Original Research Article

Observer variability in the grading of oral epithelial dysplasia

Sonia Gupta[1*], Suheel Hamid Latoo[2], Owais Gowhar[3]

[1]Tutor, [2]Associate Professor, [3]Tutor
Department of Oral Pathology, Govt. Dental College & Hospital, Srinagar, India
[*]Corresponding author email: **soniathegupta@gmail.com**

International Archives of Integrated Medicine, Vol. 6, Issue 11, November, 2019.

Available online at http://iaimjournal.com/
ISSN: 2394-0026 (P) **ISSN: 2394-0034 (O)**
Received on: 26-10-2019 **Accepted on:** 01-11-2019
Source of support: Nil **Conflict of interest:** None declared.
How to cite this article: Sonia Gupta, Suheel Hamid Latoo, Owais Gowhar. Observer variability in the grading of oral epithelial dysplasia. IAIM, 2019; 6(11): 35-41.

Abstract

Background: Various grading systems have been suggested in the literature by different authors to determine the severity of dysplastic features. Histopathological grading is subjective with low reproductivity and lacks sensitivity.
Aim: To determine the inter observer and intra observer variability in diagnosing oral epithelial dysplasia using different grading systems.
Materials and methods: Three oral pathologists from the Department the Oral pathology, GDC Srinagar observed the same 30 consecutive sections of oral epithelial dysplasia. Each reviewing pathologist asked to grade each case on the basis of WHO (2005) and binary grading system (2006) at different time intervals and was repeated twice.
Results: The inter observer variability ranged from poor to slight in WHO system and slight to fair in binary system whereas the intra observer variability ranged from slight to fair in WHO system and fair in the binary system.
Conclusion: Grading of oral epithelial dysplasia is subjective and has been shown not to be highly reproducible. The binary grading system verified to have better inter observer and intra observer agreement in the present study than the WHO grading system.

Key words

Potentially malignant disorder, Epithelial dysplasia, Binary system, Histological grading.

Introduction

Head and neck squamous cell carcinoma is a major health problem in several parts of the world. Although its incidence is rather low in most Western countries; the incidence in the Indian subcontinent and in the other parts of Asia, still, remains one of the most often

encountered malignancies [1]. To improve survival, the alarming features highlight the urgent need for an early diagnosis and careful evaluation of the oral potentially malignant disorders (OPMD) that are considered as precursors for malignancy [2]. The WHO in 2017 defined OPMDs as "clinical presentations that carry a risk of cancer development in the oral cavity, whether in a clinically definable precursor lesion or in clinically normal mucosa" [3]. OPMD is a clinical diagnosis for which the histological diagnosis may be hyperplasia, hyperkeratosis, oral epithelial dysplasia (OED) or oral squamous cell carcinoma (OSCC). OED is characterized by cytological and architectural alterations reflecting the loss of normal maturation and stratification pattern of surface epithelium [4]. The diagnosis of epithelial dysplasia requires uniform evaluation criteria and the prognosis of these lesions requires careful evaluation of the patient. The early diagnosis of these disorders may prevent their transformation to squamous cell carcinoma and according to various studies, thus provides a better prognosis [5-7].

The histopathological grading of oral epithelial dysplasia (OED) remains the most important predictors for assessing the malignant potential. Various grading systems have been put forward by different authors for histopathological assessment of OED and these grading systems utilize several histologic features as well as scoring criteria. It is subjective and lacks intra- and inter-observer agreement due to the inadequacy of validated morphological criteria and the biological nature of dysplasia [8-11].

The WHO in 2005 classified epithelial dysplasia into five histopathological stages as squamous hyperplasia, mild dysplasia, moderate dysplasia, severe dysplasia and carcinoma in situ (CIS). The Binary system in 2006 [12] categorized oral epithelial dysplasia into low risk and high risk. The aim of the present study was to determine the inter observer and intra observer variability in diagnosing oral epithelial dysplasia using WHO (2005) and binary (2006) grading systems.

Table - 1: WHO architectural and cytological criteria [11] to classify OED.

