table 1)[13,14] were explained, and the dentists were trained. Subsequently, all 100 DBW radiographs and NIRT images were analysed in two evaluation cycles by all investigators with a minimum interval of 4 weeks to ensure blinding between the diagnostic decisions. Thus, a total of 2400 diagnoses were obtained for each method. All participants were encouraged to perform the evaluation without any help by other colleagues within 2 weeks. The image evaluation was performed on a standard calibrated monitor in a darkened room.

Statistical analysis

All scores from the dentists were entered in an Excel spreadsheet (Excel 2016, Microsoft, Redmond, WA). The data analysis was performed using SAS 9.3 (SAS Institute, Cary, NC) and R 3.3.2 (R Core Team, Vienna , Austria).[16] The descriptive analysis included the calculation and illustration of the percentage of agreement. Simple Kappa (κ) and weighted Kappa ($w\kappa$) coefficients were computed as measure of agreement to determine the inter- and intraexaminer agreement among the included investigators and in relation to the reference diagnoses (Tables 2 and 3).[17,18] Furthermore, to determine the reliability with reference diagnoses (Table 4), the values from all dentists ($N = 12$) were analysed in relation to the reference diagnoses and the κ coefficients were calculated. Fleiss–Cohen weights were used for the calculation of weighted κ coefficients.[18] κ-values were categorized as low (≤0.40), moderate (0.41 to 0.60), good (0.61 to 0.80), and excellent agreement (0.81 to 1.00).[19]

Table 2 Inter- and intraexaminer reliability values for digital bitewing radiography, which was calculated among all dentists and in relation to the reference diagnoses

Weighted κ for digital bitewing radiography		*Firstst Eevaluation cycle of each dentist*												
		Reference diagnoses	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Second Eevaluation cycle of each dentist	Reference diagnoses	1	0.83	0.74	0.71	0.45	0.67	0.71	0.74	0.77	0.73	0.79	0.80	0.65
	1	0.79	0.81	0.77	0.68	0.44	0.62	0.67	0.72	0.72	0.73	0.72	0.72	0.69
	2	0.72	0.76	0.72	0.58	0.40	0.55	0.66	0.67	0.67	0.68	0.68	0.66	0.64
	3	0.77	0.80	0.74	0.56	0.49	0.66	0.78	0.80	0.75	0.78	0.73	0.72	0.70
	4	0.74	0.77	0.73	0.53	0.49	0.65	0.71	0.78	0.74	0.76	0.73	0.70	0.68
	5	0.71	0.76	0.73	0.51	0.49	0.66	0.73	0.78	0.80	0.78	0.77	0.73	0.66
	6	0.79	0.79	0.74	0.57	0.52	0.66	0.79	0.74	0.80	0.76	0.83	0.76	0.66
	7	0.68	0.71	0.75	0.52	0.45	0.63	0.74	0.77	0.73	0.76	0.73	0.67	0.58
	8	0.77	0.80	0.72	0.56	0.45	0.59	0.76	0.75	0.83	0.79	0.86	0.75	0.62
	9	0.79	0.77	0.75	0.63	0.44	0.64	0.75	0.71	0.73	0.77	0.73	0.72	0.71
	10	0.84	0.83	0.75	0.71	0.46	0.66	0.70	0.71	0.75	0.72	0.85	0.77	0.65
	11	0.84	0.82	0.74	0.63	0.51	0.66	0.71	0.75	0.81	0.77	0.80	0.84	0.66
	12	0.74	0.79	0.75	0.65	0.52	0.67	0.70	0.73	0.73	0.78	0.75	0.68	0.81

Italic numbers illustrate intra-examiner values.

Furthermore, logistic regression analysis using a backward elimination model was performed for the outcome (correct/incorrect diagnosis in relation to consensus) *vs* measure. The analysis was adjusted for diagnostic method, evaluation cycle, experience and tooth.

Results

An overview of all diagnostic decisions in relation to the reference diagnoses is presented in Figure 2. Most diagnostic decisions of DBW radiography and NIRT were registered on the diagonals. Incorrectly classified diagnoses exhibited a deviation in most cases of only plus/minus one diagnostic score and were found in one of the neighbouring diagnostic categories. An insignificant number of incorrect diagnoses differed by up to three categories.

