RESEARCH ARTICLE

Dentin reactions to caries are misinterpreted by histological “gold standards” [version 1; referees: 1 approved, 2 approved with reservations]

Priscila Florentino Silva¹, Danilo Augusto de Holanda Ferreira², Kássia Regina Simões Meira², Franklin Delano Soares Forte³, Ana Maria Barros Chaves¹, Frederico Barbosa de Sousa¹,²

¹Department of Morphology, Health Science Center, Federal University of Paraíba, Cidade Universitária, Paraiba, 58051-900, João Pessoa, Brazil
²Laboratory of Microscopy and Biological Image, Health Sciences Center, Federal University of Paraíba, Cidade Universitária, Paraiba, 58051-900, João Pessoa, Brazil
³Department of Clinical and Social Dentistry, Health Sciences Center, Federal University of Paraíba, Cidade Universitária, Paraiba, 58051-900, João Pessoa, Brazil

Abstract

Dentin reactions to caries, crucial for pathogenesis and for the determination of the severity of caries lesions, are believed to be reasonably detected by stereomicroscopy (SM) and polarized light microscopy in quinoline (PLMQ), but accuracies are not available. Here, stereomicroscopy of wet (SW) and dry (SD) ground sections of natural occlusal caries lesions resulted in moderate (0.7, for normal dentin) and low accuracies (< 0.6, for carious and sclerotic dentin) as validated by contrast-corrected microradiography. Accuracies of PLMQ were moderate for both normal (0.71) and carious dentin (0.71). The hypothesis that detection of dentin reactions by SM and PLMQ would be influenced by the contrast quality of micrographic images was rejected. Dentin reactions were scored by SW, SD, PLMQ, and three types of microradiographic images with varying contrast qualities and each technique was compared against the one that resulted in the highest number of scores for each dentin reaction. Large differences resulted, mainly related to the detection of sclerotic dentin by both SW and SD, and normal and carious dentin by PLMQ. It is concluded that contrast-corrected microradiography should be preferred as the gold standard and SM and PLMQ should be avoided, but the relationship of PLMQ with dentin mineralization deserves further investigation.
Introduction
Dentin reactions to caries are crucial for the pathogenesis and severity determination of caries lesions. Since caries is mainly a demineralization process, the high ratio of X-ray absorbance between calcium and the chemical elements of the organic content\(^1\), and the fact that the density of the mineral content is higher (more than 2 times) than that of the organic content\(^2\), radiography with microscopic resolution (microradiography, MR) is considered as a highly reliable gold standard for detecting variations in dentin mineral content. Also widely accepted as gold standards for dentin reactions are stereomicroscopy (SM; commonly referred as histology)\(^3,4\) and (to a lesser extent) polarized light microscopy with quinoline as the immersion medium (PLMQ)\(^5\). The acceptance of SW, the currently most used “gold standard”, is based on studies reporting opaque and translucent dentin under SM related to radiolucent and radiopaque dentin, respectively\(^6-8\). It is lacking, however, accuracy in numbers. To our knowledge, data regarding the accuracy of SM are available from only one study that included unerupted teeth and no data on translucent/sclerotic dentin\(^9\). In addition, early studies with MR reported cases classified as translucent dentin by transmitted light microscopy (where the interaction of light with dentin is similar to that under SM) that were then classified as demineralized dentin by the use of MR\(^10-14\).

Currently the only evidence for the detection of dentin reactions by PLMQ is only qualitative\(^1\). Thus, research into the accuracy of SM and PLMQ is needed. Regarding the use of MR as a gold standard, it must be considered that images of microradiographic plates taken using transmitted light microscopy are commonly biased by the effect of heterogeneous illumination, which is inversely proportional to the objective magnification\(^15\). Heterogeneous illumination is expected to influence judgment of brightness\(^16\) (a procedure required for diagnosis from MR images), possibly including bias when other techniques are validated using MR.

The aim of this study was three fold: to test the accuracy of both SM and PLMQ in detecting dentin reactions to natural caries; to test the hypothesis that elements of accuracy are influenced by the quality of the microradiographic image contrast; and to test the hypothesis that SM, PLMQ and MR (regardless of image contrast quality) detect dentin reactions equally.

Materials and methods
Diagnosis of occlusal caries lesions using ICDAS II
Forty three erupted third molars with various stages of natural occlusal caries were collected from volunteers who signed consent terms (as approved by the Ethical Committee of the Federal University of Paraiba; certificate of ethical appreciation number 4125.000.126-10). All teeth were gently cleaned with 1% hypochlorite solution (Vetec, Brazil), mounted in a wax base and surrounded by a rubber dam isolator prior to analysis of their occlusal surfaces using the ICDAS II scoring system\(^17\). Before obtaining final ICDAS scores for analysis, examiners were calibrated using a sub-set of the whole sample. Thirty occlusal sites were scored by two calibrated examiners (Kappa’s intra-examiner’s scores of 0.9 and 0.89, and 0.85 for inter-examiner agreement), with a one week interval, in order to test intra- and inter-examiner reproducibilities. Final scores were those obtained by a consensus between examiners. In nine teeth, two sites on the occlusal surface were selected, yielding a total of 52 occlusal caries lesions.

Ground section preparation
All teeth were cut longitudinally to their crowns (through their occlusal surfaces) using a diamond disc mounted (Kavo Sorensen, Brazil) in a low-speed handpiece under water irrigation, so that a section of the selected site with a given ICDAS score was obtained. All cuts were then ground using a customized metallic (brass) lapping jig and silicon carbide paper (granulations of 240–1200) under water irrigation to achieve a final thickness of ~100 ± 20 μm. Prepared ground sections (n = 52) were kept in a 0.02% sodium azide aqueous solution until examination.

