

Journal section: Oral Medicine and Pathology
Publication Types: Research

doi:10.4317/jced.61073
<https://doi.org/10.4317/jced.61073>

Challenges in the assessment of epithelial dysplasia in oral lichen planus and oral lichenoid lesion: Inter and intra-observer variability of the WHO criteria and binary system

Letícia-Côgo Marques ¹, Pâmella-de Pinho Montovani ², Lúcio-Souza Gonçalves ³, Karin-Soares Cunha ¹, Arley-Silva Junior ¹, Danielle-Castex Conde ¹

¹ Ph.D. Postgraduate Program in Pathology, School of Medicine, Universidade Federal Fluminense (UFF), Niterói, RJ, Brazil

² M.Sc. Postgraduate Program in Pathology, School of Medicine, Universidade Federal Fluminense (UFF), Niterói, RJ, Brazil

³ Ph.D. Postgraduation Program in Dentistry, Faculty of Dentistry, Universidade Estácio de Sá, Rio de Janeiro-RJ, Brazil

Correspondence:

Antônio Pedro University Hospital
School of Medicine
Department of Pathology
Universidade Federal Fluminense
Marquês do Paraná, 303, 4^o floor
CEP 24033-900, Niterói
Rio de Janeiro, Brazil
daniellecstex@yahoo.com.br

Marques LC, Montovani PP, Gonçalves LS, Cunha KS, Junior AS, Conde DC. Challenges in the assessment of epithelial dysplasia in oral lichen planus and oral lichenoid lesion: Inter and intra-observer variability of the WHO criteria and binary system. J Clin Exp Dent. 2024;16(1):e62-70.

Article Number: 61073 <http://www.medicinaoral.com/odo/indice.htm>
© Medicina Oral S. L. C.I.F. B 96689336 - eISSN: 1989-5488
eMail: jced@jced.es
Indexed in:
Pubmed
Pubmed Central® (PMC)
Scopus
DOI® System

Received: 29/09/2023

Accepted: 14/11/2023

Abstract

Background: Assessment of oral epithelial dysplasia is the gold standard for investigating the risk of malignant progression. The World Health Organization (WHO) methods and the binary system have limitations. This study assess the inter- and intra-observer variability of the architectural and cytological criteria and the classification of the presence and degree of epithelial dysplasia in oral lichen planus (OLP) and oral lichenoid lesion (OLL), using both the 2017 WHO criteria and the binary system.

Material and Methods: The sample consisted of 65 biopsies from lesions classified as OLP and OLL according to the American Academy of Oral and Maxillofacial Pathology (AAOMP) criteria. The histological slides were reevaluated by two oral pathologists.

Results: The individual alterations with most inter-observer disagreement were atypical mitotic figures (43.1%), loss of cohesion between epithelial cells (38.5%) and drop shape rete ridges ridges (38.5%). Inter-observer agreement analysis did not show statistically significant agreement regarding the classification of epithelial dysplasia grade by WHO criteria, only regarding the binary system classification ($k=0.257$; $p=0.035$). Intra-observer agreement analysis by evaluator 1 showed that the classification of epithelial dysplasia grade according to both methods had statistically significant agreement ($k=0.546$; $p=0.004$, $k=0.861$; $p<0.001$). Considering evaluator 2, only the evaluation of the WHO system classification showed statistically significant agreement ($k=0.593$; $p=0.010$).

Conclusions: The evaluation of epithelial dysplasia is subjective and focal changes and inflammatory infiltrate, characteristic of OLP and OLL, can increase the degree of disagreement among evaluators. The binary system presents better inter-observer agreement, while the WHO system presents better intra-observer agreement.

Key words: Oral lichen planus, oral lichenoid lesion, oral lichenoid disease, dysplasia, inter-observer variation, intra-observer variation.

Introduction

Oral lichen planus (OLP) and oral lichenoid lesion (OLL) represent a heterogeneous group of inflammatory diseases that are characterized by similar clinical manifestations and histopathological features (1). Recently, Aguirre-Urizar *et al.* (2) proposed grouping OLP and OLL under the term “oral lichenoid disease”. However, the American Academy of Oral and Maxillofacial Pathology (AAOMP) believes that distinguishing between the two lesions is important to validate the studies that investigate the biological behavior of these diseases (3). The AAOMP (3) and the World Health Organization (WHO) (4) classify OLL as a potentially malignant disorder (PMD). With regard to OLP, there is still no consensus on its potential for malignancy, and the AAOMP (3) includes the absence of epithelial dysplasia in the set of criteria for its classification. However, the WHO (4) classifies OLP as a PMD.

Oral epithelial dysplasia is a set of architectural and cytological epithelial alterations considered as the gold standard for assessing the risk of progression of pre-malignant disorders to oral squamous cell carcinoma (5). Many diagnostic methods have been proposed for evaluating and classifying epithelial dysplasia, with the most commonly used being the WHO method (6) and the binary system (7). However, these methods have some limitations due to subjectivity and the lack of guidelines for their classifications.

