Head and Neck Pathology (2020) 14:156–165
https://doi.org/10.1007/s12105-019-01035-z

ORIGINAL PAPER

Inter-observer Variability in the Diagnosis of Proliferative Verrucous Leukoplakia: Clinical Implications for Oral and Maxillofacial Surgeon Understanding: A Collaborative Pilot Study

Jasbir D. Upadhyaya[1] · Sarah G. Fitzpatrick[2] · Donald M. Cohen[2] · Elizabeth A. Bilodeau[3] · Indraneel Bhattacharyya[2] · James S. Lewis Jr.[4] · Jinping Lai[5] · John M. Wright[6] · Justin A. Bishop[7] · Marino E. Leon[5] · Mohammed N. Islam[2] · Raja Seethala[8] · Ricardo J. Padilla[9] · Roman Carlos[10] · Susan Müller[11] · Lester D. R. Thompson[12]

Received: 12 February 2019 / Accepted: 1 April 2019 / Published online: 10 April 2019

Abstract

The use of diverse terminology may lead to inconsistent diagnosis and subsequent mistreatment of lesions within the proliferative verrucous leukoplakia (PVL) spectrum. The objectives of this study were: (a) to measure inter-observer variability between a variety of pathologists diagnosing PVL lesions; and (b) to evaluate the impact of diverse terminologies on understanding, interpretation, and subsequent treatment planning by oral and maxillofacial surgeons (OMFS). Six oral pathologists (OP) and six head and neck pathologists (HNP) reviewed 40 digitally scanned slides of PVL-type lesions. Inter-observer agreement on diagnoses was evaluated by Fleiss' kappa analysis. The most commonly used diagnostic terminologies were sent to ten OMFS to evaluate their resulting interpretations and potential follow-up treatment approaches. The overall means of the surgeons' responses were compared by Student *t* test. There was poor inter-observer agreement between pathologists on the diagnosis of PVL lesions ($\kappa = 0.270$), although there was good agreement ($\kappa = 0.650$) when diagnosing frankly malignant lesions. The lowest agreement was in diagnosing verrucous hyperplasia (VH) with/without dysplasia, atypical epithelial proliferation (AEP), and verrucous carcinoma (VC). The OMFS showed the lowest agreement on identical categories of non-malignant diagnoses, specifically VH and AEP. This study demonstrates a lack of standardized terminology and diagnostic criteria for the spectrum of PVL lesions. We recommend adopting standardized criteria and terminology, proposed and established by an expert panel white paper, to assist pathologists and clinicians in uniformly diagnosing and managing PVL spectrum lesions.

Keywords Inter-observer variability · Proliferative verrucous leukoplakia · Verrucous hyperplasia · Papillary squamous cell carcinoma · Atypical epithelial proliferation

Introduction

Proliferative verrucous leukoplakia (PVL) is an aggressive, recalcitrant form of leukoplakia requiring special clinical attention due to the high rate of malignant transformation [1]. The lesions are slow-growing yet persistent, multifocal, and resistant to all forms of treatment. The effects of field cancerization lead to multifocal and recurrent lesions in PVL patients [2, 3]. Hansen et al. described the wide histopathologic spectrum of PVL with lesions ranging from benign hyperkeratosis (HK) to frank squamous cell carcinoma (SCC) [1]. As such, there is no single defining histopathologic feature for PVL, and the term is generally limited to clinical use only. The lack of specific histologic criteria

This study was presented at the American Academy of Oral and Maxillofacial Pathology (AAOMP) annual meeting in June 2018 at Vancouver, British Columbia, Canada.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12105-019-01035-z) contains supplementary material, which is available to authorized users.

✉ Jasbir D. Upadhyaya
jupadhy@siue.edu

Extended author information available on the last page of the article

necessitates a combination of clinical and histopathologic findings for appropriate diagnosis. This may impede early diagnosis and create diagnostic and management challenges for both pathologists and clinicians.

Over time, different diagnostic terms have been used by pathologists for PVL lesions. This diverse terminology may lead to a lack of consistency in both diagnosis and subsequent treatment. For these reasons, there is a need to clarify the terminology for better prediction of biologic behavior, selection of appropriate treatment, and development of a standardized set of diagnostic criteria for PVL. The objectives of this study were twofold: (a) to discern the inter-observer variability between pathologists in the diagnosis of PVL lesions; and (b) to evaluate the impact of diverse histologic terms on interpretation and treatment planning by oral and maxillofacial surgeons (OMFS). To our knowledge, this is the first study to assess inter-observer variability between pathologists in the diagnosis of PVL lesions.

Materials and Methods

Approval was obtained from the University of Florida Institutional Review Board (IRB). A retrospective search for PVL patients, seen between 1994 and 2017, was performed in the UF College of Dentistry Oral Medicine Clinic Database. Forty lesions from 24 patients, comprising various stages of the PVL spectrum, were selected for the study. The inclusion criteria were: (a) patients with a clinical diagnosis of PVL which required multifocal lesions with clinical progression over time and high recurrence rates; (b) patients with at least one biopsy proven lesion supportive of the clinical diagnosis of PVL; and (c) at least 3 years of clinical follow up or progression to malignancy before that time. The interpretation was based on a combination of clinical and histopathologic findings. Exclusion criteria were fragmented or small biopsies or lack of sufficient clinical information. Slide and case selection was done by JDU and SGF who did not participate as reviewing pathologists. The slides were digitally scanned at 20× magnification using an Aperio Scanscope digital scanner (Leica Biosystems, San Diego, CA, USA). Six board-certified oral pathologists (OP) and six board-certified pathologists with expertise in head and neck pathology (HNP) participated in the study. The slides were shared with the 12 pathologists electronically via the Aperio eSlide Manager application (Leica Biosystems). All pathologists agreed to participate in a study about PVL criteria and diagnoses; however, no other clinical information or patient data (like age, sex, history of smoking, biopsy location, clinical appearance of lesions, or recurrence, if any) was provided to the participants. A survey was created which required the pathologists to supply the following information: (1) pathology specialist type (oral or head and neck); (2) type of clinical service (academic, hospital, or both); (3) number of years in practice (< 10 years, 10–30 years, or more than 30 years); (4) diagnosis for each of the 40 cases; and (5) opinion on the likelihood of malignant transformation for each case.

