Original Article

Reproducibility of tissue autofluorescence for screening potentially malignant disorders

Jean-Christophe Fricain1,2, Mathilde Fenelon1,2,*, Louise Baschet3, Sylvain Catros1,2, Nicolas Glock1, Sébastien Leric1, Rémi Gaston1, Paul Perez4, Adélaide Doussau4

1 Pôle d’Odontologie et de Santé Buccale, CHU de Bordeaux, Bordeaux, France
2 Inserm, U1026, Tissue Bioengineering, University Bordeaux Segalen, Bordeaux Cedex, France
3 Unité de Soutien Méthodologique à la recherche clinique et épidémiologique, Pôle de Santé Publique, CHU de Bordeaux, Bordeaux, France
4 CHU de Bordeaux, Pôle de Santé Publique, Unité de Soutien Méthodologique à la recherche clinique et épidémiologique, University of Bordeaux, ISPED, Centre INSERM U897-Epidemiologie-Biostatistique, Bordeaux, France

Abstract

Introduction: Direct tissue autofluorescence (AF) visualization devices such as VELscope® are gaining interest to improve early detection of oral potentially malignant disorders (OPMD) and cancers. The main objective of this study was to assess inter-observer reproducibility of incandescent light (IL) and AF observations for OPMD and cancer.

Materials and methods: High risk patients (exposed to alcohol or tobacco) were screened by two independent operators with a conventional oral examination (IL) followed by AF examination. The primary endpoint was the inter-observer agreement on the decision to biopsy assessed by kappa coefficients. Accuracy of IL and AF were estimated by the relative true positive rate (RTPR, increase of sensitivity), relative false positive rate (RFPR, loss of specificity) and their ratio.

Results: 179 patients were included. 833 lesions were identified after IL and AF. Indication for biopsy was retrieved for 41 patients (61 lesions). Inter-observer agreement on the indications for biopsy was 93.3% after IL (Kappa coefficient 0.88 [0.80, 0.97]) and 96.1% after IL and AF (Kappa coefficient 0.78 [0.66, 0.90]). RTPR was 1.2, RFPR was 1 and their ratio was 1.2.

Conclusion: IL and AF examination has shown good inter-observer reproducibility. Adjunction of AF allowed diagnosing more leukoplakia without dysplasia.

Keywords:
oral cancer early detection / oral mucosa / autofluorescence

Introduction

Oral cancer is one of the 10 most common cancers in the world [1,2]. Cancers of the oral cavity and the lips have a world standardized incidence rate of 3.8 for 100,000 person-years (5.3 for men and 2.6 for women) [3]. These cancers account for about 125,000–145,000 deaths per year [2,4], and there is a rising trend in oral and oropharynx cancers observed in young patients worldwide [5]. In 2005, in France, cancers of the oral cavity and the lips were ranked 5th among men and 15th among women. The estimation of the cancer of oral cavity and lips incidence is 6500 new cases per year [6]. The five-year survival rates of oral and oropharyngeal cancer are approximately 50% [7]. In France, the five-year relative survival rate of oral cavity cancer varies from 33% to 38% for men and 45% to 53% for females (only mortality due to oral cancer was taken into account). This survival was not improved for cases diagnosed in recent years [6]. These data underline a poor prognosis, which could be attributed to the delayed diagnosis of oral cancer, related to poor recognition of symptoms or missed diagnosis [8,9]. A pilot study revealed recently a low awareness of oral cancer, and a poor knowledge of its signs and symptoms and risk factors in the spanish general population [10]. The main risk factors of oral cancer are tobacco and alcohol [11,12]. In a population using alcohol and tobacco, it has been demonstrated that a screening program can reduce significantly the mortality rate of 43% for men and of 22% for women [13]: this Indian study emphasizes the importance of a systematic visual oral examination with incandescent light (IL) in the populations at risk.