The aim of this study was to assess the inter- and intra-observer variability of the architectural and cytol-

gical criteria and the classification of the presence and degree of epithelial dysplasia in OLP and OLL, using both the 2017 WHO criteria⁶ and the binary system (7). Furthermore, the study aimed to discuss the difficulties faced by pathologists during the evaluation process of epithelial dysplasia.

Material and Methods

This study was approved by the Research Ethics Committee of the Antônio Pedro University Hospital of the Fluminense Federal University (APUH-UFF). It is an observational, cross-sectional, and retrospective study. The sample selection was made through a search in the electronic medical records of the Oral Diagnosis Outpatient Clinic of APUH-UFF for patients with oral lichen planus (OLP), oral lichenoid lesion (OLL), oral lichenoid mucositis, and oral lichenoid reaction between the years 2005 and 2019. The lesions were reclassified as OLP and OLL according to criteria of the AAOMP³, with the exception of the criterion of absence of epithelial dysplasia for OLP (Fig. 1).

Lesions with inadequate slides or paraffin blocks for analysis, or those associated with other PMD or candidiasis, were excluded by histopathological analysis through evaluation of stained hematoxylin and eosin (HE) and periodic acid-Schiff (PAS) sections. Cases with candidiasis were excluded due to inflammatory architectural and cytological changes, making the evaluation of epithelial dysplasia difficult (8).

The sample consisted of 65 biopsies (54 of OLP and 11

Oral lichen planus	Oral lichenoid lesion
<p>Clinical criteria:</p> <ul style="list-style-type: none"> • Multifocal symmetric distribution; • White and red lesions exhibiting one or more of the following forms: Reticular/papular, atrophic (erythematous), erosive (ulcerative), plaque, and bullous; • Lesions are not localized exclusively to the sites of smokeless tobacco placement and adjacent to and in contact with dental restorations; Lesion onset does not correlate with the start of a medication and correlate with the use of cinnamon-containing products. 	<p>Clinical criteria:</p> <ul style="list-style-type: none"> • Absence of multifocal symmetric distribution or; • Single lesions or; • Lesions localized exclusively to the sites of smokeless tobacco placement and adjacent to and in contact with dental restorations or; Lesion correlate with the start of a medication and correlate with the use of cinnamon-containing products.
<p>Histopathologic criteria:</p> <ul style="list-style-type: none"> • Band-like or patchy, predominately lymphocytic infiltrate in the lamina propria confined to the epitheliumlamina propria interface; • Basal cell liquefactive (hydropic) degeneration; • Lymphocytic exocytosis; • Absence of verrucous epithelial architectural change. 	<p>Histopathologic criteria:</p> <ul style="list-style-type: none"> • Band-like or patchy, predominately lymphocytic infiltrate in the lamina propria confined to the epitheliumlamina propria interface (it may have other types of inflammatory cells as well as the infiltrate being in deeper and perivascular areas); • Basal cell liquefactive (hydropic) degeneration.

Fig. 1: Clinical and histopathologic criteria for the classification of oral lichenoid diseases. Adapted from the American Academy of Oral and Maxillofacial Pathology (3). * For the diagnosis of OLP, all criteria must be included, and for the diagnosis of OLL, only one or more of the criteria must be present.

of OLL) from 56 participants. Seven participants had more than one biopsy from different areas and/or taken at different times.

- Assessment of the presence and grade of epithelial dysplasia

For the analysis of the presence and degree of epithelial dysplasia, histological slides stained with HE were scanned by the computerized system Aperio Technologies Inc.® and evaluated by the Aperio ImageScope® software (Leica Biosystems, Vista, CA, USA). The histological slides of each case was independently reviewed by two experienced oral pathologists (DCC and KSC) to evaluate the individual architectural and cytological criteria, according to the 2017 WHO (6) classification, and the presence or absence of epithelial dysplasia and its degree of severity, according to the 2017 WHO (6) criteria and the binary system (9). The sections were analyzed in their entirety, but the margins of the lesions were not considered to avoid artifacts. The evaluation was performed without information about the original histopathological evaluation and clinical data of the participants.

To assess the presence of epithelial dysplasia, the cytological and architectural criteria recommended by the WHO (6) were investigated. Regarding the degree of epithelial dysplasia according to the WHO (6), mild epithelial dysplasia was considered when changes were restricted to the lower third of the epithelium; moderate dysplasia when such changes extended to the middle third; and severe or intense dysplasia when the upper third was affected by the alterations (6).

For the evaluation of epithelial dysplasia according to the binary system criteria (7), the lesions were categorized as either “low risk” or “high risk”. In this case, the architectural and cytological alterations recommended by the WHO (6) were counted, and lesions that presented at least four architectural alterations and five cytological alterations were classified as high risk. The remaining lesions were considered low risk.

