Clinical Effectiveness of Laser Fluorescence, Visual Inspection and Radiography in the Detection of Occlusal Caries

V. Angnes\textsuperscript{a} G. Angnes\textsuperscript{a} M. Batistella\textsuperscript{a} R.H.M. Grande\textsuperscript{b} A.D. Loguercio\textsuperscript{a} A. Reis\textsuperscript{a}

\textsuperscript{a}Department of Dental Materials and Operative Dentistry, University of Oeste de Santa Catarina (UNOESC), Joaçaba, and \textsuperscript{b}Department of Dental Materials, University of São Paulo (USP), São Paulo, Brazil

Key Words
Diagnosis \cdot Laser fluorescence \cdot Occlusal caries \cdot Radiographs \cdot Visual inspection

Abstract
The aim of this in vivo study was to compare a laser fluorescence (LF) device with Ekstrand’s visual scoring system and radiographic assessment for detection of occlusal caries. Thirty-eight adults aged 19–35 years participated in the study; a total of 57 third molars with macroscopically intact occlusal surfaces were selected. Two examiners assessed 110 sites by visual inspection (VI), bitewing radiography (BW) and LF. Teeth were then extracted and caries extent assessed by histology. The detection methods were compared by means of sensitivity, specificity, inter-examiner reproducibility (kappa statistics) and area under the ROC curve. VI and LF had similar (p > 0.05) and superior sensitivities than BW (p < 0.05). VI and BW showed similar specificities, which was superior to LF. The inter-examiner reproducibility was good for VI and BW and moderate for LF. The area under ROC curves showed that VI was better than LF. It was concluded that Ekstrand’s visual scoring system is the most valid method for caries diagnosis. LF should be considered an adjunct to caries diagnosis.

New non-invasive, instrument-based and non-ionizing techniques have been developed and investigated for detection and quantification of demineralization [Egbertsson et al., 1999; Lussi et al., 1999, 2001]. A recent introduction is laser fluorescence (LF), which measures the fluorescence of lesions illuminated by red laser light. As the carious lesion progresses, an increase in fluorescence occurs and may be an indicator of the extent of the lesion.

LF seems to be promising for the detection and quantification of caries lesions on occlusal and smooth surfaces [Lussi et al., 1999; Shi et al., 2000; Sheehy et al., 2001; Başeren and Gokalp, 2003; Rocha et al., 2003]. However, most studies have been conducted in vitro [Lussi et al., 1999; Shi et al., 2000; Attrill and Ashley, 2001; Chong et al., 2003; Francescut and Lussi, 2003; Mendes et al., 2003]. The few clinical studies available on the LF performance have not been validated histologically because of ethical concerns [Lussi et al., 2001; Heinrich-Weltzien et al., 2002, 2003; Anttonen et al., 2003]. Consequently negative results provided by any diagnostic method cannot be confirmed. Therefore more clinical studies are required in order to provide clinicians with a more comprehensive evaluation of the LF device.
Clinical Performance of Diagnostic Methods

For a method to be considered innovative and groundbreaking for caries diagnosis its performance must be superior to the other available methods. Thus, the comparison of this device with conventional diagnostic methods, such as visual and radiographic inspection, seems to be of ultimate importance.

Therefore, the purpose of this clinical study was to compare the performance of LF with the visual-ranked assessment method and bitewing (BW) radiographs for occlusal caries detection.

Materials and Methods

Subject Selection
The sample consisted of 38 adult volunteers (19–35 years old) from Joaçaba, SC, Brazil, who had at least one third molar indicated for extraction. The Ethics Committee in Research, University of São Paulo (44/03), approved the study before the start. Inclusion criteria for teeth in this study were the apparent absence of occlusal restorations and fissure sealants, absence of hypoplastic pits, advanced degree of fluorosis, frank occlusal cavitation and large carious lesions on smooth and approximal surfaces.