Architectural/ tissue changes	Cytological/ cellular changes
• Loss of polarity • Disordered maturation from basal to squamous cells • Includes top-to-bottom change of carcinoma in situ • Increased cellular density • Basal cell hyperplasia • Dyskeratosis (premature keratinization and keratin pearls deep in epithelium) • Bulbous drop shaped rete pegs • Secondary extensions (nodules) on rete tips	• Abnormal variation in nuclear size and shape (anisonucleosis and pleomorphism) • Abnormal variation in cell size and shape (anisocytosis and pleomorphism) • Increased nuclear/cytoplasmic ratio • Enlarged nuclei and cells • Hyperchromatic nuclei • Increased mitotic figures • Abnormal mitotic figures (abnormal in shape or location) • Increased number and size of nucleoli

Materials and methods

The study was conducted on 30 histopathologically diagnosed cases of oral epithelial dysplasia in the Department of Oral Pathology, GDC & H Srinagar. The study group included 10 cases originally signed out as mild dysplasia, 10 cases signed out as moderate dysplasia and 10 cases signed out as severe dysplasia. Three oral pathologists were participating in the study. All the slides were blinded and graded independently twice at an interval of two months so as to eliminate bias for

the second round of grading using WHO (2005) and binary (2006) grading systems. The score sheets were made for all the two grading system to confirm the calibration of reporting. The criteria for diagnosing epithelial dysplasia were based on architectural/ tissue and cytological/ cellular changes (**Table - 1**). On the basis of these criteria, WHO (2005) graded epithelial dysplasia as shown in **Table – 2** [11].

Table - 2: WHO (2005) grading system [11] of oral epithelial dysplasia.

Grade	Levels involved	Cytological changes	Architectural changes
Hyperplasia	N/A	1. None	1. Thickened epithelium 2. Hyperkeratosis 3. Normal maturation
Mild(I)	Lower third	1. Cell and nuclear pleomorphism 2. Nuclear hyperchromatism	1. Basal cell hyperplasia
Moderate (II)	Up to middle	1. Cell and nuclear pleomorphism 2. Anisocytosis and anisonucleosis 3. Nuclear hyperchromatism 4. Increased and abnormal mitotic figures	1. Loss of Polarity 2. Disordered maturation from basal to squamous cells 3. Increased cellular density 4. Basal Cell hyperplasia 5. Bulbous drop shaped rete pegs
Severe (III)	Up to upper third	1. Cell and nuclear pleomorphism 2. Anisocytosis and anisonucleosis 3. Nuclear hyperchromatism 4. Increased and abnormal mitotic figures 5. Enlarged nuclei and cells 6. Hyperchromatic nuclei 7. Increased number and size of nucleoli 8. Apoptotic bodies	1. Disordered maturation from basal to squamous cells 2. Increased cellular density 3. Basal cell hyperplasia 4. Dyskeratosis (premature Keratinization and keratin pearls deep in epithelium) 5. Bulbous drop shaped rete pegs 6. Secondary extensions (nodules) on rete tips 7. Acantholysis
Carcinoma-in situ	Full thickness	1. All changes may be present	1. Top-to-bottom change 2. Loss of stratification

Kujan, et al. in 2006 [12] proposed a new grading system known as binary system which was based on the same architectural and cytological criteria used by WHO (2005) for grading epithelial dysplasia. The lesions were graded as:

- **High risk lesions** (with potential susceptibility for malignant transformation): were based on observing at least four architectural changes and five cytological changes.
- **Low risk lesions** (did not have the potential susceptibility for malignant transformation): were associated with observation of less than four architectural changes or less than five cytological changes.

Statistical analysis

The data was collected and statistically analysed with the help of SPSS software (statistical package for social sciences) version 21.0 using

Kappa statistics for the determination of intra observer and inter observer variability. A probability value of $\leq$ 0.05 was considered to be statistically significant. Value of k was considered as < 0.00 = Poor, 0.00-0.20 = Slight, 0.21-0.40 = Fair, 0.41-0.60 = Moderate and 0.61-0.80 = Good.

Results

Inter observer agreement

WHO grading system

Using WHO grading system, the first observer graded 11, 11 and 8 cases as mild, moderate and severe dysplasia respectively during the first observation. The second observer graded 11, 14 and 5 cases as mild, moderate and severe dysplasia respectively as well as the third observer graded 7, 12 and 11 cases as mild, moderate and severe dysplasia respectively. The inter observer agreement between observer 1 and 2, 1 and 3, 2 and 3 showed a kappa score of 0.024 (slight), 0.033 (slight) and -0.042 (poor) respectively. The p value was found to be statistically non-significant between all the observers (**Table - 3**). During the second observation, the first observer graded 9, 13 and 8 cases as mild, moderate and severe dysplasia respectively. The second observer graded 11, 10 and 9 cases as mild, moderate and severe dysplasia respectively as well as the third observer graded 13, 9 and 8 cases as mild, moderate and severe dysplasia respectively. The inter observer agreement between observer 1 and 2, 1 and 3, 2 and 3 showed a kappa score of 0.125 (slight), 0.024 (slight) and -0.090 (poor) respectively. The p value was found to be statistically non-significant between all the observers (**Table - 3**).