- Evaluation of intra and inter-observer agreement

After histopathological evaluation and classification of epithelial dysplasia, the results generated by the two oral pathologists were compared for inter-observer variability. To evaluate intra-observer agreement, 1/3 of OLP histological slides and 1/3 of OLL histological slides were selected and re-evaluated by the same oral pathologists six months after the first evaluation. The sample selection was made randomly using the website Dice throw (11), which generates real-time random numbers by measuring quantum vacuum fluctuations (12).

- Statistical Analysis

The Statistical Package for the Social Sciences® (SPSS), version 23.0, was used for the analyses and the significance level was set at 5% ($p \leq 0.05$). The intraclass correlation coefficient and the Kappa test were used for

the validation and reliability of the evaluation of the presence and severity of epithelial dysplasia. The Landis and Koch (13) evaluation were used to characterize the different ranges of Kappa, with a value less than zero indicating no agreement; between 0.0 and 0.20, indicating slight agreement; between 0.21 and 0.40, indicating fair agreement; 0.41 to 0.60 indicating moderate agreement; 0.61 to 0.80 indicating substantial agreement; and 0.81 to 1.00 indicating almost perfect agreement.

Results

- Inter-observer variability

Regarding individual evaluation, according to the first evaluation of the 65 biopsies, the histopathological characteristics most frequently observed by evaluator 1 were: increased number and size of nucleoli ($n=60/92.3\%$), loss of polarity of basal layer cells ($n=58/89.2\%$), nuclear hyperchromatism ($n=57/87.7\%$), and abnormal variation in cell shape ($n=55/84.6\%$) (Table 1). The histopathological characteristics most frequently observed by evaluator 2 were: loss of polarity of basal layer cells ($n=63/96.9\%$), increased number and size of nucleoli ($n=61/93.8\%$), nuclear hyperchromatism ($n=60/92.3\%$), and irregular epithelial stratification ($n=60/92.3\%$) (Table 1).

Regarding inter-observer disagreement of histopathological characteristics for the evaluation of epithelial dysplasia, according to the first evaluation of the 65 biopsies, the changes that most diverged in terms of their presence were: atypical mitotic figures ($n=28/43.1\%$), loss of cohesion between epithelial cells ($n=25/38.5\%$), and drop shaped rete ridges ($n=25/38.5\%$) (Table 2).

Statistically significant agreements were found in inter-observer variability regarding evaluations of the presence of: increased number of mitotic figures ($k=0.349$; $p=0.003$; fair agreement), superficial mitoses ($k=0.253$; $p=0.023$; fair agreement), premature individual keratinization ($k=0.327$; $p=0.003$; fair agreement), abnormal variation in nuclear shape ($k=0.272$; $p=0.028$; fair agreement), abnormal variation in cell size ($k=0.410$; $p=0.001$; moderate agreement), abnormal variation in cell shape ($k=0.460$; $p<0.001$; moderate agreement), and nuclear hyperchromatism ($k=0.405$; $p=0.001$; moderate agreement) (Table 2).

Regarding the presence of epithelial dysplasia according to WHO criteria, evaluator 1 classified 15 (23.1%) biopsies as without epithelial dysplasia and 50 (76.9%) with the presence of epithelial dysplasia, with the majority of biopsies with mild dysplasia ($n=44/67.7\%$). Evaluator 2 did not classify any biopsy as without epithelial dysplasia, and the majority of biopsies were classified as mild epithelial dysplasia ($n=49/75.4\%$) (Table 1).

Regarding classification according to the binary system, evaluator 1 classified 47 (72.3%) biopsies as low risk and 18 (27.7%) biopsies as high risk. Evaluator 2 clas-

Table 1: Distribution of classification regarding the presence and absence of individual histopathological criteria and the presence and degree of dysplasia according to each observer.

Histopathologic criteria	Observer 1				Observer 2			
	Yes		No		Yes		No	
Architectural changes								
Irregular epithelial stratification	47	72.3%	18	27.7%	60	92,3%	5	7.7%
Loss of polarity of the basal cells	58	89.2%	7	10.8%	63	96,9%	2	3.1%
Drop-shaped rete ridges	21	32.3%	44	67.7%	12	18,5%	53	81.5%
Increased number of mitotic figures	14	21.5%	51	78.5%	24	36,9%	41	63.1%
Abnormally superficial mitotic figures	5	7.7%	60	92.3%	2	3,1%	63	96.9%
Premature keratinization in single cells	32	49.2%	33	50.8%	46	70,8%	19	29.2%
Keratin pearls within rete ridges	0	0.0%	65	100.0%	0	0,0%	65	100.0%
Loss of epithelial cell cohesion	24	36.9%	41	63.1%	3	4,6%	62	95.4%
Cytological changes								
Abnormal variation in nuclear size	54	83.1%	11	16.9%	58	89,2%	7	10.8%
Abnormal variation in nuclear shape	51	78.5%	14	21.5%	51	78,5%	14	21.5%
Abnormal variation in cell size	54	83.1%	11	16.9%	56	86,2%	9	13.8%
Abnormal variation in cell shape	55	84.6%	10	15.4%	50	76,9%	15	23.1%
Increased N:C ratio	52	80.0%	13	20.0%	58	89,2%	7	10.8%
Atypical mitotic figures	26	40.0%	39	60.0%	36	55,4%	29	44.6%
Increased number and size of nucleoli	60	92.3%	5	7.7%	61	93,8%	4	6.2%
Hyperchromasia	57	87.7%	8	12.3%	60	92,3%	5	7.7%
WHO grade dysplasia	n		%		n		%	
Without dysplasia	15		23.1%		0		0.0%	
Mild dysplasia	44		67.7%		49		75.4%	
Moderate dysplasia	5		7.7%		14		21.5%	
Severe dysplasia	1		1.5%		2		3.1%	
Binary system								
Low-grade dysplasia	47		72,3%		42		64,6%	
High-grade dysplasia	18		27,7%		23		35,4%	