It has been estimated that the sensitivity of screening by a general practitioner compared to a specialist (gold standard) to diagnose oral cancer was 84.8% and the specificity was 96.5% [14].
To improve the early detection of potentially malignant disorders and cancers, different adjunctive techniques have been proposed. Some are used to predict the malignancy of a lesion (transepithelial brushing, DNA testing, molecular markers), the other allow improving visual detection of malignant lesions (toluidine blue staining, chemiluminescence, fluorescence). These adjunctive methods are still controversial and biopsy remains the “gold standard” [15]. However, several studies have shown that dysplastic and cancerous lesions may be present while the mucosa appears clinically healthy [16]. Few case reports have shown that toluidine blue, chemiluminescence or autofluorescence have detected these lesions invisible with IL [16–21]. Compared to other techniques, visualization of tissue autofluorescence with VELscope system® (AF) is particularly suited to identify lesions of the oral mucosa. Clinical examination is simple, non-invasive and inexpensive with this technique. Moreover, preliminary studies have shown that the examination of tissue autofluorescence (AF) could be used to distinguish the healthy mucosa, dysplasia and cancer with a sensitivity of 98% and a specificity of 100% [22]. Autofluorescence may also improves the detection of precancerous and cancerous lesions in the periphery of malignant tumors, and might be useful to better define surgical margins [21]. Some clinical case reports have shown that the AF allowed diagnosing dysplastic and cancerous lesions not visible with IL [20,23].

More recently, AF has been evaluated in patients having oral mucosa lesions: these patients were referred to a specialist in order to diagnose oral cancer and/or dysplasia. All of these studies have concluded that VELscope system® enhanced the visibility of the lesions but was unable to distinguish high and low risk lesions [19,21–26]. Only five studies have determined the usefulness of the AF in a screening program. Four of these studies were done in the general population, with patients consulting for dental treatments [27–30]. One study was performed in a high-risk population, having a history of smoking, alcohol use and/or head and neck cancer [31]. In general practice, one study has shown an increase of epithelial dysplasia diagnosis using AF [27] and another study suggested that IL is more efficient than AF to differentiate benign and malignant lesions [29]. One study concluded that AF can help with the detection of the site of biopsy [30] and another study stated that AF is useful in facilitating the clinical decision to refer suspect lesions [28]. In high-risk population, the authors have concluded that AF improved the sensitivity of IL and that it could be a useful tool in high-risk patients screening [31]. However, some questions remain unanswered concerning the reproducibility and the accuracy of IL followed by AF (IL and AF) in the screening of high-risk populations.

Therefore, the main objective of this diagnostic study was to assess the inter-observer reproducibility of IL and AF in a screening setting of potentially malignant disorders and cancers of the oral mucosa, in a population of alcohol and tobacco users. The secondary objective was to assess the potential contribution of AF added to IL to improve the diagnostic accuracy in this population.

Material and methods

Design of the study

This was a prospective cross-sectional paired diagnostic study. All patients were observed twice, in a random order for each subject, by two independent operators blinded to each other. Conventional oral examination was first carried out under IL. This was followed by AF examination using the VELscope system®. The two operators examining each patient had a different experience. The junior was the less experienced and the senior was the more experienced. Five operators participated to the study. The study was approved by institutional Research and Ethics Committees of University Hospital of Bordeaux, France (2010-A00262-37) and it was registered on ClinicalTrials.gov (NCT01167790).

Study population

180 consecutive patients with a history of alcohol use and smoking habit were screened in the department of Oral Medicine of Bordeaux University Hospital (Bordeaux, France). The eligibility criteria were: age greater than 18 years, affiliation to the French medical plan, use of alcohol exceeding 14 drinks per week for women and 21 drinks per week for men and active smoking or smoking cessation for less than one year. Exclusion criteria were known oral precancerous condition, dysplasia or malignant lesions of the oral mucosa, hemostasis impairment (thrombopenia <50,000, international normalized ratio >3, haemophilia and willebrand disease) and recent use of acetyl salicylic acid (>3 g during the previous week). The patients received information about the study, provided signed informed to participate. An interview was conducted to collect age, gender, level of education, date of last visit to their general dentist, tobacco smoking habit (number of pack per year, including smoked and non-smoked tobacco), alcohol use (number of drinks per day), and the Alcohol Use Disorder Identification Test (AUDIT).