Table - 3: Kappa values with its strength of agreement, percentage and probability values for all inter observer observations.

Observer pairs	Grading system	First observation			Second observation		
		K	P	Percentage	K	P	Percentage
1st observer versus 2nd observer	WHO	0.024 (Slight)	0.724	42.26	0.125 (Slight)	0.137	48.03
1st observer versus 3rd observer		0.033 (Slight)	0.625	27.57	0.024 (Slight)	0.638	24.44
2nd observer versus 3rd observer		-0.042 (Poor)	0.524	28.57	-0.090 (Poor)	0.482	20.04
1st observer versus 2nd observer	Binary	0.220 (Fair)	0.078	69.25	0.250 (Fair)	0.080	52.25
1st observer versus 3rd observer		0.098 (Slight)	0.072	38.50	0.092 (Slight)	0.069	38.04
2nd observer versus 3rd observer		0.021 (Slight)	0.689	42.75	0.028 (Slight)	0.690	43.78

Binary grading system

Using binary grading system, the first observer graded 15 and 15 cases as low risk and high risk lesions respectively during the first observation. The second observer graded 13 and 17 cases as low risk and high risk lesions respectively as well as the third observer graded 7 and 23 cases as low risk and high risk lesions respectively. The inter-observer agreement between observer 1 and 2, 1 and 3, 2 and 3 showed a kappa score of 0.220 (fair), 0.098 (slight) and 0.021 (slight) respectively. The p value was found to be statistically non-significant between all the observers (**Table - 3**). During the second observation, the first observer graded 14 and 16 cases as low risk and high risk lesions respectively. The second observer graded 13 and 17 cases as low risk and high risk lesions respectively as well as the third observer graded 9 and 21 cases as low risk and high risk lesions

respectively. The inter-observer agreement between observer 1 and 2, 1 and 3, 2 and 3 showed a kappa score of 0.250 (fair), 0.092 (slight) and 0.028 (slight) respectively. The p value was found to be statistically non-significant between all the observers (**Table - 3**).

Intra observer agreement

WHO grading system

The intra observer agreement for observer 1, 2 and 3 between the first and second observation was 59.75%, 48.03% and 61.31% respectively. The kappa score for observers 1, 2 and 3 were 0.225 (fair), 0.132 (slight) and 0.246 (fair) respectively. The p value was found to be statistically significant for observers 1 and 3 (**Table - 4**).

Table - 4: Kappa values with its strength of agreement, percentage and probability values for all intra observer observations.

Grading system	Value	1st observer	2nd observer	3rd observer
WHO	K	0.225 (Fair)	0.132 (Slight)	0.246 (Fair)
	P	0.004	0.162	0.017
	Percentage	59.75	48.03	61.31
Binary	K	0.245 (Fair)	0.224 (Fair)	0.382(Fair)
	P	0.058	0.062	0.002
	Percentage	78.95	68.07	76.32

Binary grading system

The intra observer agreement for observer 1, 2 and 3 between the first and second observation was 78.95%, 68.07% and 76.32% respectively. The kappa score for observers 1, 2 and 3 were 0.245 (fair), 0.224 (fair) and 0.382 (fair) respectively. The p value was found to be statistically significant for observer 3 (**Table - 4**).

Discussion

Dysplastic features of a stratified epithelium are characterized by cellular atypia and loss of normal maturation as well as stratification. The subjectivity in evaluating OED has often been raised which is mainly due to lack of well-defined criteria that can be recommended for grading. Though various oral pathologists determine and accept the criteria for grading epithelial dysplasia, there is great variability in their interpretation of the presence, degree and significance of the individual criteria [8, 13].

Numerous studies have shown great variability in inter observer and intra observer agreement in the diagnosis and grading of oral epithelial dysplasia as well as results ranged from poor to substantial agreement using different statistical methods [8, 13, 14]. In the present study, two grading systems namely WHO (2005) and Binary (2006) grading systems have been used for determining the inter observer and intra observer variability.

