ORIGINAL ARTICLE

Grading of oral epithelial dysplasia: Points to ponder

KM Geetha, M Leeky¹, TV Narayan², S Sadhana¹, J Saleha¹

Department of Oral Pathology and Microbiology, Mamata Dental College, Giriprasad Nagar, Khammam, Andhra Pradesh, ¹Department of Oral and Maxillofacial Pathology, The Oxford Dental College, ²Private Dental Practitioner, Impressions Dental Care, No. 102, Elegant Desire, Bengaluru, Karnataka, India

Address for correspondence:

Dr. M Leeky, Department of Oral and Maxillofacial Pathology, The Oxford Dental College, Bengaluru - 560068, Karnataka, India. E-mail: leekymohanty@yahoo.com

Received: 03-02-2014 Accepted: 06-07-2015

ABSTRACT

Background: Over the years many grading systems have been put forward in an attempt to obtain objectivity in grading oral epithelial dysplasia (OED). However, despite these efforts variability remains unresolved. Our study aimed to evaluate the intra- and inter-observer variability in grading OED, using World Health Organization (WHO), Smith and Pindborg and Ljubljana grading systems and discuss the possible reasons for this variability if any. Materials and Methods: Three oral pathologists graded 50 slides of OED independently twice at a time interval of 3 months. Variability was evaluated by multivariate kappa analysis. Results: Intra-observer reproducibility ranged from moderate to good in WHO system, fair to moderate in Smith and Pindborg system and moderate to poor in Ljubljana grading system. Inter-observer agreement was found to be fair in WHO, poor in Smith and Pindborg system and poor to fair in Ljubljana grading systems. Intra-observer reproducibility of the dysplastic features in WHO system was good for all except the loss of polarity and basilar hyperplasia for first observer and enlarged nucleoli for the third observer. Inter-observer agreement was good for increased number of mitosis and nuclear hyperchromatism. Intra-observer reproducibility and inter-observer agreement were found to be best in the WHO grading system though variability within this system still existed. Conclusion: There is a need for an International body of pathologists to come to a consensus on a more definable grading system to resolve the issue of variability in grading dysplasia. Key words: Grading systems, intra- and inter-observer variability, oral epithelial dysplasia

INTRODUCTION

Oral squamous cell carcinomas are often preceded by readily detectable changes within the oral mucosa. The World Health Organization (WHO) working group of oral pathologists grouped such clinical presentations that carry an increased risk of cancer under the term "potentially malignant disorders (PMD)."^[1] The most common PMD is leukoplakia,^[2] the histopathology of which may show epithelial dysplasia.^[3]

Epithelial dysplasia has been a subject of much debate owing to increased intra- and inter-observer variability in its grading.^[4-8] In this study, we examined for any intra- and inter-observer variability in already established grading systems.^[9-11]

Access this article online			
Quick Response Code:	Website: www.jomfp.in		
	DOI: 10.4103/0973-029X.164533		

MATERIALS AND METHODS

Fifty histopathologically diagnosed cases of oral epithelial dysplasia (OED) were retrieved from the archives of the department. It included 17 cases originally signed out as mild dysplasia, 9 signed out as moderate dysplasia, 3 as severe dysplasia and 21 as hyperkeratosis with no features of dysplasia [Figures 1-4]. Three oral pathologists participated in the study. All the slides were blinded and were graded independently using - WHO, Smith and Pindborg and Ljubljana grading systems.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: 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.



Figure 1: Photomicrograph of epithelium showing mild dysplasia (H&E stain, ×200)



Figure 3: Photomicrograph of epithelium showing severe dysplasia (H&E stain, ×200)

World Health Organization grading system

In its report in 1978, WHO listed out 12 histologic characteristics that characterized the epithelial dysplasia:

- Loss of polarity of basal cells
- Presence of more than one layer of cells having basaloid appearance
- Increased nuclear-cytoplasmic ratio
- Drop shaped rete-pegs
- Irregular epithelial stratification
- Increased number of mitotic figures, mitotic figures that are abnormal in form may be present
- Presence of mitotic figures in the superficial half of epithelium
- Cellular polymorphism
- Nuclear hyperchromatism
- Enlarged nucleoli
- Reduction of cellular cohesion
- Keratinization of single cells or cell groups in the prickle cell layer.^[12]



Figure 2: Photomicrograph of epithelium showing moderate dysplasia (H&E stain, ×200)



Figure 4: Photomicrograph of epithelium showing hyperkeratosis without dysplasia (H&E stain, ×200)

It graded epithelial dysplasia as:

- Mild dysplasia
- Moderate dysplasia
- Severe dysplasia.