Ten OMFS were selected for the second phase of the study. A survey was created and sent to five OMFS in an academic setting and five in a private practice setting. The survey included the following questions: (a) number of years in practice (< 10 years, 10–30 years, or more than 30 years); (b) type of practice (academic/hospital or private); (c) type of training (4-year, 6-year, or other); and (d) opinion on the helpfulness of an additional comment in the biopsy report to the interpretation of the diagnosis (minimally, moderately, or very helpful, or depends on the diagnosis). We provided 15 descriptive non-malignant diagnostic terms which were most commonly used by pathologists and asked the surgeons for their suggested treatment options based on these descriptors. We further queried if the diagnosis was clear and unambiguous in terms of interpretation and management. Many of these terms started with the word "atypical" (see Table 4). Terms of clear-cut malignancy were not included in the survey, as the treatment of SCC is well established and not within the scope of this study. The treatment options included:

1. Observation and re-evaluation in < 6 months, with or without possible additional biopsy;
2. Observation and re-evaluation between 6 and 12 months, with or without possible additional biopsy;
3. Complete surgical excision of remaining lesional tissue and continued close follow up;
4. Laser removal of remaining lesional tissue and continued close follow up;
5. Surgical cold scalpel excision followed by laser ablation of remaining tissue and continued close follow up;
6. Referral to cancer center or academic oral surgery facility for treatment.

For statistical analysis, we collapsed these six treatment options into three main categories: level 1 (observation): options 1 and 2; level 2 (excision of remaining lesional tissue): options 3, 4 and 5; and level 3 (referral): option 6.

Statistical Analysis

Inter-observer variability between pathologists on histologic diagnoses was evaluated by Fleiss multi-rater Kappa analysis using Microsoft Excel 2013. Kappa (κ) score is commonly used to evaluate reliability of paired agreements against pure chance agreement [range 0 (random

agreement) to 1 (perfect agreement)] [4]. The following grading of κ values was used: < 0.40: poor agreement; 0.40–0.75: fair to good agreement; > 0.75: excellent agreement [4, 5]. The means of responses from OMFS were compared by Student *t*-test using IBM SPSS version 25, considering $p < 0.05$, at 95% confidence intervals, as statistically significant.

Results

Survey 1

Seven of the 12 participating pathologists were in practice for 10–30 years, two for < 10 years, and three had been practicing for more than 30 years. Five pathologists participated in a dental school biopsy service, three were hospital-based, and the remaining four worked in both academic and hospital-based services. The histologic diagnoses rendered from the pathologists were compiled into six broad categories as follows:

Category 0	Squamous papilloma;
Category 1	Simple HK with or without low-grade dysplasia;
Category 2	Verrucous hyperplasia (VH)/keratosis with or without low-grade dysplasia;
Category 3	High-grade dysplasia or carcinoma-in situ, with or without verrucous surface change;
Category 4	Verrucous carcinoma (VC) or atypical epithelial proliferation (AEP) suggestive of VC/SCC;
Category 5	Papillary or conventional SCC

Examples of lesions belonging to the five major categories are illustrated in Fig. 1. Squamous papilloma (category 0) does not fall in the spectrum of PVL lesions, but this category was created to accommodate the histologic diagnosis since it was received from a few pathologists.

Our study demonstrated an overall poor inter-observer agreement (κ = 0.270) between pathologists when evaluating the PVL spectrum lesions (Table 1). All pathologists agreed in only two cases, whereas 10 of 12 pathologists agreed on seven of the 40 cases. Not surprisingly, the best agreement (κ = 0.650, fair to good agreement) between pathologists was for category five lesions (Table 1). There was poor agreement on diagnosis for categories 1 (κ = 0.312), 2 (κ = 0.150),

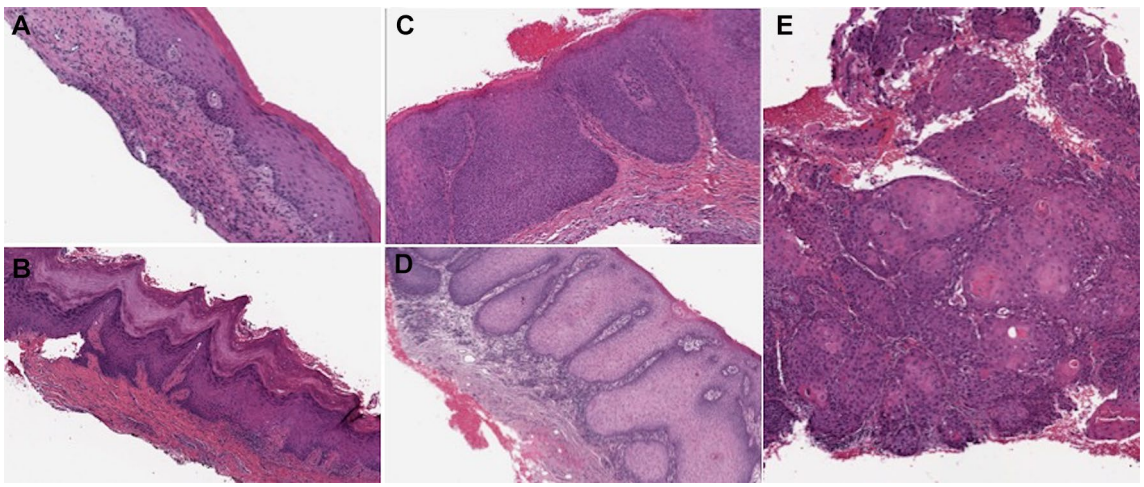

Fig. 1 Classic examples of proliferative verrucous leukoplakia cases belonging to the five major categories. **a** Category 1—hyperkeratosis and mild lichenoid mucositis. **b** Category 2—verrucous hyperplasia. **c** Category 3—acanthosis and moderate to severe epithelial dysplasia. **d** Category 4—verrucous carcinoma. **e** Category 5—well-differentiated squamous cell carcinoma

Table 1 Inter-observer level of agreement, evaluated by Fleiss kappa (κ), between pathologists for diagnosis of PVL lesions

Group	Overall (κ)	Cat 0 (κ)	Cat 1 (κ)	Cat 2 (κ)	Cat 3 (κ)	Cat 4 (κ)	Cat 5 (κ)
All pathologists	0.270	0.177	0.312	0.150	0.192	0.156	0.650
OP only	0.225	None	0.231	0.179	0.111	0.146	0.602
HNP only	0.344	0.392	0.410	0.197	0.328	0.111	0.719