In the present study, the inter observer reliability using WHO system showed slight and poor agreement for both the observations. These results were similar to study carried out by Krishnan, et al. [14] but a study done by Kujan, et al. [12] showed slight and fair agreement.

The intra observer agreement in WHO grading system in the present study was found to be fair and slight as similar to study done by Krishnan, et al. [14]. In the earlier studies, the Kappa varied from 0.30–0.83 to 0.05–0.49 [15, 16].

In binary grading system (based on the WHO 2005 classification), there was slight and fair inter observer agreement for both the observations. These results were similar to study carried out by Krishnan, et al. [14] but in contrast to study carried out by Kujan, et al. [12] who revealed moderate agreement. The intra observer agreement in binary grading system in the

present study was found to be fair as similar to study done by Krishnan, et al. [14]. The intra observer agreement was better in binary system than in the WHO grading system. The binary system has graded as two-point scoring system than many scoring system used in the other grading systems. Though binary system may simply categorize the disease and thus decreases observer variability.

Conclusion

Oral epithelial dysplasia is a potentially malignant disorder of the oral cavity characterized histopathologically by varying degree of cytological atypia and an abnormality in the maturation of cells within a tissue. Morphologic assessment of epithelial dysplasia has usually been used as an indicator of malignant transformation. Grading of oral epithelial dysplasia is subjective and has been shown not to be highly reproducible. The binary grading system verified to have better inter observer and intra observer agreement in the present study than the WHO grading system.

References

1. van Monsjou HS, Balm AJ, van den Brekel MM, Wreesmann VB. Oropharyngeal squamous cell carcinoma: A unique disease on the rise? Oral Oncol., 2010; 46: 780-785.
2. Silverman S, Gorsky M, Kaugars GE. Leukoplakia, dysplasia and malignant trans-formation. Oral Surg Oral Med Oral Radiol Endod., 1996; 82: 117.
3. Reibel J, Gale N, Hille J, et al. Oral potentially malignant disorders and oral epithelial dysplasia. In: El-Naggar AK, Chan JKC, Grandis JR, Takata T, Slootweg PJ, eds. WHO Classification of Tumours of the Head and Neck. 4th edition, Lyon, France: IARC Press; 2017.
4. Ho PS, Chen PL, Warnakulasuriya S, Shieh TY, Chen YK, Huang IY, et al. Malignant transformation of oral potentially malignant disorders in males: A retrospective cohort study. BMC Cancer, 2009; 9: 260.
5. Speight M. P. Update on Oral Epithelial Dysplasia and Progression to Cancer. Head and Neck Pathol., 2007; 1: 61–66.
6. Sharma N, Hosmani JV, Tiwari V. Epithelial dysplasia: different grading system and its applications. J. Int Oral Health, 2010; 2(1): 1-16.
7. Jayaraj G, Ramani P, Sherlin HJ, Premkumar P, Anuja N. Inter-observer agreement in grading oral epithelial dysplasia – A systematic review. J Oral Maxillofac Surg Med Pathol., 2014; 30: 1-5.
8. 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: 127-133.
9. Bosman FT. Dysplasia classification: pathology in disgrace? J Pathol., 2001; 194: 143-144.
10. McLaren KM, Burnett RA, Goodlad JR, Howatson SR, Lang S, Lee FD, Lessells AM, Ogston S, Robertson AJ, Simpson JG, Smith GD, Tavadia HB, Walker F. Scottish Pathology Consistency Group: Consistency of histopathological reporting of laryngeal dysplasia. The Scottish Pathology Consistency Group. Histopathology, 2000; 37: 460-463.
11. Gale N, Pilch BZ, Sidransky D, El-Naggar AK, Westra W, Califano J, et al. Tumours of the oral cavity and oropharynx (Epithelial precursor lesions). In World Health Organization Classification of Tumours. Pathology & genetics. Head and neck tumours. International Agency for Research on Cancer (IARC) Edited by: Barnes L, Eveson JW, Reichart P, Sidransky D. Lyon: IARC Press; 2005, p. 177-179.
12. 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: 987-993.

13. 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: 224-231.
14. 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: 261-268.
15. 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: 300-305.
16. Abbey LM, Kaugars GE, Gunsolley JC, Burns JC, Page DG, Svirsky JA, et al. The effect of clinical information on the histopathologic diagnosis of oral epithelial dysplasia. Oral Surg Oral Med Oral Pathol Oral Radiol Endod., 1998; 85: 74-77.