Calibration of examiners of ground sections and selection of histological sites
At each ground section, histological sites (area of ~150 μm × 150 μm) presenting suggestive signs of normal, carious, or sclerotic dentin were selected. All examinations of ground sections (SM of wet and dry samples, three types of MR images, and PLMQ) were performed by the same two examiners, whose intra and inter-examiner reproducibilities were determined (using Kappa’s statistics) from their scores of all histological sites (of all samples) from each technique obtained with a one week interval. Examiners agreed on the final scores by consulting with each other.

Histological sites were selected from the outer half of the dentin layer, including the area adjacent to the deepest enamel lesion; at least one dentin reaction type per sample was included where possible. Cases were included in the sample when two sites had the same type of dentin reaction detected by SM, while showing different types of dentin reactions when detected using MR. Thus, up to 6 histological sites were selected per sample, yielding a total sample size of 168 sites.

Stereomicroscopy (SM)
SM (10x magnification) with reflected light was used to analyze ground sections under two conditions: wet (SW) and dried (SD; after exposure to 25°C and 50% relative humidity for 2 hours). Temperature and relative humidity were measured just adjacent to the samples. Digital photomicrographs (digital camera Nikon D80) of wet and dried samples were obtained. Dentin reactions were scored as normal, carious discolored (white/yellow/brown), and translucent (“sclerotic”).

Microradiography (MR)
All samples were mounted in a microradiographic plate (resolution of 2000 lines/mm; AGHD plates, Microchrome Technology, San Jose, USA) and exposed to X-rays in a PCBA Inspector (tungsten anode filtered with a 0.25 mm-thick beryllium window, GE, Germany) for 25 minutes using 40 keV and 0.25 mA. Digital photomicrographs of the microradiographic plate were obtained in a transmitted light microscope (2x objective) under different conditions:

Condition 1: using the condenser aligned according to the principles of Kohler illumination for low magnification objectives\(^16\);
Condition 2: no condenser and using a light shaping filter (Luminit, USA) above the field diaphragm.
The possible scores for dentin reactions using MR were: normal dentin, demineralized (radiolucent) dentin, and hypermineralized (highly radiopaque; sclerotic) dentin. Digital images were analysed (using the freeware program ImageJ, NIH, USA) with the following contrast conditions:

Image obtained with aligned condenser, without any adjustment of brightness and contrast from ImageJ, and no light shaping filter (NFNBC image);

Image obtained with aligned condenser, no light shaping filter, but with adjustment of brightness and contrast from ImageJ (NFBC image);

Image obtained without a condenser, with both a light shaping filter and adjustment of brightness and contrast from ImageJ (FBC image).

Such conditions created an ordinary scale of heterogeneous illumination of the field of view. Images without light shaping filter (NF) and FBC images presented a Gaussian normalized light intensity ($R^2 = 0.87$ for both) across the field of view with heights of 0.13 and 0.05 (lower heterogeneity), respectively. Brightness and contrast adjustment (according to a consensus from both examiners) allowed this difference to be easily detected by the naked eye.

**Polarized light microscopy in quinoline (PLMQ)**

Ground sections were dried at room temperature for 24 hours, immersed in quinoline (Vetec, Brazil) for 24 hours, and then were positioned with the dentin tubules at – 45º on the stage of a polarizing microscope (Axioskop, Carl Zeiss, Germany) equipped with a Red I filter, 2× objective, and digital camera (Nikon D80, Japan). Dentin reactions were scored as either negatively (carious) or positively birefringent (normal). Since the technique of PLMQ is not intended to diagnosis dentin sclerosis, no diagnosis of sclerotic dentin was attempted. Color digital images were split into color channels using ImageJ, resulting in a sharp demarcation of negatively and positively birefringent areas.

**Accuracy of SM and PLMQ**

We tested the accuracies of the SW, SD, and PLMQ techniques for detecting dentin reactions using the FBC MR image as the gold standard in all cases. Total positive (TP), total negative (TN), false positive (FP), and false negative values (FN) were obtained and used to calculate accuracy (AC) from:

$$AC = \frac{TN + TP}{TN + TP + FN + FP}$$  \hspace{1cm} (1)

**Effect of MR contrast on the detection of dentin reactions from SM and PLMQ**

Positive (PPV) and negative predictive values (NPV) of each dentin reaction detected by SM were calculated using the following combinations:

1. SW × NFNBC;
2. SW × NFBC;
3. SW × FBC;
4. SD × NFNBC;
5. SD × NFBC;
6. SD × FBC.

The number of a given dentin reaction was that obtained from agreement from both the SM and MR images of each combination. Each SM technique was validated against all MR images types. The PPV was calculated from the ratio of the TP (determined by each MR image type) by the test outcome positive, and NPV was calculated from the ratio of the TN (determined by each MR image type) by the test outcome negative.

In order to test whether PPV and NPV were altered by dentin hydration and or MR contrast, combinations 1–5 were compared with combination 6 (considered as the gold standard combination) and Cohen’s effect size for proportions ($h$) was calculated from:

$$h = \phi_x - \phi_y$$  \hspace{1cm} (2)

and

$$\phi = 2 \cdot \arcsin \sqrt{P}$$  \hspace{1cm} (3)

Where $\phi_x$ and $\phi_y$ are the arcsines (in radians) of the proportions “P” (PPV or NPV), of a given dentin reaction from the test combination (1–5) and gold standard combination, respectively. The same test was performed with PLMQ, using:

1. PLMQ × NFNBC;
2. PLMQ × NFBC;
3. PLMQ × FBC.

Equation (2) was used to calculate the difference between combinations 7–8 and combination 9 (gold standard).

**Effect of “temporary gold standard” on the proportions of dentin reactions**

Dentin reactions were summed separately as detected from each of the six techniques tested and converted to proportions in relation to the number of sites detected by the technique that yielded the highest number (this later ascribed as the “temporary gold standard”). Equation (2) was used to calculate the difference between the “temporary gold standard” proportion ($\phi$) and the test proportion ($\phi$).