Mild dysplasia

Slight nuclear abnormalities, most marked in the basal third of the epithelial thickness and minimal in the upper layers, where the cells show maturation and stratification. A few, but no abnormal mitoses may be present, usually accompanied by keratosis and chronic inflammation.

Moderate dysplasia

More marked nuclear abnormalities and nucleoli tend to be present with changes most marked in the basal two-third of the epithelium, nuclear abnormalities may persist up to the surface, but cell maturation and stratification are evident in the upper layers. Mitoses are present in the parabasal and intermediate layers, but none is abnormal.

Severe dysplasia

Marked nuclear abnormalities and loss of maturation involve more than two-third of the epithelium with some stratification of the most superficial layers. Mitoses some of which are abnormal may be present in the upper layers.^[9]

Smith and Pindborg grading system

They used 13 histologic features, which were standardized by a set of photographs. For each feature, they suggested a score and features were graded as "none," "slight" or "marked" after comparing with photographic standards [Table 1]. The scores are eventually added to achieve the epithelial atypia index (EAI). The maximum possible index value is 75.

Depending upon EAI:

- 0–10: No dysplasia
- 11–25: Mild
- 26–45: Moderate
- 46–75: Severe.^[13]

Ljubljana grading system

Lesions were categorized according to the Ljubljana classification into four different groups:

Simple hyperplasia

A benign hyperplastic process with retention of the normal pattern of epithelium was thickened because of increased prickle cell layer. The cellular components of basal and parabasal region of epithelium (1–3 layers) remain unchanged. There is no cellular atypia.

Table 1	1:	Smith	and	Pindborg	grading	system
---------	----	-------	-----	----------	---------	--------

Type of change	Severity of dysplasia			
Drop shaped rete pegs	None (0)	Slight (2)	Marked (4)	
Irregular epithelial	None (0)	Slight (2)	Marked (5)	
stratification				
Keratinization of cells below	None (0)	Slight (1)	Marked (3)	
keratinized layer				
Basal cell hyperplasia	None (0)	Slight (1)	Marked (4)	
Loss of intercellular	None (0)	Slight (1)	Marked (5)	
adherence				
Loss of polarity	None (0)	Slight (2)	Marked (6)	
Hyperchromatic nuclei	None (0)	Slight (2)	Marked (5)	
Increased	None (0)	Slight	Marked	
nucleo-cytoplasmic ratio in		increase (2)	increase (6)	
basal and prickle cell layers				
Anisocytosis and	None (0)	Slight (2)	Marked (6)	
anisonucleosis				
Pleomorphic cells and nuclei	None (0)	Slight (2)	Marked (6)	
Mitotic activity	Normal (0)	Slight	Marked	
		increase (1)	increase (5)	
Level of mitotic activity	Normal (0)	Slight (3)	Marked (10)	
Presence of bizarre mitoses	None (0)	Slight (6)	Marked (10)	

Abnormal hyperplasia

A benign augmentation of basal and parabasal layers. Basal and parabasal cells are augmented to a degree, which constitutes up to one-half of epithelial thickness. It is important that stratification is fully retained. Occasionally, more than this proportion of the epithelium may be involved by the hyperplastic cells without significant atypical nuclear changes. Nuclei in cells of the augmented basal and parabasal layers may be moderately enlarged but still maintain a uniform distribution of nuclear chromatin. Occasional typical mitoses may be found in the basal layer. Small numbers of epithelial cells <5% are dyskeratotic.