Study procedures

The standardized examination of the oral cavity (IL and AF) was performed according to the diagram of Roed-Petersen and Renstrup [32] adopted by World Health Organization (WHO). First, elementary lesions and suspicious lesions were recorded after IL examination. Suspicous lesions were potentially malignant disorders [33], or lesions suggesting a malignant process as described by Scully C [34]. Second, intraoral examination was done through the VELscope® handpiece and all measurements were repeated. According to the manufacturer, abnormal oral mucosa appeared dark (loss of autofluorescence), orange, or with increase fluorescence, whereas healthy tissue showed an apple green fluorescence. Any suspicious lesion detected was noted IL+ or AF+ when identified using IL or AF respectively. Conversely when IL or AF were not detecting a suspicious lesion, lesions were categorized IL− or AF−, respectively. A suspicious lesion detected by
at least one operator, and identified as IL+ or AF+, had to be systematically biopsied to get a histological confirmation (gold standard) and the decision was made after completion of the examination of the two observers to enable blinded assessment between observers. Biopsies were made under local anesthesia, and referred to the pathology laboratory following the conventional procedure of the University Hospital of Bordeaux. The diagnosis of potentially malignant disorders, dysplasia and cancer were conducted according to the workshop coordinated by the WHO Collaborating Centre for Oral Cancer and Precancer in the UK [35].

The patients diagnosed with cancers were referred to a center specialized in the treatment of oral cancer and a systematic follow up was established for the potentially malignant disorders.

**Statistical methods**

The primary endpoint was the inter-observer agreement on the decision to biopsy at least one lesion after examining with IL and AF (Cohen’s kappa coefficient and its 95% confidence interval). Secondarily, kappa coefficients of the agreement on the decision to biopsy after examination with IL or with IL followed by AF were compared [36,37].

Relative accuracy was studied using the results of the senior observer. IL and AF accuracy was estimated by the relative true positive rate (RTPR, increase of sensitivity; a RTPR >1 reflects the ability of AF do detect more cancers/precancers that IL), relative false positive rate (RFPR, loss of specificity; a RFPR >1 would indicate that AF leads to unnecessary biopsies) and their ratio [38–41]. A result was considered true positive when the test identified it as suspicious and the lesion was confirmed positive by the histological examination (presence of actinic cheilitis, dysplasia or carcinoma). Conversely a result was considered false positive when a test indicated it as suspicious but it turned out to be benign on histopathological findings. Leukoplakia without dysplasia were first excluded. The estimations were also performed in a complementary analysis including compatible histological results with the clinical diagnosis of leukoplakia without dysplasia as defined by Warnakulasuriya et al. [33].

The sample size was computed in order to test whether concordance between the two observers after AF examination was superior to 0.6 (expected kappa coefficient of 0.8, 80% power, alpha risk of 5%). Given that the prevalence of positive AF lesions was unknown, we computed two estimates: for an expected prevalence of positive AF lesions of 25%, 165 patients would be needed (335 patients for a prevalence of 10%). Sample size was calculated with the software N Query Advisor v6.0. The statistical unit was the patient. Quantitative variables were expressed in terms of mean, standard deviation (SD), median, first and third quartiles (Q1–Q3). Results for qualitative variables were expressed in terms of frequency, percentage, and 95% confidence interval around percentage. Statistical analysis were performed using SAS V9.2.

**Results**

Among the 180 patients enrolled between July 2010 and March 2012, 179 patients were included (one patient was secondarily excluded for not respecting eligibility criteria regarding curatorship). The average age was 48.4 ± 10.5 years and 80% were men. 34% of the people had completed high school education. The percentage of patients hospitalized for alcohol withdrawal was 77.5%. Regarding tobacco smoking, the average number of cumulated packs-a-year was 32.6 ± 20.3 and 24% of the patients declared using cannabis. The median number of alcohol drinks consumed per day was 15 (Q1–Q3: 7–25). The average total score on AUDIT test was 29.1 ± 7.8. The median time since the last visit to a dentist was 16 months (Q1–Q3: 3–42).

After examination of the two independent observers, a total of 752 different lesions were identified using IL and 721 lesions were identified with AF. When combining both techniques and both examiners, 833 lesions were observed. The number of lesions reported with IL and AF was approximately the same. However the adjudication of AF allowed to diagnose 68 additional lesions for seniors and 69 additional lesions for juniors. Among the 23 anatomical sites identified for the study, the most frequent localizations were cheeks, hard palate, dorsum of the tongue and lower lip vermilion border (regrouping 56.7% and 60.1% of lesions observed by the junior with IL and AF, respectively). The topographic distribution was similar for the senior. Elementary lesions are summarized in Table I. “White patch” was the more frequently diagnosed elementary lesion with IL, and the juniors noticed more “white patches” than seniors. Observations made with AF were mainly “loss of fluorescence”. Juniors more frequently observed loss of fluorescence than seniors (Tab. II).