**Correlation between translucency and radiopacity**

We tested the hypothesis that translucency under SM was related to the radiopacity in dentin. Five ground sections (with ICDAS scores ranging from 0 to 2) were microradiographed in a digital X-ray machine (Faxitron model MX20, Tucson, USA; tungsten anode...
filtered with a 0.25 mm thick beryllium window) using 20 keV, 0.3 mA and digitally photomicrographed under SM (with dry dentin; SD) with a dark background created with a neutral filter. Under SM, samples presented normal and translucent dentin sites only. Faxitron images had an almost flat illumination across the field of view (Gaussian fit: $R^2 = -0.0004$ and height of 0.028; and linear fit: $R^2 = 0.18$). Eighty histological sites (16/sample) were selected and gray levels were measured on both SM and Faxitron digital images using image analysis software (ImageJ, NIH, USA). The maximum intensity of both image histograms was lower than 80% of maximum intensity allowed. Translucency and radiopacity were measured by:

$$I_{\text{ext}} - \Delta I$$

where $\Delta I$ is the difference (always converted to a positive value) between the intensity of the histological site and the maximum intensity of image histogram ($I_{\text{ext}}$; maximum intensity for radiopacity, and minimum intensity for translucency). Translucency and radiopacity were normalized and then analyzed using Pearson product-moment correlation coefficient.

**Results**

The occlusal caries ICDAS scores for the samples analysed were: 10 with score 0, 2 with score 1, 23 with score 2, 10 with score 3, 1 with score 4, 4 with score 5, and 2 with score 6. For microscopy, intra-examiner agreements were (examiner 1/examiner 2): 0.896/0.914 (SW), 0.952/0.944 (SD), 0.909/0.899 (NFNBC), 0.90/0.91 (NFBC), and 0.919/0.898 (FBC). The inter-examiner agreements were 0.914 (SW), 0.953 (SD), 0.918 (NFNBC), 0.949 (NFBC), and 0.979 (FBC).

Typical aspects of normal, carious and sclerotic/translucent histological points under the six techniques tested here are shown in Figure 1. The number of dentin reactions detected varied among techniques (Table 1). Detection of normal dentin had a moderate accuracy for both SW (0.7976) and SD (0.7976), while both techniques presented a low accuracy for detecting carious (0.5952 for SW; 0.631 for SD) and sclerotic dentin (0.5833 for both SW and SD). Regarding PLMQ, the accuracy was moderate for both carious (0.7092) and normal dentin (0.7902).

PPV and NPV values of SW, SD, and PLMQ for dentin reactions from combinations 1–9 are shown in Table 2. For normal dentin, SW and SD had moderate PPV and NPV, but for carious and sclerotic dentin they ranged from very low to high values. This later behavior was observed for PLMQ from all combinations.

NFNBC, PLMQ, and SD were the “temporary gold standards" for normal, carious, and sclerotic dentin, respectively (Figure 2). Regarding both normal and carious dentin, all comparisons of the “temporary gold standard" with other techniques resulted in large effect sizes. For sclerotic dentin, SM techniques had a small difference, while large effects sizes were measured for comparisons with MR techniques (Figure 2).

The effects sizes of MR image contrast on the PPV and NPV of different combinations of SM and MR was negligible (Table 3). The corresponding values for PLMQ were low: (i) for normal dentin, effect sizes of 0.154 (PLMQ×NFNBC) and 0.043 (PLMQ×NFBC) for PPV, and 0.044 (PLMQ×NFNBC) and 0.0 (PLMQ×NFBC) for NPV; and (ii) for carious dentin, effect sizes of 0.044 (PLMQ×NFNBC) and 0.0 (PLMQ×NFBC) for PPV, and 0.154 (PLMQ×NFNBC) and 0.043 (PLMQ×NFBC) for NPV. Analysis of the relationship between translucency and radiopacity of dentin showed that they were not correlated (Figure 3).
Dentin reactions detected from SM and PLMQ

2 Data Files

http://dx.doi.org/10.6084/m9.figshare.895737

Table 2. PPV and NPV of SM and PLMQ for dentin reactions from all combinations of test and gold standard outcomes.

<table>
<thead>
<tr>
<th>PPV and NPV</th>
<th>Dentin reactions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td>SW × NFNBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>SW × NFBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>SW × FBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>SD × NFNBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>SD × NFBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>SD × FBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>PLMQ × NFNBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>PLMQ × NFBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
<tr>
<td>PLMQ × FBC</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
</tr>
</tbody>
</table>

Discussion

Here we used Cohen’s effect size\(^{19}\) to test differences between outcomes of a given gold standard and a test technique. Effect sizes > 0.5 mean that differences are higher than the acceptable amount to neglect them. In some cases, more than one section per tooth was analyzed. There is conflicting data on whether demineralization can spread laterally to the main dentinal tubules (close to the enamel-dentin junction) or not in non-cavitated carious lesions\(^{20,21}\). If lateral spread does occur, dentin demineralization and sclerosis might occur at random in different sections of a same lesion. In addition, two sites located at different fossae on the same occlusal surface are expected to present independent dentin reactions. This is why we considered more than one section per tooth as independent samples.

This study showed that SM has a low accuracy for detecting carious and sclerotic dentin. This agrees with early qualitative reports showing that translucent dentin can be either caries or sclerosis\(^ {11-14}\). The explanation is that translucency is related to discontinuities in refractive indexes\(^ {22}\), but is not necessarily related to radiopacity (Figure 3). When we tested the assumption that any technique is likely to measure dentin reactions equally, pronounced differences were obtained for all dentin reactions, with the most pronounced difference being that between SM and the other techniques with regard to sclerotic dentin (Figure 2). PLMQ, another technique regarded (without evidence) to be useful for detecting normal and carious dentin\(^ {5}\), showed the most pronounced differences for the detection of normal and carious dentin. To our knowledge, there are no data explaining the relationship between birefringence and dentin mineral content, and this gap impedes further consideration of using PLMQ for detecting dentin reactions. Its use should be avoided until its relationship with dentin mineral content is clarified. The large effect sizes obtained when PLMQ was the temporary gold standard and moderate accuracies of PLMQ are intriguing, and worth further investigation.