OP Oral pathologists, *HNP* Head and Neck pathologists, *Cat* category, *Cat 0* Squamous papillomas, *Cat 1* hyperkeratosis, with or without low-grade dysplasia, *Cat 2* verrucous hyperplasia/verrucous keratosis, with or without low-grade dysplasia, *Cat 3* high-grade dysplasia, *Cat 4* verrucous carcinoma or atypical epithelial proliferation, *Cat 5* papillary or conventional squamous cell carcinoma

3 (κ = 0.192) and 4 (κ = 0.156) lesions. Overall, HNP had slightly better agreement (κ = 0.344) than the OP (κ = 0.225) on all categories except category 4 (OP, κ = 0.146; HNP, κ = 0.111). HNP had fair agreement (κ = 0.41) on diagnosis of category 1 lesions, whereas poor inter-observer agreement was noted among OP (κ = 0.231). HNP agreed on the diagnosis in eight of the 40 cases. By contrast, OP were in agreement for only three cases. Five of the six HNP agreed on a diagnosis in eight cases and four HNP were in agreement on 11 cases. These cases were either a completely benign lesion (category 1) or invasive SCC (category 5). Five of the six OP agreed on a diagnosis in six of 40 cases, whereas only four OP were in agreement for 10 cases. These cases were either benign lesions (category 1) or frank malignancies (category 5).

The most diverse diagnoses, using many different histologic terms, were received for two cases (represented by Figs. 2 and 3 respectively). Diagnoses ranged from a benign reactive lesion to carcinoma. The various diagnoses received for the first case (Fig. 2) were squamous papilloma (n = 3), verrucous squamous proliferation (n = 1), papillary keratosis with mild to moderate dysplasia (n = 3), PVL/low-grade keratinizing dysplasia with verrucous features (n = 1), atypical papillary hyperplasia (n = 1), carcinoma cuniculatum (n = 1), papillary epithelial proliferation suggestive of papillary SCC (n = 1), and VC (n = 1). The opinions on likelihood of the lesion to transform into malignancy ranged from unlikely to transform (n = 3), to moderately likely (n = 2), and likely to transform (n = 5). Diagnoses for the second case (Fig. 3) were squamous hyperplasia with candidiasis (n = 1), verrucous squamous proliferation (n = 1), VH with focal mild dysplasia (n = 1), PVL (n = 1), mild to moderate dysplasia (n = 1), high-grade dysplasia with acanthosis and hyperkeratosis (n = 1), atypical verrucoid proliferation, consistent with VC (n = 1), VC (n = 4), and verrucous SCC (n = 1).

Two additional verrucoid lesions were also given a number of diagnoses. Interpretation of the lesion in Fig. 4 were marked HK (n = 1), verruciform epithelial HK (n = 1), verruciform squamous epithelial proliferation, PVL if multifocal (n = 1), VH ± dysplasia, keratosis (n = 3), PVL/low-grade dysplasia with verrucous features (n = 1), mild dysplasia (n = 1), moderate/severe dysplasia with verrucous HK (n = 1), atypical verrucous HK (n = 1), and VC (n = 1). The verrucoid lesion in Fig. 5 included the following interpretations: hyperorthokeratosis with dysplasia (n = 3), low-grade keratinizing dysplasia (n = 1), VH with dysplasia (n = 1), verrucous/verrucoid keratosis with dysplasia (n = 3), PVL (n = 1), atypical papillary proliferation with severe HK, consistent with VC (n = 1), and

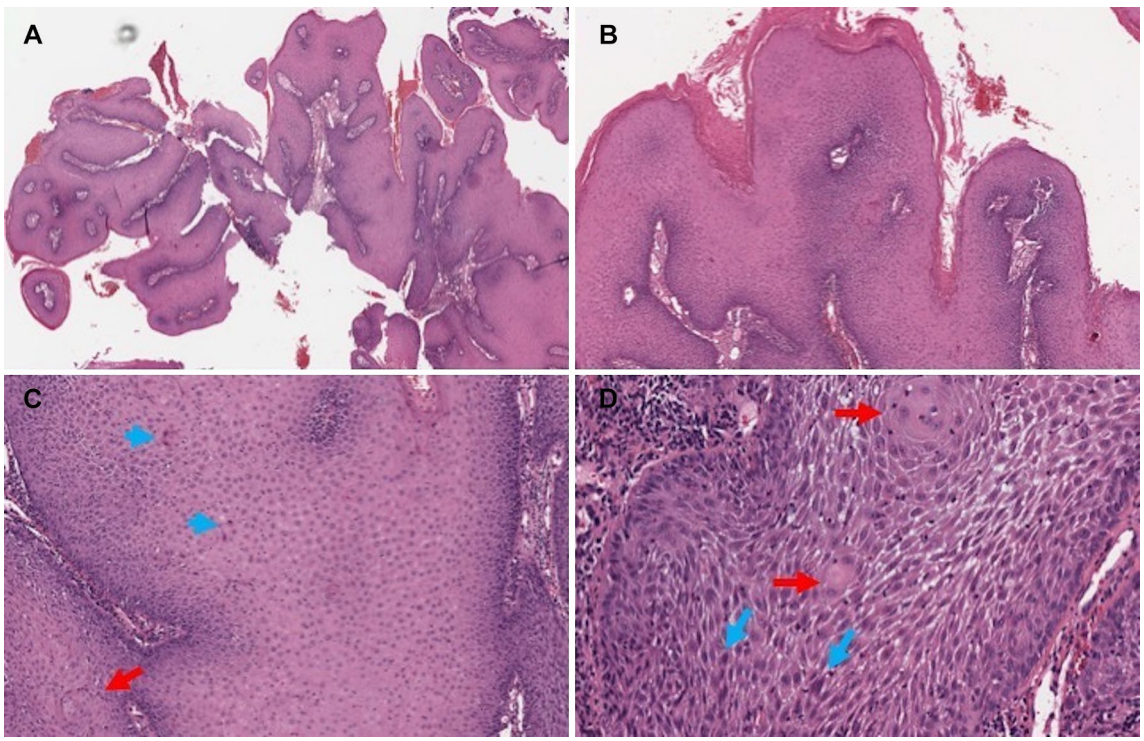

Fig. 2 **a** Biopsy of a proliferative verrucous leukoplakia (PVL) lesion demonstrating papillary proliferation (H&E, ×2). **b** Papillary epithelial proliferation exhibiting prominent keratosis and keratin clefts (H&E, ×4). **c** Multiple foci of epithelial swirls (red arrow) and cells demonstrating hyperchromatic nuclei and pleomorphism (blue arrows, H&E, ×10). **d** Higher magnification of epithelial swirls (red arrows) and pleomorphic cells (blue arrows) noted in the papillary epithelial proliferation (H&E, ×20)