A total of 61 lesions found in 41 patients were identified as IL+ and/or AF+, and had an indication for biopsy. Thirty-six biopsies were performed (the 25 other were not be biopsied due to participant refusal). Eight biopsies revealed a histological

**Table I.** Distribution of the number of elementary lesions observed with IL by senior and junior operators. One lesion could be described by several elementary types of lesions.

<table>
<thead>
<tr>
<th>Elementary lesion</th>
<th>Senior</th>
<th>Junior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erythema</td>
<td>102</td>
<td>98</td>
</tr>
<tr>
<td>Erosion</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>Ulceration</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>White patches</td>
<td>155</td>
<td>190</td>
</tr>
<tr>
<td>Pigmentation</td>
<td>72</td>
<td>74</td>
</tr>
<tr>
<td>Nodule</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Vegetation</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Pseudo-membranous</td>
<td>66</td>
<td>81</td>
</tr>
<tr>
<td>Crust</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Other</td>
<td>56</td>
<td>93</td>
</tr>
<tr>
<td>TOTAL</td>
<td>485</td>
<td>588</td>
</tr>
</tbody>
</table>
Table II. Tissue autofluorescence characteristics of all lesions observed by senior and junior operators.

<table>
<thead>
<tr>
<th>Tissue autofluorescence</th>
<th>Senior</th>
<th>Junior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of fluorescence</td>
<td>196</td>
<td>205</td>
</tr>
<tr>
<td>Increase of fluorescence</td>
<td>114</td>
<td>180</td>
</tr>
<tr>
<td>Orange fluorescence</td>
<td>77</td>
<td>86</td>
</tr>
<tr>
<td>Loss and increase of fluorescence</td>
<td>81</td>
<td>93</td>
</tr>
<tr>
<td>Loss and orange fluorescence</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Increase and orange fluorescence</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>472</td>
<td>574</td>
</tr>
</tbody>
</table>

Table III. Inter-observer agreement on the decision to biopsy at least one lesion after examining the patient with IL and AF.

<table>
<thead>
<tr>
<th></th>
<th>Senior+</th>
<th>Senior—</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior+</td>
<td>34 (19%)</td>
<td>3 (1.7%)</td>
</tr>
<tr>
<td>Junior—</td>
<td>4 (2.2%)</td>
<td>138 (77.1%)</td>
</tr>
</tbody>
</table>

A relative false positive rate of 1 (9/9). The ratio of RTPR/RFPR was therefore 0.9 for malignant lesions and 1.2 for the definition of the lesions to screen extended to leukoplakia without dysplasia. Only one biopsy of normal tissue was wrongly performed due to adjunction of AF in a patient that was correctly screened otherwise for another lesion.

For the juniors, the adjunction of AF led to diagnose two additional leukoplakia, one actinic cheilitis and one moderate dysplasia in four patients. The lesions of moderate dysplasia observed only with autofluorescence by a junior were not seen by the senior (Fig. 1). In five participants a treatment was recommended, and a treatment was documented for one participant.

Discussion

This is the first study designed to assess the inter-observer reproducibility of IL followed by AF in a screening situation of potentially malignant disorders and cancer of the oral mucosa, in a population of alcohol and tobacco users. Only one study has mentioned 60.7% agreement on the VELscope® diagnostic made in a population of patients referred to an oral medicine specialist for white or white and red lesion [23]. In our study, the inter-observer agreement on the indications for biopsy at least one lesion after IL and AF was 96.1% with a Kappa coefficient of 0.88 [0.80, 0.97]. It was significantly superior to the minimum threshold of 0.6 to consider the method reproducible, although adjunction of AF didn’t improve significantly the reproducibility compared to IL. The difference in agreement between the two studies was probably due to differences in evaluation criteria.

In our study, junior and senior were compared because VELscope® was designed for all practitioners, specialized or not in oral medicine. The Kappa coefficient indicates an agreement on biopsy decision between juniors and seniors nearly perfect according to Landis and Koch scale [42]. This result indicates a good interobserver agreement on the indication of biopsying a lesion, and is in favour of using VELscope® for oral precancerous lesions and cancer screening. Moreover VELscope® is easy to use and safe.
However, reliability and validity of the VELscope® to diagnose precancerous and cancerous lesions is discussed [15]. VELscope® has been used as a screening tool in general practice [27,29] but there is no robust evidence to support the use of autofluorescence in primary care environment [43,44]. Another study has used VELscope® to detect malignant and pre-malignant lesions in a high risk population having a history of alcohol and tobacco abuse as well as a suspicious lesion or with a history of head and neck cancer [31]. This study was performed on a heterogeneous population of patients with known lesions, thus it is more a diagnostic study than a true screening on a population at risk. Other studies have been made on population with known lesions such as potentially malignant disorders [23,26], innocuous lesions [45], white and red patches suspicious of potentially malignant disorders [24,46], or benign and suspicious lesions [25,47,48]. All these studies were designed to assess whether the VELscope® allowed differentiating benign and malignant lesions already known, which is quite different from screening.

