Article:
Ng, G.T.E., Phang, S.C., Yu, K.S. et al. (4 more authors) (2024) Understanding interobserver variability of pathologists to improve oral epithelial dysplasia grading. Oral Diseases. ISSN 1354-523X

https://doi.org/10.1111/odi.15078

Received: 25 May 2024 | Revised: 30 June 2024 | Accepted: 9 July 2024

DOI: 10.1111/odi.15078

ORIGINAL ARTICLE

ORAL DISEASES Leading in Oral, Maxillofacial, Head & Neck Medicine WILEY

Understanding interobserver variability of pathologists to improve oral epithelial dysplasia grading

Grace Tze Ern Ng[1] | Sarah Carmen Phang[1] | Kae Shyang Yu[1] | Lalima Tiwari[1] | Syed Ali Khurram[2] | Philip Sloan[3,4,5] | Omar Kujan[1]

[1]UWA Dental School, The University of Western Australia, Nedlands, Western Australia, Australia

[2]Unit of Oral & Maxillofacial Pathology, School of Clinical Dentistry, University of Sheffield, Sheffield, UK

[3]School of Dental Sciences, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

[4]Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

[5]AMLo Biosciences, Newcastle upon Tyne, UK

Correspondence
Omar Kujan, UWA Dental School, The University of Western Australia, 17 Monash Avenue, Nedlands, 6009 WA, Australia.
Email: omar.kujan@uwa.edu.au

Abstract

Objective: This study aimed to understand reasons for interobserver variability in the grading of oral epithelial dysplasia (OED) through a survey of pathologists to provide insight for improvements in the reliability and reproducibility of OED diagnoses.

Methods: The study design included quantitative and qualitative methodology. A pre-validated 31-item questionnaire was distributed to general, head and neck, and oral and maxillofacial histopathology specialists worldwide.

Results: A total of 132 pathologists participated and completed the questionnaire. Over two-thirds used the three-tier grading system for OED, while about a third used both binary and three-tier systems. Regular reporters of OED preferred the three-tier system and grading architectural features. Continuing education significantly aided recognition of architectural and cytological changes. Irregular epithelial stratification and drop-shaped rete ridges had the lowest prognostic value and recognition scores, while loss of epithelial cell cohesion had the highest. Most participants used clinical information and often sought a second opinion when grading OED.

Conclusion: Our study has found that frequency of OED reporting and attendance of CME/CPD can play an important role in grading OED. Variations in the prognostic value of individual histological features and the use of clinical information may further contribute to interobserver variability.

KEYWORDS
grading, interobserver variability, intra-observer, oral epithelial dysplasia, survey

1 | INTRODUCTION

Grading of epithelial dysplasia is deemed the most important predictor of the risk of malignant transformation in oral potentially malignant disorders (OPMD; Odell et al., 2021; Speight et al., 2018). Oral epithelial dysplasia (OED) refers to histopathological changes in the oral epithelium that may transform into carcinoma (Odell et al., 2021). Interestingly, the hallmarks of carcinogenesis suggest that the process of a normal cell or tissue transforming into an abnormal and subsequently invasively destructive cell/tissue involves a continuous acquisition of mutations and genetic aberrations (Hanahan & Weinberg, 2011). Therefore, any grading system of OED does not accurately represent the carcinogenesis process and is technically an arbitrary measure of malignant transformation

risks, as the natural history of oral carcinogenesis has not been established, and the grades are subjective estimates of the observed changes at a point in time (Odell et al., 2021). Advances in molecular tests continue to emerge to decipher the risk of the development of OED within OPMD more objectively (El-Sakka et al., 2018; Kujan et al., 2021, 2023; Morais et al., 2020). However, before these molecular tests can be plausibly carried out clinically, histopathological assessment of OPMD remains the current standard used for predicting the risk of malignant transformation (Ranganathan et al., 2020; Reibel, 2003).