Atypical hyperplasia

This demonstrates the recognizable alteration of epithelial cells towards malignancy, but not to such a degree as seen in carcinomatous cells. Stratification is still preserved in general epithelial structure. Nuclei are enlarged and nuclear contour may be irregular with marked variations in staining intensity. Nuclear/cytoplasmic ratio is increased. Mitotic figures are increased but not numerous and are found within two-thirds of epithelium above the basement membrane. They are rarely, if ever, abnormal. Dyskeratotic cells are frequent. Civatte bodies may be present.

Carcinoma *in situ* shows the features of carcinoma without invasion. Stratification of the epithelium as a whole is lost. Marked cellular alterations of the type found in atypical hyperplasia are present to a considerably greater degree. Many mitotic figures are present throughout the epithelium, including its upper one-third and abnormal mitoses are frequently found.^[11]

In addition, intra- and inter-observer reproducibility of individual features of dysplasia using WHO grading system was also assessed. Slides were assessed twice by the oral pathologists with a time interval of 3 months between the two assessments. No clinical data or information about the cases was provided and none of the observers were involved in the original sign-out diagnosis. All the data were entered in the score sheets and statistically analyzed.

The following statistical methods were employed to evaluate the observer variability:

- Inter-examiner variation was assessed for finding the agreement by Fleiss kappa statistics (K)
- Intra-class correlation co-efficient (ICC) was computed for finding the correlation among the three observers
- Cronbach-α was computed to know the consistency or reliability of the three observers
- Simple matching coefficient (SMC) was used to assess the presence or absence of distributions on a set of sites which counts the number of sites having the same status (presence or absence) in both distributions.

The statistical softwares namely SPSS 15.0 [IBM], Stata 8.0 [Stata Corp], MedCalc 9.0.1 [Siemens] and Systat 11.0 [Systat software] were used for the analysis of the data and Microsoft Word and Excel have been used to generate tables and graphs.

RESULTS

Intra-observer agreement

In WHO grading system, moderate to good agreement was found by all the observers. In Smith and Pindborg photomicrographic grading system fair to moderate agreement was found and in Ljubljana grading system poor to moderate agreement was observed [Table 2].

Inter-observer agreement

WHO grading system showed a fair agreement, with moderate to very large correlation and good consistency was noticed at both times of grading. Smith and Pindborg grading system showed a poor agreement with moderate correlation and average consistency both the times. Ljubljana grading system showed a poor agreement with moderate correlation and average to good consistency for the first and second times [Table 3].

Intra-observer reproducibility of the individual histological parameters in WHO grading system

The individual histological parameters showed good

Table 2: Overall intra-observer agreement in threegrading systems (Kappa values)

Observer	WHO	Smith and Pindborg	Ljubljana	
1	0.664	0.243	0.451	
2 0.552 3 0.540		0.460	0.558	
		0.474	0.154	

WHO: World Health Organization

Table 3: Inter-observer agreement in three grading systems

Diagnosis	Grading system				
	WHO	Smith and Pindborg	Ljubljana		
First time					
κ*	0.263	0.197	0.192		
ICC^{\dagger}	0.781	0.435	0.384		
C-α [‡]	0.781	0.706	0.653		
Second time					
κ*	0.220	0.168	0.227		
ICC^{\dagger}	0.489	0.406	0.423		
C-α [‡]	0.742	0.702	0.734		

* κ : Kappa, [‡]C- α : Cronbach- α , [†]ICC: Intra-class correlation coefficient. Kappa values (agreement) - <0.20=Poor agreement, 0.21-0.40=Fair, 0.41-0.60=Moderate, 0.61-0.80=Good, 0.81-1.00=Very good. C- α (consistency) - <0.50=Poor, 0.50-0.70=Average, >0.70=Good. ICC (correlation) - <0.30=Small, 0.31-0.50=Moderate, 0.51-0.70=Large, 0.71-0.90=Very large, >0.91=Good. WHO: World Health Organization reproducibility for all the parameters except for loss of polarity of basal cells and basilar hyperplasia for the first observer, whereas the above two parameters showed a perfect reproducibility for the second observer and a good reproducibility in the remaining features. The third observer obtained good reproducibility for all the parameters except for enlarged nucleoli [Table 4].