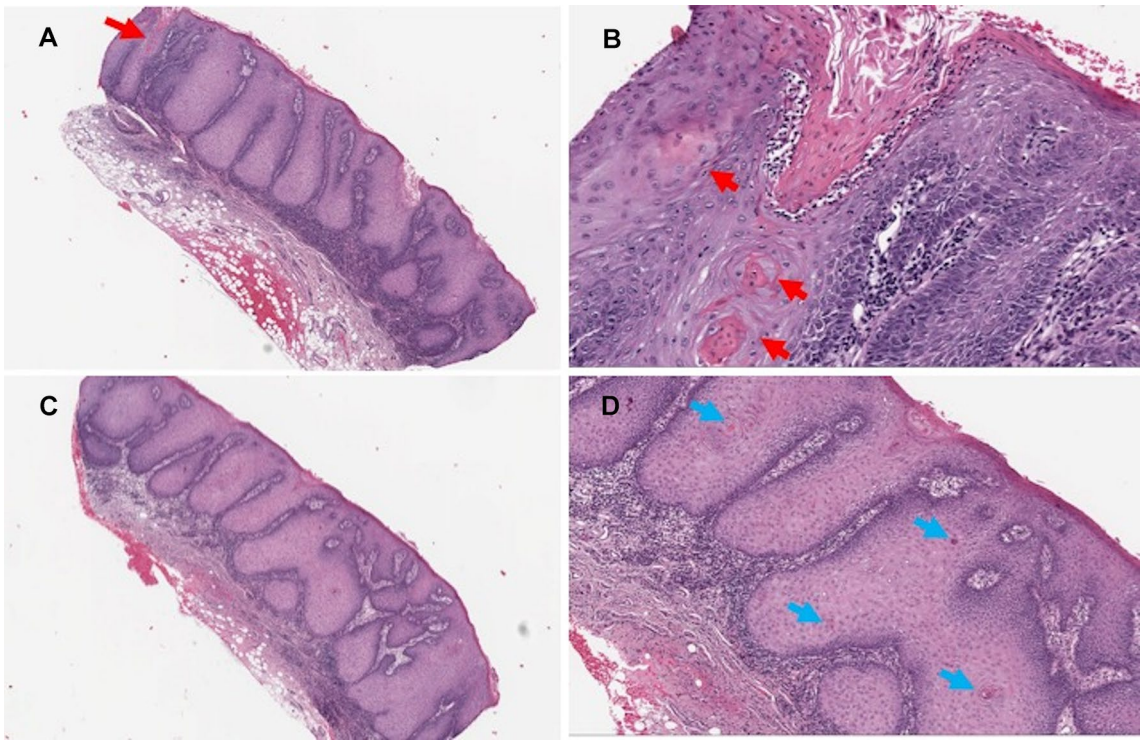

Fig. 3 **a** Histologic features of biopsy of PVL lesion exhibiting hyperplastic elongated epithelial rete ridges and a prominent band of lymphocytic infiltrate beneath the rete ridges. Red arrow marks the area magnified in (**b**) (H&E, ×2.5). **b** Cytologic features demonstrating keratin clefting and keratin pearls marked by red arrows (H&E, ×20). **c** Histologic appearance of the additional section of tissue on the same slide as part A (H&E, ×2.5). **d** Hyperplastic elongated rete ridges showing a glassy cytoplasm and cells exhibiting dyskeratosis (blue arrows) (H&E, ×7)

VC (n = 2). The evaluation of likelihood of the lesion to transform into malignancy ranged from unlikely to transform (n = 1), to moderately likely (n = 6), and likely to transform (n = 4).

For overall levels of agreement over "expected risk for malignant transformation", the risk levels were categorized into three risk levels; risk level 1 (low), risk level 2 (moderate), and risk level 3 (high). Only 11 pathologists participated in this exercise, because one believed all PVL lesions are likely to transform to malignancy. The pathologists achieved poor overall agreement for expected risk for malignant transformation of PVL lesions with κ values of 0.289 (Table 2). The pathologists were in fair to good agreement for high-risk lesions (κ = 0.524). This agreement was consistent among OP and HNP with κ values of 0.564 and 0.572 respectively (Table 2). Lowest agreement between pathologists was seen in lesions with a moderate potential of malignant transformation (κ = 0.108). The HNP had no consensus among themselves for moderate-risk level lesions (κ = -0.009), whereas they had slightly better agreement (κ = 0.322) for low-risk lesions in comparison to OP (κ = 0.278).

Survey 2

Of the ten participating OMFS, six had been in practice for 10–30 years, two < 10 years, and two for more than 30 years. Four attended a 4-year oral surgery residency program, and six received a 6-year MD/oral surgery residency training. When asked about the importance of an additional comment on the biopsy report, seven reported that it was very important to them. Two surgeons stated that the importance of a comment depended on the diagnosis, and it was of moderate importance to one surgeon. For analytical purposes, the 15 most commonly used histologic terminologies for the PVL spectrum lesions (see Table 4) were categorized into five subclasses: (a) HK, with or without atypia (n = 3); (b) verrucopapillary hyperkeratosis (VPHK)/VH (n = 2); (c) atypical VPHK/VH or AEP (n = 4); (d) low-grade dysplasia, with or without VPHK/VH/AEP (n = 3); and (e) high-grade dysplasia, with or without VPHK/VH/AEP (n = 3). The highest agreement between OMFS was for the treatment of HK, with or without atypia, with 83% preferring to observe and 17% wanting to excise the tissue (Fig. 6). The lowest agreement was on treatment of VPHK/VH and atypical VPHK/VH or AEP lesions. Observation was the preferred treatment for VPHK/VH by 55% surgeons, whereas 45%

