Reliability of a Fluorescence-aided Identification Technique (FIT) for detecting tooth-colored restorations: an ex vivo comparative study

C. Meller 1 · T. Connett 1,2 · C. Löst 1 · A. ElAyouti 1

Received: 29 August 2015 / Accepted: 14 March 2016 © Springer-Verlag Berlin Heidelberg 2016

Abstract

Objective The aim of the study was to compare the diagnostic predictive values of a Fluorescence-aided Identification Technique (FIT) with those of the conventional diagnostic method regarding the identification of resin composite restorations.

Materials and methods Twenty examiners, 10 students, and 10 dentists were asked to identify composite restorations in a full-mouth model using both the FIT (405-nm light source) and the conventional method in combination with a common diagnostic light source. Each dental examination was repeated three times to calculate inter-/intra-operator agreement, repeatability, and reproducibility using kappa statistics. Predictive values were calculated using a filling prevalence of 42 % and the sensitivity and specificity of each method. Pearson’s test and the 99 % confidence interval (CI) were used for comparison.

Results The sensitivity (97 %, CI 94–99 %) and specificity (100 %, CI 98–100 %) of the FIT were significantly higher than those of the conventional method (27 %, CI 21–31 %, and 65 %, CI 58–72 %, respectively). Consequently, the positive (100 %) and negative (98 %) predictive values of the FIT were significantly higher than those of the conventional method (35 and 55 %, respectively). As expressed by the kappa statistics, the repeatability (0.96) and reproducibility (0.95) of the FIT were significantly better than those of the conventional method (0.49, CI 0.42–0.56, and 0.34, CI 0.26–0.43, respectively).

Conclusion Compared to the conventional technique, the FIT was significantly more reliable as shown by higher sensitivity, specificity, repeatability, and reproducibility values.

Clinical relevance The FIT should be considered as a reliable and practicable alternative in contrast to the conventional method, which was hardly sufficient as a diagnostic procedure.

Keywords Composite resin · Restoration · Filling · Diagnosis · Fluorescence

Introduction

Today, restorative materials are available in various shades and translucencies. In combination with contemporary restoration techniques, tooth-colored restorations can match the color of the natural tooth better than ever before. Because of this, misdiagnosis of tooth-colored restorations occurs more often nowadays. Identification of restorations is getting more complicated, more time consuming, and less reliable. Regardless of good illumination and drying of the teeth during diagnosis, false-positive results are not unusual [1–4] due to the high-quality esthetics of restorations. An increase in false diagnoses of restorations is also expected in case of epidemiological or forensic screening under circumstances where less time and equipment are available [5]. Clinicians’ uncertainty about tooth-colored restorations when recording dental findings of patients presumably leads to a more false-positive diagnosis of a restoration, so that intact teeth with slight enamel defects can more often be wrongly diagnosed as having restorations in situ. The argument that there is no harm done
because insufficient restorations with open or stained margins are easily diagnosed is not acceptable, as false patient records have far more consequences than just a false diagnosis. Drawbacks include false caries risk estimation within the framework of epidemiological data [6] and excessive or unnecessary removal of tooth structure during restorative procedures [1]. Leaving composite remnants in the preparation will lower the quality and the durability of further adhesive restorations [1]. In forensic medicine, false patient records will make the identification of human remains according to dental findings impossible [5, 7–9]. On the other hand, when the restorations are easily identified and differentiated from tooth substances, open and overhanging margins can be better and more efficiently diagnosed and treated. A summary of the described drawbacks resulting from the inability to differentiate composite from tooth structure is listed in Table 1.

A Fluorescence-aided Identification Technique (FIT) could be a powerful diagnostic tool for the identification of restorative materials and can help to easily differentiate them from tooth substance, because the great majority of commercially available modern composites fluoresce differently than the tooth substance [2, 10]. Thus, the FIT allows for a reliable, non-invasive, less time-consuming diagnostic procedure. In a recent study analyzing the single fluorescence properties of a vast sample of available composite shades of several different brands, the excitation wavelength allowing for the best detection of composites was observed at 400 ± 5 nm [10]. Although the FIT has been tested for the identification of composite restorations [3, 5, 7–9], there is scarce information about the reliability of the FIT in terms of clinical accuracy, repeatability, and reproducibility. Moreover, inter-operator agreement when using the FIT is still unclear.

The aim of the present study was to compare the reliability and operator agreement of the FIT with those of the conventional method. The null hypothesis is that there is no difference between the FIT and the conventional diagnostic method regarding the identification of resin composite restorations.

