

# Evaluation of oral mucosal diseases: inter- and intra-observer analyses

Yehuda Zadik<sup>1</sup>, Hadas Orbach<sup>1</sup>, Amy Panzok<sup>2</sup>, Yoav Smith<sup>3</sup>, Rakefet Czerninski<sup>1</sup>

<sup>1</sup>Department of Oral Medicine, Hebrew University-Hadassah School of Dental Medicine, Jerusalem, Israel; <sup>2</sup>University at Buffalo School of Dental Medicine, Buffalo, NY, USA; <sup>3</sup>Genomic Data Analysis Unit, Hebrew University School of Medicine, Hadassah Medical Center, Jerusalem, Israel

**BACKGROUND:** Essential to the diagnosis and management of oral mucosal diseases are visual evaluation and monitoring of the oral cavity. Digital photography has recently become an important clinical tool. The aims of this study were to (i) evaluate inter- and intra-observer differences between oral medicine and other dental specialists when assessing changes in oral mucosal pathology and (ii) assess the influence of calibration labels incorporated into the clinical images on the reliability and consistency of evaluation.

**METHODS:** Ten oral medicine specialists (OM) and 10 other dental specialists (DS) participated in the study. Pairs of images captured with an intraoral camera at two time points from 17 cases of mucosal diseases were presented to the participants. Each pair of photographs was presented with a calibration label (showing length and white to black hues) and again without one. The participants were asked to respond to a questionnaire evaluating changes in size, color, location, and severity for each image pair.

**RESULTS:** Oral medicine specialists had better absolute agreement and consistency than DS when evaluating the changes in the images; however, these parameters did not exceed 52% in either group. The incorporation of a calibration label in the images increased agreement and consistency of evaluations in both groups.

**CONCLUSIONS:** Follow-up evaluations of oral mucosal lesions are performed better by oral medicine practitioners compared to other dental specialists. The incorporation of a calibration label in the clinical images seems to enhance evaluation.

J Oral Pathol Med (2012) 41: 68–72

**Keywords:** digital photography; education; oral diagnosis; oral diseases; oral mucosa; oral pathology

## Introduction

Visual examination is a primary diagnostic and clinical tool in dentistry and oral medicine. The management of potentially malignant and other mucosal disorders in the oral cavity includes periodic examination and observation. Thus, evaluation of lesions activity, treatment outcome, and treatment decisions, are often based on clinical observation (1). Accurate and reproducible assessment of the lesion parameters including exact location, texture, color, and size is critical to patient care and follow-up. Changes in one or more of these parameters often indicate the need for intervention. Documentation using color photographs and videos of the lesion are invaluable aids to long-term follow-up and are recommended in the management of oral lichen planus and other potentially malignant disorders (2, 3).

During the last decade, the intraoral camera has been introduced as an auxiliary tool for clinical diagnosis and follow-up of mucosal lesions. Regarding skin lesions, clinical images providing baseline information seem to improve diagnostic accuracy at follow-up evaluations (4); however, the clinical evaluation of oral mucosal lesions has hardly been studied.

Torres-Pereira et al. (5) found that when two practitioners evaluated photographs of oral mucosal diseases, they only gave the correct diagnosis from a clinical image in 60% of the cases. Scully (6) notes that the clinical evaluation of potentially malignant disorders is difficult and emphasizes the importance of clinical experience for accurate diagnosis. Moreover, despite the consensus regarding the importance of follow-up of oral mucosal lesions, the literature does not address the question of whether these periodic clinical observations need to be performed by skilled professionals or specialists (such as oral medicine/pathology specialists) or whether they can be performed by other dental practitioners as well (3).

The aims of this study were to (i) evaluate inter- and intra-observer differences between oral medicine and other dental specialists when assessing changes in oral mucosal pathology using clinical images and (ii) assess

Correspondence: Rakefet Czerninski, Department of Oral Medicine, Hebrew University-Hadassah School of Dental Medicine, PO Box 12272, Jerusalem 91120, Israel. Tel: 9722 6776141, Fax: 9722 6447919, E-mail: rakefetc@hadassah.org.il

Accepted for publication July 11, 2011

the influence of calibration labels in the clinical images on the reliability and consistency of evaluation.

