Received: 8 June 2023 | Revised: 23 November 2023 | Accepted: 5 February 2024
DOI: 10.1002/hed.27687

ORIGINAL ARTICLE

WILEY

Improving detection of oral lesions: Eye tracking insights from a randomized controlled trial comparing standardized to conventional approach

Behrus Puladi MD, DMD[1,2] | **Beatrice Coldewey MSc**[2] | **Julia S. Volmerg MSc**[2] | **Kim Grunert**[1,2] | **Jeff Berens**[1] | **Ashkan Rashad MD, DMD**[1] | **Frank Hölzle MD, DMD, PhD**[1] | **Rainer Röhrig MD, PhD**[2] | **Myriam Lipprandt PhD**[2]

[1]Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, Aachen, Germany

[2]Institute of Medical Informatics, Medical Faculty, RWTH Aachen University, Aachen, Germany

Correspondence
Behrus Puladi, Department of Oral and Maxillofacial Surgery & Institue of Medical Informatics, University Hospital RWTH Aachen, Pauwelsstraße 30, 52074 Aachen, Germany.
Email: bpuladi@ukaachen.de

Funding information
Medical Faculty RWTH Aachen

Section Editor: Steven Chang

Abstract

Background: Early detection of oral cancer (OC) or its precursors is the most effective measure to improve outcome. The reasons for missing them on conventional oral examination (COE) or possible countermeasures are still unclear.

Methods: In this randomized controlled trial, we investigated the effects of standardized oral examination (SOE) compared to COE. 49 dentists, specialists, and dental students wearing an eye tracker had to detect 10 simulated oral lesions drawn into a volunteer's oral cavity.

Results: SOE had a higher detection rate at 85.4% sensitivity compared to 78.8% in the control ($p = 0.017$) due to higher completeness ($p < 0.001$). Detection rate correlated with examination duration ($p = 0.002$).

Conclusions: A standardized approach can improve systematics and thereby detection rates in oral examinations. It should take at least 5 min. Perceptual and cognitive errors and improper technique cause oral lesions to be missed. Its wide implementation could be an additional strategy to enhance early detection of OC.

KEYWORDS
education, eye tracking, oral cancer, oral examination, visual screening

1 | INTRODUCTION

Oral cancer (OC) is one of the most common cancers in humans. The annual global incidence is estimated to be 377 713 cases, resulting in 177 757 deaths in 2020.[1] OC can develop spontaneously or from oral potentially malignant disorders (OPMDs), such as leukoplakia, erythroplakia, oral submucous fibrosis, or actinic cheilitis.

Abbreviations: COE, conventional oral examination; LMM, linear mixed-effects model; NNS, number needed to screen; OC, oral cancer; OPMD, oral potentially malignant disorders; RCT, randomized controlled trial; SOE, standardized oral examination.

With an estimated global prevalence of 4.47%, OPMDs represent some of the most common lesions in the human oral cavity,[2] and regardless of OPMD subtype, it has a potential for malignancy with a reported overall transformation rate of 7.9%.[3]

Both OC and OPMD can be detected by conventional oral examination (COE), with early detection of OC considered the most important measure for successful treatment, reduction of mortality, and improvement of overall prognosis, while at the same time being cost-effective.[4–7] For example, the odds ratio (OR) for the occurrence of advanced-stage OC (T3, T4, or $N > 0$) with a delay of >1 month in diagnosis is 1.69,[8] resulting in a serious outcome decline.[9] Consequently, one of the primary objectives of the World Health Organization regarding OC is the improvement of early detection rate.[10]

Although, guidelines clearly recommend a COE in which all relevant areas of the oral mucosa, including the lips, labial mucosa, buccal mucosa, gingiva, floor of the mouth, tongue, palate, retromolar, and throat, are systematically viewed and critically evaluated,[11,12] the sensitivity of COE conducted by general dentists compared to specialists ranges from 60% to 81%, and the specificity from 94% to 99%. In the world's only randomized controlled trial (RCT) conducted in India between 1994 and 2009, despite 4 cycles of COE to detect OPMD and OC, the achieved sensitivity was only 67.4%.[13,14] As a result, 2–4 out of 10 OCs could be missed during the COE, even if repeated 4 times. It is reported, that approximately 3 out of 10 OPMDs are missed by general dentists.[15] This is further complicated by the existence of more than 20 types of OPMD.[16] About 85% of dentists report difficulty diagnosing oral mucosal lesions.[17] At the same time, large deficiencies in dental students' knowledge of OC and how to perform a COE have been extensively reported worldwide.[18–24] To make matters worse, the reported sensitivity of oral self-examination by patients is only at 24.6%.[25] All this has implications not only for the early detection of OC/OPMD, but also for the follow-up of OC survivors, who suffer a recurrence rate of up to 37% within 10 years,[26] and may lead to a delayed diagnosis of recurrence.[27]

In recent years, auxiliary methods have been introduced to improve the detection of high-risk oral lesions, such as direct and imaging fluorescence-based methods[28] or real-time dynamic optical microscopy.[29] However, these devices increase the cost of the oral examination and require more complex training and experience. In fact, despite population-based screening programmes,[13,30] improvements in the early detection of OC have not yet been achieved in individual patient encounters.[14] At the same time, early detection of other cancer entities, such as breast and cervical cancer has led to improvements in their survival rates in recent decades.[31] Paradoxically, the survival rates for OC and colorectal cancer are comparable, even though the former could be easily diagnosed with a COE, and the latter is inside the body and requires a colonoscopy for diagnosis.[1,31–33]

In 1914, the surgeon Walter G. Spence stated that early diagnosis of OC is the most effective measurement for outcome improvement.[34] More than 100 years later with regard to COE, we can only speculate whether a lack of systematics (i.e., incompleteness of the examination), visual perception, or cognitive errors lead to the missing of oral lesions. We do not know what the true sensitivity of COE is because the reported values have no true reference due to the study design and could therefore be biased.[14,30] However, this is very important for two reasons. First, patients expect us to identify oral lesions with a high degree of certainty on an individual level. The impact of late or missed diagnoses is dramatic at all levels and for maintaining public trust in the profession.[35] Second, we need to understand the underlying reasons for missing oral lesions in order to take appropriate countermeasures.

To address these issues, we conducted an RCT to compare COE with a standardized approach to oral examination (SOE) in an experimental setting among dental students, dentists, and specialists at an individual level.

2 | MATERIALS AND METHODS

2.1 | Study design

In this parallel RCT, dentists (including specialists) and dental students had to perform oral examinations on a simulated patient. The study was conducted in the Department of Oral and Maxillofacial Surgery at the University Hospital RWTH Aachen from May to June 2022 and followed the CONSORT statement (Figure 1).[36] The study was approved by the Ethics Committee of RWTH Aachen University (approval number EK 076/22 from 28.04.2022) and registered in the German Clinical Trials Register (DRKS00027450). No changes were made to the trial design after approval.

The inclusion criteria were dentists from different specialties and dental students in the clinical part of their studies. The acquisition of participants was conducted by B.P. and J.B. The randomization sequence was generated by B.C. After obtaining informed consent, randomization was performed according to an allocation rule by K.G. The participants were unaware of the true objective of the experiment and of their assigned group until their participation in the experiment was over.