A total of 110 sites were selected from 57 third molars (generally at the mesial, central and/or distal fossae). After careful cleaning of the occlusal surfaces with rotating bristle brush and water, a simple drawing was made of the occlusal surface to be analysed, in order to reproduce the groove-fossa system and allow the sites to be identified. BW radiographs were taken for each tooth, using Siemens Heliodent 60B apparatus (Siemens, Munich, Germany) set at 60 kVp and 10 mA. The radiographs were taken with BW positioners (Jon Han-Shin PF 682, Jon Ind., São Paulo, Brazil) and Kodak Insight No. 2 radiographic films (Eastman Kodak, Rochester, N.Y., USA) with an exposure time of 0.6 s. Radiographs were developed consecutively using an automatic processor. An individual drawing was used to locate the precise investigation site on the radiograph in a mesial-distal plane. To record their findings the examiners used individual and separate sheets with the teeth sites identified.

Examination Methods
Caries lesions in the selected sites were assessed by KaVo DIAGNOdent (Kaltenbach & Voigt GmbH & Co., Biberach, Germany), visual inspection (VI) and radiography. Two examiners participated in this study; one of them trained the other on diagnostic procedures using several representative teeth for each visual scoring system and several representative radiographs for each radiographic scoring system according to Ekstrand’s visual scoring system [Ekstrand et al., 1997] described in table 1. The examiners were trained on how to use the LF device, according to the manufacturer’s directions. Both examiners performed the analyses for each individual of the sample on the same day, independently.

VI was performed with patients positioned in a dental chair, with the aid of a light reflector, air/water spray and plane buccal mirror using the criteria shown in table 1 [Ekstrand et al., 1997]. Before the start of the examination, a third person examined the tooth, making a drawing and selecting the site. This drawing was shown each examiner and the findings of each one were recorded by this third person on a separate sheet. Soon after VI, teeth were re-hydrated with water. LF device was calibrated with a ceramic standard provided, as recommended by the manufacturer. The standard value for each tooth was also obtained by measuring the fluorescence in a region of sound teeth located in the buccal surfaces. Sites were evaluated under cotton roll isolation and after

| Table 1. Criteria used for visual, radiographic and histological examination |
|-----------------|-----------------|-----------------|-----------------|
| Score | Visual | Radiographic | Histological |
| 0 | No or slight change in enamel after prolonged air-drying (10 s) | No radiolucency visible | No enamel demineralization or a narrow surface zone of opacity (edge phenomenon) |
| 1 | Opacity or discoloration hardly visible on the wet enamel, but distinctly visible after air-drying | Radiolucency visible in enamel | Enamel demineralization limited to the outer 50% of the enamel layer |
| 2 | Opacity or discoloration in enamel distinctly visible without air-drying | Radiolucency visible in dentine, but restricted to the outer third of dentine | Demineralization involving between 50% of the enamel and third of dentine |
| 3 | Localized enamel breakdown in opaque or discolored enamel and/or greyish discoloration from the underlying dentine | Radiolucency extending to the middle third of dentine | Demineralization involving the middle third of dentine |
| 4 | Cavitation in opaque or discolored enamel exposing dentine | Radiolucency in the pulpal third of dentine | Demineralization involving the inner third of dentine |
tooth-drying with a 3-second air spray. The laser tip was positioned on the target site and rotated around its long axis to allow the recording of the highest value, with no time restriction. Two measurements for each site were made and averaged for statistical purposes. No intra-examiner reproducibility was performed because the teeth were examined on the same day they were scheduled to be extracted.

To carry out the radiographic examination, the examiners placed the films on a viewing box, and a black card was placed around each film to cut out extraneous light. The evaluation occurred in a darkened room without any magnification.

**Histological Validation**

After extraction of the teeth, the sites were hemi-sectioned in a buccal to lingual direction using a 0.3-mm-thick diamond saw mounted in a microtome (Labcut 1010, Extec Co., Conn., USA). Only after sectioning of all sites were they evaluated histologically. Meanwhile, the tooth sections were maintained in saline solution approximately for 5 days. One of the examiners evaluated blindly the two sections of each site under a stereomicroscope with magnification of $\times 25$ and reflected light (SZPT Olympus, Tokyo, Japan). The side with more extensive alterations was re-examined under $\times 40$ magnification and classified according to the adopted criteria (table 1). The depth of enamel demineralization was measured at the area showing the greatest extension of opacity along the direction of the rods. The depth of dentine demineralization was measured at the area where the color changed from brownish/yellowish to grey discoloration along a line at right angles to the enamel-dentine junction towards the pulp. The examiner was not aware of previous diagnostic results.