## Methods

### Participants

Twenty specialists participated in the study, 10 oral medicine specialists formed the Oral medicine (OM) group. The dental specialists (DS) group included three periodontists, three pedodontists, two endodontists, and two specialists in public health dentistry. All participants had at least 3 years of postgraduate training and were active in their fields of specialization at the time of the study.

### Image presentation and questionnaire

Pairs of images obtained using an intraoral camera (ViperCam; Integra Medical, 311 Enterprise Drive, Plainsboro, NJ, USA) at two appointments from 17 cases of mucosal diseases were presented. Each pair of photographs was presented twice, once with a calibration label (designed to calibrate size and light saturation) adjacent to the area in question, and again without one. The label consisted of a white background with three different sized and colored lines (black, gray, and dark gray; Fig. 1). The images included white lesions (oral lichen planus, frictional hyperkeratosis, and leukoplakia), ulcerations (erosive lichen planus, traumatic ulceration, and mucous membrane pemphigoid), pigmented lesions, and erythroleukoplakia.

In all the pairs of images, one was labeled 'first visit' and the other 'second visit' (Fig. 2). The participants evaluated the images by responding to a questionnaire comparing five features in the 'second visit' photograph to the 'first visit' photograph: (i) change in size of lesion (5-step scale from '0', no change, to '4', pronounced change); (ii) type of size change (increase or decrease); (iii) presence of color change (5-step scale); (iv) change in location of most erythematous area; and (v) clinical improvement or worsening of the lesion.

Statistical analysis included the intra-class correlation (ICC) to assess agreement between observers. Two types of ICC were calculated for each question: absolute agreement and consistency. *Absolute agreement* measures whether the scoring was identical among the

observers, and *consistency* measures whether the scoring was similar throughout the questionnaire. An ICC of 100% indicates perfect agreement; the lower the ICC the weaker the agreement.

## Results

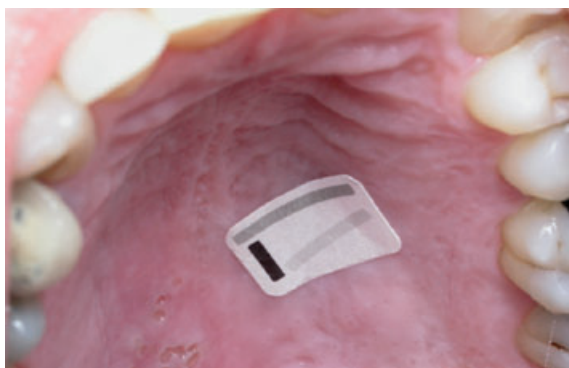
The ICC absolute agreement from both groups for each of the five questions is presented in Table 1. Higher agreement in four of five questions was found for the images with the calibration label in both groups. The greatest improvement was in the assessment of image size-question 2 (improvement of 11.8% and 8.5% in the OM and DS groups, respectively). The label did not enhance agreement within the OM group for color change (question 3; a 10.0% reduction) and within DS group for location of the change (question 4; a 15.2% reduction). The lowest congruency in both groups was in assessment of clinical improvement/worsening, without the label (26.9% and 11.7% for OM and DS groups, respectively; Table 1).

The ICC consistency for each group is presented in Table 2. The answers were more consistent in four of the five questions when the reference label was present; the color change interpretation by OM and the location of most erythematous feature by DS are the exception in which the labeled images did not yielded higher consistency (reduction of 8.8% and 19.0%, respectively). The consistency values were highest among the OM group for color changes of non-labeled images (question 3; 51.7%), while the lowest consistency was within the DS group for the evaluation of clinical improvement/worsening of non-labeled images (question 5; 12.3%; Table 2).