Over the years, various grading systems for OED have evolved (Odell et al., 2021; Ranganathan & Kavitha, 2019). At the time of the survey, the current grading system for OED grading was the WHO 2022 classification that is based on the 2017 version of the 3-tiered system: mild, moderate and severe with an expansion of the descriptive cytological and architectural changes (Muller & Tilakaratne, 2022). A revised classification has been published recently (Lingen et al., 2023). Like the previous WHO edition (2017), the WHO 2022 classification states that the binary OED grading system introduced in 2006 is helpful but still needs further validation (Lingen et al., 2023; Muller & Tilakaratne, 2022). Nonetheless, intra- and interobserver agreement variability of OED reporting has been widely reported (Dost et al., 2014; Kujan et al., 2007; Mahmood et al., 2022; Manchanda & Shetty, 2012; Nankivell et al., 2013; Ranganathan et al., 2020). The relatively poor and moderate reproducibility of OED grading has shadowed its clinical utility. Therefore, understanding why pathologists perceive the microscopic features of OED differently is a major step in improving the reliability of OED reporting. The lack of agreement on grading oral dysplastic lesions has been attributed to subjectivity in the evaluation of the established criteria of grading, arbitrary division of the gradings, lack of calibration of the used criteria and grading, and a lack of sufficient knowledge of which criteria are significant for the prediction of malignant potential (Kujan et al., 2007; Odell et al., 2021; Ranganathan et al., 2020). In this study, we aimed to report the opinions and practices of pathologists regarding their reporting of OED worldwide in an attempt to understand the factors that might be contributing towards the intra-and interobserver variability in the diagnosis and severity of OED. By gaining a further understanding of the diagnostic process of pathologists, we could improve the reliability and reproducibility of OED grading.

2 | MATERIALS AND METHODS

2.1 | Study design

This is a cross-sectional online survey conducted with approval by the Human Research Ethics Committee of The University of Western Australia (UWA) (Re: 2021/ET000685) and reported in accordance with CHERRIES guidelines (Eysenbach, 2004).

2.2 | Participants

In this investigation, a convenience sample strategy was used to invite eligible pathologists to participate in the study anonymously and willingly. The target population involved registered and practising pathologists with specialisations in General histopathology, Head and Neck histopathology and Oral and Maxillofacial histopathology. Inclusion criteria were the following (i) participants over the age of 18, and (ii) registered and practicing pathologists in their designated country. No associated incentives were provided for participants, and only one survey reminder email was sent. The questionnaire, along with a cover letter that explained the study's goals and methods and assured that participation was voluntary, anonymous and that all information provided would be kept confidential and used only for research purposes, was sent to the secretary of the following associations to be distributed to their members:

- International Association of Oral and Maxillofacial Pathologists
- Royal College of Pathologists Australasia
- Royal College of Pathologists (UK)
- British Association for Oral and Maxillofacial Pathology
- American Academy of Oral and Maxillofacial Pathology
- North American Society of Head and Neck Pathology
- Japanese Society of Oral and Maxillofacial Pathology
- Asian Society of Oral and Maxillofacial Pathology
- Indian Association of Oral and Maxillofacial Pathologists
- Chinese Stomatological Association
- Spanish Society of Odontostomatology
- Italian Society of Oral Pathology and Medicine

The questionnaire was distributed to pathologists using Qualtrics XM® (Qualtrics Survey Platform, WA, USA) between September 2021 and March 2022.

2.3 | Questionnaire and data collection

The 31-item questionnaire (Material S1) utilised in this study was developed following multiple discussions with leading experts in OED diagnosis and grading (OK, SAK and PS). To ensure its clarity and simplicity, the questionnaire was pilot-tested on a group of four histopathologists. The test–retest method was used to examine reliability, with eight students and interns completing the questionnaire twice within 2 weeks. The outcomes were compared two times using Pearson's correlation coefficient, which revealed a substantial stability coefficient indicating strong test–retest reliability. Internal consistency was tested using the coefficient alpha 'Cronbach's alpha' between survey items. Cronbach's alpha = 0.857 was obtained, indicating acceptable internal consistency.