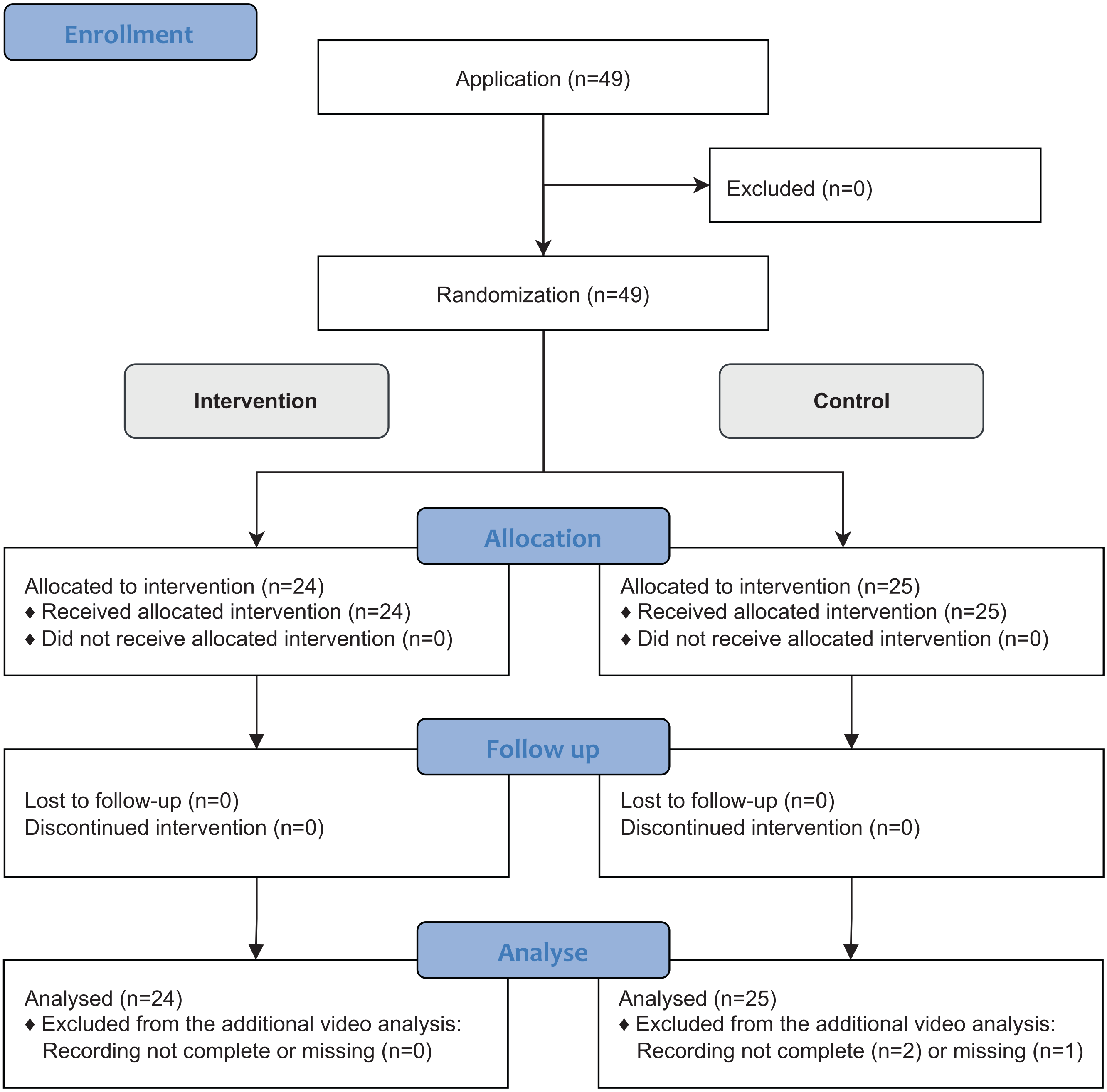

FIGURE 1 CONSORT flow diagram. [Color figure can be viewed at wileyonlinelibrary.com]

For the intervention group (SOE), a video (2 min; https://www.youtube.com/watch?v=hDG1mRFFusc) explaining the function of the eye tracker and training to standardize the COE (5 min) with a systematic and step-wise examination (13 steps) of all relevant regions of the oral mucosa was provided (Video S1, Supporting Information). The training program for standardization was designed according to the recommendations of Reichert and Philipsen[37] and the National Institute of Dental and Craniofacial Research (Figure 2A).[38] For the control group with individual approaches (COE), a video explaining only the function of the eye tracker was shown (2 min; Eye tracker part in Video S1).

Subsequently, both the intervention and control groups had to perform oral screenings for simulated oral lesions on a volunteer (J.B.). For this purpose, 10 water-proof, standardized, and evident black dots were placed at different locations in the oral cavity of the simulated patient (Figure 2B). The participants were asked to examine the oral cavity for the aforementioned dots and report verbally their location without knowing the number of dots or their positions. During the oral screening, videos

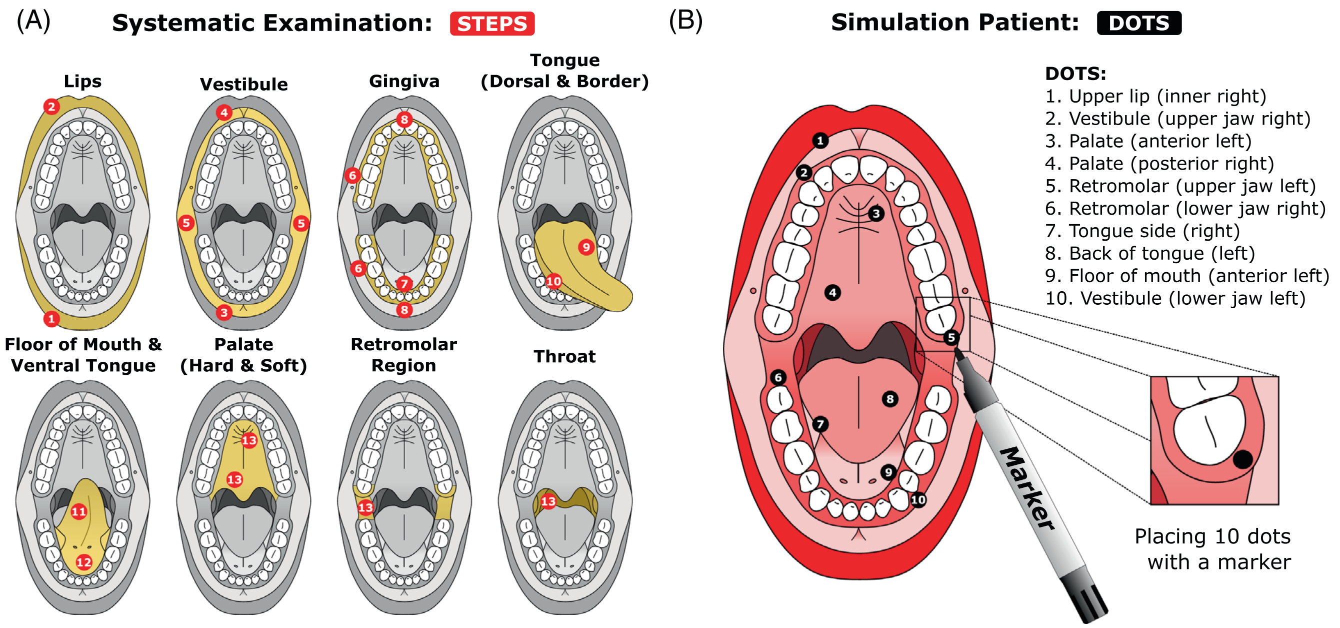

FIGURE 2 (A) The different regions of the oral cavity highlighted in yellow (lips, vestibule, gingiva, tongue, floor of mouth, palate, retromolar region, and throat). A modified recommended systematic examination scheme for the oral cavity with 13 steps is marked in red.[37] (B) The simulated patient was prepared by drawing dots using a surgical marker in the oral cavity at 10 different anatomical sites: upper lip (inner right), vestibule (upper jaw right), palate (anterior left), palate (posterior right), retromolar (upper jaw left), retromolar (lower jaw right), tongue side (right), back of tongue (left), floor of mouth (anterior left), and vestibule (lower jaw left). [Color figure can be viewed at wileyonlinelibrary.com] [Color figure can be viewed at wileyonlinelibrary.com]