Previous studies that have reported qualitative evidence of translucent dentin under SM related to radiopaque dentin\(^ {23}\), and have been cited as the basis for regarding SM as the gold standard\(^ {23}\), performed their analysis with images showing more than 2 mm × 2 mm of the tooth crown. Such a field of view size can only be obtained from low magnification objectives\(^ {24}\). Heterogeneous illumination of the field of view in light microscopy images is inversely related to objective magnification\(^ {15,24}\). Thus, heterogeneous illumination may have been a common factor in their analyses, and most probably our NF MR images (NFNBC and NFBC) are the ones that more closely resemble the images obtained in these older studies. The hypothesis that MR image contrast (influenced by heterogeneous illumination) could explain variations in the PPV and NPV of dentin reactions...
layer combined with translucent dentin in the outer 300 microns of the dentin layer are interpreted as an indication that the carious lesion is confined to the outer enamel and that the dentin has already reacted to it by producing sclerotic dentin. Our results and previous evidence\textsuperscript{11–14} show that this translucent dentin might actually be carious demineralization. This alters the interpretation of the lesion depth and also the understanding of how carious demineralization propagates in the hard dental tissues. In this context, SM data without quantitative evidence of mineral content through the carious lesion might be misleading. In addition, or alternatively, it might be misleading to ignore the possibility that caries formation is a result of two relatively independent events: initial acid infiltration and late demineralization, as shown experimentally\textsuperscript{29}. Cariogenic acid could infiltrate (from the tooth surface) down to the inner enamel and outer dentin prior to demineralization of the entire enamel layer located more externally. Acid could infiltrate into the tooth crown following enamel sheaths (the main pathways for transport of materials in enamel)\textsuperscript{30}, which are large nanochannels at the boundaries of enamel prisms, under the influence of an osmotic gradient\textsuperscript{31} created by the higher organic content found in inner enamel\textsuperscript{32}. Demineralization at the early stages of caries lesion formation could take place at two locations: (i) the enamel surface, and (ii), a bit later than at the enamel surface, the region near the enamel-dentin junction (involving both enamel and dentin). At the enamel surface because it is closer to the source of cariogenic acid. At the region near the enamel-dentin junction because there the osmotic gradient would be minimized so that acid could move more slowly and find more favourable conditions to diffuse to the surfaces of the mineral crystallites surrounding the main pathways in enamel and dentin.

New optical techniques for caries diagnosis have been validated with SM only\textsuperscript{4,33}. The assumption behind such validation is that SM has an acceptable accuracy for detecting dentin demineralization and sclerosis. Our results show that the probability of a correct diagnosis of both carious and sclerotic dentin using SM is low. Clinical visual caries diagnostic systems validated primarily by SM\textsuperscript{34,35} should have their validation tested using MR corrected with regard to heterogeneous illumination. The nature of dental caries should be studied without the bias related to the aspect of dentin under SM.

### Table 3. Comparisons (effect sizes) between tests (1–5) and gold standard (6) combinations of stereomicroscopy and microradiography on PPV and NPV of dentin reactions.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SW×NFNC</td>
</tr>
<tr>
<td>Normal dentin</td>
<td></td>
</tr>
<tr>
<td>PPV</td>
<td>-0.174</td>
</tr>
<tr>
<td>NPV</td>
<td>0.194</td>
</tr>
<tr>
<td>Carious dentin</td>
<td></td>
</tr>
<tr>
<td>PPV</td>
<td>0.233</td>
</tr>
<tr>
<td>NPV</td>
<td>-0.122</td>
</tr>
<tr>
<td>Sclerotic dentin</td>
<td></td>
</tr>
<tr>
<td>PPV</td>
<td>0.190</td>
</tr>
<tr>
<td>NPV</td>
<td>0.044</td>
</tr>
</tbody>
</table>

**Figure 3.** Plot of normalized radiopacity against normalized translucency showing no correlation (black line = Pearson’s correlation fit; $R^2 = -0.013$). N = 80 histological sites.
We conclude that, except for normal dentin, SM has low accuracy for detecting dentin reactions related to caries, and SM and PLMQ accuracies are not influenced by the quality of MR image contrast. FBC microradiographic images should be preferably used as the gold standard for judging dentin reactions.

**Data availability**
Figshare: Dentin reactions detected from SM and PLMQ, doi:10.6084/m9.figshare.895737

**Author contributions**
FBS, AMBC, and FDSF designed the study. PFS collected samples, performed analysis of ICDAS scores, prepared ground sections and performed a pilot study. DAHF and KRSM performed analysis with microradiography, stereomicroscopy and polarized light microscopy. FBS analyzed data. FBS wrote the manuscript, which was revised by AMBC and FDSF.

**Competing interests**
No competing interests were disclosed.

**Grant information**
The first author received a master degree scholarship from CNPq (Brazilian Ministry of Science, Innovation and Technology). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**References**

1. National Institute of Standards and Technology (NIST) X-ray attenuation databases. Reference Source


**Open Peer Review**

**Current Referee Status:** ? ? ✓

---

**Version 1**

Referee Report 13 November 2014
doi:10.5256/f1000research.3418.r6314

✓ Chris Deery
The School of Clinical Dentistry, The University of Sheffield, Sheffield, UK

This is an interesting paper which asks some interesting questions. The methodology is well described.

The title is appropriate. The abstract should be expanded to make the findings clearer. The main text has a minor problem with the use of abbreviations, this makes the paper hard to read particularly pages 5 and 6.

In the introduction there is no mention of other techniques such as micro-CT, ICDAS is just ICDAS not ICDAS II.