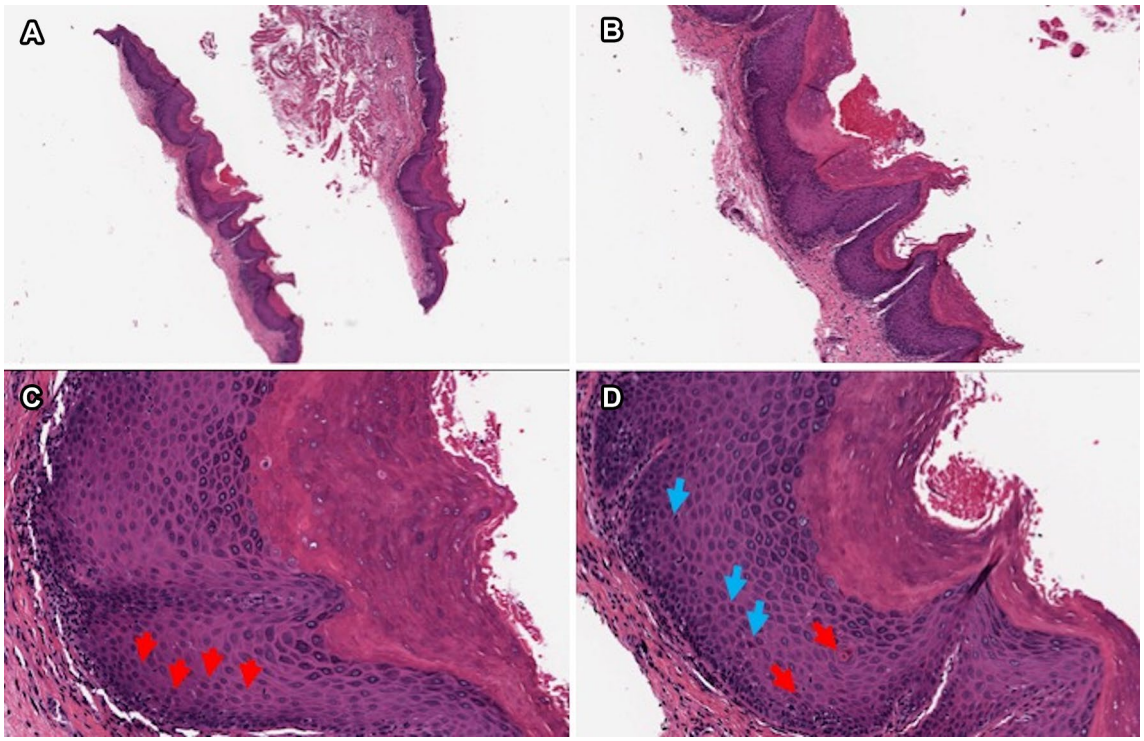

Fig. 4 **a** Histologic findings of a PVL lesion displaying verrucous architecture and prominent keratosis (H&E, ×2). **b** Verrucous hyperplasia demonstrating keratin clefting (H&E, ×4). **c** Prominent hyperkeratosis with cells exhibiting premature keratinization in lower third of the epithelium (red arrows, H&E, ×20). **d** Dyskeratosis (red arrows) and pleomorphic cells (blue arrows) (H&E, ×20)

preferred to remove the remaining tissue. Further, surgeons recommended varied treatments for atypical VPHK/VH or AEP lesions: 50% recommended complete removal of tissue, 40% preferred to monitor and observe the lesion, and 10% would refer the patient to a cancer center or academic facility for treatment (Fig. 6). For low-grade dysplasia ± VPHK/VH/AEP cases, 13% of clinicians preferred to monitor the lesion, 67% recommended excision, and 20% would refer the patient to a cancer center or academic facility for treatment. For high-grade dysplasia ± VPHK/VH/AEP, 56% clinicians preferred to remove the tissue, and 37% said they would send the patient to a referral center. Interestingly, 7% of surgeons opted to observe and monitor the lesion.

Importantly, when we evaluated if the diagnoses were unambiguous in terms of interpretation, OMFS reported that 10 of the 15 diagnoses were unclear to them (Table 3 and 4). Of these, seven diagnostic terminologies started with "atypical". Diagnoses beginning with "atypical" (n = 7) lacked clarity to 47% surgeons. This lack of clarity came from 54.3% academic surgeons and 40% private practice surgeons (Table 3). For diagnoses without "atypical" term (n = 8), 30% OMFS reported that the diagnosis was not clear. Of these, 40% were academic surgeons and 20% private practice surgeons. The number of unclear responses for "atypical" diagnoses when compared to "non-atypical" responses was statistically significant ($p < 0.004$). Among the "atypical" diagnoses, the most ambiguity was for atypical squamous/epithelial proliferation (60%), followed by 50% each for atypical verrucous/papillary hyperplasia, atypical HK, atypical verrucous/papillary HK, and atypical verrucous/papillary proliferation. Forty and thirty percent, respectively, found AEP with low-grade or high-grade dysplasia ambiguous. Of the diagnoses without "atypical" terminology, verrucous/papillary hyperplasia with low-grade or high-grade dysplasia (40% each) and verrucous/papillary HK with high-grade dysplasia (40%) were the most unclear.

Discussion

Several studies have determined the inter-observer variation in pathologic diagnosis of oral lesions, but literature assessing variability in diagnosis of PVL lesions is lacking [6–15]. The various evolutionary stages within the histopathologic spectrum of PVL render a definitive diagnosis difficult. Studies published in the past followed the diagnostic criteria developed by Hansen et al. [1]. Later, reformulated criteria for early diagnosis and better management of PVL was recommended by many authors [16–20]. Understanding the biologic potential of VH,

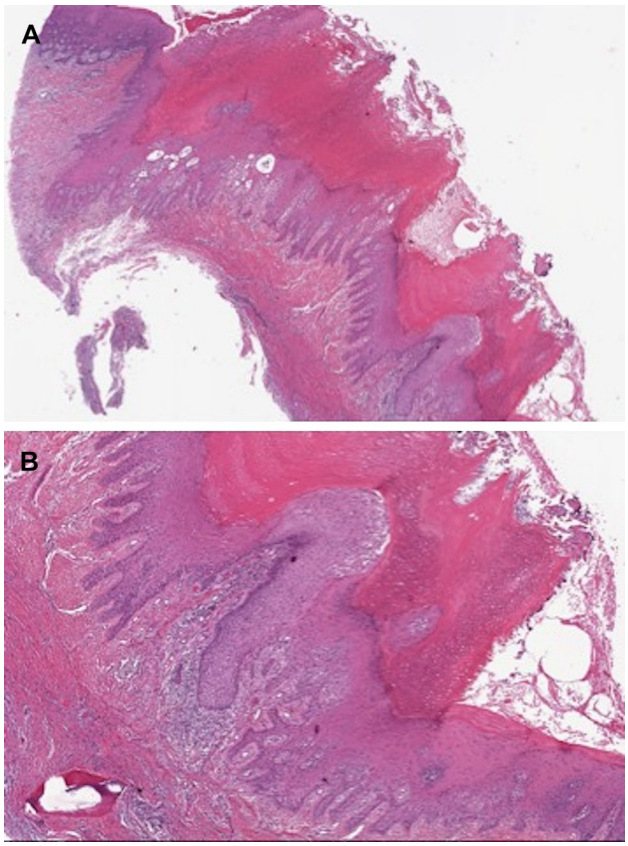

Fig. 5 **a** PVL lesion demonstrating an atypical verrucoid architecture and elongated epithelial rete ridges (H&E, ×2). **b** Higher magnification showing significant keratosis and hyperchromatic nuclei in lower third of the epithelium (H&E, ×5)