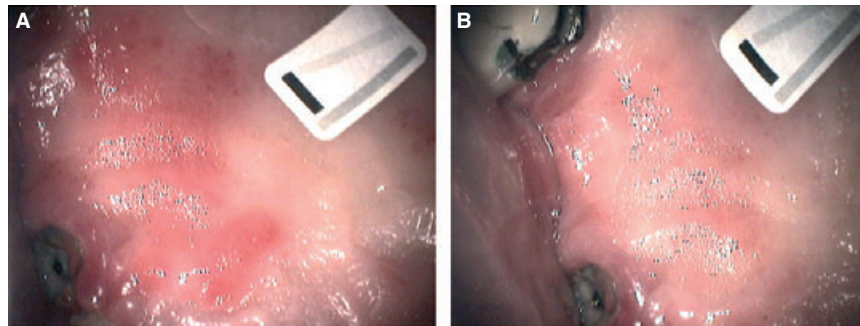
Overall, the OM group had higher consistency and absolute agreement results in nine of 10 questions presented in the study (five questions with the calibration label and four questions without). The only question in which the DS group had higher absolute agreement and consistency values was for question 4 regarding location change of the most erythematous feature without the label (Tables 1 and 2).

## Discussion

The current study examined the inter-observer differences in the evaluation of clinical images of oral lesions taken using an intraoral camera. The long-term monitoring of changes in the oral mucosa cavity is essential because specific changes can be a sign of disease progression, malignant transformation, or treatment response (7, 8). According to Wood and Goaz (9), the diagnosis of oral lesions is fundamentally a study of changes. For many years, however, the primary means of recording oral mucosal conditions included written descriptions and/or drawings. Recent technological advances allow more clinicians to include digital images in their records (2) with numerous advantages including: (i) improved communication between patients and practitioners (10); (ii) better patient feedback regarding treatment outcome (11); (iii) more accurate follow-up of lesions with a natural history of fluctuations over time



**Figure 1** Calibration label for size and color reference. Black line length is 5 mm.



**Figure 2** Intra-oral images of the same patient at two time points before (A) and after (B) treatment.

**Table 1** Intra-class correlation (ICC) – absolute agreement value (%) of oral medicine and other dental specialist groups, with and without the incorporation of a calibration label

Question	ICC w/o label (%)	ICC with label (%)
Oral medicine		
1. Size change	30.9	40.9
2. Size change trend <sup>a</sup>	30.0	41.8
3. Color change	49.5	39.5
4. Location of most erythematous area	25.7	40.5
5. General improvement or worsening	26.9	30.3
Other dental specialists		
1. Size change	20.3	30.9
2. Size change trend <sup>a</sup>	20.6	29.1
3. Color change	21.5	24.4
4. Location of most erythematous area	27.4	12.2
5. General improvement or worsening	11.7	24.0

w/o, without.

<sup>a</sup>One participant's data are missing. ICC cannot be calculated.

**Table 2** Intra-class correlation (ICC) – consistency (intra-observer agreement) value (%) of oral medicine and other dental specialist groups, with and without incorporation of a calibration label

Question	ICC w/o label (%)	ICC with label (%)
Oral medicine		
1. Size change	39.2	40.9
2. Size change trend <sup>a</sup>	31.6	44.6
3. Color change	51.7	42.9
4. Location of most erythematous area	29.4	43.9
5. General improvement or worsening	30.7	31.3
Other dental specialists		
1. Size change	26.7	36.7
2. Size change trend <sup>a</sup>	25.2	42.4
3. Color change	25.8	29.9
4. Location of most erythematous area	34.4	15.4
5. General improvement or worsening	12.3	24.8

w/o, without.

<sup>a</sup>One participant's data are missing. ICC cannot be calculated.

(11); (iv) aiding in research and education (12); (v) enabling 'teledentistry' for long-distance oral medicine consultations (5).

Inter- and intra-observer differences are important when visual parameters play a crucial diagnostic role. This issue has been studied in other fields of medicine

and dentistry involving subjective visual diagnoses such as dermatology (13) and oral radiology (14). Subjective histopathologic evaluation of oral biopsies by oral pathologists has been studied with repeated findings of inter- and intra-observer differences (15).