### Materials and methods

#### Tooth models

A full-mouth tooth model was fabricated by mounting 32 extracted teeth in a mandibular and maxillary model at their respective anatomical position. In order to simulate clinical conditions, tooth models were fixed in a dental mannequin (KaVo Dental GmbH, Biberach, Germany). A total of 21 light-cured composite restorations were placed in 16 of the 32 teeth (Fig. 1, left). Cavity preparation was performed using water-cooled diamond burs in a high-speed handpiece. Placement of the restorations followed a standardized filling protocol using phosphoric acid etching (Ultra-Etch, Ultradent Products, South Jordan, USA) and a bonding agent (OptiBond™ FL, Kerr, Scafati, Italy). Composite restorations were hand modeled using a multiple-layering technique and polymerized with a light-curing lamp (Bluephase®, Ivoclar Vivadent AG, Schaan, Liechtenstein). Afterwards, the composite surface was polished using rubber polishers (yellow Identoflex® composit pre polisher, KerrHawe SA, Bioggio, Switzerland) and small self-polishing cup-shaped brushes (Occlubrush®, KerrHawe SA, Bioggio, Switzerland). The teeth were stored in 0.9 % sterile saline solution at 37 °C (±2 °C) for 3 months prior to the study. Widely used, light-curing resin composites were utilized. A detailed description of the materials and shades applied are listed in Table 2. The dentist who placed the composite restorations was not involved in the diagnostic part of the study. The 21 composite restorations were assessed and categorized according to the

### Table 1

Drawbacks resulting from the inability to differentiate the composite from the tooth structure

<table>
<thead>
<tr>
<th>Drawbacks</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>False dental records may lead to improper treatment planning</td>
</tr>
<tr>
<td>2.</td>
<td>In forensic medicine, false patient records may render the identification of human remains difficult</td>
</tr>
<tr>
<td>3.</td>
<td>False epidemiological data for the estimation of caries indices, e.g., deft(s)/DMFT(S) when screening for filled tooth/surface</td>
</tr>
<tr>
<td>4.</td>
<td>False caries risk assessment by means of past caries experience (number and localization of restored teeth)</td>
</tr>
<tr>
<td>5.</td>
<td>False diagnosis of a filled tooth as a sound tooth might cause new pathological findings to remain undetected</td>
</tr>
<tr>
<td>6.</td>
<td>Excessive or unnecessary removal of the tooth structure during restorative procedures and removal of orthodontic appliances or tooth splints</td>
</tr>
<tr>
<td>7.</td>
<td>Composite remnants in the preparation may reduce the quality and the durability of further adhesive restorations</td>
</tr>
<tr>
<td>8.</td>
<td>Overfillings and excess material beyond the margins may not be detected</td>
</tr>
</tbody>
</table>
degree of difficulty concerning their identification with either “easy to moderate” (12 restorations) or “difficult” (9 restorations). The categorization was solely based on visual perception.

As soon as the tooth models were fixed in the dental mannequin, they were kept moist using custom-made trays (*Erkodur*, Erkodent, Erich Kopp GmbH, Pfalzgrafenweiler, Germany) filled with moistened polyester fiber wipes (*Desco Wipes*, Dr. Schumacher GmbH, Malsfeld, Germany). During the examinations, the teeth were rewetted with water spray at intervals of 1 min to prevent dehydration and changes in the color of the teeth.

**Examination procedure**

Twenty examiners, composed of 10 undergraduate students and 10 dentists, were involved in the study. The students were in their final semester of dentistry, and the dentists were working at the Department of Conservative Dentistry of the Eberhard-Karls University, Tübingen, Germany. Participation in the study was entirely voluntary, but color blindness and color weakness assessed by an Ishihara test led to exclusion. The examiners received standardized instructions in the form of an information sheet. Furthermore, the examiners were informed that the results will not be anonymous to ensure that they performed the examination to their utmost ability. Also, it was not allowed amongst the examiners to exchange any information throughout the entire study period (e.g., the number or position of composite fillings). Examiners were asked to record the dental findings of the dental models by drawing the extent of the diagnosed fillings on a dental chart. All examinations were supervised by the same operator to make sure that the procedures were performed according to the study protocol. Examinations were performed under identical light conditions in a dark room illuminated only by artificial light. The instruments for the examination procedures were a dental mirror (rhodium, \(\Theta = 22\) mm), a three-way triple air water syringe for blow-drying, and a sharp double-ended cowhorn explorer (*EXD3CH*, Hu-Friedy Mfg. Co., Chicago, USA).