N:C=Nucleo-citoplasma; WHO=World Health Organization. Architectural and cytological changes according to the 2017 WHO criteria.

sified 42 (64.6%) biopsies as low risk and 23 (35.4%) biopsies as high risk (Table 1).

The relationship of inter-observer disagreement in the classification of the presence and degree of epithelial dysplasia according to WHO criteria and the binary system is described in Table 2. No statistically significant agreement was found regarding the classification of the degree of epithelial dysplasia according to WHO criteria. However, there was statistically significant agreement regarding the classification of the binary system ($k=0.257$; $p=0.035$; fair agreement) (Table 2).

- Intra-observer variability

Evaluator 1 presented four disagreements regarding the presence of epithelial dysplasia and seven regarding the grade of epithelial dysplasia according to WHO criteria (Table 3). Concerning the seven disagreements, it was observed that in three (42.8%) biopsies, the classifica-

tions changed from no epithelial dysplasia to mild epithelial dysplasia; in two (28.6%) biopsies, from mild to moderate dysplasia; in one (14.3%) biopsy, from moderate to mild dysplasia; and in another (14.3%) biopsy, from mild dysplasia to no epithelial dysplasia after discussion. It was found that only the individual criteria of increased nucleus-to-cytoplasm ratio and hyperchromatism did not show statistically significant agreement. Abnormal variation in nucleus shape ($k=0.861$; $p<0.001$) and increased number and size of nucleoli ($k=1.000$; $p<0.001$) showed almost perfect agreement (Table 3). Both the classification of the grade of epithelial dysplasia according to WHO criteria and the binary system showed statistically significant agreement ($k=0.546$; $p=0.004$, moderate agreement, and $k=0.861$; $p<0.001$, almost perfect agreement) (Table 3).

Evaluator 2 presented three disagreements regarding

Table 2: Inter-observer agreement according to individual histopathological criteria for the presence and degree of epithelial dysplasia.

	Disagreement		Agreement		Kappa	SE	p*
Architectural changes							
Irregular epithelial stratification	19	29.2%	46	70.8%	0.061	0.105	0.522
Loss of polarity of the basal cells	7	10.8%	58	89.2%	0.183	0.182	0.069
Drop-shaped rete ridges	25	38.5%	40	61.5%	0.010	0.117	0.933
Increased number of mitotic figures	18	27.7%	47	72.3%	0.349	0.117	0.003
Abnormally superficial mitotic figures	5	7.7%	60	92.3%	0.253	0.225	0.023
Premature keratinization in single cells	22	33.8%	43	66.2%	0.327	0.106	0.003
Keratin pearls within rete ridges	0	0.0%	65	100.0%	a	a	a
Loss of epithelial cell cohesion	25	38.5%	40	61.5%	-0.009	0.065	0.895
Cytological changes							
Abnormal variation in nuclear size	14	21.5%	51	78.5%	0.104	0.142	0.384
Abnormal variation in nuclear shape	16	24.6%	49	75.4%	0.272	0.139	0.028
Abnormal variation in cell size	10	15.4%	55	84.6%	0.410	0.153	0.001
Abnormal variation in cell shape	11	16.9%	54	83.1%	0.460	0.136	<0.001
Increased N:C ratio	16	24.6%	49	75.4%	0.070	0.130	0.548
Atypical mitotic figures	28	43.1%	37	56.9%	0.157	0.116	0.185
Increased number and size of nucleoli	7	10.8%	58	89.2%	0.165	0.193	0.180
Hyperchromasia	7	10.8%	58	88.2%	0.405	0.181	0.001
WHO grade dysplasia							
Presence of dysplasia	16	24.6%	49	75.4%	a	a	a
Grade dysplasia	34	52.3%	31	47.7%	-0.107	0.049	0.126
Binary system							
Classification	21	32.3%	44	67.7%	0.257	0.124	0.035

SE= Standard error; N:C=Nucleo-citoplasmia; WHO=World Health Organization. Architectural and cytological changes according to the 2017 WHO criteria. *p (p<0.05). aNo statistical analysis, as data are constant.

the grade of epithelial dysplasia according to WHO criteria (Table 4), all of which changed from moderate to mild dysplasia after discussion. No statistically significant agreement was found in the individual criteria of loss of polarity of basal cells layer, loss of cohesion between epithelial cells, abnormal variation in nucleus size, abnormal variation in cell shape, and increased nucleus-to-cytoplasm ratio. The evaluation of loss of polarity of basal cells layer (k=1.000; p<0.001) and hyperchromatism (k=1.000; p<0.001) showed almost perfect agreement. Only the evaluation of the WHO system classification showed statistically significant agreement (k=0.593; p=0.010, moderate agreement) (Table 4).