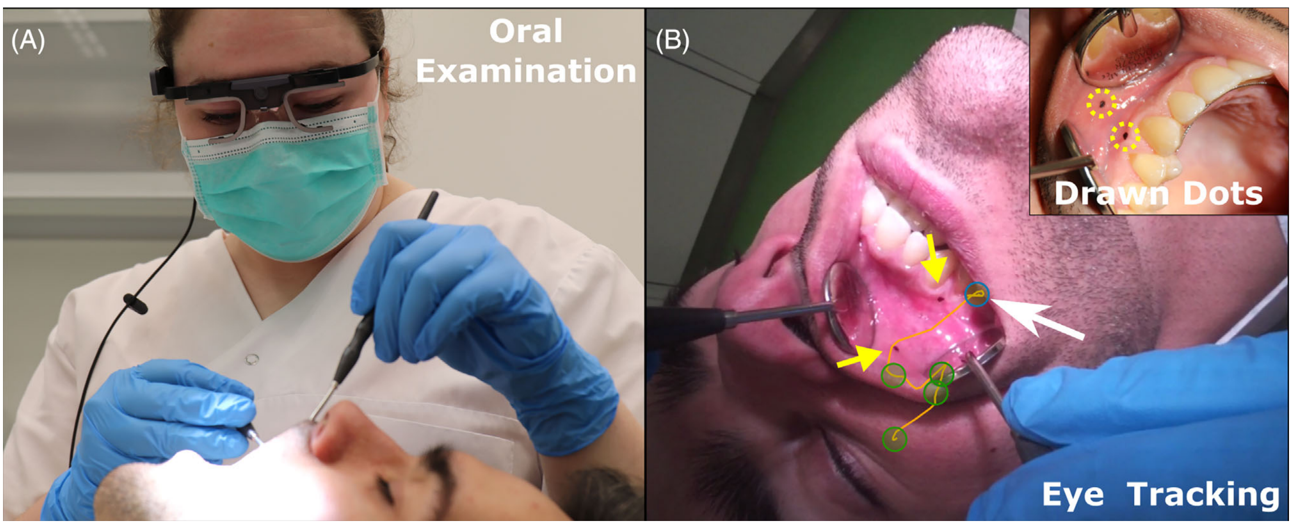

FIGURE 3 (A) During the oral examination, an eye tracker (Tobii 2) was worn, which recorded a video that included the visual gaze. (B) Representation of the view from the eye tracker. The black dot (yellow arrow) and the gaze position (white arrow) were both visible during the examination of the oral cavity of the simulated patient. In the upper right corner, the drawn points are surrounded by yellow circles. [Color figure can be viewed at wileyonlinelibrary.com]

were recorded from both a first-person perspective, which involved eye tracking using Tobii Pro Glasses 2 (Tobii AB, Danderyd, Stockholm, Sweden), and an external perspective (Figure 3). Finally, a questionnaire (including 4-point Likert scale questions; Table 5) was completed by the participants, who were informed of the true objective of the study.

The primary endpoint of the study was the number of simulated oral lesions detected. Secondary endpoints were the frequency of detection of each dot, the systematics of the examination, the examination duration, the use of aids, and the self-assessment of the quality of the examination.

2.2 | Data acquisition

The video recordings with eye tracking were analyzed using the iMotions software (iMotions A/S, Copenhagen, Denmark) by two investigators (K.G. and J.B.). The results, including the other collected data, were stored in REDCap (Vanderbilt University, Nashville, TN). Disagreements were resolved by two additional investigators (B.P. and B.C.).

The videos were used to determine the detection rate, which was also documented in writing (as a backup in case of technical problems). The examination duration was determined and the presence of the 13 steps (Table 3) of the SOE according to the training (intervention group) or a comparable approach (control group) was assessed (Figure 2A).[37,38] To account for interrater variability, the examination duration was calculated as the average determined by two investigators (K.G. and J.B.). Disagreements and discrepancies in the examination duration of more than 15 s were resolved by two additional investigators. Furthermore, the videos were used to determine whether participants used mouth mirrors, adjusted the light, used additional aids (e.g., suction or compresses), sat or stood, and whether the simulated patient was examined in a sitting, semi-recumbent, or recumbent position.

2.3 | Sample size calculation

Sample size was determined according to the recommendations of a statistician using G*Power (Version 3.1; Heinrich Heine University Düsseldorf, North Rhine-Westphalia, Germany) based on the detection rate from one pretest (7.5 ± 1.12 [mean ± SD]; $n = 4$). With a significance level of $\alpha = 0.05$, 80% power and a mean difference of 1 oral lesion detected between two independent groups, this resulted in a sample size of 24 participants (including 3 dropouts) for each arm, resulting in a total of 48 participants.

2.4 | Statistical analysis

Statistical analysis was performed using R (version 4.2.2). $p < 0.05$ was considered significant. The difference in the primary endpoints between the intervention and control groups was analyzed using a linear mixed-effects model (LMM) with the lme4 package to take into account the influence of experience and different professional groups. The dependent variable was the number of drawn dots detected, and the independent variables were intervention (yes/no), work experience/study progress (mixed effect), and specialty (mixed effect). However, specialty showed zero variance, so this parameter was changed to a fixed effect in the final model without having impact on the significance. Furthermore, differences in detection rates between professional groups (specialists, dentists, dental students) were examined using the Kruskal–Wallis test. The degree to which the examination was conducted in a systematic fashion and the duration of the examination were compared as secondary endpoints using a *t* test or a Mann–Whitney *U* test, depending on the presence of a normal distribution determined by the Shapiro–Wilk test.

3 | RESULTS

A total of 49 participants were recruited, including 24 dental students and 25 dentists, of which 24 were assigned to the intervention group and 25 to the control group. One subject was additionally included after reaching the planned sample size. Due to technical reasons, the video/eye tracking recording was incomplete ($n = 2$) or failed ($n = 1$) for three participants. However, this did not negatively affect the analysis of the detection rate, as it was recorded both in written and video formats, resulting in 49 participants included in the detection rate analysis and 46 participants included in the eye tracking analysis. No losses or exclusions occurred after randomization (Figure 1).

The gender distribution was almost balanced (25 women vs. 24 men). Regarding age, 22 subjects were between 18 and 25 years old, 22 were between 26 and 35, and only 5 were 36 years or older. Most students (79.2%) were in the second half of their 4th year (3rd clinical semester) or the first half of their 5th year (4th clinical semester) of a total of 5 years of dental training, while most of the dentists had four or more years of professional experience (68%) (Table 1). Sixty percent of the dentists were specialists with >3 years of training in oral surgery or oral and maxillofacial surgery with an additional degree in medicine and are routinely involved in the detection, treatment, and follow-up of OC and OPMD in our department.

The detection rate was significantly higher in the SOE/intervention group (8.54 ± 1.25 out of 10) than in the COE/control group (7.88 ± 1.39; LMM; $\beta = 0.88$, $p = 0.017$) adjusted for professional group and experience; Figure 4A). Regardless of the oral examination approach (SOE/COE), the detection rate was 8.13 ± 1.36

TABLE 1 Cohort.