Inter-observer agreement in World Health Organization grading system for individual histological parameters

A consistent agreement was seen both times: For increased number of mitotic figures (moderate agreement); nuclear hyperchromatism and irregular stratification (fair agreement) and poor agreement for loss of polarity of basal cells, basilar hyperplasia, increased nuclear cytoplasmic ratio and individual cell keratinization on both the times. All three observers showed good agreement for abnormal mitosis for the first time, however, showed a poor agreement for the second time. There was a fair agreement among the observers for enlarged nucleoli and moderate agreement for mitotic figures in the superficial half of the epithelium; and cellular and nuclear pleomorphism for the first time and poor agreement for the second time. Poor agreement for the first time and fair agreement for the second time for drop shaped rete ridges. Fair agreement for the first time and moderate agreement for the second time for loss of intercellular adherence [Table 5].

When individual observer diagnosis of WHO was compared with sign-out diagnosis, a fair agreement was obtained for all observers [Table 6].

DISCUSSION

Leukoplakia is the most common PMD of the oral mucosa, the malignant transformation rates of which have varied from 0.13% to 17.5%.^[14] Extensive research done in this field currently attributes the malignant transformation to be mainly dependent on the type of leukoplakia,^[15] site of leukoplakia and the severity of dysplasia.^[16]

The severity of dysplasia is arrived at based on the grading of dysplasia in leukoplakia. However, it is an established fact that dysplasia grading suffers from intra- and inter-observer variability.^[4-8] More and more grading systems have evolved over the years in an attempt to bring in a quantifiable, consistently reproducible, objective measure in grading dysplasia to reduce or eliminate intra and inter observer variability, but in vain. The analysis we wanted to make was whether the intra- and inter-observer variability is indeed dependent on the subjectivity or objectivity of the parameters or does it depend on any other factors.

In our study, we used three established grading systems namely the 1978 WHO grading system, Smith and Pindborg

Histological parameter	Observer 1:	Observer 1:	Observer 2:	Observer 2:	Observer 3:	Observer 3:
	Kappa	SMC	Kappa	SMC	Kappa	SMC
Loss of polarity of basal cells	0	0.44	-	1.0	0	0.50
Basilar hyperplasia	0.12	0.32	-	1.0	-	0.96
Drop shaped rete ridges	0.02	0.54	0.04	0.52	0.198	0.64
Increased N/C ratio	0.05	0.66	0.07	0.88	0.264	0.62
Nuclear hyperchromatism	0.15	0.58	0.07	0.54	0	0.52
Enlarged nucleoli	0.04	0.66	0.14	0.66	0	0.44
Increased number of mitotic figures	0	0.74	0	0.64	0	0.56
Abnormal mitosis	0	0.84	0	0.94	0.146	0.86
Mitotic figures in the superficial half of the epithelium	0	0.92	0	0.94	-	0.96
Cellular and nuclear pleomorphism	0.04	0.50	0	0.64	0.152	0.58
Irregular epithelial stratification	0.03	0.54	0.02	0.54	0.217	0.74
Loss of intercellular adherence	0	0.60	0.02	0.78	0	0.62
Individual cell keratinization	0	0.76	0	0.92	0.035	0.76

Table 4: Intra-observer correlation of the individual histopathological parameters in WHO grading system of epithelial dysplasia by Kappa and SMC

SMC: Simple matching coefficient, WHO: World Health Organization, N/C: Nuclear-cytoplasmic

Table 5: Inter-observer correlation of the individualhistological features in the WHO grading system