Tables 2 and 3 present the intra- and interexaminer reliability based on w*κ* values for DBW radiography and NIRT (the *κ*-values exhibited a similar distribution; data not shown). Analyses of all interexaminer data of the first evaluation cycle revealed that 69.8% of

Table 3 Inter- and intraexaminer reliability values for near-infrared transillumination, which were calculated among all dentists and in relation to the reference diagnoses

Weighted Kappa for near-infrared transillumination		*First Evaluation cycle of each dentist*												
		Reference diagnoses	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Second Eevaluation cycle of each dentist	Reference diagnoses	1	0.80	0.87	0.96	0.76	0.92	0.64	0.88	0.84	0.92	0.97	0.78	0.92
	1	0.85	0.72	0.80	0.88	0.69	0.84	0.56	0.83	0.83	0.89	0.83	0.76	0.85
	2	0.83	0.66	0.78	0.84	0.66	0.82	0.56	0.80	0.80	0.88	0.80	0.75	0.82
	3	0.83	0.69	0.77	0.85	0.66	0.83	0.60	0.85	0.81	0.83	0.84	0.75	0.87
	4	0.67	0.57	0.64	0.69	0.58	0.68	0.47	0.69	0.76	0.65	0.65	0.67	0.67
	5	0.84	0.69	0.81	0.87	0.67	0.83	0.54	0.82	0.82	0.89	0.82	0.75	0.84
	6	0.63	0.30	0.57	0.62	0.55	0.64	0.72	0.60	0.59	0.62	0.65	0.62	0.62
	7	0.85	0.66	0.79	0.86	0.66	0.85	0.54	0.90	0.85	0.85	0.83	0.78	0.86
	8	0.85	0.72	0.83	0.85	0.71	0.85	0.58	0.88	0.96	0.86	0.85	0.82	0.83
	9	0.87	0.67	0.80	0.87	0.71	0.85	0.56	0.80	0.81	0.88	0.85	0.75	0.87
	10	0.86	0.71	0.81	0.86	0.70	0.84	0.60	0.85	0.84	0.88	0.85	0.79	0.86
	11	0.76	0.59	0.77	0.78	0.67	0.77	0.60	0.78	0.79	0.75	0.76	0.84	0.78
	12	0.95	0.74	0.83	0.92	0.72	0.88	0.62	0.87	0.84	0.95	0.93	0.77	0.90

Italic numbers illustrate intra-examiner values.

the DBW radiographs and 79.0% of the NIRT images were correctly assessed. The second evaluation cycle presented a concordant diagnosis of 72.4% for the DBW radiographs and 71.2% for the NIRT images. The level of agreement of the diagnoses in the second evaluation cycle increased for DBW radiographs but was reduced for NIRT images (Table 4). The interexaminer wκ increased from 0.74 to 0.76 for the evaluation of DBW radiographs and decreased from 0.86 to 0.82 for NIRT images.

According to the binomial logistic regression, the diagnostic method and the evaluation cycle significantly influenced the reliability. NIRT images [adjusted odds ratio (aOR) 0.82] had a higher likelihood of being correctly diagnosed, whereas NIRT diagnoses of the second evaluation cycle (aOR 1.12) were more likely incorrect. Clinical experience or the tooth type did not have any significant effect on the accuracy of the diagnosis.

Discussion

In this study, we aimed to investigate the inter- and intraexaminer reliability of DBW radiography and NIRT for caries detection and assessment on proximal surfaces. We hypothesized that both methods exhibited similar results for inter- and intraexaminer reliability. In general, both diagnostic methods revealed good to excellent reliability (Figure 2 and Tables 2–4). When comparing the reliability of both methods, good $w\kappa$ values were obtained for inter- and intraexaminer reliability of DBW radiographs. Good and even excellent wκ values were more frequently obtained for NIRT images. The analysis of inter- and intraexaminer reliability of DBW radiography and NIRT revealed results in or better than the typical range of previously published studies.[15, 20–22]

Considering the results from all κ statistics and the logistic regression analysis, the initially formulated hypothesis must be rejected. NIRT exhibited significantly better inter- and intraexaminer reliability than DBW radiography (Tables 4 and 5). A possible interpretation for this finding might be that enamel caries lesions on proximal sites imaged with NIRT exhibited unambiguous characteristics. Therefore, it might be argued that little room for interpretation existed, as often occurs on DBW radiographs with early lesions. Previous studies have demonstrated that DBW radiographs have less contrast, resulting in relatively low sensitivity values.[23] This observation is consistent with recent experiences from dental practice. For example, enamel caries lesions are easily detected on NIRT images, but the corresponding DBW image does not exhibit any clear indications of the presence of proximal demineralizations. This notion might be a possible explanation for the slightly weaker reliability of DBW radiography in this study.