**Statistical Analysis**

Inter-examiner reproducibility of visual and radiographic ranked scoring systems was assessed using unweighted kappa statistics [Fleiss, 1981]. The same procedure was performed for DIAGNOdent values after categorization of the measurements. Sensitivity and specificity were calculated using the threshold between 2 and 3 for visual and radiographic inspection (table 1). McNemar’s change test was applied to compare the performance of the diagnostic methods for each examiner. Receiver operating characteristic (ROC) analysis was also used to compare the diagnostic performance of the three methods for occlusal caries diagnosis. In addition, a non-parametric statistical test was applied to estimate the significance of areas under ROC curves [Hanley and McNeil, 1983].

**Results**

Histological examination showed that 20 sites were classified as score 0, 24 as score 1, 50 as score 2, 14 as score 3 and 2 as score 4. So, 16 out of 110 sites were classified as ‘carious’ at the dentine level.

The continuum scale that correlates the lesion extent and the range of the DIAGNOdent values was obtained by performing three ROC analyses. This analysis was performed after dichotomization of the histological scores into three cuts-offs: H1 (histological grades 0 and 1), H2 (histological grade 2) and H3 (histological grades 3 and 4). The best corresponding cut-off readings observed were: H1 < 15; H2: readings between 15 and 19 and H3 > 19. LF threshold was set between H2 and H3.

Unweighted kappa statistics showed good inter-examiner reproducibility for VI (0.667) and BW (0.629) and moderate for LF (0.531). Sensitivities, specificities and areas under the ROC curve ($A_z$) are shown in table 2. Both examiners reached the highest sensitivities for VI and LF ($p < 0.05$). The radiographic method was the least reliable in this respect ($p < 0.05$). The specificity values were similar for VI and BW ($p > 0.05$) and both were superior to LF ($p < 0.05$). The area under the ROC curves demonstrated that the VI followed by LF was the most accurate methods for caries diagnosis.

**Discussion**

All detection took place on occlusal surfaces. However, multiple sites were selected on the same surfaces instead of selecting only a specific site or the entire occlusal surface. This approach has been employed in some studies [Pereira et al., 2001; Tonioli et al., 2002; Côrtes et al., 2003], but is not a consensus among researchers [Bader et al., 2002]. In fact during routine clinical practice, the entire occlusal surface is not ‘scored’ per se, since some sites may present caries lesion in different stages. It is
common to make an operative intervention in the distal fossae, with demineralization in the middle third of dentine and only monitor the central and mesial fossa with only enamel demineralization in a lower molar.

According to Bader et al. [2002], one of the most limiting aspects of the literature is that a majority of the studies have been performed in vitro. Clinical studies are therefore encouraged since they increase the strength of the evidence. However, it should be kept in mind that in vivo studies have also certain limitations such as the inclusion only of teeth that can be scheduled for extraction like premolars and/or third molars. This limitation reduces the external validity of the study because first and second molars differ from the premolars and third molars in ways that may affect the performance of diagnostic methods, for instance the extension of well-coalesced fissures [Bader et al., 2002].

Apart from the above limitations, the present study could confirm the excellent reliability of Ekstrand’s visual scoring system for assessment of occlusal caries. Although the sensitivity and specificity were not as high in a Brazilian sample as those published by the originator of the ranked scoring system [Ekstrand et al., 1995, 1997], they are within the range of the highest values ever published for VI [Ie and Verdonschot, 1994]. This is in agreement with an in vitro study that evaluated this visual scoring system [Attrill and Ashley, 2001].

However, the application of this ranked scoring system requires extensive training of representative macroscopic occlusal signs of each score beforehand. This previous calibration procedure is of paramount importance in order to achieve a high accuracy with this method. In fact, one may consider that the bad performance of Ekstrand’s visual scoring system in other published papers [Pereira et al., 2001; Heinrich-Weltzien et al., 2002], including a former study performed by some of the authors of the present investigation [Reis et al., 2004], could conceivably be attributed to the lack of experience and training in using this visual scoring system.