Table 2 Overall level of agreement between pathologists over the "expected risk of malignant transformation" of PVL lesions

Group	Overall agreement (κ)	Risk level 1 (low) (κ)	Risk level 2 (moderate) (κ)	Risk level 3 (high) (κ)
All pathologists (n = 11)[a]	0.289	0.220	0.108	0.524
OP (n = 6)	0.391	0.278	0.279	0.564
HNP (n = 5)	0.303	0.322	−0.009	0.572

OP oral pathologist; *HNP* head and neck pathologist

[a]One pathologist did not participate in this exercise because he thought all PVL lesions are likely to transform to malignancy

atypical verrucous lesions, and atypical epithelial lesions is perplexing in the current literature [19, 21]. A lack of specific histopathologic diagnostic criteria, compounded with diverse histologic terminology, makes the diagnosis of PVL lesions challenging and open to different interpretations. This is in agreement with the results of our study where the pathologists had the least agreement when interpreting category 2 (VH/keratosis ± low-grade dysplasia), 3 (high-grade dysplasia or carcinoma-in situ, ± verrucous surface change), and 4 (VC or AEP suggestive of VC/SCC) lesions.

It is obviously disconcerting that interpretations for the same biopsy ranged from reactive to malignant in four of the test cases (Figs. 2, 3, 4, 5). This lack of agreement will clearly have therapeutic consequences for the patient. A wide variety of histologic terms like verruciform/verrucous/papillary HK, verruciform epithelial HK, verrucous/papillary/squamous/epithelial proliferation/hyperplasia, papillary keratosis, atypical epithelial/epitheliomatous/verrucous/verrucoid/papillary HK/proliferation/hyperplasia, VPHK, atypical/verruciform squamoproliferative lesion, atypical VH, verrucous SCC were interchangeably and inconsistently applied by pathologists for diagnosing the cases in our study (see supplemental file). Many of these terms lacked clarity for many of the surgeons (Table 4). The disparate and confusing terminology leads to inconsistency in the diagnosis of PVL and may result in ineffective patient management.

PVL is often a retrospective diagnosis based on both clinical and histologic findings that develop over years, and the clinical features are not always provided for the histologic diagnosis. Like oral epithelial dysplasia, the progression of PVL may not be step-wise, thus, an early lesion may be long-standing and a lesion exhibiting marked verrucous keratosis may be recent. Therefore, it may be better to provide a descriptive diagnosis of the evolutionary stage so that a clearer picture is presented to the surgeon enabling adequate treatment. Differing diagnostic terms result in unacceptable ambiguity in the characterization of these lesions.

Our results demonstrate poor overall agreement between pathologists diagnosing PVL lesions. This is likely due to a lack of well-described diagnostic features which results in use of subjective terminology. Importantly, the clinical information for each case, such as the patient's age, anatomic location(s), previous diagnoses (if any), or clinical impression, would probably have favorably influenced the accuracy of histologic diagnosis had they been known. When κ-values for inter-observer agreement were evaluated for each category, good agreement was only achieved on category 5 lesions (frank carcinoma). Not surprisingly, one of the categories with least agreement between pathologists and most ambiguity in interpretation by OMFS comprises VH/keratosis (Table 1), a known challenging and difficult diagnosis. In an attempt to develop standardized diagnostic criteria for VH in Asian patients, a consensus meeting of clinicians and pathologists was held in Malaysia. Rosnah et al. developed standardized criteria for the diagnosis of exophytic VH so that its potential for malignant transformation could be ascertained [19]. A study by Karabulut et al. evaluated the impact of education and professional training

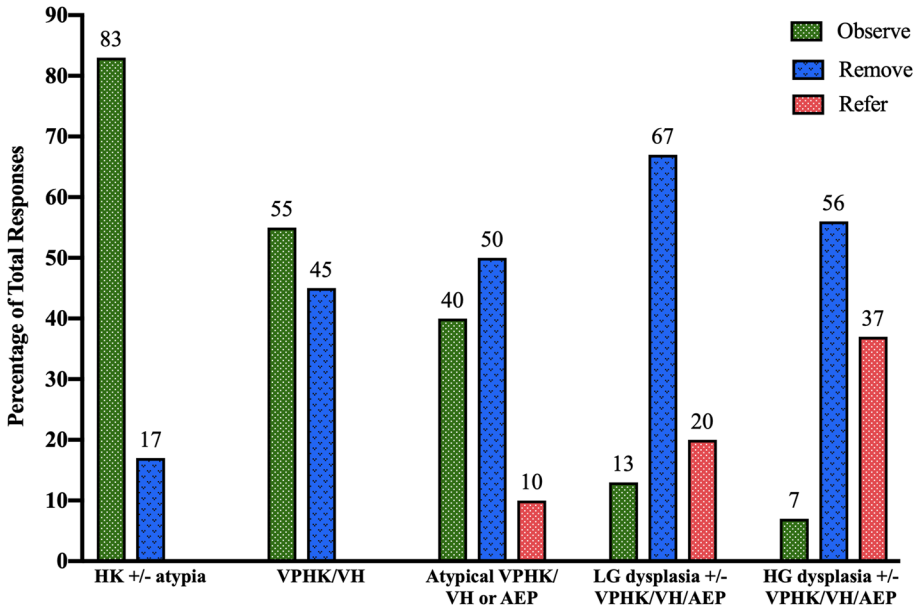

Fig. 6 Overall agreement between the ten OMFS on treatment options for the most commonly used histologic terminologies. The total responses received for treatment of each diagnosis are displayed as percentage. *HK* hyperkeratosis, *VPHK* verrucopapillary hyperkeratosis, *VH* verrucous hyperplasia, *LG* low-grade, *HG* high-grade, *AEP* atypical epithelial proliferation