Each examiner recorded the dental findings using both the conventional method and the FIT, whereby the conventional method had to be completed first. Each method was repeated 20 min later to calculate the repeatability and after 2 weeks to calculate the reproducibility. This resulted in a total of 120 dental records (20 examiners \(\times\) 2 methods \(\times\) 3 repetitions).

The conventional method included the use of illumination from a dental unit lamp (*Kavolux 1410*, KaVo, Biberach, Germany). The FIT setup (Fig. 2) uses light from a fluorescent light-emitting diode (\(\lambda = 405 \pm 7\) nm) transmitted via fiber-optic cable through a reflection lens mounted like a headlamp. The headlight produced a sharply outlined spot light, large enough to illuminate the entire oral cavity from a 40-cm working distance. The examiners wore yellow-tinted eyeglasses (filter lenses \(\lambda \geq 520\) nm) to enhance the contrast of the fluorescence inducing blue-violet light. The FIT setup employed was assembled with the help of the Karl Storz GmbH & Co. KG, Tuttlingen, Germany. The time needed for each examination was recorded.
Statistical analysis

For each method, the sensitivity and the specificity were calculated together with the 99 % confidence intervals (CI) and Pearson’s test for comparison. The positive and negative predictive values for each method were calculated using Bayes’ theorem and a filling prevalence of 42 % [11].

The repeatability and the reproducibility of each method as well as the inter- and intra-operator agreement were evaluated using the kappa statistics. Operator agreement analysis resulted in multiple comparisons of the 20 operators. Therefore, for a comprehensive presentation of the results, the smallest kappa value of FIT and the highest kappa value of the conventional method were presented. For each method, the mean time in minutes needed by the examiners to complete the
examination procedure was calculated and compared using the 99 % CI and the $t$ test.

Contingency analysis was performed for each recorded parameter, and both methods were compared using the Pearson test.

Results

The accuracy of restoration detection (sensitivity) using FIT was 94 % (CI 93–96 %), which was statistically significantly higher than that of the conventional method with 20 % (CI 17–23 %); the Pearson test resulted in $p < 0.0001$. Moreover, using FIT, none of the teeth surfaces were falsely diagnosed as restorations, whereas in the conventional method, 36 % of intact teeth were diagnosed as filled (false-positive results; 1-specificity). Detailed analysis according to the method and the number of filled surfaces is presented in Fig. 3.

Using the conventional method, a detected restoration had the probability of 0.35 or 35 % to be correct (positive predictive value). For the FIT method, detecting a restoration was always 100 % accurate (truly a restoration). Diagnosing intact teeth using the conventional versus the FIT method was correct (negative predictive value) in 55 versus 98 % of the cases, respectively. These values were calculated using a filling prevalence of 42 %. Other predictive values corresponding to a different prevalence rate can be depicted in Fig. 4.

The time the operators needed when using FIT (mean 6:34 min, 99 % CI 6:01–7:07) was statistically significantly shorter than when using the CDM (mean 11:24 min, CI 10:27–12:22). Box and whisker plots showing the duration of each method are presented in Fig. 5.

There was a statistically significant difference between the dentists and the students regarding the accuracy of restoration detection using the conventional method ($p = 0.015$). However, when FIT was applied, no differences between the dentists and the students were detected ($p = 0.17$).

The repeatability of FIT was better than that of the conventional method as expressed by kappa statistics 0.96 (CI 0.94–0.98) versus 0.49 (CI 0.42–0.56), respectively, and similarly was the reproducibility 0.95 (CI 0.93–0.97) versus 0.34 (CI 0.26–0.43), respectively.

The intra- and inter-operator agreement for FIT (minimum kappa 0.89) was significantly higher than that of the conventional method (kappa 0.59). When using the conventional method, students had less agreement than dentists.

Discussion

The FIT method had a significantly higher accuracy than the conventional method in detecting composite fillings (sensitivity) and intact teeth (specificity). Despite a different study design, the results correlated with that of a previous study by Pretty et al., in which photographs of 12 intact and 12 class V-restored premolars were sent to 10 examiners and a similar accuracy of FIT was calculated.