Discussion

This is the first study to evaluate the intra- and inter-observer agreement of individual architectural and cytological features of epithelial dysplasia in OLP and OLL according to the 2017 WHO criteria and binary system, and to discuss the challenges faced by pathologists in

diagnosing epithelial dysplasia and its severity in these lesions.

Applying the “thirds” epithelial classification proposed by the WHO is difficult due to the oral mucosa’s higher degree of heterogeneity, with a wide variation in size, thickness, and complexity of epithelial architecture, making it harder to define the level of “thirds,” particularly in very thin epithelia, in contrast to what occurs in lesions in the cervix, for example (15).

Thus, defining the degree of dysplasia based solely on the number of altered epithelial thirds oversimplifies the complexity of classification, as the presence of cytological atypia in only the basal third may be sufficient for a diagnosis of severe dysplasia, depending on individual characteristics such as bulbous epithelial ridges, and marked cytological atypia (16). Alternatively, lesions with mildly atypical features that extend into the middle third of the epithelium may warrant classification as mild dysplasia (17).

In our study, highest inter-observer agreements among

Table 3: Agreement of the intraobserver assessment by observer 1.

Observer 1	Disagreement		Agreement		Kappa	SE	p*
Architectural changes							
Irregular epithelial stratification	2	9.1%	20	90.9%	0.773	0.149	<0.001
Loss of polarity of the basal cells	2	9.1%	20	90.9%	0.463	0.305	0.010
Drop-shaped rete ridges	4	18.2%	18	81.8%	0.596	0.167	0.002
Increased number of mitotic figures	5	22.7%	17	77.3%	0.353	0.194	0.030
Abnormally superficial mitotic figures	2	9.1%	20	90.9%	0.463	0.305	0.010
Premature keratinization in single cells	5	22.7%	17	77.3%	0.567	0.153	0.003
Keratin pearls within rete ridges	0	0.0%	22	100.0%	a	a	a
Loss of epithelial cell cohesion	5	22.7%	17	77.3%	0.433	0.184	0.014
Cytological changes							
Abnormal variation in nuclear size	3	13.6%	19	86.4%	0.582	0.215	0.006
Abnormal variation in nuclear shape	1	4.5%	21	95.5%	0.861	0.135	<0.001
Abnormal variation in cell size	3	13.6%	19	86.4%	0.492	0.251	0.019
Abnormal variation in cell shape	2	9.1%	20	90.9%	0.614	0.249	0.004
Increased N:C ratio	5	22.7%	17	77.3%	0.396	0.222	0.062
Atypical mitotic figures	6	27.3%	16	72.7%	0.421	0.164	0.015
Increased number and size of nucleoli	0	0.0%	22	100.0%	1.000	<0.001	<0.001
Hyperchromasia	5	22.7%	17	77.3%	0.304	0.240	0.150
WHO grade dysplasia							
Presence of dysplasia	4	18.2%	18	81.8%	0.546	0.196	0.009
Grade dysplasia grade	7	31.8%	15	68.2%	0.462	0.170	0.002
Binary system							
Classification	1	4.5%	21	95.5%	0.861	0.135	<0.001

SE= Standard error; N:C=Nucleo-citoplasmia; WHO=World Health Organization. Architectural and cytological changes according to the 2017 WHO criteria. Re-evaluation of 1/3 of the OLP slides and 1/3 of the OLL slides. *p ($p \leq 0.05$). aNo statistical analysis, as data are constant.

architectural criteria were an increase in the number of mitotic figures, premature keratinization in single cells, and superficial mitoses. When evaluating cytological criteria, that abnormal variation in cell shape, abnormal variation in cell size, and hyperchromatism showed higher levels of agreement among observers. On the other hand, the most discordant criteria were the loss of cohesion between epithelial cells and drop shaped rete ridges. The loss of cohesion between epithelial cells was described as an easily recognizable feature (18). However, we observed a limitation in its differentiation from spongiosis caused by inflammation, which is common in oral epithelium, as well as liquefactive degeneration of basal layer cells, characteristics of OLP and OLL. Additionally, an important characteristic found in these lesions is the presence of saw teeth rete ridges (19), which may have made it difficult for evaluators to assess drop shaped projections.