Histological parameter	First time Kappa values	Second time Kappa values
Loss of polarity of basal cells	0.0	0.005
Basilar hyperplasia	0.094	0
Drop shaped rete ridges	0.173	0.279
Increased N/C ratio	0.148	0.146
Nuclear hyperchromatism	0.224	0.321
Enlarged nucleoli	0.248	0.123
Increased number of mitotic figures	0.489	0.576
Abnormal mitosis	0.638	0
Mitotic figures in the superficial half of the epithelium	0.479	0
Cellular and nuclear pleomorphism	0.406	0
Irregular epithelial stratification	0.247	0.212
Loss of intercellular adherence	0.341	0.417
Individual cell keratinization	0.004	0.005

WHO: World Health Organization, N/C ratio: Nuclear-cytoplasmic ratio

Table 6: Sign-out diagnosis versus observers' diagnosis (WHO) (Kappa values)

Sign-out diagnosis			Obse	erver		
	First time			Second time		
	1	2	3	1	2	3
Kappa value	0.371	0.357	0.262	0.385	0.247	0.386

WHO: World Health Organization

photomicrographic standards system and Ljubljana classification of epithelial hyperplastic laryngeal lesions and looked for intra and inter-observer variability. As the observers were well experienced in reporting on dysplasias using the 1978 WHO grading system, it was used instead of the 2005 grading system. We found a better intra-observer and inter-observer agreement in the WHO system compared

to the other two. This is an interesting finding suggesting that objective grading systems need not always lead to a better agreement. We found the variability to be more in the two objective grading systems (SP and Lj) and least in WHO grading system, which is a system that relies on subjective criteria for grading. Thus, there may be multiple factors why a particular grading system has less or more observer variability.

The reason we found best agreeability in the WHO system could be because this was the one that the three oral pathologists were accustomed to use in their routine practice over the years while the other two were new systems. The observers were also from the same institution. The understanding and the way they interpreted each individual dysplastic feature may be similar and this may have had a bearing on the less variability seen in this system. There were a few drawbacks in the other two systems, which may also have contributed to our findings.

Smith and Pindborg photomicrographic standards system^[10] uses a photomicrograph, which was difficult to obtain. This grading system was found to be tedious and time consuming. Despite claiming to be an objective method, the numerical scores allotted to individual features were given subjectively by the authors.

The Lj system was applied for OED^[11] because of the anatomical similarity between laryngeal and oral mucosa. Features like atypical hyperplasia indicate a risky epithelium without stating the risk clearly. Also terms like "frequent," "increased" and "numerous," lack precision when used in the context of "mitosis increased but not numerous" and "dyskeratotic cells are frequent" leading to possible differences in interpretation of such terminologies among observers leading to variability in this grading system.

As we found the WHO grading system to be the best with good agreement, we decided to check for intra and inter-observer

variability in analyzing the individual dysplastic features of this system. Surprisingly, here the inter-observer variability was more. Except for features of increased mitosis and nuclear hyperchromatism, there was no agreement seen among observers. Intra-observer variability was less barring a few features such as loss of polarity, basilar hyperplasia and enlarged nucleoli. Hence, our assumption that the better agreement was seen in WHO system among the observers was because they worked in the same institution and interpreted each individual feature, in the same way, turned out to be wrong. This suggests there may be an unknown factor, or it may be just the common usage of the system that is responsible for the good agreement.

With so much variability seen among observers for individual parameters of the WHO system, it was difficult to understand how the WHO system gave the best agreement overall while grading. This led us to some questions. Does this mean the three observers formed an opinion on the grade of dysplasia without looking at the 13 dysplastic features? and did they use any crude way to grade dysplasia? None of the three observers were involved in the original sign-out diagnosis. This ruled out any existing bias.

Subsequent analysis revealed that extent of distribution of features like the nuclear hyperchromatism, altered nuclear cytoplasmic ratio and loss of stratification within the arbitrarily divided three-thirds of the epithelium led the observers to form an opinion on the grade of dysplasia while screening through the slide itself. The other dysplastic features were then observed under higher magnification before arriving at a final diagnosis. Hence, in some cases, the predetermined grade may have stayed as the final diagnosis and in others it may have changed. This may explain why there was a good agreement in WHO system among observers even though there was poor agreeability while assessing the individual dysplastic features of the same system by the same observers.