Parameters		SOE: Intervention ($n = 24$)	COE: Control ($n = 25$)	Total ($n = 49$)
Gender	Female	10 (41.7%)	15 (60.0%)	25 (51.0%)
	Male	14 (58.3%)	10 (40.0%)	24 (49.0%)
Age	18–25	12 (50.0%)	10 (40.0%)	22 (44.9%)
	26–35	9 (37.5%)	13 (52.0%)	22 (44.9%)
	36–45	3 (12.5%)	1 (4.0%)	4 (8.2%)
	≥46	0 (0.0%)	1 (4.0%)	1 (2.0%)
Profession	Dental student	12 (50.0%)	12 (48.0%)	24 (49.0%)
	Dentist	12 (50.0%)	13 (52.0%)	25 (51.0%)
	[Specialists]	[6 (25.0%)]	[9 (36.0%)]	[15 (30.6%)]
Study progress (clinical semester)	1st (3rd year)	0 (0.0%)	1 (8.3%)	1 (4.2%)
	2nd (4th year)	1 (8.3%)	0 (0.0%)	1 (4.2%)
	3rd (4th year)	4 (33.3%)	3 (25.0%)	7 (29.2%)
	4th (5th year)	4 (33.3%)	8 (66.7%)	12 (50.0%)
	5th (5th year)	3 (25.0%)	0 (0.0%)	3 (12.5%)
Work experience (years)	<1	2 (16.7%)	0 (0.0%)	2 (8.0%)
	1–3	4 (33.3%)	2 (15.4%)	6 (24.0%)
	4–5	2 (16.7%)	7 (53.8%)	9 (36.0%)
	6–9	3 (25.0%)	2 (15.4%)	5 (20.0%)
	>10	1 (8.3%)	2 (15.4%)	3 (12.0%)

among all specialists, 8.00 ± 1.26 among dentists (including specialists), and 8.42 ± 1.44 among dental students (Kruskal–Wallis test, $p = 0.25$). Stratified for oral examination approach (SOE/COE), the detection rate was 8.50 ± 1.05 (SOE) and 7.89 ± 1.54 (COE) among specialists, among dentists (including specialists) 8.33 ± 1.07 (SOE) and 7.69 ± 1.38 (COE), and 8.75 ± 1.42 (SOE) and 8.08 ± 1.44 (COE) among dental students.

Interestingly, these differences did not show up in the qualitative self-assessment ($n = 49$). Participants in the COE group rated themselves slightly higher (4-point Likert scale; 1 = *disagree*, 4 = *agree*; 3.4 ± 0.5) than the SOE group (3.2 ± 0.8) (Figure 4D).

Despite this, the rate of detection of the drawn dots varied depending on their anatomic site. The highest detection rate (93.9%) was found for the vestibular region of the maxilla and mandible and for the retromolar region of the mandible, while the lowest detection rate was found for the right inner side of the upper lip (57.1%), the floor of the mouth (63.3%), and the anterior palate (71.4%) (Table 2).

The number of steps (total of 13) performed (i.e., degree of systematic) according to the recommended steps was significantly higher among the SOE group (12.50 ± 0.78 steps) compared to the COE group (10.64 ± 1.89 steps) (Mann–Whitney *U* test, $p < 0.001$; Figure 4B and Table 3). During the intraoral examination, 37% of the participants stood, 60.9% sat, and 2.2% did both. In 63% of the cases, the simulated patient was examined in a recumbent position, 21.7% were in a semi-recumbent position, and 15.2% were in a sitting position. The oral cavity was illuminated in 91.3% of cases and only half of the time in 8.7% of cases. The majority of investigators wore gloves (95.7%), only two investigators did not wear gloves (4.3%). Compresses were used in 50% of the cases, and dental suction were used in 8.7%. Of the examiners, 95.7% used two mirrors, and 4.3% used only one mirror and their own fingers during the examination (Table 4).

The duration of the examination was 4:14 ± 1:13 (min:s) for the intervention group and 3:36 ± 1:22 for the control group, which was not significantly different (*t* test, $p = 0.108$). However, there was a significant correlation between detection rate and examination duration ($\beta = 0.467$, $p = 0.002$) with an increase starting at 3:00 and reaching a plateau at 4:30 (Figure 4C). Consistent with this, participants with a detection rate of ≥9 (4:25 ± 1:14) were linked to a significantly longer examination compared to those with a detection rate ≤8 (3:33 ± 1:17; *t* test, $p = 0.025$).

Both groups reported following one strategy during the examination (SOE: 3.2 ± 0.8 vs. COE: 3.4 ± 0.8). Similarly, both groups reported that the approach was suitable for the detection of oral mucosal changes (SOE: 3.0