Furthermore, we analysed the influence of certain co-variables on reliability. Apart from the above-mentioned, significance of the diagnostic methods, the logistic regression model revealed that only the cycle of examination had a significant influence on reliability (Table 5). The first cycle of examination (aOR 1.0) was associated with an increased likelihood of a correct diagnosis compared to the second examination (aOR 1.14). This result was predominately supported by the data of the NIRT method (Table 4). In contrast to this finding, κ and $w\kappa$ values were slightly but significantly higher in the case of DBW radiography. When explaining these findings, it should be noted that one could expect a moderate increase in the reliability between the first and second evaluation cycles in such a study. This effect could be attributed to learning and memory effects, given that the examiners can remember previously observed images. Furthermore, a training course and participation in a study could increase the awareness of each investigator, which may result in better clinical diagnostics. When considering such effects, the results from DBW radiography are consistent with this assumption. In contrast to this finding, the significantly lower reliability values of NIRT images are difficult to explain. An unproven hypothesis could be that all examiners made greater efforts to assess the NIRT images correctly during the first investigation and neglected the importance of the second course. However, this assumption contrasts with the findings of the DBW radiography images.

Table 4 Overall inter- and intraexaminer reliability values for digital bitewing radiography and near-infrared transillumination

Evaluation cycle	*% agreement*	*Reliability with reference diagnoses*			*Intra-examiner reliability*		
		Kappa	*Weighted Kappa*	*p-value*	*Kappa*	*Weighted Kappa*	*p-value*
DBW first evaluation	69.8%	0.60 (0.57–0.64)	0.74 (0.72–0.77)	<0.0001	0.64 (0.61–0.67)	0.77 (0.75–0.80)	0.0001
DBW second evaluation	72.4%	0.63 (0.60–0.67)	0.76 (0.74–0.79)	<0.0001			
NIRT first evaluation	79.0%	0.74 (0.71–0.77)	0.86 (0.84–0.88)	<0.0001	0.68 (0.65–0.71)	0.84 (0.82–0.85)	0.004
NIRT second evaluation	71.2%	0.64 (0.61–0.67)	0.82 (0.80–0.84)	<0.0001			

DBW, digital bitewing; NIRT, near-infrared light transillumination. The percentages of agreement and κ-values were calculated for all investigators in relation to the reference diagnoses.

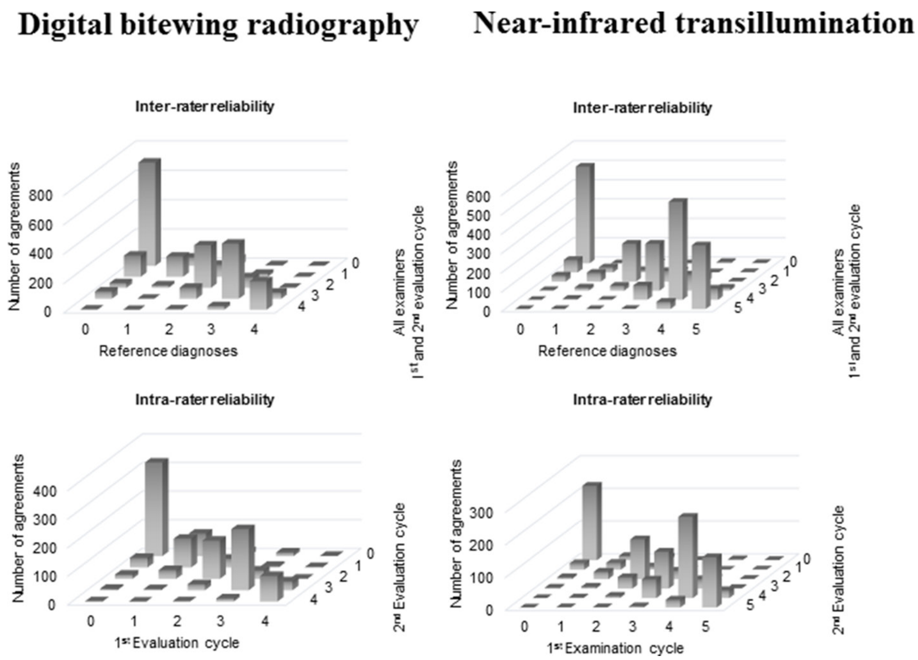

Figure 2 Illustration of inter- and intraexaminer agreement for digital bitewing radiographs and near-infrared transillumination images between the first and second evaluation cycles and the relationship to the reference diagnoses.