Thus, from our study we conclude that subjective parameters were not solely the source of variation in all grading systems.

Based on our study, we would like to make a few recommendations:

- An international body of oral pathologists after debate and consensus need to suggest a particular grading system that should be followed universally
- Individual features of that grading system should then be clearly stated both in written and using quality photomicrographs
- There is a need for deliberation on whether we need so many parameters in grading dysplasia and if so do all features carry the same degree of importance or are there some features that need to be given more weightage
- There should be proper training given to all oral pathologists regarding this. It should be a grading system that works well for oral pathologists with even the least

number of years of experience

- If variability still persists, a consensus among the observers in a particular setup may be needed to arrive at a final diagnosis
- In the grading of dysplasia, it would be advisable to come up with new grading systems only after considerable research has been done using that system, preferably as multicentric studies. Else, we will have far too many grading systems, which may dilute the usefulness of research in this field
- Finally, good record maintenance and follow-up of patients is needed to validate usefulness of any grading system in the treatment of leukoplakia or in predicting possible malignant transformation.

Grading of dysplasia, on its own, may not predict malignant transformation of PMD like leukoplakia but in combination with the site and type of clinical presentation it does play a significant role. We, thus, cannot afford to have such extreme variability in grading epithelial dysplasia. Oral pathologists need to come together to overcome the hurdles presently seen in the utility of various grading systems so as to be able to benefit a large number of patients.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

- 1. Warnakulasuriya S, Johnson NW, van der Waal I. Nomenclature and classification of potentially malignant disorders of the oral mucosa. J Oral Pathol Med 2007;36:575-80.
- van der Waal I, Schepman KP, van der Meij EH, Smeele LE. Oral leukoplakia: A clinicopathological review. Oral Oncol 1997;33:291-301.
- 3. Sciubba JJ. Oral leukoplakia. Crit Rev Oral Biol Med 1995;6:147-60.
- 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.
- 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.
- Brothwell DJ, Lewis DW, Bradley G, Leong I, Jordan RC, Mock D, et al. Observer agreement in the grading of oral epithelial dysplasia. Community Dent Oral Epidemiol 2003;31:300-5.
- Tilakaratne WM, Sherriff M, Morgan PR, Odell EW. Grading oral epithelial dysplasia: Analysis of individual features. J Oral Pathol Med 2011;40:533-40.
- 8. Manchanda A, Shetty DC. Reproducibility of grading systems in oral epithelial dysplasia. Med Oral Patol Oral Cir Bucal

2012;17:e935-42.

- Warnakulasuriya S. Histological grading of oral epithelial dysplasia: Revisited. J Pathol 2001;194:294-7.
- Katz HC, Shear M, Altini M. A critical evaluation of epithelial dysplasia in oral mucosal lesions using the Smith-Pindborg method of standardization. J Oral Pathol 1985;14:476-82.
- Zerdoner D. The Ljubljana classification Its application to grading oral epithelial hyperplasia. J Craniomaxillofac Surg 2003;31:75-9.
- Pindborg JJ, editor. Definitions of terms related to oral cancer and precancer. In: Oral Cancer and Precancer. 1st ed. Great Britain: Dorset Press; 1980. p. 16.
- 13. Smith CJ, Pindborg JJ. Histological grading of oral epithelial

atypia- by using photographic standards. Copenhagen WHO reference Centre for Oral Precancerous Conditions; 1969.

- 14. Reibel J. Prognosis of oral pre-malignant lesions: Significance of clinical, histopathological, and molecular biological characteristics. Crit Rev Oral Biol Med 2003;14:47-62.
- 15. Axéll T, Pindborg JJ, Smith CJ, van der Waal I. Oral white lesions with special reference to precancerous and tobacco- related lesions: Conclusions of an international symposium held in Uppsala, Sweden, May 18-21 1994. International Collaborative Group on Oral White Lesions. J Oral Pathol Med 1996;25:49-54.
- 16. Speight PM. Update on oral epithelial dysplasia and progression to cancer. Head Neck Pathol 2007;1:61-6.