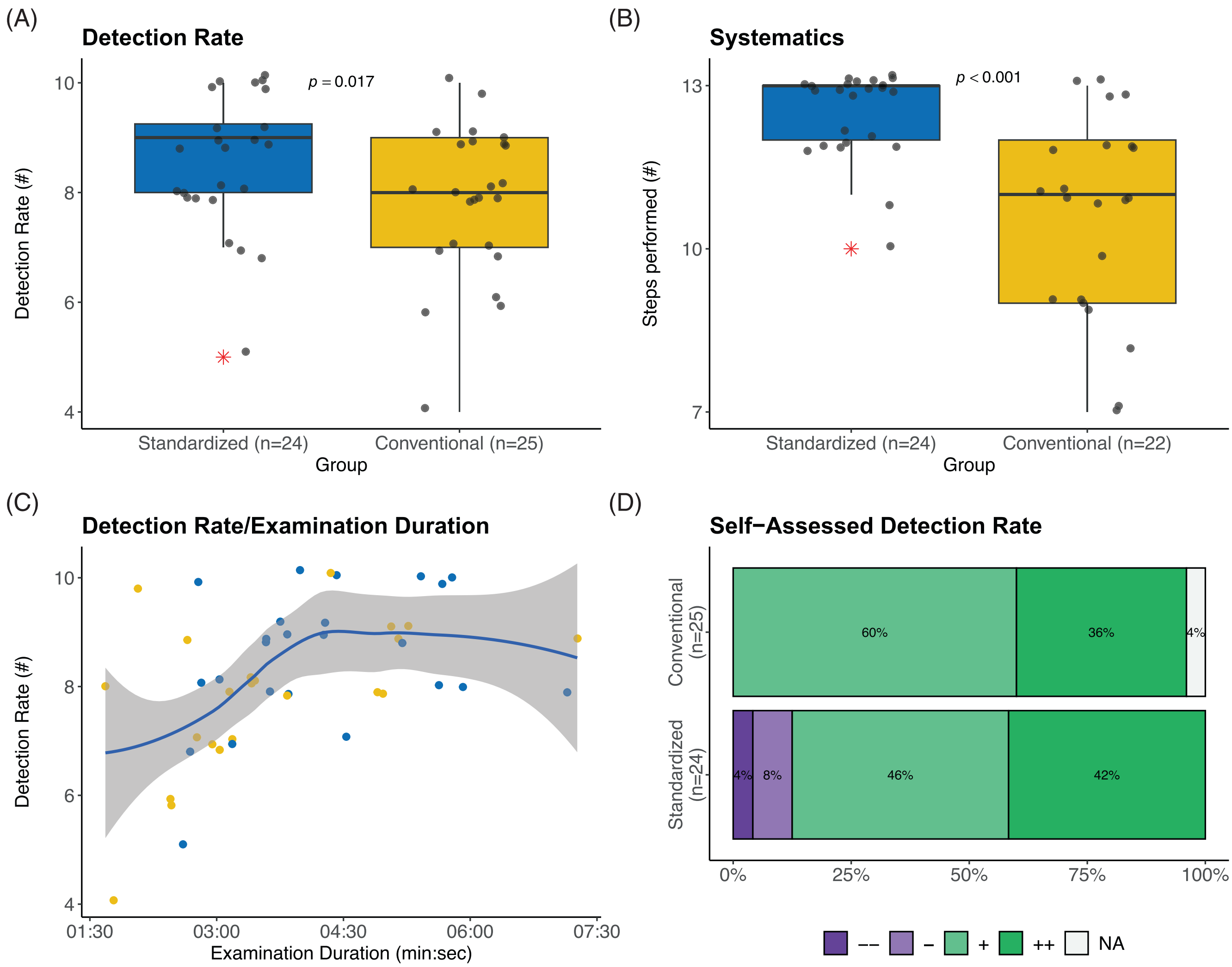

FIGURE 4 (A) The intervention group (SOE) that underwent training detected significantly more points (mixed linear effects model, $p = 0.017$) than the control group (COE). (B) The intervention group (SOE) completed more steps (i.e., degree of systematic) than the control group (COE) (Mann–Whitney U test, $p < 0.001$). (C) A temporal relationship visualized with the locally estimated scatterplot smoothing. There was a significant increase in the number of points detected when the examination duration exceeded 4 min. The intervention group is shown in blue, and the control group is shown in yellow. (D) Intrinsic evaluation of the detection rate between the intervention (SOE) and control group (COE). [Color figure can be viewed at wileyonlinelibrary.com] [Color figure can be viewed at wileyonlinelibrary.com]

± 0.8 vs. COE: 3.2 ± 0.6) and for malignant oral mucosal lesions (SOE: 3.2 ± 0.8 vs. COE: 3.3 ± 0.6). The optimal use of aids was confirmed by both groups (SOE: 3.2 ± 0.8 vs. COE: 3.4 ± 0.6). In the intraoral examination, the eye tracker was not perceived as disturbing (SOE: 1.5 ± 0.7 vs. COE: 1.6 ± 0.7), while participants were able to imagine wearing the eye tracker in a real-world setting (SOE: 2.7 ± 0.8 vs. COE: 2.9 ± 0.7) (Table 5; non-translated version as Table S1).

4 | DISCUSSION

Although tremendous efforts have been made to improve the diagnosis of evident lesions by vital staining, oral cytology, light-based detection, or oral spectroscopy[7] and the recommendation that any obvious oral lesion that persists for more than 2 weeks should be biopsied,[39] all these measures are in vain if the patient, dentist, oral specialist, head and neck surgeon, or any other healthcare provider is unaware of the oral lesion. Therefore, COE is still the most important step for detection of OPMD or OC.

To our knowledge, this study is the first RCT to examine the effectiveness of COE in an experimental setting using eye tracking. We originally assumed that almost all 10 simulated oral lesions would be detected due to their obvious appearance compared to normal mucosa. Surprisingly, the sensitivity of the COE group was 78.8% and within the upper range of COE sensitivity reported in the literature (60%–81%) when dentists were compared to specialists.[15] Interestingly, the sensitivity of the SOE group was actually higher at 85.4%.

Both the sensitivity of COE and the increase due to SOE are remarkable and should not be overlooked. First, we do not know what the true sensitivity of COE is because either specialists are used as the reference[7,15] or it is estimated from follow-up data,[14,30] both of which are subject to error and are only an approximation.[30] In the experimental setting of our study, we do not have this issue because the number of the oral lesions (black dots) is given. In light of this, reported high sensitivity rates (>90%) where non-specialists are used as reference[40] are probably an interrater correlation between health care workers and specialists, rather than a true sensitivity. This is also supported by the fact that in the largest population-based screening program to date, the sensitivity was estimated to be 88.29% in one subpopulation, despite multiple screening, and using a previous negative screening as a reference.[30] We thus have for the first time the possibility to verify the performance of a COE in an experimental setting, which can be used to test sensitivity baselines for investigators, as required in population-based methods.[14]

Second, the increased sensitivity rate of the SOE approach without consideration of mixed effects leads to a number needed to screen (NNS) of 339 for OPMD ($NNS = \frac{1}{(4.47\% \times 6.6\%)}$) (assuming a prevalence for OPMD of 4.47% and sensitivity increase of 6.6%). This means that one in 339 patients would have an additional OPMD detected by SOE, given the OPMD is recognizable as such. Including mixed effects, the NNS would be even lower at 260. In risk groups with a higher prevalence, the NNS may be much lower. It is estimated that population-based oral screening programs could prevent at least 37 000 deaths per year worldwide at current COE sensitivities.[41] A nationwide screening in Taiwan involving over 2 million individuals showed a 26% reduction in mortality (Risk ratio 0.74; 95% CI: 0.72–0.77).[30] Therefore, an improvement in the sensitivity rate due to SOE could have a sustainable impact in daily clinical practice and could be applied outside of population-based programs to all patients being orally examined.

TABLE 2 Detection rate.

Location (*n* = 49)	Detected, *n* (%)	Not detected, *n* (%)
Upper lip (inner right)	28 (57.1)	21 (42.9)
Floor of mouth (anterior left)	31 (63.3)	18 (36.7)
Palate (anterior left)	35 (71.4)	14 (28.6)
Retromolar (upper jaw left)	41 (83.7)	8 (16.3)
Palate (posterior right)	42 (85.7)	7 (14.3)
Back of tongue (left)	43 (87.8)	6 (12.2)
Tongue side (right)	44 (89.8)	5 (10.2)
Upper jaw vestibule (right)	46 (93.9)	3 (6.1)
Lower jaw vestibule (left)	46 (93.9)	3 (6.1)
Retromolar of lower jaw (right)	46 (93.9)	3 (6.1)

TABLE 3 Performed steps.

Steps (*n* = 46)		Performed, *n* (%)	Missing, *n* (%)
#1	Lower lip (red lip and adjacent skin)	40 (87.0)	6 (13.0)
#2	Upper lip (labial red and adjacent skin)	40 (87.0)	6 (13.0)
#3	Mucosa of lower lip and sulcus	46 (100.0)	0 (0.0)
#4	Mucosa of upper lip and sulcus	46 (100.0)	0 (0.0)
#5	Commissures, buccal mucosa, and sulcus (upper and lower jaw)	45 (97.8)	1 (2.2)
#6	Alveolar processes, buccal (2 mirrors)	45 (97.8)	1 (2.2)
#7	Alveolar processes, lingual (2 mirrors)	34 (73.9)	12 (26.1)
#8	Alveolar processes, frontal, and tongue in resting position	46 (100.0)	0 (0.0)
#9	Tongue extended, held by examiner	39 (84.8)	7 (15.2)
#10	Tongue margins held by the examiner	34 (73.9)	12 (26.1)
#11	Tongue (ventral surface) and floor of mouth	34 (73.9)	12 (26.1)
#12	Floor of mouth with mouth wide open	38 (82.6)	8 (17.4)
#13	Hard and soft palate, mouth wide open, head tilted back, including pharynx and retromolar region	46 (100.0)	0 (0.0)

TABLE 4 Examination characteristics.

Parameters		SOE: Intervention ($n = 24$)	COE: Control ($n = 25$)	Total ($n = 49$)	p-value
Video/eye-tracking data	Missing/ incomplete	0	3	3	
Investigator position	Standing	15 (62.5%)	13 (59.1%)	28 (60.9%)	0.572
	Sitting	9 (37.5%)	8 (36.4%)	17 (37.0%)	
	Both	0 (0.0%)	1 (4.5%)	1 (2.2%)	
Patient position	Sitting	4 (16.7%)	3 (13.6%)	7 (15.2%)	0.782
	Recumbent	14 (58.3%)	15 (68.2%)	29 (63.0%)	
	Semi-recumbent	6 (25.0%)	4 (18.2%)	10 (21.7%)	
Light	Yes	21 (87.5%)	21 (95.5%)	42 (91.3%)	0.350
	No	3 (12.5%)	1 (4.5%)	4 (8.7%)	
Gloves	Yes	24 (100.0%)	20 (90.9%)	44 (95.7%)	0.137
	No	0 (0.0%)	2 (9.1%)	2 (4.3%)	
Dental suction	Yes	0 (0.0%)	4 (18.2%)	4 (8.7%)	0.029
	No	24 (100.0%)	18 (81.8%)	42 (91.3%)	
Compresses	Yes	19 (79.2%)	4 (18.2%)	23 (50.0%)	<0.001
	No	5 (20.8%)	18 (81.8%)	23 (50.0%)	
Mirrors	Yes	24 (100.0%)	20 (90.9%)	44 (95.7%)	0.131
	No	0 (0.0%)	2 (9.1%)	2 (4.3%)	

TABLE 5 Likert Questionnaire.