Other studies (20-22) have reported the histopathological characteristics that showed higher and lower agreement in PMD. Kujan *et al.* (21) evaluated 68 cases of epithelial dysplasia by four observers using the WHO

(23) criteria of 2005, and demonstrated that increase of mitotic figures and drop shaped rete ridges had the highest levels of agreement among the observers when assessing architectural changes. Among cytological changes, there was higher agreement in the increase of nucleus size and variation in cell shape. Irregular stratification of the epithelium, loss of polarity, variation in nucleus size, atypical mitosis, and hyperchromatism had the greatest disagreements among observers (21).

Krishnan *et al.* (22), in the evaluation of 63 leukoplakia slides by three observers using the 2005 WHO (23) criteria, reported excellent agreement in the assessment of irregular epithelial stratification, abnormal variation in nucleus size and shape, and increased nucleus-to-cytoplasm ratio. However, they found greater disagreement in the evaluations of superficial mitosis, atypical mitosis, premature keratinization, and increased number of mitotic figures (22).

In Ranganathan *et al.*'s study (20), using the 2005 WHO (23) criteria, six evaluators analyzed 72 cases of PMD (leukoplakia, erythroplakia, proliferative verrucous leukoplakia, lichen planus, and submucosal fibrosis). There

Table 4: Agreement of the intraobserver assessment by observer 2.

Observer 2	Disagreement		Agreement		Kappa	SE	p*
Architectural changes							
Irregular epithelial stratification	2	9.1%	20	90.9%	a	a	a
Loss of polarity of the basal cells	0	0.0%	22	100.0%	1.000	<0.001	<0.001
Drop-shaped rete ridges	5	22.7%	17	77.3%	0.179	0.230	0.334
Increased number of mitotic figures	5	22.7%	17	77.3%	0.545	0.172	0.008
Abnormally superficial mitotic figures	4	18.2%	18	81.8%	*	*	*
Premature keratinization in single cells	3	13.6%	19	86.4%	0.637	0.190	0.003
Keratin pearls within rete ridges	0	0.0%	22	100.0%	a	a	a
Loss of epithelial cell cohesion	6	27.3%	16	72.7%	0.132	0.203	0.449
Cytological changes							
Abnormal variation in nuclear size	8	36.4%	14	63.6%	0.129	0.121	0.219
Abnormal variation in nuclear shape	7	31.8%	15	68.2%	0.364	0.153	0.027
Abnormal variation in cell size	9	40.9%	13	59.1%	0.182	0.121	0.138
Abnormal variation in cell shape	8	36.4%	14	63.6%	0.313	0.141	0.044
Increased N:C ratio	10	45.5%	12	54.5%	0.091	0.088	0.306
Atypical mitotic figures	8	36.4%	14	63.6%	0.279	0.201	0.184
Increased number and size of nucleoli	1	4.5%	21	95.5%	a	a	a
Hyperchromasia	0	0.0%	22	100.0%	1.000	<0.001	<0.001
WHO grade dysplasia							
Presence of dysplasia	0	0.0%	22	100.0%	a	a	a
Grade dysplasia grade	3	13.6%	19	86.4%	0.593	0.200	0.002
Binary system							
Classification	7	31.8%	15	68.2%	0.280	0.187	0.125

SE= Standard error; N:C=Nucleo-citoplasmia; WHO=World Health Organization. Architectural and cytological changes according to the 2017 WHO criteria. Re-evaluation of 1/3 of the OLP slides and 1/3 of the OLL slides. *p (p<0.05). aNo statistical analysis, as data are constant.

was a higher level of agreement regarding the increase of mitotic figures, loss of cohesion of epithelial cells, and increased nucleus-to-cytoplasm ratio (20). The greatest disagreements were observed in irregular stratification of the epithelium, loss of polarity of basal layer cells, abnormal variation in nucleus shape, and abnormal variation in cell size.

With these controversial results (20-22), it is difficult to synthesize any guidance on which microscopic features are reported among oral pathologists²⁰. However, despite the increase in the number of mitotic figures being a subjective analysis and no study indicating how many figures need to be present, our study and the studies by Kujan *et al.* (21) and Ranganathan *et al.* (20) showed some agreement among evaluators in this analysis. Furthermore, although both evaluators in our study reported difficulties in differentiating Civatte bodies from dyskeratosis in some cases when located in the lower third of the epithelium during individual evaluations, we found considerable inter-observer agreement with respect to this criterion.

The only previous study that evaluated inter-observer agreement of epithelial dysplasia in oral lichenoid diseases, according to the 2017 WHO criteria (6). These authors evaluated 84 cases of proliferative verrucous leukoplakia (PVL) and 28 cases of OLL and found a wide variability in the interpretation of epithelial dysplasia, with low inter-observer reliability among the four examiners (24). The Kappa was classified as very low in the two repeated evaluations among the examiners. Regarding the grading of epithelial dysplasia according to the WHO criteria⁶, in our study, we also did not find a statistically significant agreement among the evaluators, and the Kappa value was considered very low.