Contemporary composite materials next to the tooth structure are susceptible to illuminant metameric fluorescence failure. A recent study showed that restorations can be easily detected when illuminated with fluorescence-inducing light, because the majority of resin composites achieve a maximum fluorescence higher than that of natural teeth. Even when composites have a maximum fluorescence similar to that of teeth, there are still differences in the corresponding excitation and emission wavelengths [2]. It has been shown that even changes in the UV component of daylight (direct and indirect sunlight) and in artificial light can make resin restorations visible, because of the generated contrast to the adjacent natural tooth structure [12]. This optical fluorescence metamerism has the particular advantage of identifying a restoration or a persistent rest of it, for instance in case of retreatment, allowing a never before accomplished level of diagnostic accuracy as demonstrated in the present study by FIT. Differentiating the filling from the tooth was evident using the FIT (Fig. 1, right). Even though the composites and shades employed in the present study exhibit different fluorescence intensities [10], the identification of the restored teeth was always reliable when using FIT. This was due to the fact that all employed composites fluoresced much more intensely than the natural hard tooth substance. Some concern has been published assuming that the optical fluorescence intensity of resin composites may decrease as the material is aging [13]. Nevertheless, studies analyzing the several different composite brands and their different shades need to be performed first, in order to get further understanding about the consequences,
Fig. 3 Number of identified fillings according to the method and examiner.

Fig. 4 Predictive values for a correctness of detecting a filling in relation to the prevalence as presented by the positive predictive values (PPV) of the FIT (green line) and the conventional method (red line). b Correctness of detecting sound teeth in relation to the prevalence as presented by the negative predictive values (NPV) of the FIT (green line) and the conventional method (red line).
if any. It may well be that even with some reduction of the fluorescence signal, the filling material fluoresces still more than the natural hard tooth substance. Furthermore, in the unlikely case that the filling material loses its fluorescence signal completely, it will still stand out against the natural autofluorescence of the tooth. The sensitivity and specificity are the most commonly published values for diagnostic methods and are based on the pre-known status (filled or intact) of the specimen. However, clinically, the sensitivity and specificity do not give adequate information about the correctness of the diagnosis because the true status is unknown. The predictive values, which are calculated on the basis of the prevalence of the filling in the population, are more informative, since they provide the probability of correctness when we clinically, and not experimentally, diagnose a filling (positive predictive value) or an intact tooth (negative predictive value). The predictive values calculated in the present study were based on a filling prevalence of 42% [11]; however, the filling prevalence can be demographically different. Therefore, predictive plots (Fig. 5) have been provided to obtain the corresponding predictive values with different filling prevalences. The predictive values of FIT showed that a diagnosis of filled or intact teeth (98 and 100%, respectively) is always truly correct. Actually, the main reason for two small-sized class V fillings to be largely undetected was their positions, in one case the buccal surface and in the other the lingual one. These surfaces might be more likely neglected, because examiners tend to look for fillings more likely on occlusal than on cervical surfaces. Other reasons might be that accessibility or illumination is possible only by mirror, which might be neglected especially by inexperienced examiners. In contrast, using the conventional method, every third diagnosed restoration was actually an intact tooth and every second diagnosed intact tooth was actually filled. Apart from the predictive values, another advantage of FIT is the evident recognition of the margins and extension of the filling; this resulted in examiners’ certainty, and it may help to better detect composite overhangs or remnants, e.g., during the removal of old restorations. Using the conventional method, even when a restoration is detected, it is still a challenge to trace the extension/margins of the filling, partially due to the thin margins of the composite and good matching shade of the filling.

The reliability of a method can be described by the accuracy, repeatability, and reproducibility of the diagnosis. Accuracy is the correctness of the method expressed by sensitivity and specificity. The repeatability expresses the ability of the method to convey similar results within a short period of time without changing the examination circumstances, e.g., using similar equipment. The reproducibility gives information about differences in the results after changing the examination circumstances over a long period of time. The high repeatability of the FIT indicated that the method does not need to be repeated to attain higher accuracy. Also, the high reproducibility indicated that the results tend to have the same accuracy at different demographic locations over a longer period of time. In contrast, the reproducibility of the conventional method was not acceptable (kappa 0.34) and it may be more dramatic in circumstances where an optimal setup for the examination is not feasible.