Likert questions (1 = *disagree*, 4 = *agree*)	Intervention ($n = 24$)	Control ($n = 25$)
I follow a strategy during the investigation	3.2 ± 0.8	3.4 ± 0.8
My approach is suitable for the detection of all oral mucosal changes	3.0 ± 0.8	3.2 ± 0.6
My procedure is suitable for the detection of malignant oral mucosal lesions	3.2 ± 0.8	3.3 ± 0.6
I selected all the necessary examination aids (mouth mirror, lighting conditions, and patient position on chair) optimally for the examination	3.2 ± 0.8	3.4 ± 0.6
I was disturbed by the eye-tracking glasses when performing the intraoral examination	1.5 ± 0.7	1.6 ± 0.7
I can imagine wearing the eye tracker in a real-world setting	2.7 ± 0.8	2.9 ± 0.7

Three important considerations arise from the fact that sensitivity is not yet at 100%. First, despite the best intentions of healthcare providers, they must always be aware of this sensitivity during oral examinations or cancer follow-up. Second, we need to understand why oral lesions are overlooked in order to take countermeasures. Third, if countermeasures to improve oral examinations are not sufficient, we need to think of other, auxiliary approaches, such as artificial intelligence[42] or examiner-independent screening approaches like salivary markers.[14] In the following, we will take a closer look at all three aspects.

From theoretical considerations, the observed results of non-detection of the drawn dots could be caused by two factors. First, perception is based on brain processes, not just what crosses our visual field. Perceptual filters, gestalt, and inattentional blindness may affect the detection of skin changes during visual inspection.[43] Perceptual filters, including confirmation biases and search satisfying biases, limit our perceptions.[44] Gestalt is a psychological theory that describes the cognitive processes involved in perception. This includes, for example, the fact that we group objects that are closer to each other or that missing information is filled in by the brain.[45] In addition, inattentional blindness and change blindness cause an object to be visually present but not perceived.[46] Second, the surface area of the oral cavity of an adult is estimated to be about 214.7 cm^2 (of which 45.3 cm^2 is occupied by the teeth).[47] Respectively, about 169.4 cm^2 needs to be scanned by the examiner for any mucosal changes, such as the dots drawn in our simulated patient. Although the skin has a much larger surface area (estimated to be 1.16–2.17 m^2),[48] the oral cavity is not a "flat" surface; instead, it corresponds to the inner surface of a sphere that has "hills" and "valleys," causing large parts of the surfaces to touch each other. Therefore, a brief look inside the oral cavity is not sufficient; instead, it is necessary to look closely at all areas in a systematic fashion.

The relevance of the aforementioned factors in oral examinations is supported by the analysis of our eye-tracking data. In some cases, we observed that the drawn dots were not cognitively perceived despite their clear visibility and the participants gazing at them. This can be explained by inattentional blindness. In addition, three dots, specifically the ones on the upper lip, anterior palate, and floor of the mouth, showed a considerably reduced detection rate (Table 2). Our eye-tracking evaluation revealed that the dot on the inner right side of the upper lip was often hidden by a dental mirror or fingers. Another drawn dot was in close proximity (see Figure 2B), which may have caused the search to stop due to search satisfaction bias. The dot on the anterior palate was hidden by a palatal plica. To detect this, it was necessary to look into the mouth with a mirror or at least tangentially, which was not done by every examiner. The drawn dot on the floor of the mouth was visible only when the floor of the mouth was pressed down slightly, pulling the mucosa apart, which was also an issue for some examiners.

Without the use of eye tracking technology, we would not have been able to obtain these insights. Eye tracking enables very precise determination of the gaze point and eye movement. This can be done either remotely or with a head-mounted eye tracker, which uses cameras or sensors to monitor the eyes. In recent years, eye trackers have been increasingly used in research,[49] and for example revealed that the duration of fixation are significantly associated with the correct diagnosis of caries.[50]

Despite these insights, sensitivity is not set in stone, as shown by the comparison between a standardized (SOE) and an individual approach to conventional oral examination (COE). The reasons were, that on average, the SOE group performed more steps (i.e., higher degree of systematic) according to the recommended examination steps than the COE group with an individual approach (12.50 ± 0.78 vs. 10.64 ± 1.89, $p < 0.001$) (Figure 4B). This was associated with more drawn dots detected. At the same time, the detection rate correlated with the duration of the examination (Figure 4C). Although this was not significantly different between the two groups ($4{:}14 \pm 1{:}13$ vs. $3{:}36 \pm 1{:}22$ [min:s]), the examination duration was significantly longer for participants who recognized 90% or more drawn dots ($4{:}25 \pm 1{:}14$ vs. $3{:}33 \pm 1{:}17$). This substantiates the phrase that 5 min of oral examination can already save a patient's life[33] and highlights the importance of minimizing time pressure during oral examination. Interestingly, only the SOE group self-assessed their detection rate negatively (Figure 4D). The reason for this could be the Dunning–Kruger effect, which describes that incompetent people tend to overestimate their abilities.[51] This effect was also demonstrated in dental students who rated themselves as less competent at midterm than at the beginning of the semester.[52] In this respect, although COE is considered to be a systematic examination, our results show that there is a gap between the assumed and actual completeness of the examination. Interestingly, dental students tended to have a higher mean detection rate than dentists and specialists ($p = 0.25$), but without significance. This is in line with other studies where dental students tended to have a higher sensitivity in diagnosing caries on radiographs or oral pathology on images.[53,54] This may be because they see participation in the study as an opportunity to demonstrate their skills, or it may be because they are over-cautious.

Consequently, it is crucial to ensure that the entire surface of the oral cavity is examined in a standardized fashion. This helps to ensure that fewer examination steps are omitted. Furthermore, oral changes should not be missed due to cognitive errors, while cognitive debiasing strategies already exist to minimize such errors. These include developing insight/awareness, considering alternatives, metacognition, reducing dependence on memory, targeted training, simulation, cognitive constraint strategies, task facilitation, minimizing time pressure, accountability, and feedback.[44]

But even with these measures, we are unlikely to capture all oral lesions in every clinical encounter. In fact, the detection of an oral change is further complicated by the fact that oral changes can be more subtle or can be misinterpreted.[16,32] Therefore, in addition to the detection rate, the clinical evaluation has an impact, especially on early diagnosis. As a result, we should consider complementary approaches. One possibility would be to photograph the entire intraoral surface.[55] For example, sequential total body photography can help detect new skin changes.[56] A combination of the deep learning methods developed in recent years could be combined with an entire intraoral scan. In fact, deep learning models trained on about 44 000 clinical images of OC showed an accuracy of 92.3% (95% CI: 90.2–94.3%) on a clinical test dataset, which is comparable to OC experts at 92.4% accuracy (95% CI: 91.2–93.6%).[42] In addition to optical imaging technologies, saliva markers could be used. However, to date, none of these methods have been established in primary care.[14]

Another possibility is the use of fluorescence visualization, which seems promising.[57,58] However, this would require additional training and incur costs.