Our research exclusively investigated the reliability of DBW radiography and the new NIRT method for proximal caries detection and assessment with inclusion of different co-variables (Table 5). Another strength of our study design is the inclusion of 12 examiners and 100 image pairs. Each evaluation cycle was randomly sorted to eliminate recognizability and memorability effects. Each evaluation cycle was blinded and randomized with a 4-week time interval before the next cycle to avoid any active influence by the examiners. Another unique feature of this study was the use of a logistic regression model to weight the influence of any co-variables. We had performed the power calculation for the DBW data alone using the R package kappaSize, because the sample size calculation for reliability studies is clear when the number of raters are two or more, with an equal number of ratings per subject and a maximum of five categories (DBW with five *vs* NIRT with six categories).[24,25] Having 100 specimens in this study would have a power over 80% even after considering the presence of six categories with the NIRT technique. In addition, one critical argument is that the principal investigators chose all DBW and NIRT images subjectively based on strict quality criteria. We aimed to select unambiguous cases that were not negatively influenced by fuzziness, overlapping or any additional diagnoses. Therefore, this selection strategy might not be representative of clinical practice, where often less-than-perfect images need to be assessed. It could also be argued that the image selection had a positive influence on the documented reliability data. No validation of caries extension was performed as this would require an invasive evaluation of healthy surfaces in vivo or those with non-cavitated caries lesions which can not be justified due to ethical reasons.

Conclusions

In addition to its strengths and limitations, this study revealed a good to excellent inter- and intraexaminer reliability for proximal caries detection using DBW and NIRT images. The logistic regression analysis revealed significantly better reliability for NIRT. In addition, the first evaluation cycle was more reliable than the second according to the reference diagnoses.

Table 5 Adjusted odds ratio with corresponding 95% CI; p-values were computed according to the logistic regression model using backward elimination

Co-variables	*Groups*	*aOR*	*95% CI*	*p-value*
Diagnostic method	DBW	1	-	-
	NIRT	**0.82**	**0.72–0.93**	**0.002**
Evaluation cycle	First evaluation	1	-	-
	Second evaluation	**1.14**	**1.00–1.30**	**0.04**
Clinical experience	0–1 year	1	-	-
	2–5 years	0.96	0.75–1.23	0.76
	>5 years	0.93	0.74–1.18	0.55
Tooth type	Molars	1	-	-
	Premolars	0.94	0.79–1.13	0.51

aOR, adjusted odds ratio; CI, confidence interval; DBW, digital bitewing; NIRT, near-infrared light transillumination.
Bold numbers illustrate a significant influence.