Our questionnaire referred to the lesion's color and size. Lesion color (vascular, pigmentation changes, etc.) is clinically significant. The intensity of erythema is thought to indicate the severity of inflammation (1). Changes in size may indicate progression or healing of a pathological process (16) as well as risk of malignant transformation (e.g. more than 200 mm<sup>2</sup> for leukoplakia; 17). Because accuracy is important, and changes may be slight, Cawson and Odell (11) recommend including the teeth or a scale in clinical images. We included a calibration label in the clinical image that provided references for both size and color. In the current study, the calibration label enhanced the clinicians' ability to compare photographic records and thereby increased the accuracy and consistency of clinical evaluation in both study groups. The exception to this finding was decreased agreement and consistency regarding color change in the OM group when examining images with the calibration label. The contribution of a grayscale reference to the evaluation of redness of the lesions may be limited; furthermore, because oral medicine practitioners routinely evaluate mucosal conditions without reference labels, they may have relied on their past clinical experience rather than on the label. In contrast, the other specialists do not routinely perform such evaluations, and therefore, the calibration label was beneficial. The authors believe that if calibration labels are routinely incorporated into clinical images, the evaluations of the oral medicine practitioners will also improve.

In both the OM and DS groups, the highest ICC value for agreement and consistency was in response to the questions referring to changes in size. This may be due to the fact that size is more easily assessed in well-defined lesions. This finding is in agreement with a previous dermatological study, showing that even patients could accurately detect changes in skin mole size, especially with the use of a baseline image (4). The lowest agreement and consistency was in response to the question about the general condition of the lesion. The purpose of this question was to summarize the clinical impression of size and color and other parameters, which are difficult to name but may be expressed



as the clinicians intuition regarding whether the patient is improving or not. Thus, the lowest agreement was found in the most subjective question.

For each of the questions, the absolute agreement and consistency was higher among the oral medicine practitioners than the other DS. This is congruent with other studies showing that specialists, and practitioners who deal with these conditions regularly, have an advantage over non-specialists, and inexperienced practitioners, regarding clinical decision making in their specific field (18).

Our findings may contribute to the dilemma regarding whether general dental practitioners or oral medicine practitioners should perform follow-ups of potentially malignant disorders. It seems that the patient benefits from periodic evaluations by experienced practitioners. Practically speaking, almost half the patients seen by oral medicine practitioners suffer from oral mucosal diseases (19), whereas general dental practitioners routinely deal with dental conditions. Considering that the diagnosis and management of oral mucosal lesions lies within the scope oral medicine practitioners (20) and is included in all postgraduate programs in oral medicine worldwide (21), the advantages of follow-up by oral medicine practitioners are clear. For example, a recent study noted that correct diagnosis of oral mucosal lesions was achieved in 46–80% of cases by English general dental practitioners, whereas 78–87% of the cases were correctly diagnosed by a specialist (22). However, the present study also reflects the advantages of knowledge and experience, which can be gained by all practitioners who study the relevant material and gain clinical experience.

Potential limitations of the study are the small groups (10 practitioners each) and the non-homogenous DS group. However, considering that other publications studied fewer practitioners, often without control group, the present study, with a total of 20 practitioners and a control (DS) group, can be considered quite large. The DS group was selected to match postgraduate education with the OM group (because a group of general practitioners may also differ in this aspect, and thus be harder to compare as a control group).

## Summary

Owing to the increased use of photography as a clinical aid for diagnosis, follow-up, and communication, we decided to study the agreement between clinicians interpreting images of oral mucosal lesions. Using a relatively large group of specialists, we found that oral medicine practitioners, who routinely deal with oral mucosal conditions, are more consistent and accurate in detecting changes in oral mucosal diseases than other dental specialists. In addition, the incorporation of a calibration label within clinical images may help to increase agreement and consistency in image evaluation.