Our study is the first made in a screening purpose, in a high risk population for oral cancer, but without known lesion. Twenty-three per cent of patients had at least one lesion with an indication of biopsy. The high prevalence of suspicious lesions in this study is due to the choice of the study population. A similar prevalence of potentially malignant lesions was observed in another study where a group having a high risk of oral cancer was studied [49].

Other older studies report prevalence of potentially malignant lesions ranging from 0.4% to 19% [50–54]. The high prevalence of potentially malignant lesions and cancers observed in the population of patients hospitalized for alcohol withdrawal and the easiness of an oral examination are arguments in favor of conducting a systematic examination of the oral cavity of these patients during their hospitalization. The advantage of adding fluorescence to visual examination in screening programmes had not been assessed in the literature before [43,44]. All the previous studies using autofluorescence for oral mucosa examination had the objective to evaluate sensibility and specificity of examination with AF versus IL. In this study we have made a comparison between IL and IL associated with AF because VELscope® is a complementary examination of IL.

In all the studies published before, the gold standard to evaluate VELscope® was the result of biopsy because those were diagnostic studies. In case of screening studies, the gold standard is usually the specialist in oral medicine because it is not possible to practice a biopsy of all elementary lesions as research biopsies without some expectations of potential clinical benefit could be considered ethically problematic [15]. In this study we have used a screening tool to detect lesions in a subclinical stage, so, the gold standard after clinical screening remains to recommend a confirmatory biopsy, and because it was not acceptable to biopsy tissues that did not look suspicious by any technique, we might have missed some subclinical lesions and the true rate of non-diseased patients cannot be ascertained. This is the reason why IL and IL associated with AF were compared by estimating relative sensibility, and relative specificity [38–41], given that sensitivity or specificity cannot be ascertained in this case. Using a very restrictive definition of the lesions to screen, the addition of AF to IL did not show an increase in relative sensitivity. However the addition of AF to IL allowed diagnosing more potentially malignant disorders for juniors and seniors, when the definition of the lesions to diagnose was extended with inclusion on leucoplakia without dysplasia. Nevertheless, it was not possible to conclude whether this result is due to hazard or not because the study was not powered to test this secondary objective. Another limitation of this study is that 25 clinically suspicious lesions were not biopsied, thus our estimate of the prevalence of cancerous/precancerous lesions could be underestimated. This might represent challenges in any screening procedure, and might be exacerbated in alcoolo-tabagic patients who might be reluctant to undergo preventive procedures. Likewise, out of five patients for which a treatment was indicated, we could document a treatment procedure for only one of them. We observed a higher number of lesions after AF examination. This could be due to the AF technique, but the longer time dedicated to the observation of oral mucosa when AF was used could have contributed, independently of the technique used. This study seems to indicates that the addition of fluorescence might increases the sensitivity of the examination, as assessed in the literature [31]. In this study, only one biopsy was performed incorrectly due to the addition of AF for the seniors. The vast majority of additional lesions diagnosed with increased fluorescence were leuoplakia without dysplasia. So, as others we believe that VELscope® can enhance the visibility of the lesions but was unable to distinguish between high risk and low risk lesions [23–26].

Conclusion

In conclusion, VELscope® showed to have a good interobserver reliability to decide which lesions to biopsy, as a screening tool of potentially malignant disorders and cancer in a high risk population. Although VELscope® seems to enhance the detection of certain lesions, mainly leuoplakia, the performance to diagnose potentially malignant disorders and cancer in a screening situation remains to be demonstrated in a specific study.

Acknowledgements. The authors acknowledge D Maldonado for data management, H Savel for participation to the study design, M Meddeb as clinical research assistant, The “Direction de Recherche Clinique du CHU de Bordeaux” for technical assistance and the French National Cancer Institute (InCa) for funding of the study (Convention de recherche n°2009-348-Project FLUO-K).

References