Page 4: accuracy is not the best method to use to assess the results, as it tends to inflate the results.

I think on page 4 1.SWxNFNBC etc could be better explained.

The Kappa results on page 5 paragraph 3 might be better presented as a table.

The authors should make a stronger explanation of why they believe microradiology is in fact more valid. Perhaps the problems they suggest are present with stericomicroscopy due to training or interpretation issues. They do not mention any training only calibration, these authors must be aware that reproducibility is very different to validity. This issue should be addressed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

**Competing Interests:** No competing interests were disclosed.

Author Response 03 Dec 2014

**Frederico Sousa**, Federal University of Paraiba, Brazil

The report of approved is greatly acknowledged. The answers to the comments are provided below item by item. In the near future, we will publish a revised version containing the changes resulted from all referee reports.

**Answer to comment on abstract:** The abstract will be changed in the revised version and here follows how it will be written:
“Dentin reactions to caries, crucial for pathogenesis and for the determination of the severity of caries lesions, are believed to be reasonably detected by stereomicroscopy (SM) and polarized light microscopy in quinoline (PLMQ), but accuracies are lacking. We aimed at testing (i) accuracies of SM and PLMQ, (ii) the effect of image contrast of microradiography (MR) on accuracy, and (iii) the hypothesis that SM, PLMQ and MR (regardless image contrast) equally detect dentin reactions. Ground sections (n = 63) of natural occlusal caries (ICDAS scores 0 to 6) were analyzed using MR as the gold standard. SM of wet and dry ground sections of natural occlusal caries lesions resulted in moderate (0.7, for normal dentin) and low accuracies (< 0.6, for carious and sclerotic dentin) as validated by contrast-corrected MR. Accuracies of PLMQ were moderate for both normal (0.71) and carious dentin (0.71). The main differences were related to the detection of sclerotic dentin by both SW and SD, and normal and carious dentin by PLMQ. The hypothesis that detection of dentin reactions by SM and PLMQ would be influenced by the contrast quality of MR images was rejected. In conclusion, contrast-corrected MR should be preferred as the gold standard and SM and PLMQ should be avoided, but the relationship of PLMQ with dentin mineralization deserves further investigation.”

Answer to comments on introduction: We did not mention micro-CT because we understand that micro-CT contrasts sound from carious tissues based on the mass attenuation coefficient of X-rays of the chemical elements, similarly to microradiography. We have responded to referee 1 describing the reasons why microradiography is the best gold standard for detecting variations in mineral content in hard dental tissues. As all responses are available to all, we think this solves the problem satisfactorily.

We used ICDAS II because this term has been used previously in major dental journals (Shoaib et al., 2009).

Answer to comments on materials and methods: in the revised version (the combinations SW x NFNBC etc. will be better explained.

We believe that accuracy, PPV and NPV are reasonable parameters to test the validity of SM to detect dentin reactions to caries. This is in agreement with previous reports that tested micro-CT for detecting demineralization in hard dental tissues affected by dental caries (Soviero et al., 2012).

Answer to comments on results: the reproducibility values of each examiner with regard to SM and MR are presented in the first paragraph of the “Results” section. Examiners were trained in the lab using ground sections of carious teeth before performing analysis with SM and MR. In addition: (i) it can be see in Figure 1 that some points with translucent dentin are radiolucent; (ii) previous publications reported cases of translucent dentin that represented demineralization when examined by microradiography (references 11-14).

Competing Interests: No competing interests were disclosed.
The present paper concerns a relevant topic related to dentin reactions to dental caries lesions. However, the authors’ intention, testing accuracies of methods for detecting these reactions, could be valorized/stressed in this paper. I believe that it is important to readers to know how important is to go beyond in this topic. In addition, considering the rationale, the authors stated that “only evidence for the detection of dentin reactions is qualitative”. I would recommend taking care to avoid misunderstandings, since the methods tested were qualitative, but I suppose they are looking for figures about accuracy of these methods. I suggest clarifying that. Another point to be clarified is the use of microradiography as a reference method. In the text, authors mentioned that “radiography with microscopic resolution (microradiography, MR) is considered as highly reliable gold standard for detecting variations in dentin mineral content.” Is that an authors’ affirmation or is based on a previous background? Another aspect regarding reference methods is that the authors proposed to use the microradiography as gold standard, but, actually, they believed the microradiography was similar to other methods tested (according to the 3rd objective). Why not choose an actual quantitative method for that? The rationale for that could be better developed in the introduction. Was that the idea of authors initially or was that a concept formulated a posteriori (after discussing and analyzing present findings)? Maybe, showing the reasons why opting for posteriori testing methods using a reference you “suspect” that is similar to other methods. I suggest these ideas could be explored in the Introduction.

Concerning methodology, one point remained unclear for this reviewer. In the paragraph “Histological sites were selected from the outer half of the dentin layer, including the area adjacent to the deepest enamel lesion; at least one dentin reaction type per sample was included where possible. Cases were included in the sample when two sites had the same type of dentin reaction detected by SM, while showing different types of dentin reactions when detected using MR. Thus, up to 6 histological sites were selected per sample, yielding a total sample size of 168 sites.” it seems that only specimens in which there was some dentin reaction were selected. Is that true? If no, please, clarify that in the text and do not consider the next part. If yes, I believe this step could have led to spectrum of disease bias and selection bias, as well. If you try to detect the dentin reaction in a pool of histological sites, you tend to have a different performance than if you only selected sites in which the reaction is present. In addition, if you examined the teeth by one of the methods tested subsequently, you tend to favor this method in further examinations (because the selection was based on that). Based on these points, if you really selected only the sites with dentin reactions, I suggest including some information about those points in the Discussion.