A recent multicenter study (20) evaluated the intra- and inter-observer agreement of six observers who examined PMD (leukoplakia, erythroplakia, proliferative verrucous leukoplakia, lichen planus, and submucosal fibrosis) using the 2005 WHO (23) classification. The study found substantial agreement between two evaluators. However, the other observers had poor to fair (20). Speight *et al.* (14) highlight that the highest levels of

disagreement are found when categorizing low-grade lesions due to the fact that these lesions present more subtle changes, which can be observed in reactive lesions as a result of an inflammatory infiltrate (14). Therefore, we believe that the low values of inter-observer agreement in our study may also be related to the fact that the majority of our sample of OLP and OLL presented subtle cytological and architectural changes, as well as the intense inflammatory infiltrate, which is characteristic of these lesions.

Regarding intra-observer agreement, the study by Zohdy *et al.* (24) conducted in oral lichenoid diseases reported a slight improvement in intra-observer reliability during the evaluation of the epithelial dysplasia grading by the WHO system (6), which varied among examiners.

In our study, we found better intra-observer agreement between the two evaluators. The evaluator 1 showed higher agreement in evaluating cytological criteria, while evaluator 2 showed higher agreement in evaluating architectural criteria. Regarding architectural criteria, we found that the highest agreements were for loss of polarity of basal cells layer, increased number of mitotic figures, and individual premature keratinization of the cell. On the other hand, abnormal variation in nucleus shape and abnormal variation in cell shape were the cytological criteria with the highest agreements.

In an attempt to minimize inter- and intra-observer disagreements, many authors (15,20) recommend the use of the binary system (7), as they believe this approach is standardized and could overcome subjectivity in reporting epithelial dysplasia (20). However, in our study, despite finding better inter-observer agreement with the binary system compared to the WHO system (6), evaluator 2 showed better reliability in assessing the classification of epithelial dysplasia according to the 2017 WHO criteria.

These results reinforce that grading of epithelial dysplasia is therefore subjective, lacks reproducibility, and may be influenced by evaluators' experience, fatigue, and emotional factors (25). It has been shown that more anxious individuals tend to behave more negatively when making professional decisions (26), suggesting potential consequences in microscopic interpretation.

Despite their limitations, the evaluation of epithelial dysplasia offers the pathologist the best opportunity to convey the overall risk of malignancy to the clinician (14). Additionally, high concordance of a method does not indicate that it is the most correct for use, only that it is reproducible. Thus, we emphasize the importance of defining guidelines to improve the interpretation of criteria involving the evaluation of epithelial dysplasia and reduce inter-observer variability. We also highlight the need for studies that seek to correlate the classification of epithelial dysplasia in both systems with the risk of malignant transformation.

This study has some methodological limitations due to its retrospective design. In addition, despite the pathologists having training in Pathology from the same institution and working together for years, there was no prior standardization before the evaluations. Furthermore, it has been questioned in the literature whether the Kappa statistic, used to measure reproducibility in a dichotomous decision model in pathology, is the most appropriate (27), as it is believed that these verbal descriptors of reproducibility can be arbitrary (15). However, it is still the most widely used in the majority studies to date (20-22,24).

The evaluation of both the presence of individual criteria and the classification of the presence and degree of epithelial dysplasia is subjective and dependent on the individual experience of the evaluators. The most discordant criteria among evaluators in the assessment of OLP and OLL are the evaluation of loss of cohesion of the epithelial cells and drop shaped rete ridges. The subtle and focal changes and the inflammatory infiltrate, characteristic of these lesions, can increase the degree of disagreement between evaluators. The binary system shows better inter-observer agreement, while the 2017 WHO system shows better intra-observer agreement.

References

1. Carrozzo M, Porter S, Mercadante V, Fedele S. Oral lichen planus: A disease or a spectrum of tissue reactions? Types, causes, diagnostic algorithms, prognosis, management strategies. *Periodontol* 2000. 2019;80:105-25.
2. Aguirre-Urizar JM, Alberdi-Navarro J, Lafuente-Ibáñez de Mendoza I, Marichalar-Mendia X, Martínez-Revilla B, Parra-Pérez C, et al. Clinicopathological and prognostic characterization of oral lichenoid disease and its main subtypes: A series of 384 cases. *Med Oral Patol Oral Cir Bucal*. 2020;25:e554-62.
3. Cheng Y-SL, Gould A, Kurago Z, Fantasia J, Muller S. Diagnosis of oral lichen planus: a position paper of the American Academy of Oral and Maxillofacial Pathology. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2016;122:332-54.
4. Warnakulasuriya S, Kujan O, Aguirre-Urizar JM, Bagan JV, González-Moles MÁ, Kerr AR, et al. Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. *Oral Dis*. 2020;odi.13704.
5. Barnes L, JW E, P R, D S. Pathology and Genetics of Head and Neck Tumours [Internet]. 2005 [cited 2019 Aug 30]. Available from: <https://publications.iarc.fr/Book-And-Report-Series/Who-Iarc-Classification-Of-Tumours/Pathology-And-Genetics-Of-Head-And-Neck-Tumours-2005>
6. EL-Naggar A, Chan JK, Grandis J, Takata T, Slootweg P. WHO Classification of Head and Neck Tumours. Fourth edition - WHO - OMS - [Internet]. 2017 [cited 2019 Aug 30]. Available from: <https://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=70&codcch=4009>
7. Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. *Oral Oncol*. 2006;42:987-93.
8. Loss R, Sandrin R, França BHS, de Azevedo-Alanis LR, Grégio AMT, Machado MÂN, et al. Cytological analysis of the epithelial cells in patients with oral candidiasis. *Mycoses*. 2011;54:e130-135.
9. Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of a new binary system of grading oral epithelial dysplasia