The questionnaire had 29 closed-ended questions and two open questions divided into three sections: demographics, practices and opinions about OED reporting. The first section contained

demographic information such as age, gender, educational level, histopathology training type and duration, and continuing education activity. The second part comprised 21 questions about the participants' reporting practices of OED and the perceived prognostic value of OED histopathological characteristics.

2.4 | Statistical analysis

Quantitative data were analysed using the SPSS software (IBM SPSS Statistics for Windows, Version 27.0). One-way ANOVA was used for categorical data, and the chi-squared test was used for nominal and ordinal data. Descriptive statistics and frequency tables were generated from the data analysis. Incomplete questionnaire responses were excluded from the statistical analysis of results. Where survey questions required ranking, mean scores were calculated from the range of 1 (*best*) to 10 (*worst*). The participants were asked to rank, with lower mean scores depicting greater ease of recognition and prognostic values and vice versa. Statistical significance was set at $p < 0.05$.

3 | RESULTS

This survey was completed by 132 participants (Male: 70, Female: 62), most of whom were 40–59 years old (59.9%). Their primary qualifications were mostly dental (78.8%). Most of the participants were involved in a full-time role (62.1%) and were based at an academic/university institution (57.6%).

With regard to post-graduate qualifications and working experiences of our participants, we found that the majority (78%) of our participants had their speciality training in Oral and Maxillofacial Histopathology as compared to the other majors such as head and neck histopathology and general histopathology and reported on both biopsies and excisions/resections of both OED and OSCC (Table 1). More than two-thirds (68.9%) of our participants reported OED cases weekly (Table 1). Additionally, the previous history of oral cancer was regarded as the most important clinical, demographic factor for clinicopathological correlation with OED (mean rank = 2.29). The use of digital pathology or whole slide images in the evaluation of OED is uncommon since 50% reported they never used it, 35% used it a few times, 9.5% reported very often use, and nearly 6% used it every time.

When asked about the most recent time they attended CME/CPD course, more than half of the participants reported that it was within the past 12 months, with the remaining participants reporting attending such courses more than a year ago (Table 1). Significant differences were found between groups when assessing the ease of recognition of architectural and cytological changes based on their last attended course (Table 1).

Over two-thirds (67.5%) of our participants adopted the three-tier grading system, while approximately a third (27.3%) of them utilised both the binary and three-tier grading systems (Table 1). Most participants (93.2%) used the WHO 2017 Head and Neck Tumour Classification for the final grading of OED (Table 1). Participants reporting OED cases weekly noted increased satisfaction when using the 3-tier grading system and in grading architectural changes (Table 1). Furthermore, there was significant ease in recognising architectural and cytological changes, including irregular epithelial stratification, abnormal superficial mitosis and anisocytosis (Table 1).

When architectural changes of epithelial cells were being scored by participants for their perceived prognostic values and ease of recognition, irregular epithelial stratification and drop-shaped rete ridges were found to have the lowest mean scores (4.52, 3.91) and (4.71, 2.89), respectively (Table S2). On the other hand, loss of epithelial cell cohesion was reported with the highest mean scores of 5.46 and 4.32, respectively, for its perceived prognostic value and ease of recognition (Table S2). When scored by participants for perceived prognostic values and ease of recognition of cytological changes, anisonucleosis had the lowest mean scores of 4.70 and 3.82, respectively, whereas atypical mitotic figures had the lowest mean scores of 3.40 for the latter (Table S2). The increased number and size of nucleoli had the highest mean scores of 5.56 and 4.85, respectively, for its perceived prognostic value and ease of recognition (Table S2). Additionally, the satisfaction of the pathologist with the available clinical information led to significant increases in recognition of specific architectural changes (Table 1).