Although the use of fluorescence for the diagnosis of composite restorations has been proposed [3, 5, 7–9], the present work compared, for the first time, the FIT with the conventional diagnostic method with regard to intra- and inter-examiner agreement. Examiners using the FIT showed high agreement among themselves, indicating operator substantial certainty during the identification of the fillings. This also was expressed by the observed marginal differences between the students and the dentists. In contrast, using the conventional method, the dental findings of the second examination deviated greatly from those of the first one regardless of the experience of the operator (student/dentist). Even less agreement was found when the examination was repeated after 2 weeks. However, dentists were more accurate and were able to correctly diagnose more fillings than students.

The poor inter-examiner agreement and the low reliability of the conventional method question the accuracy of previously recorded dental findings. Furthermore, the problem of false diagnosis of restorations is continuously growing with the rise of the number of tooth-colored restorations and the decline of metal restorations. For instance, according to the American Dental Association, in 1990, when amalgam was still the most used filling material in posterior teeth, the estimated annual number of posterior composite restorations placed in private practices in the USA was 13.8 million [14]. In contrast, 15 years later—and in spite of the generalized caries decline observed in the USA—this number increased more than five and a half times to 76.4 million posterior composite restorations and to approximately 122.6 million composite resin restorations (anterior and posterior) placed in private practices annually in the USA alone [15].

Given the problematic use of the conventional diagnostic method, it would be reasonable to speculate that surveys concerning epidemiological data collection of teeth restored with resin composites are nowadays becoming more and more
inaccurate each day. This implies, for instance, that the factor “f/F” of the de/f/s/DMFT/S caries index cannot be reliably recorded. It has to be noted that the calculated accuracy of the conventional method was performed in the present study under optimal diagnostic conditions. Therefore, higher accuracy of the conventional method is expected in epidemiological studies where frequently the examinations are performed in mobile dental clinics or when time pressure may further affect the quality of the examination. In these situations, the FIT could play an essential role, as the method is not only more accurate but also faster. As shown in the results, compared to the FIT, examiners needed almost twice as much time when using the conventional method. Moreover, the FIT is very convenient as no dental explorer (for visual-tactile examination) and/or prior drying of teeth is required, and what’s more, no previous training is necessary. Another advantage of the employed FIT setup is that it can be used as a conventional headlight system that projects a clearly defined light spot, large enough to illuminate the entire oral cavity. Although very bright ambient lighting can reduce the contrast between the restoration and the tooth, the FIT setup is powerful enough that neither saliva nor even plaque should have much impact on the detection of composite restorations. Devices using fluorescence like, for instance, the QLF™ system, SoproLife®, or the VistaCam® which are primarily used for caries detection are not only expensive but also tedious for identifying composite restorations as the inspection of each tooth must be captured with an intraoral camera and viewed on a computer monitor. The use of the FIT does not necessarily require complex and expensive devices, since there are many simple and practical light sources available in a variety of wavelengths. For example, some light-emitting diodes with 400 nm wavelength are not expensive and can be efficiently used for the detection of tooth-colored restorations.

According to the principles of good clinical practice, all available means for better diagnostics and treatment should be utilized. This investigation revealed that by using the conventional method, even a meticulous visual inspection under ideal circumstances did not achieve the diagnostic power needed. Therefore, the conventional method should be considered insufficient as a standard diagnostic tool. The high reliability of the FIT qualifies it, under ethical considerations, to be the standard diagnostic procedure and to be included in undergraduate and postgraduate educational programs.

Conclusions

• The accuracy of detecting composite restorations using the FIT was significantly higher than that of the conventional method as shown by the increased sensitivity and specificity.

• In contrast to the conventional method, the FIT showed high repeatability and reproducibility as well as inter- and intra-examiner agreement.

• Examiners’ experience was not relevant to the high accuracy of the FIT.

• Less time was needed with the FIT.

• The conventional method presented a clinically unacceptable accuracy, and therefore, its use should be critically scrutinized.

Acknowledgments We would like to acknowledge Dr. Klaus Irion and Dr. André Ehrhardt, Karl Storz GmbH & Co. KG, Tuttingen, for their help in assembling the fluorescence illumination device used in this study. Also, many thanks to all the examiners for their meticulous work and to Carmen Buckley for her help with the manuscript and the images.

Compliance with ethical standards The intellectual property of the FIT setup used in the present work is owned by the first author of the present work. The FIT setup used in the present work was selected by the jury of the German Foundation Innovative Dental Medicine (Stiftung Innovative Zahmedizin) as the winner of the Dental Innovation Award 2015 in category 2: “Innovative ideas.”

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

Funding The study was performed and funded by the Department of Restorative Dentistry, Eberhard-Karls University Tübingen.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

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

References