References

1. Poorterman JH, Aartman IH, Kieft JA, Kalsbeek H. Value of bite-wing radiographs in a clinical epidemiological study and their effect on the DMFS index. *Caries Res* 2000; **34**: 159–63. doi: https://doi.org/10.1159/000016584
2. Kidd EA, Pitts NB. A reappraisal of the value of the bitewing radiograph in the diagnosis of posterior approximal caries. *Br Dent J* 1990; **169**: 195–200. doi: https://doi.org/10.1038/sj.bdj.4807325
3. Poorterman JH, Aartman IH, Kalsbeek H. Underestimation of the prevalence of approximal caries and inadequate restorations in a clinical epidemiological study. *Community Dent Oral Epidemiol* 1999; **27**: 331–7. doi: https://doi.org/10.1111/j.1600-0528.1999.tb02029.x
4. Espelid I, Mejàre I, Weerheijm K. EAPD. EAPD guidelines for use of radiographs in children. *Eur J Paediatr Dent* 2003; **4**: 40–8.
5. American Dental Association. Dental radiographic examinations: recommondations for patient selection and limiting radiation exposure council on scientific affairs and U.S. food and drug administration. 2017. Available from: http://www.ada.org/~/media/ADA/Publications/ADA%20News/Files/Dental_Radiographic_Examinations_2012.pdf?la=en [accessed 03.07.2017]
6. European Union Commission. Radiation Protection 136. In: *European guidelines on radiation protection in dental radiology*. Luxembourg: Office for official publications of the EC; 2004.
7. Fried D, Featherstone JD, Darling CL, Jones RS, Ngaotheppitak P, Bühler CM. Early caries imaging and monitoring with near-infrared light. *Dent Clin North Am* 2005; **49**: 771–93. doi: https://doi.org/10.1016/j.cden.2005.05.008
8. Bühler C, Ngaotheppitak P, Fried D. Imaging of occlusal dental caries (decay) with near-IR light at 1310-nm. *Opt Express* 2005; **13**: 573–82. doi: https://doi.org/10.1364/OPEX.13.000573
9. Zijp JR, ten Bosch JJ. Angular dependence of HeNe-laser light scattering by bovine and human dentine. *Arch Oral Biol* 1991; **36**: 283–9. doi: https://doi.org/10.1016/0003-9969(91)90098-F
10. Vaarkamp J, ten Bosch JJ, Verdonschot EH. Propagation of light through human dental enamel and dentine. *Caries Res* 1995; **29**: 8–13. doi: https://doi.org/10.1159/000262033
11. Fried D, Glena RE, Featherstone JD, Seka W. Nature of light scattering in dental enamel and dentin at visible and near-infrared wavelengths. *Appl Opt* 1995; **34**: 1278–85. doi: https://doi.org/10.1364/AO.34.001278
12. Kühnisch J, Söchtig F, Pitchika V, Laubender R, Neuhaus KW, Lussi A, et al. In vivo validation of near-infrared light transillumination for interproximal dentin caries detection. *Clin Oral Investig* 2016; **20**: 821–9. doi: https://doi.org/10.1007/s00784-015-1559-4
13. Söchtig F, Hickel R, Kühnisch J. Caries detection and diagnostics with near-infrared light transillumination: clinical experiences. *Quintessence Int* 2014; **45**: 531–8. doi: https://doi.org/10.3290/j.qi.a31533
14. Marthaler TM. A standardized system of recording dental conditions. *Helv Odontol Acta* 1966; **10**: 1–18.
15. Abogazalah N, Eckert GJ, Ando M. In vitro performance of near infrared light transillumination at 780-nm and digital radiography for detection of non-cavitated approximal caries. *J Dent* 2017; **63**: 44–50. doi: https://doi.org/10.1016/j.jdent.2017.05.018
16. R Development Core Team. *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2016. http://www.R-project.org.
17. Cohen J. A Coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; **20**: 37–46. doi: https://doi.org/10.1177/001316446002000104
18. Fleiss JL, Cohen J. The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33**: 613–9. doi: https://doi.org/10.1177/001316447303300309
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74. doi: https://doi.org/10.2307/2529310
20. Langlais RP, Skoczylas LJ, Prihoda TJ, Langland OE, Schiff T. Interpretation of bitewing radiographs: application of the kappa statistic to determine rater agreements. *Oral Surg Oral Med Oral Pathol* 1987; **64**: 751–6. doi: https://doi.org/10.1016/0030-4220(87)90181-2
21. Senneby A, Elfvin M, Stebring-Franzon C, Rohlin M. A novel classification system for assessment of approximal caries lesion progression in bitewing radiographs. *Dentomaxillofac Radiol* 2016; **45**: 20160039. doi: https://doi.org/10.1259/dmfr.20160039
22. Kositbowornchai S, Basiw M, Promwang Y, Moragorn H, Sooksuntisakoonchai N. Accuracy of diagnosing occlusal caries using enhanced digital images. *Dentomaxillofac Radiol* 2004; **33**: 236–40. doi: https://doi.org/10.1259/dmfr/94305126
23. Kühnisch J, Ifland S, Tranaeus S, Heinrich-Weltzien R. Comparison of visual inspection and different radiographic methods for dentin caries detection on occlusal surfaces. *Dentomaxillofac Radiol* 2009; **38**: 452–7. doi: https://doi.org/10.1259/dmfr/34393803
24. RM A. Sample size estimation functions for studies of interobserver agreement. 2013. Available from: https://CRAN.R-project.org/package=kappaSizekappaSize2013
25. Fleiss JL, Nee JC, Landis JR. The large variance of kappa in the case of different sets of raters Psychological bulletin. *Psychological Bulletin* 1979; **86**: 974–7.