Table 3 Number of histologic diagnoses which were unclear in terms of interpretation for oral & maxillofacial surgeons

Type of diagnosis	Academic surgeons (n = 5)	Private practice surgeons (n = 5)	Total number of unclear responses (n = 10)
"Atypical" diagnoses (n = 7)	19/35	14/35	33/70
"Non-atypical" diagnoses (n = 8)	16/40	8/40	24/80
"Atypical" versus "Non-atypical" (*t*-test)	p = 0.091	p = 0.004	p = 0.004

Table 4 Histologic diagnoses unclear in terms of interpretation for the oral & maxillofacial surgeons

Diagnosis no.	Diagnoses starting with "atypical" (n = 7)[a]	No. of unclear responses	Diagnosis no.	Diagnoses not starting with "atypical" (n = 8)[b]	No. of unclear responses
1.	Atypical hyperkeratosis	50% (n = 5)	8.	Epithelial/squamous hyperplasia	20% (n = 2)
2.	Atypical verrucous/papillary hyperplasia	50% (n = 5)	9.	Hyperkeratosis with atypia	20% (n = 2)
3.	Atypical verrucous/papillary hyperkeratosis	50% (n = 5)	10.	Verrucous/papillary hyperkeratosis	20% (n = 2)
4.	Atypical verrucous/papillary proliferation	50% (n = 5)	11.	Verrucous/papillary hyperplasia	30% (n = 3)
5.	Atypical squamous/epithelial proliferation	60% (n = 6)	12.	Papillary/verrucous hyperkeratosis with low-grade dysplasia	30% (n = 3)
6.	Atypical epithelial proliferation with low-grade dysplasia	40% (n = 4)	13.	Papillary/verrucous hyperplasia with low-grade dysplasia	40% (n = 4)
7.	Atypical epithelial proliferation with high-grade dysplasia	30% (n = 3)	14.	Papillary/verrucous hyperkeratosis with high-grade dysplasia	40% (n = 4)
	–	–	15.	Papillary/verrucous hyperplasia with high-grade dysplasia	40% (n = 4)

Student t-test shows significant difference between unclear responses for "Atypical" and Non- "Atypical" diagnoses. $P = 0.004$ (significance = $p < 0.05$ at 95% CI)

[a]Total 35 unclear responses of total possible 70 (50%) for "Atypical"

[b]Total 24 unclear responses of total possible 80 (30%) for Non "Atypical"

differences for pathologists on agreement in grading oral epithelial dysplasia [14]. They concluded that inter-observer variability is based on individual differences rather than on factors such as training background. Similar studies assessing diagnostic variability in oral dysplasias reported that the process by which a pathologist makes a diagnosis is to some extent, a subjective interpretation [11, 22].

A limitation of our study is the relatively small number of lesions (40 cases) and cohort of pathologists and surgeons participating. This resulted in small sample size when subgroups (numbers of years in practice, type of training, etc.) were evaluated. Additionally, the participating private practice OMFS were biopsy providers for UF. They may be more attuned to some of the histopathologic terminologies used in this particular biopsy service. Other studies evaluating agreement among pathologists have been criticized for inherent bias based on case selection [23]. Sources of bias include non-random selection of slides or inclusion of only challenging cases. We tried to lessen this bias by incorporating lesions from different evolutionary stages of the PVL spectrum, with attempts to include a similar number of lesions from each stage. The diagnoses varied from easy to difficult.

It is worth mentioning the current study was performed using digitally scanned hematoxylin and eosin-stained slides for more consistent and standardized material for review. There is no potential for different diagnoses based on different levels or deeper tissues as serial slides would create; no breakage; no difference in staining; identical material for review independent of time (i.e., all reviews could be done simultaneously if multiple users were logged in at the same time). Digital pathology is validated and approved for primary diagnosis, along with being utilized for quality assurance purposes, second opinion consults, clinical conferences, and teaching. Many studies have demonstrated that primary histopathologic diagnoses are essentially identical when made digitally using whole slide imaging (WSI) or physical glass slide review [24–26]. Discrepancies in diagnoses between digital and glass slides in previous studies were attributed either to poor image resolution, missed tissue on the digital slide, or individual pathologist's lack of experience using the WSI system. All pathologists who participated in this review were experienced digital pathology users, and thus the use of virtual microscopy was not considered a confounding influence.

In conclusion, an assessment of PVL lesions showed poor reproducibility between pathologists in our study. The ability to reliably recognize PVL is critical because of the high potential for malignant transformation and follow-up ramifications vis-à-vis patient outcome. Therefore, we suggest that standardized diagnostic criteria and histologic terminologies be proposed to assist pathologists and clinicians in uniform diagnosis and management of PVL spectrum lesions.

Acknowledgements The authors are thankful to Drs. John Hardeman, University of Florida College of Dentistry; Andrew Salama, Boston University Henry M. Goldman School of Dental Medicine; Bradley Szutz, Alaska Oral and Facial Cosmetic Surgery, Anchorage, AK; Michael McDermott, Midwest Oral & Maxillofacial Surgery, Omaha, NE; Ravi Chandran, University of Mississipi Medical Center, Jackson, MS; George Kushner, University of Louisville School of Dentistry; Eric Fox, Oral Surgey Coral Springs, FL; Michael Langan, Central Florida Oral & Maxillofacial Surgery; Danielle Freburg-Hoffmeister, University of Florida College of Dentistry; and Carl Kimbler, Northern Plains Oral & Maxillofacial, Aberdeen, SD, for their participation in the oral and maxillofacial surgeon survey.

Compliance with Ethical Standards

Conflict of interest All authors declare that they have no conflict of interest as it relates to this research project.

Ethical Approval All procedures performed in this retrospective data analysis involving human participants were in accordance with the ethical standards of the institutional review board, which did not require informed consent.