Another important difference in the Ekstrand scoring system is that the validation criteria differ from other authors, since they do not differentiate between histological demineralization limited to the enamel only or when the demineralization extended to the dentine. The histological score 2 combined deep enamel demineralization with initial dentine demineralization [Ekstrand et al., 1997]. This attempt was made based on the close relationship between progressing stages of enamel demineralization and reaction in the underlining dentine, where zones of reactive translucency were observed before establishment of contact between enamel lesion and the enamel-dentine junction [Ekstrand et al., 1995].

LF can be considered a good adjunct to occlusal caries diagnosis, mainly when examiners are not well trained in Ekstrand’s visual scoring system. However, this system cannot be used as a substitute for VI. Although similar results were obtained in terms of sensitivities, LF was not as good as VI in the detection of non-affected sites, as shown by the lower specificities of LF. This was also observed in another clinical evaluation [Alwas-Danowska et al., 2002].

According to Ekstrand’s visual scoring system, the presence of discoloration, visible without air-drying, is an indicator of demineralization involving between 50% of the enamel and the outer third of dentine [Ekstrand et al., 1997], which is considered a ‘sound site’ at the dentine caries level. When discolored sites are evaluated by LF there is a high tendency for sites to be overscored, resulting in a high rate of false-positives. A possible explanation could be attributed to the presence of calculus or organic debris on the occlusal surfaces [Krause et al., 2003]. Nevertheless, as great care was taken to clean the sites beforehand, it is unlikely that this is the reason. Recent evidence has shown that LF tends to overscore discolored sites. At similar lesion depths, discolored fissures resulted in LF median values about 5–7 units higher than in opaque or not discolored fissures [Sheehy et al., 2001; Francescut and Lussi, 2003; Heinrich-Weltzien et al., 2003]. As professional cleaning with rotating brushes cannot remove intrinsic or extrinsic stains or discoloration in the fissure system, this could be the source of the high rate of false-positives obtained with LF and the consequent reduction in specificity.

Had the intrinsic or extrinsic stains been removed by proteolytic solutions, such as sodium hypochlorite, the mean LF values should be lower for discolored fissures [Mendes et al., 2004]. These observations along with the inability of LF to detect in vitro remineralization [Mendes et al., 2003] and the poor correlation of LF readings with mineral loss [Shi et al., 2001] are further evidence that the device does not measure adequately small changes in the mineral content. However, it might reflect changes in organic content [Hibst et al., 2001; Mendes et al., 2003], since the molecules responsible for the caries fluorescence seem to be porphyrins, which are synthesized by several microorganisms in caries lesions [Hibst et al., 2001], and are deposited in discolored fissures [Galil and Gwinnett, 1975].

Thus, LF is not reliable when stained occlusal surfaces are the target of diagnosis and if a treatment decision is solely based on the LF reading there will be a tendency to
overestimate the lesion stage and a higher risk of over-
treatment [Heinrich-Weltzien et al., 2003]. Possibly VI is
more appropriate to the clinician, or at least the dentist
using the LF system must be aware of its limitations be-
fore reaching a treatment decision. Other in vitro and in
vivo studies are also in agreement that VI alone is sig-
ificantly better than LF with respect to specificity [Al-
was-Danowska et al., 2002; Chong et al., 2003].

Another difficulty with LF is the variation in clinical
guidelines from the distributor (KaVo Clinical Guide-
lines) and from in vivo [Lussi et al., 2001; Heinrich-Welt-
zien et al., 2003] and in vitro studies [Lussi et al., 1999;
Shi et al., 2000; Pereira et al., 2001]; this confuses dentists
daily practice. In particular, the performance of LF seems
to be influenced by the cut-off limits chosen to detect den-
tine caries [Heinrich-Weltzien et al., 2003], so a rather
small difference on the LF scale can be a shift from non-
operative to operative intervention.