## References

1. Reichart PA, Philipsen HP. Oral erythroplakia – a review. *Oral Oncol* 2005; **41**: 551–61.
2. Terry DA, Snow SR, McLaren EA. Contemporary dental photography: selection and application. *Compend Contin Educ Dent* 2008; **29**: 432–42.
3. Jack H, Lee K, Polonowita A. Dilemmas in managing oral dysplasia: a case report and literature review. *N Z Med J* 2009; **122**: 89–98.
4. Oliveria SA, Chau D, Christos PJ, Charles CA, Mushlin AI, Halpern AC. Diagnostic accuracy of patients in performing skin self-examination and the impact of photography. *Arch Dermatol* 2004; **140**: 57–62.
5. Torres-Pereira C, Possebon RS, Simões A, et al. Email for distance diagnosis of oral diseases: a preliminary study of teledentistry. *J Telemed Telecare* 2008; **14**: 435–8.
6. Scully C. Erythroplakia, leukoplakia, keratosis and other potentially malignant lesions. In: Scully C, ed. *Oral and maxillofacial medicine*, 2nd edn. Edinburgh: Elsevier, Churchill Livingstone, 2008; 211–24.
7. Mignogna MD, Fedele S, Lo Russo L. Dysplasia/neoplasia surveillance in oral lichen planus patients: a description of clinical criteria adopted at a single centre and their impact on prognosis. *Oral Oncol* 2006; **42**: 819–24.
8. Hsue SS, Wang WC, Chen CH, Lin CC, Chen YK, Lin LM. Malignant transformation in 1458 patients with potentially malignant oral mucosal disorders: a follow-up study based in a Taiwanese hospital. *J Oral Pathol Med* 2007; **36**: 25–9.
9. Wood NK, Goaz PW. Correlation of gross structure with clinical features. In: Wood NK, Goaz PW, eds. *Differential diagnosis of oral and maxillofacial lesions*, 5th edn. St. Louis: Mosby, 1997; 14–38.
10. Shorey R, Moore KE. Clinical digital photography today: integral to efficient dental communications. *J Calif Dent Assoc* 2009; **37**: 175–7.
11. Cawson RA, Odell EW. Principles of investigation and diagnosis. In: Cawson RA, Odell EW, eds. *Cawson's essentials of oral pathology and oral medicine*, 8th edn. Edinburgh: Elsevier, 2008; 1–17.
12. Piboonniyom SO, Treister N, Pitiphat W, Woo SB. Scoring system for monitoring oral lichenoid lesions: a preliminary study. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2005; **99**: 696–703.
13. Chen S, Wang Q, Chu T, Zheng M. Inter-observer reliability in assessment of sensation of skin lesion and enlargement of peripheral nerves in leprosy patients. *Lepr Rev* 2006; **77**: 371–6.
14. Baksi BG, Alpöz E, Sogur E, Mert A. Perception of anatomical structures in digitally filtered and conventional panoramic radiographs: a clinical evaluation. *Dentomaxillofac Radiol* 2010; **39**: 424–30.
15. Kujan O, Khattab A, Oliver RJ, Roberts SA, Thakker N, Sloan P. Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia: an attempt to understand the sources of variation. *Oral Oncol* 2007; **43**: 224–31.
16. Rozycki TW, Rogers RS Iii, Pittelkow MR, et al. Topical tacrolimus in the treatment of symptomatic oral lichen planus: a series of 13 patients. *J Am Acad Dermatol* 2002; **46**: 27–34.
17. Van Der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral Oncol* 2009; **45**: 317–23.
18. Bigras BR, Johnson BR, Begole EA, Wenckus CS. Differences in clinical decision making: a comparison between specialists and general dentists. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008; **106**: 139–44.
19. Hodgson TA, Buchanan JA, Garg A, Ilyas SE, Porter SR. An audit of the UK national cancer referral guidelines for

- suspected oral mucosal malignancy. *Br Dent J* 2006; **201**: 643–7.
20. Stoopler ET, Shirlaw P, Arvind M, et al. An international survey of oral medicine practice: Proceedings from the 5th World Workshop in Oral Medicine. *Oral Dis* 2011; **17**(Suppl 1): 99–104.
21. Rogers H, Sollecito TP, Felix DH, et al. An international survey in postgraduate training in oral medicine. *Oral Dis* 2011; **17**(Suppl 1): 95–8.
22. Brocklehurst PR, Baker SR, Speight PM. Factors which determine the referral of potentially malignant disorders by primary care dentists. *J Dent* 2010; **38**: 569–78.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.