References

1. Hansen LS, Olson JA, Silverman S Jr. Proliferative verrucous leukoplakia. A long-term study of thirty patients. Oral Surg Oral Med Oral Pathol. 1985;60:285–98.
2. Bagan JV, Murillo J, Poveda R, Gavalda C, Jimenez Y, Scully C. Proliferative verrucous leukoplakia: unusual locations of oral squamous cell carcinomas, and field cancerization as shown by the appearance of multiple OSCCs. Oral Oncol. 2004;40:440–3.
3. Feller L, Wood NH, Raubenheimer EJ. Proliferative verrucous leukoplakia and field cancerization: report of a case. J Int Acad Periodontol. 2006;8:67–70.
4. Fleiss JL. Statistical methods for rates and proportions. 1st ed. London: John Wiley & Sons; 1981.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
6. Shubhasini AR, Praveen BN, Hegde U, Uma K, Shubha G, Keerthi G, et al. Inter- and intra-observer variability in diagnosis of oral dysplasia. Asian Pac J Cancer Prev. 2017;18:3251–4.
7. Krishnan L, Karpagaselvi K, Kumarswamy J, Sudheendra US, Santosh KV, Patil A. Inter- and intra-observer variability in three grading systems for oral epithelial dysplasia. J Oral Maxillofac Pathol. 2016;20:261–8.
8. Geetha KM, Leeky M, Narayan TV, Sadhana S, Saleha J. Grading of oral epithelial dysplasia: points to ponder. J Oral Maxillofac Pathol. 2015;19:198–204.
9. Manchanda A, Shetty DC. Reproducibility of grading systems in oral epithelial dysplasia. Med Oral Patol Oral Cir Bucal. 2012;17:e935–42.
10. 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.
11. Abbey LM, Kaugars GE, Gunsolley JC, Burns JC, Page DG, Svirsky JA, et al. Intraexaminer and interexaminer reliability in the diagnosis of oral epithelial dysplasia. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1995;80:188–91.
12. Speight PM, Abram TJ, Floriano PN, James R, Vick J, Thornhill MH, et al. Interobserver agreement in dysplasia grading: toward

an enhanced gold standard for clinical pathology trials. Oral Surg Oral Med Oral Pathol Oral Radiol. 2015;120:474–82 **e2**.
13. Fischer DJ, Epstein JB, Morton TH, Schwartz SM. Interobserver reliability in the histopathologic diagnosis of oral pre-malignant and malignant lesions. J Oral Pathol Med. 2004;33:65–70.
14. Karabulut A, Reibel J, Therkildsen MH, Praetorius F, Nielsen HW, Dabelsteen E. Observer variability in the histologic assessment of oral premalignant lesions. J Oral Pathol Med. 1995;24:198–200.
15. van der Meij EH, Reibel J, Slootweg PJ, van der Wal JE, de Jong WF, van der Waal I. Interobserver and intraobserver variability in the histologic assessment of oral lichen planus. J Oral Pathol Med. 1999;28:274–7.
16. Cerero-Lapiedra R, Balade-Martinez D, Moreno-Lopez LA, Esparza-Gomez G, Bagan JV. Proliferative verrucous leukoplakia: a proposal for diagnostic criteria. Med Oral Patol Oral Cir Bucal. 2010;15:e839–45.
17. Carrard VC, Brouns ER, van der Waal I. Proliferative verrucous leukoplakia; a critical appraisal of the diagnostic criteria. Med Oral Patol Oral Cir Bucal. 2013;18:e411–3.
18. Garcia-Chias B, Casado-De La Cruz L, Esparza-Gomez GC, Cerero-Lapiedra R. Diagnostic criteria in proliferative verrucous leukoplakia: evaluation. Med Oral Patol Oral Cir Bucal. 2014;19:e335–9.
19. Rosnah BZ, Thomas GK, Anand R, Jin K, Wm T, Takashi T, et al. Exophytic verrucous hyperplasia of the oral cavity—application of standardized criteria for diagnosis from a consensus report. Asian Pac J Cancer Prev. 2016;17:4491.
20. Villa A, Menon RS, Kerr AR, De Abreu Alves F, Guollo A, Ojeda D, et al. Proliferative leukoplakia: proposed new clinical diagnostic criteria. Oral Dis. 2018;24:749–60.
21. Muller S. Oral epithelial dysplasia, atypical verrucous lesions and oral potentially malignant disorders: focus on histopathology. Oral Surg Oral Med Oral Pathol Oral Radiol. 2018;125:591–602.
22. Pindborg JJ, Reibel J, Holmstrup P. Subjectivity in evaluating oral epithelial dysplasia, carcinoma in situ and initial carcinoma. J Oral Pathol. 1985;14:698–708.
23. Henson DE. Studies on observer variation. Should the rules be changed? Arch Pathol Lab Med. 1991;115:991–2.
24. Al-Janabi S, Huisman A, Nap M, Clarijs R, van Diest PJ. Whole slide images as a platform for initial diagnostics in histopathology in a medium-sized routine laboratory. J Clin Pathol. 2012;65:1107–11.
25. Bauer TW, Schoenfield L, Slaw RJ, Yerian L, Sun Z, Henricks WH. Validation of whole slide imaging for primary diagnosis in surgical pathology. Arch Pathol Lab Med. 2013;137:518–24.
26. Indu M, Rathy R, Binu MP. "Slide less pathology": fairy tale or reality? J Oral Maxillofac Pathol. 2016;20:284–8.

Affiliations

Jasbir D. Upadhyaya[1] · Sarah G. Fitzpatrick[2] · Donald M. Cohen[2] · Elizabeth A. Bilodeau[3] · Indraneel Bhattacharyya[2] · James S. Lewis Jr.[4] · Jinping Lai[5] · John M. Wright[6] · Justin A. Bishop[7] · Marino E. Leon[5] · Mohammed N. Islam[2] · Raja Seethala[8] · Ricardo J. Padilla[9] · Roman Carlos[10] · Susan Müller[11] · Lester D. R. Thompson[12]

1 Section of Diagnostic Sciences, Department of Applied Dental Medicine, Southern Illinois University School of Dental Medicine, 2800 College Avenue, Alton, IL 62002, USA

2 Department of Oral and Maxillofacial Diagnostic Sciences, University of Florida College of Dentistry, Gainesville, FL, USA

3 Department of Diagnostic Sciences, University of Pittsburgh School of Dental Medicine, Pittsburgh, PA, USA

4 Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA

5 Department of Pathology, Immunology and Laboratory Medicine, University of Florida College of Medicine, Gainesville, FL, USA

6 Department of Diagnostic Sciences, Texas A&M College of Dentistry, Dallas, TX, USA

7 Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA

8 Department of Pathology, University of Pittsburgh, Pittsburgh, PA, USA

9 Department of Diagnostic Sciences, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

10 Pathology Division, Centro Clínico de Cabeza y Cuello, Guatemala City, Guatemala

11 Emory University School of Medicine, Atlanta Oral Pathology, Decatur, GA, USA

12 Southern California Permanente Medical Group, Woodland Hills Medical Center, Woodland Hills, CA, USA