Although there are only a few scientific and clinical
studies about the cut-off points of LF, a careful examina-
tion of them reveals that LF readings higher than 19–20
indicate dentinal involvement and are currently used as
a threshold between ‘sound’ and ‘caries’ sites [Verdon-
schot et al., 1999; Lussi et al., 2001; Alwas-Danowska et
al., 2002], which is in agreement with the present inves-
tigation. Therefore when LF readings are higher than 20,
it is likely that the lesion has extended into dentine. From
a clinical standpoint, this does not mean necessarily op-
erative intervention since other factors like patient risk
and the LF cannot evaluate the activity of the disease in
the site.

There are other proposed cut-offs for LF in the litera-
ture [Lussi et al., 1999; Pereira et al., 2001], but clinicians
should rely primarily on clinical data, since several fac-
tors may influence the results of in vitro studies, such as
the storage medium [Shi et al., 2001; Takamori et al.,
2001; Mendes et al., 2004] or treatment with agents such
as sodium hypochlorite [Takamori et al., 2001; Mendes
et al., 2004].

The kappa values for inter-examiner reproducibility
are relatively low, particularly for LF. This technique was
supposed to be more objective than visual examination
and therefore the kappa would be expected to be high.
Although the inter-examiner reliability tended to be
somewhat lower than intra-examiner agreement, it is still
higher in other published papers than the value shown in
the present study [Bader and Shugars, 2004]. Although
no obvious explanation can be given for this finding, it is
likely that the subjectivity in the rotation of the probe in
the long axis of the sites could have been responsible for
the moderate reliability of LF, since no attempt was made
to standardize this procedure.

In the present study, BW achieved a very low sensivi-
ty for detection of dentine lesions, while the specificity
was similar to VI and higher than LF. These findings are
not thought to be representative of radiographic perfor-
mance, because the study was only performed in third
molars, which are usually not aligned correctly in the arch
and the position of which makes the correct placement of
BW positioners difficult. As a consequence, sound enam-
el vestibular or lingual to each fissure could have effect-
ively masked any evidence of small and even medium
radiolucency in dentine, and hence impaired caries detec-
tion. A rather low sensitivity and high specificity of con-
ventional radiography was also observed in another study
conducted in third molars [Wenzel and Fejerskov, 1992],
which partially agrees with the previous statement. Be-
sides that, the caries prevalence of the sample was very
low and non-cavitated lesions were intentionally selected.
These details of the study had a profound effect on the
ability of the BW to correctly diagnose ‘caries lesions’
and may be the cause of the low sensitivity of this meth-
od. Had we included more advanced carious sites, or in-
cluded other teeth than third molars, the BW inspection
could have had a better performance [Lussi, 1996].

An advantage of ROC analysis is that it reflects the
diagnostic performance more comprehensively than sen-
sitivity and specificity, which are determined by only one
cut-off point. This analysis also provides an overall valid-
ity of the methods employed. By reference to table 2, it
can be seen that VI was the most valid method for caries
diagnosis followed by LF method. The radiographic in-
spection was considered the least accurate. In fact, al-
though from a mathematical point of view, the ROC anal-
ysis is the best one, since it does not consider any given
threshold, in clinical practice clinicians usually have to
consider a cut-off point in which the treatment options
will fall from non-invasive to invasive approaches. There-
fore, the sensitivity and specificity values are still valuable
tools for comparison of diagnostic methods.

Although the clinical diagnosis of occlusal caries is one
of the most difficult tasks in the field of dentistry, a me-
ticulous VI enables the dentist to score earliest signs of
caries changes or to differentiate between fissures with or
without discoloration, staining or opacities. Clinicians
should bear in mind that diagnosis precedes treatment
planning. As far as carious lesions are concerned, diagno-
sis implies deciding whether demineralization is present,
the depth of the lesion and whether it is progressing rap-
idly or slowly (active caries) or whether it is already ar-
rested. Therefore, the diagnosis of a site as being ‘carious’ does not necessarily mean that operative intervention is required. Even if a microcavity is present, a simple sealing may give room for cleaning and this can arrest the demineralization progression [Handelman, 1991].

Compared to BW, the non-invasive LF device is more reliable for caries detection, as detected by other studies [Shi et al., 2000; Anttonen et al., 2003]. However because of its relatively high cost compared to VI, it should be considered just as a supplementary method for occlusal caries diagnosis.

Acknowledgements
The PIBIC/CNPq (UNOESC/2003) sponsored this study.

References