Regarding the results, I suggest including the confidence intervals to facilitate the comprehension of possible overlapping of marginal values. Besides, I would include some discussion about the reference standard used and the relevance of detecting this level of demineralization in clinical setting or other applications. If the purpose is choosing the best treatment to the patients, some caution should be had in inferring data about very precise methods for demineralization quantification (sometimes these differences could lead to no difference in clinical approach to be adopted). But, certainly, other purposes exist and the investigation is valid for them. Moreover, I would explore more deeply how the figures showed for methods performance impact on inferences (e.g. which is the contribution of testing predictive values instead of showing only accuracy. Considering these points, I suggest some caution in writing the title.

The authors stated that “If lateral spread does occur, dentin demineralization and sclerosis might occur at random in different sections of a same lesion. In addition, two sites located at different fossae on the
same occlusal surface are expected to present independent dentin reactions. This is why we considered more than one section per tooth as independent samples. I believe that it could be true, but, to reinforce this assertive, I suggest testing the influence of the sampling in the present results or using a statistical approach that consider the clustering effect for the specimens evaluated.

Finally, I would consider separating the analysis for dentin reactions related to caries lesions. Despite being reactions in different ways (desmineralization vs. hypermineralization), caries lesion formation and sclerotic dentin formation are not actually opposite actions. They involve different mechanisms of pathogenesis and not only the quantity of minerals. That is why I believe the purpose of identifying each one of these alterations has its importance independently of the other. This separation, in my point of view, may help the authors in explaining the findings of this study. Additionally, based on this observation, I would also suggest reinforcing that in other parts of this paper.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Competing Interests:** No competing interests were disclosed.

---

**Author Response 28 Jan 2015**

**Frederico Sousa**, Federal University of Paraiba, Brazil

The authors would like to thank the reviewer for her report and the valuable comments in it. As a consensus among all authors, the answers to her comments are provided below. Changes in the manuscript text will be included when a revised version is published in the near future.

**Answer to the first comment:** For the statement “only evidence for the detection of dentin reactions is qualitative” we meant to say that the available studies indicating that SM and PLMQ could be used as gold standards lacked accuracy in numbers. The referee is correct by stressing that we should state more clearly that we were looking for accuracy. This will be included in the revised version.

**Answer to the question on microradiography as the gold standard:** We kindly ask the reviewer to refer to the answer to a similar comment made in the first referee report (from Dr. Soviero), where it was explained the reasons why microradiography is considered a reliable gold standard for changes in the mineral content of dentin.

**Answer to the comment that “the authors proposed to use the microradiography as gold standard, but, actually, they believed the microradiography was similar to other methods tested”**: The belief that SM, PLMQ and MR are similar is found in the literature (see references 3-10). We considered that is important to have their accuracies tested in order to make a scientifically robust decision on whether SM and PLMQ are similar to MR or not. In the Introduction, we stressed that MR is a reliable gold standard (and reasons were given; for more on those reasons, please refer to our answer to the comment of the first referee report on why MR was chosen as a gold standard) and that accuracies for SM and PLMQ were lacking. In summary, SM and PLMQ are currently believed to be similar to MR (a proven reliable gold standard) by almost all of those in the field of histopathology of dental caries, but accuracies were lacking. We hope that highlighting this here might help the reviewer (and the readers, as all comments and responses are open) to understand the reasons why we tested SM and PLMQ against MR.
Answer to the comment “It seems that only specimens in which there was some dentin reaction were selected. Is that true?”: It were selected sites presenting aspects of both normal dentin and dentin affected by caries. This will be clarified in the revised version.

Answer to the comment asking for confidence intervals: The confidence intervals will be included in the revised version.

Answer to the comment asking for cluster analysis: The authors are aware that cluster analysis is receiving growing attention in dental research, and some guidelines for statistical methodology in dental research have been published stressing that cluster analysis (or the reason why clustering was ignored) should be taken into account, mainly when different sites from the same subject are used (Hannigan et al., 2013). In this study we ignored clustering, mainly because the correlation between dentin translucency under SM and dentin radiopacity was very weak (R² = -0.013; see our Fig. 3). It has been stressed that the consequences of ignoring the clustering depends on the correlation that exists in the data (Galbraith et al., 2010). Considering that the two main features of clusters are compactness and isolation, and that correlation coefficient is a similarity index (see pages 11 and 188 in Jain and Dubes, Algorithms for clustering data. Prentice Hall: Engelwood Cliffs, NJ, 1988) – any clusters in our sample would not have enough compactness to be attain validity -, it can be reasonably considered that ignoring the clustering is justified in our study. These considerations will be included in the revised version.

Reviewer Comment: “Finally, I would consider separating the analysis for dentin reactions related to caries lesions. Despite being reactions in different ways (desmineralization vs. hypermineralization), caries lesion formation and sclerotic dentin formation are not actually opposite actions. They involve different mechanisms of pathogenesis and not only the quantity of minerals. That is why I believe the purpose of identifying each one of these alterations has its importance independently of the other. This separation, in my point of view, may help the authors in explaining the findings of this study. Additionally, based on this observation, I would also suggest reinforcing that in other parts of this paper.”

Answer to the comment described above: We did not understand this comment well because all the results are already separated by the type of dentin reaction (normal, sclerotic and carious). We would be happy to read in more details the differences between what is already reported and what the reviewer expects to see.

Competing Interests: No competing interests

Vera Mendes Soviero
Department of Preventive and Community Dentistry, State University of Rio de Janeiro, Rio de Janeiro, Brazil

The manuscript is related to an interesting and relevant issue. The analysis of different gold standards for caries assessment than stereomicroscopy is of great interest, as we all know that the “histological gold
standard” has limitations. However, in my view, certain points need to be clarified. I have carefully revised the manuscript and I hope the comments below are helpful to improve the quality of its final version.

ABSTRACT
I do recommend rewriting the abstract. It should be structured with Aim, Methods, Results and Conclusion in this order. It is not necessary to mention the titles, but it is important to follow this sequence. The aims are not mentioned in the abstract and results are shown before methodology.