for prediction of malignant transformation. *Oral Oncol.* 2006;42:987-93.

10. Manchanda A, Shetty D-C. Reproducibility of grading systems in oral epithelial dysplasia. *Med Oral Patol Oral Cir Bucal.* 2012;17:e935-42.

11. Dice throw - ANU QRNG [Internet]. [cited 2022 May 3]. Available from: <https://qrng.anu.edu.au/dice-throw/>

12. Symul T, Assad SM, Lam PK. Real time demonstration of high bitrate quantum random number generation with coherent laser light. *Appl Phys Lett.* American Institute of Physics; 2011;98:231103.

13. Landis JR, Koch GG. A One-Way Components of Variance Model for Categorical Data. *Biometrics.* [Wiley, International Biometric Society]; 1977;33:671-9.

14. 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;120:474-482.e2.

15. Odell E, Kujan O, Warnakulasuriya S, Sloan P. Oral epithelial dysplasia: Recognition, grading and clinical significance. *Oral Diseases.* 2021;27:1947-76.

16. Jo VY, Demicco EG. Update from the 5th Edition of the World Health Organization Classification of Head and Neck Tumors: Soft Tissue Tumors. *Head Neck Pathol.* 2022;16:87-100.

17. Warnakulasuriya S, Reibel J, Bouquot J, Dabelsteen E. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *J Oral Pathol Med.* 2008;37:127-33.

18. Woo S-B. Oral Epithelial Dysplasia and Premalignancy. *Head Neck Pathol.* 2019;13:423-39.

19. Fernandez-Gonzalez F, Vazquez-Alvarez R, Reboiras-Lopez D, Gandara-Vila P, Garcia-Garcia A, Gandara-Rey Jm. Histopathological findings in oral lichen planus and their correlation with the clinical manifestations. *Med Oral.* 2011;e641-6.

20. Ranganathan K, Kavitha L, Sharada P, Bavle RM, Rao RS, Patanshetty SM, et al. Intra-observer and inter-observer variability in two grading systems for oral epithelial dysplasia: A multi-centre study in India. *Journal of Oral Pathology & Medicine.* 2020;49:948-55.

21. 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.

22. 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.

23. Barnes, L. et al. World Health Organization Classification of tumors. *Pathology & Genetics.* Head and Neck Tumors. 2005. 2005.

24. Zohdy M, Cazzaniga S, Nievergelt H, Blum R, Suter VGA, Feldmeyer L, et al. Inter-Observer and Intra-Observer Variations in the Assessment of Epithelial Dysplasia in Oral Lichenoid Diseases. *Dermatopathology (Basel).* 2021;8:84-8.

25. Diniz PB, Sena-Filho M, Graner KM, Mariz BALA, dos Reis LA, de Almeida OP, et al. Comparison of the whole slide imaging and conventional light microscopy in the grading of oral epithelial dysplasia: a multi-institutional study. *Med Oral Patol Oral Cir Bucal.* 2021;26:e8-13.

26. Hartley CA, Phelps EA. Anxiety and decision-making. *Biol Psychiatry.* 2012;72:113-8.

27. Cross SS. Kappa statistics as indicators of quality assurance in histopathology and cytopathology. *J Clin Pathol.* 1996;49:597-9.

Ethics approval and consent to participate

This work was approved by the Research Ethics Committee of Hospital Universitário Antônio Pedro of the Universidade Federal Fluminense (#5.563.543) and the procedures performed were in accordance with the Helsinki Declaration of 1975, as revised in 1983. The informed consent for participation was obtained from all participants.

Consent for publication

The participant provided an informed consent declaration to permit

the use of images and medical information and the manuscript is in accordance with the Institutional Ethics Committee.

Funding

The authors are not receiving any research funding for this study.

Authors' contributions

LCM conceptualized and designed the study, interpreted results, performed the statistical analysis, and wrote the manuscript; PPMS wrote the manuscript; LSG performed the statistical analysis; KSC offered advice on methods, evaluated histological glass slides, and critically edited the manuscript; ASJ conceptualized and designed the study, and critically edited the manuscript; DCC conceptualized and designed the study, evaluated histological glass slides, and critically edited the manuscript.

Conflict of interest

No conflict of interest to disclose.