Finally, our developed concept of a simulated patient and the implemented training program to SOE could address several of the issues raised in this paper. This setting can be used for specific training sessions and simulations. The drawn points only have to be found without any further morphological interpretation by the examiner, which simplifies the task and facilitates the learning of a systematic examination. The proposed interventions could also be used to provide feedback to each participant regarding how well they performed. Furthermore, these interventions could include eye tracking, feedback, and metacognition support to further enhance their efficacy. In this regard, eye tracking has already been successfully used to facilitate the teaching of objective, structured clinical examinations in medical school.[59]

However, our study has some limitations and weaknesses. First, we used only one simulated patient. Although this had the advantage that we did not have additional confounding variance to examine the effect between SOE/COE, the disadvantage is that we cannot estimate the effect of different simulated patients. In addition, the simulated oral lesions of black dots were highly evident, probably increasing the sensitivity rate. They could be changed to a less obvious color such as white or reddish, which would be closer to real oral lesions, but might make the analysis with the eye-tracker more difficult. Another aspect is that we had a fixed number of 10 simulated oral lesions. This could be varied to get a good estimate of specificity, positive and negative predictive values. These aspects should be addressed in a follow-up study, while this study can serve as a methodological basis.

5 | CONCLUSION

Training in a standardized approach to oral examination led to a significant improvement in the level of systematics, resulting in an increased detection rate. Any oral examination should take at least 5 min. The inner lip, anterior palate, and floor of the mouth were identified as areas with low detection rates. Therefore, special attention should be paid to proper examination techniques, and examiners should be aware of the presence of various perceptual and cognitive errors. The findings of this study should be incorporated into the curricula for all healthcare providers, who examine the oral cavity on regular basis. Its widespread implementation could be an additional strategy to increase the early detection rates of OPMD and OC.

AUTHOR CONTRIBUTIONS

Conceptualization: B.P., M.L., B.C., J.V., and R. R. *Methodology*: B.C., J.V., M.L., and B.P. *Validation*: B.P., B. C., and A.R. *Formal analysis*: B.P., M.L., B.C., R.R., and F. H. *Investigation*: B.P., K.G., J.B., and B.C. *Resources*: M.L., R. R., and F.H. *Data curation*: B.C. and B.P. *Writing – original draft preparation*: B.P. *Writing – review and editing*: B.P., M. L., B.C., R.R., F.H., J.V., K.G., J.B., and A.R. *Visualization*: B. P. and B.C. *Supervision*: R.R. *Project administration*: B.P., B. C., and M.L. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

We would like to express our gratitude to all the dental students, dentists, and specialists who participated in this study. We would like to thank the Audiovisual Media Center of RWTH Aachen University for their kind support in the production of the training video. Behrus Puladi was funded by the Clinician Scientist Program of the Medical Faculty of the RWTH Aachen University. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

The study was approved by the Ethics Committee of RWTH Aachen University (approval number EK 076/22 from 28.04.2022). Informed consent was obtained from all subjects involved in the study.

ORCID

Behrus Puladi https://orcid.org/0000-0001-5909-6105
Beatrice Coldewey https://orcid.org/0000-0001-9564-9467
Julia S. Volmerg https://orcid.org/0000-0003-4945-4604
Kim Grunert https://orcid.org/0009-0007-5999-7094
Jeff Berens https://orcid.org/0009-0009-9263-7394
Ashkan Rashad https://orcid.org/0000-0003-3743-8658
Frank Hölzle https://orcid.org/0000-0002-2298-3517
Rainer Röhrig https://orcid.org/0000-0002-0032-5118
Myriam Lipprandt https://orcid.org/0000-0001-9371-0551

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209-249.

2. Mello FW, Miguel AFP, Dutra KL, et al. Prevalence of oral potentially malignant disorders: a systematic review and meta-analysis. *J Oral Pathol Med*. 2018;47:633-640.
3. Iocca O, Sollecito TP, Alawi F, et al. Potentially malignant disorders of the oral cavity and oral dysplasia: a systematic review and meta-analysis of malignant transformation rate by subtype. *Head Neck*. 2020;42:539-555.
4. Scott S, McGurk M, Grunfeld E. Patient delay for potentially malignant oral symptoms. *Eur J Oral Sci*. 2008;116:141-147.
5. Brocklehurst P, Kujan O, O'Malley LA, Ogden G, Shepherd S, Glenny A-M. Screening programmes for the early detection and prevention of oral cancer. *Cochrane Database Syst Rev*. 2013;2021:CD004150.
6. Walsh T, Liu JLY, Brocklehurst P, et al. Clinical assessment to screen for the detection of oral cavity cancer and potentially malignant disorders in apparently healthy adults. *Cochrane Database Syst Rev*. 2013;2013:CD010173.
7. Walsh T, Macey R, Kerr AR, Lingen MW, Ogden GR, Warnakulasuriya S. Diagnostic tests for oral cancer and potentially malignant disorders in patients presenting with clinically evident lesions. *Cochrane Database Syst Rev*. 2021;7:CD010276.
8. Gómez I, Seoane J, Varela-Centelles P, Diz P, Takkouche B. Is diagnostic delay related to advanced-stage oral cancer? A meta-analysis. *Eur J Oral Sci*. 2009;117:541-546.
9. Zanoni DK, Montero PH, Migliacci JC, et al. Survival outcomes after treatment of cancer of the oral cavity (1985–2015). *Oral Oncol*. 2019;90:115-121.
10. Petersen PE. Oral cancer prevention and control—the approach of the World Health Organization. *Oral Oncol*. 2009;45:454-460.
11. Olson CM, Burda BU, Beil T, Whitlock EP. *Screening for Oral Cancer: A Targeted Evidence Update for the U.S*. Preventive Services Task Force; 2013.
12. Hertrampf K, Kunkel M. S2k-Leitlinie: Diagnostik und Management von Vorläuferläsionen des oralen Plattenepithelkarzinoms in der Zahn-, Mund- und Kieferheilkunde. AWMF; 2019. https://www.awmf.org/uploads/tx_szleitlinien/007-092l_S2k_orale_Vorlaeuferlaesion_Plattenepithelkarzinom_2020-04_1.pdf. Accessed August 25, 2021.
13. Sankaranarayanan R, Ramadas K, Thara S, et al. Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India. *Oral Oncol*. 2013;49:314-321.
14. Warnakulasuriya S, Kerr AR. Oral cancer screening: past, present, and future. *J Dent Res*. 2021;100:1313-1320.
15. Downer MC, Moles DR, Palmer S, Speight PM. A systematic review of test performance in screening for oral cancer and precancer. *Oral Oncol*. 2004;40:264-273.
16. Mortazavi H, Baharvand M, Mehdipour M. Oral potentially malignant disorders: an overview of more than 20 entities. *J Dent Res Dent Clin Dent Prospects*. 2014;8:6-14.
17. Ergun S, Ozel S, Koray M, Kürklü E, Ak G, Tanyeri H. Dentists' knowledge and opinions about oral mucosal lesions. *Int J Oral Maxillofac Surg*. 2009;38:1283-1288.
18. Cannick GF, Horowitz AM, Drury TF, Reed SG, Day TA. Assessing oral cancer knowledge among dental students in South Carolina. *J Am Dent Assoc*. 2005;136:373-378.
19. Carter LM, Ogden GR. Oral cancer awareness of undergraduate medical and dental students. *BMC Med Educ*. 2007;7:44.
20. Dumitrescu AL, Ibric S, Ibric-Cioranu V. Assessing oral cancer knowledge in Romanian undergraduate dental students. *J Cancer Educ*. 2014;29:506-513.
21. de Lima Medeiros Y, de Matos Silveira G, Clemente VB, Leite ICG, Vilela EM, de Abreu Guimarães LD. Knowledge about oral cancer among dental students and primary health care dentists: a Brazilian study. *J Dent Educ*. 2022;86:1488-1497.
22. Ozdemir-Ozenen D, Tanriover O, Ozenen G, Ozdemir-Karatas M, Ozcakir-Tomruk C, Tanalp J. Dental education for prevention of oral cancer in Turkey: needs for changing the curriculum. *J Cancer Educ*. 2022;37:1496-1503.
23. Taneja P, Marya CM, Jain S, Nagpal R, Kataria S. Assessment of knowledge, attitude, and practice regarding oral cancer among dental graduates—a web-based survey. *J Cancer Educ*. 2022;37:1194-1200.
24. Rai P, Goh CE, Seah F, et al. Oral cancer awareness of tertiary education students and general public in Singapore. *Int Dent J*. 2023;73:651-658.
25. Shah A, Bhushan B, Akhtar S, Singh PK, Garg M, Gupta M. Effectiveness of mouth self-examination for screening of oral premalignant/malignant diseases in tribal population of Dehradun district. *J Family Med Prim Care*. 2020;9:4381-4385.
26. Brands MT, Smeekens EAJ, Takes RP, et al. Time patterns of recurrence and second primary tumors in a large cohort of patients treated for oral cavity cancer. *Cancer Med*. 2019;8:5810-5819.
27. Brands MT, Brennan PA, Verbeek ALM, Merkx MAW, Geurts SME. Follow-up after curative treatment for oral squamous cell carcinoma. A critical appraisal of the guidelines and a review of the literature. *Eur J Surg Oncol*. 2018;44:559-565.
28. Farah CS, Janik M, Woo S-B, Grew J, Slim Z, Fox SA. Dynamic real-time optical microscopy of oral mucosal lesions using confocal laser endomicroscopy. *J Oral Pathol Med*. 2023;52:539-547.
29. Tomo S, Miyahara GI, Simonato LE. History and future perspectives for the use of fluorescence visualization to detect oral squamous cell carcinoma and oral potentially malignant disorders. *Photodiagnosis Photodyn Ther*. 2019;28:308-317.
30. Chuang S-L, Su WW-Y, Chen SL-S, et al. Population-based screening program for reducing oral cancer mortality in 2,334,299 Taiwanese cigarette smokers and/or betel quid chewers. *Cancer*. 2017;123:1597-1609.
31. Jemal A, Ward EM, Johnson CJ, et al. Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *J Natl Cancer Inst*. 2017;109:djx030.
32. Mashberg A. Diagnosis of early oral and oropharyngeal squamous carcinoma: obstacles and their amelioration. *Oral Oncol*. 2000;36:253-255.
33. Mignogna MD, Fedele S. Oral cancer screening: 5 minutes to save a life. *Lancet*. 2005;365:1905-1906.
34. Spencer WG. Discussion on the etiology and treatment of carcinoma of the tongue. *Br Med J*. 1914;2:457-459.
35. Gibson J. Oral cancer-CPD and the GDC. *Br Dent J*. 2018;225:884-888.
36. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
37. Reichart PA, Philipsen HP. *Oralpathologie*. 3rd ed. Thieme; 1999:285.