INTRODUCTION
1st paragraph, lines 6 to 8
“… radiography with microscopic resolution (MR) is considered as a highly reliable gold standard for detecting variations in dentin mineral content.”

I miss a reference for this statement. If MR was not validated previously for the detection and assessment of dentin caries or dentin reactions to caries, the aim of the study should be changed. Instead of testing the accuracy of stereomicroscopy against MR, it should be testing the accuracy of MR against stereomicroscopy. Reference number 10 (Hintze et al. 1995) concluded that SM was more reliable than MR. References number 11 to 14 (Applebaum et al. 1933; Applebaum 1935; Gottlieb 1946; Gottlieb et al. 1946) don’t seem to refer to validation studies.

AIM
First aim: how to test the accuracy of SM if it is established as gold standard?
Second aim: I didn’t understand. Please, explain better what is “elements of accuracy”.
Third aim: In my view, the third aim summarizes everything.

MATERIALS AND METHODS
Calibration of examiners of ground sections and selection of histological sites
Please, clarify better how calibration was done. What means “Examiner agreed on the final scores by consulting with each other”? Does it mean that examinations were not done independently?

Please, clarify the following sentence: “… included in the sample when two sites had the same type of dentin reactions detected by SM, while showing different types of dentin reactions when detected using MR.” It is not clear for me. If MR had to be different from SM, the results would be completely biased by this “selection criterion”. Or not?

RESULTS
Tables in general
Abbreviations should be explained using notes underneath the tables.

Table 1
Carious and sclerotic dentin was observed in all samples. How do you explain it considering that 10 teeth were scored visually as ICDAS 0? The sums in table 1 vary from 153 (PLMQ) to 171 (NFNBC). Why?

Figure 1
The legend is not clear. Figures are not explained following the sequence of numbers. The arrows should be seen in all pictures.

Table 2
The title is not clear enough. I suggest writing in the title that microradiography was used as gold
standard.

**Table 3**
*In the title, what means (1-5) and (6)? Please, clarify the title.*

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Competing Interests:** No competing interests were disclosed.

Author Response 30 Sep 2014

**Frederico Sousa,** Federal University of Paraiba, Brazil

The authors would like to thank the reviewer for her valuable technical comments, which will certainly contribute to improve the quality of the paper. The answers provided below are listed following the sequence of the reviewer comments and were prepared in common agreement among authors. Shortly we will publish a revised version of the article including the changes resulted from this review.

**Answer to comment on abstract:** We agree with the referee. The abstract will be changed in the revised version and here follows how it will be written:

"Dentin reactions to caries, crucial for pathogenesis and for the determination of the severity of caries lesions, are believed to be reasonably detected by stereomicroscopy (SM) and polarized light microscopy in quinoline (PLMQ), but accuracies are lacking. We aimed at testing (i) accuracies of SM and PLMQ, (ii) the effect of image contrast of microradiography (MR) on accuracy, and (ii) the hypothesis that SM, PLMQ and MR (regardless image contrast) equally detect dentin reactions. Ground sections (n = 63) of natural occlusal caries (ICDAS scores 0 to 6) were analyzed using MR as the gold standard. SM of wet and dry ground sections of natural occlusal caries lesions resulted in moderate (0.7, for normal dentin) and low accuracies (< 0.6, for carious and sclerotic dentin) as validated by contrast-corrected MR. Accuracies of PLMQ were moderate for both normal (0.71) and carious dentin (0.71). The main differences were related to the detection of sclerotic dentin by both SW and SD, and normal and carious dentin by PLMQ. The hypothesis that detection of dentin reactions by SM and PLMQ would be influenced by the contrast quality of MR images was rejected. In conclusion, contrast-corrected MR should be preferred as the gold standard and SM and PLMQ should be avoided, but the relationship of PLMQ with dentin mineralization deserves further investigation."

**Answer to comments on introduction:** The following papers can be used as references for the statement that microradiography is considered as a reliable gold standard for variations in dentin mineral content: Ten Bosch & Angmar-Månsson (1991); and Arends et al. (1997). However, the fundamental reason why microradiography is a reliable gold standard lies on what was stated in the first six lines of the Introduction. Here, we explain those reasons more clearly. Basically (and this applies for the analysis of mineralized tissues by radiography in general), from the components of dentin (mineral, organic matter and water), the mineral content absorbs X-rays more intensively than the other components. The basic parameter is the linear attenuation coefficient (LAC) for X-rays, which is the product of the mass attenuation coefficient (MAC) by the density of the unit cell. Information on MAC and LAC (corresponding to X-ray peak of 24 keV, which is the emission
peak emitted by the X-ray source used here operating at 40 kV and 0.25 mA) for both the mineral and the organic components of dentin are listed below:

1. For the mineral component: with the unity cell’s empirical formula Ca$_{8.856}$Mg$_{0.088}$Na$_{0.292}$K$_{0.010}$P$_{5.592}$Cl$_{0.078}$O$_{25.561}$H$_{3.222}$C$_{0.457}$ (reference 2), the MAC is 3.83 cm$^2$•g$^{-1}$, the density is 2.99 g•cm$^{-3}$ (reference 2), and the LAC is 11.45 cm$^{-1}$.

2. For the organic component, represented by collagen: with the unity cell’s empirical formula C$_{12}$H$_{24}$N$_{3}$O$_{4}$ (Omokanwaye et al., 2010), the MAC is 0.41 cm$^2$•g$^{-1}$, the density is 1.19 g•cm$^{-3}$ (Gautieri et al., 2011), and the LAC is 0.49 cm$^{-1}$.

MACs for the chemical elements of the empirical formulas listed above can be found in reference 1. Gray levels of dentin in MR are the result of the product of the volume of each component by its respective LAC, and the sum of all products is multiplied by sample thickness. Thus, as mineral volume decreases in dentin, gray levels decrease (i.e., dentin becomes less radiopaque), and vice-versa.

With light scattering, which governs the aspect of dentin under SM, is different. By matching the refractive index of an immersion liquid with the refractive index of dentin light scattering is reduced, rendering the tissue more translucent without modifying tissue composition (Fried et al., 1995). In addition, visible light scattering in dentin is highly influenced by the density (Zijp & Bosch, 1993; Kienle et al., 2003) and angulations (Kienle et al., 2003) of dentinal tubules in the sample. And the density of dentinal tubules increase as one approaches the dental pulp (Garberoglio & Brannstrom, 1976).

**Answer to comments on the aims:** The evidences supporting SM as a gold standard for dentin reactions to caries are typically those cited in the review paper by Kidd & Fejerskov (2004) (reference 25); to quote: references 6, 8 and 9, and Johnson et al., (1969). They are not validation studies (i.e., accuracy is not provided), but qualitative studies where microradiographic images are reported without correction for heterogeneous illumination. Previous reports showed that translucent dentin could be either demineralization or sclerosis (references 11-14), and this fact means that the probability of finding similar results is high if SM and MR are compared many times. It is not excluded neither the possibility of finding translucent sclerotic dentin nor the possibility of finding a correlation between SM and MR when there is no translucent dentin. The study by Hintze et al. (reference 10) was performed with unerupted sound teeth, which lack translucent dentin in response to caries. Measurements of lesion size with SM exclude translucent dentin, resulting in underestimation of lesion size if part of the translucent dentin is demineralization. Translucent dentin comprises a significant part of dentin reactions to the carious process (Bjørndal & Thylstrup, 1995; Bjørndal et al., 1999). On this basis, in the “Aims” we considered that further studies were needed to support SM as a gold standard for detecting dentin reactions to caries. The third aim does not include the test of the accuracy of SM. With all due respect, we think that the aims should be kept as they are.

Positive and negative predictive values are considered as elements of accuracy because they describe the performance of diagnostic tests and high results are indicative of good accuracy. In order to put this more clearly in the revised version, PPV and NPV will be specified as elements of accuracy in the aims.

**Answer to comments on materials and methods:** The statement “Examiner agreed on the final scores by consulting with each other” means that the final scores were given by consensus between examiners, just like was done with the visual diagnosis with ICDAS II described in “Diagnosis of occlusal caries lesions using ICDAS II”. When disagreement occurred, the examiners
consulted each to reach a consensus. This is a common approach in the dental literature.

**Answer to comment on the statement "... included in the sample ... when detected using MR."**: This statement was included to say that SM and MR scores were not the same. Histological sites were selected on the basis of MR images. In many cases, sites scored as normal dentin by SM, for example, were not scored as normal dentin by MR. This did not apply for all cases, as stated in the previous paragraph where it is written that “at least one dentin reaction type per sample was included where possible”. This was not possible for all cases. We recognize, however, that the statement is not clear, is misleading, and it will be changed in the revised version.

**Answer to comments on results**: We accept the suggestion of the referee on abbreviations. This modification will be included in the revised version.

**Answer to comments on Table 1**: Regarding the report that 10 teeth with ICDAS score 0 presented carious and sclerotic dentin, below follow reports of cases of occlusal surfaces with ICDAS score 0 that were scored by SM (with regard to the depth of the carious lesion) as having demineralization in the inner 50% of the enamel layer (E2) or in the outer 50% of the dentin layer (D1):

1. Ekstrand *et al.* (1997) reported 05 cases of occlusal caries with ICDAS score 0 and SM score E2 (see Table 8 in their paper);
2. Ekstrand *et al.* (1998) reported 01 case of occlusal caries with ICDAS score 0 and SM score E2 (see Table 4 in their paper);
3. Jablonski-Momeni *et al.* (2008) reported 02 cases of occlusal caries with ICDAS score 0 and SM score E2 (see Table 4 in their paper);
4. Braga *et al.* (2009) reported 01 case of occlusal caries with ICDAS score 0 and SM score E2 (see Table 4 in their paper);
5. Diniz *et al.* (2009) reported 06 cases of occlusal caries with ICDAS score 0 and SM score E2 and 01 case of ICDAS score 0 and SM score D1 (see Table 3 in their paper).

SM score E2 is important because it has been related to the occurrence of translucent dentin (*Bjørndal & Thylstrup, 1995; Bjørndal et al., 1999*), and this later might be either carious or sclerotic (see our results and references 11-14), so that part or all of those cases with SM score E2 listed above could represent carious dentin. On this basis, the occurrence of carious and sclerotic dentin with ICDAS score 0 is not an unexpected result.

The totals of PLMQ and NFNBC differ because, as explained in the section “Polarized light microscopy in quinoline”, PLMQ is not used to detect sclerotic dentin. The total of PLMQ cases is the sum of carious plus normal dentin, while the total of NFNBC cases is the sum of carious plus normal plus sclerotic/translucent dentin.

**Answer to comments on Figure 1**: The fact that figures are not explained following their sequence was intentional. The legend describes the figures following the sequence of the aspects of the sites 1-4 pointed in Fig.1A (wet SM). This way, emphasis is given to the main differences between the wet SM image (commonly used in the literature) and the others. Regarding the arrows, we had prepared a version with arrows in all figures but we thought that it would disturb free interpretation of some images by the reader. This is why arrows were included in half of the images. With all due respect, we still think this way is the best one.

**Answer to comments on Table 2**: We accepted the suggestion of the reviewer and the title will be
changed (in the revised version) to: “Table 2. PPV and NPV of SM and PLMQ for dentin reactions from all combinations of test and gold standard (MR) outcomes”.

**Answer to comments on Table 3:** (1-5) and (6) are the combinations described in the Material and Methods. In the revised version, we will explain them using a footnote in Table 3.

**Competing Interests:** No competing interests were disclosed.