38. National Institute of Dental and Craniofacial Research. Detecting oral cancer: A guide for health care professionals. https://www.nidcr.nih.gov/sites/default/files/2020-10/Detecting-Oral-Cancer-Healthcare-Professionals.pdf. Accessed February 15, 2023.
39. Rethman MP, Carpenter W, Cohen EEW, et al. Evidence-based clinical recommendations regarding screening for oral squamous cell carcinomas. *J Am Dent Assoc*. 2010;141:509-520.
40. Mathew B, Sankaranarayanan R, Sunilkumar KB, Kuruvila B, Pisani P, Nair MK. Reproducibility and validity of oral visual inspection by trained health workers in the detection of oral precancer and cancer. *Br J Cancer*. 1997;76:390-394.
41. Sankaranarayanan R, Ramadas K, Thomas G, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet*. 2005;365:1927-1933.
42. Fu Q, Chen Y, Li Z, et al. A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: a retrospective study. *EClinicalMedicine*. 2020;27:100558.
43. Ko CJ, Braverman I, Sidlow R, Lowenstein EJ. Visual perception, cognition, and error in dermatologic diagnosis: Key cognitive principles. *J Am Acad Dermatol*. 2019;81:1227-1234.
44. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med*. 2003;78:775-780.
45. Wagemans J, Elder JH, Kubovy M, et al. A century of gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol Bull*. 2012;138:1172-1217.
46. Jensen MS, Yao R, Street WN, Simons DJ. Change blindness and inattentional blindness. *Wiley Interdiscip Rev Cogn Sci*. 2011;2:529-546.
47. Collins LM, Dawes C. The surface area of the adult human mouth and thickness of the salivary film covering the teeth and oral mucosa. *J Dent Res*. 1987;66:1300-1302.
48. Yu C-Y, Lin C-H, Yang Y-H. Human body surface area database and estimation formula. *Burns*. 2010;36:616-629.
49. Carter BT, Luke SG. Best practices in eye tracking research. *Int J Psychophysiol*. 2020;155:49-62.
50. Snell S, Bontempo D, Celine G, Anthonappa R. Assessment of medical practitioners' knowledge about paediatric oral diagnosis and gaze patterns using eye tracking technology. *Int J Paediatr Dent*. 2021;31:810-816.
51. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*. 1999;77:1121-1134.
52. Surdilovic D, Adtani P, Fuoad SA, Abdelaal HM, D'souza J. Evaluation of the Dunning-Kruger effects among dental students at an academic training institution in UAE. *Acta Stomatol Croat*. 2022;56:299-310.
53. Mileman PA, van den Hout WB. Comparing the accuracy of Dutch dentists and dental students in the radiographic diagnosis of dentinal caries. *Dentomaxillofac Radiol*. 2002;31:7-14.
54. Sbricoli L, Zago R, Cavallin F, Stellini E, Bacci C. Diagnostic ability in oral pathology among different population clusters. *Oral Dis*. 2023;1-7. doi:10.1111/odi.14689
55. Lin I, Datta M, Laronde DM, Rosin MP, Chan B. Intraoral photography recommendations for remote risk assessment and monitoring of oral mucosal lesions. *Int Dent J*. 2021;71:384-389.
56. Ji-Xu A, Dinnes J, Matin RN. Total body photography for the diagnosis of cutaneous melanoma in adults: a systematic review and meta-analysis. *Br J Dermatol*. 2021;185:302-312.
57. Simonato LE, Tomo S, Scarparo Navarro R, Balbin Villaverde AGJ. Fluorescence visualization improves the detection of oral, potentially malignant, disorders in population screening. *Photodiagnosis Photodyn Ther*. 2019;27:74-78.
58. Simonato LE, Tomo S, Miyahara GI, Navarro RS, Villaverde AGJB. Fluorescence visualization efficacy for detecting oral lesions more prone to be dysplastic and potentially malignant disorders: a pilot study. *Photodiagnosis Photodyn Ther*. 2017;17:1-4.
59. Grima-Murcia MD, Sanchez-Ferrer F, Ramos-Rincón JM, Fernández E. Use of eye-tracking technology by medical students taking the objective structured clinical examination: descriptive study. *J Med Internet Res*. 2020;22:e17719.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Puladi B, Coldewey B, Volmerg JS, et al. Improving detection of oral lesions: Eye tracking insights from a randomized controlled trial comparing standardized to conventional approach. *Head & Neck*. 2024;1-13. doi:10.1002/hed.27687