Furthermore, most participants (94.7%) reported using clinical information when grading OED. Amongst these participants, 82.6% requested additional clinical information, whereas almost 50% used the clinical information provided to alter their final grading (Table 1). After grading OED, most participants (77%) sought a second opinion to confirm their diagnosis (Table 1). Most participants who asked for a second opinion were in the 'occasionally' and 'very often' groups (Table 1). A similar pattern was observed for the proportion of participants who sought consensus when they encountered a disagreement in recognition of characteristics or grading (Table 1). Overall, a significant increase in satisfaction with participants' OED grading was associated with the frequency of OED reporting and use of clinical information.

4 | DISCUSSION

Pathologists have crucial responsibilities in determining the histopathological diagnosis of OED. Yet their role has been frequently overshadowed by the existing discrepancies in agreement between pathologists on the diagnosis and grading of OED. It has been suggested that variability in reporting may be due to differing levels of importance given to morphological characteristics in grading by the pathologist, variability in the observations of these characteristics and grading of these observations into the various categories of epithelial dysplasia (De Vet et al., 1992; Elmore et al., 2017). Furthermore, grading dysplasia is limited by the arbitrary division into distinct categories of a continuous progressive process. In this study, the survey method was utilised to delve into understanding

TABLE 1 Percentage of respondents to the corresponding survey questions and their statistical significance.

	Percentage (%) of respondents	Statistical significance
Specialty training major		
General	11.4	Q20* (use of grading system) $p=0.016$
Head and Neck	10.6	
Oral and Maxillofacial	78	
Last CME/CPD course attendance		
Never taken a course	6.8	Ease of recognising architectural changes Q24G* (Loss of epithelial cell cohesion) $p=0.037$ *Ease of recognising cytological changes* Q26B* (Nuclear pleomorphism) $p=0.028$
>5 years ago	4.5	
2–5 years ago	18.2	
>12 months and <24 months ago	10.6	
≤12 months ago	59.1	
Grading system used		
Three-tier system	67.4	Ease of recognising architectural changes Q25A* (Irregular epithelial stratification) $p<0.001$ Q25B* (Loss of polarity of basal cells) $p<0.001$ *Satisfaction* Q31C* (Binary grading) $p=0.008$
Binary system	4.5	
Three-tier and binary system	27.3	
Neither three-tier or binary systems	0.8	
Use of 2017 WHO classification blue book for scoring architectural and cytological features of OED		
Yes	56.8	*Emphasis on cytological features* Q24F* (Atypical mitotic figures) $p=0.038$ Ease of recognising architectural changes Q25E* (Abnormal superifical mitosis) $p=0.041$
No	43.2	
Use of 2017 WHO Classification Blue Book for Grading OED		
Yes	93.2	*Satisfaction* Q31A* (Overall reporting of OED) $p=0.046$
No	6.8	
Use of Clinical Information for OED Grading		
Yes	94.7	Ease of recognising cytological changes Q26H* (Hyperchromasia) $p<0.01$ *Satisfaction of OED reporting* Q31F* (Higher satisfaction of architectural features) $p=0.027$
No	5.3	
Request for clinical information/photographs		
Yes	82.6	Q10* (last CME/CPD course attendance) $p=0.012$
No	17.4	
Downgrade/upgrade of final grading		
Yes	48.9	No statistical significance between groups
No	51.1	
Seeking Second Opinions for OED Grading		
Yes	77	Q17* (Additional clinical photographs and information for diagnosis) $p=0.031$
No	23	
Weekly reporting of OED		
Yes	68.9	*Ease of recognition of architectural changes* Q25A* (Irregular epithelial stratification) $p=0.003$ Q25E* (Abnormal superficial mitosis) $p=0.018$ *Ease of recognition of cytological changes* Q26C* (Anisocytosis) $p=0.035$ *Satisfaction* Q31B* (Increased satisfaction with use of 3 tier grading system) $p=0.037$ Q31F* (Increased satisfaction with architectural features) $p=0.015$
No	31.1	
Types of specimen reporting		
Biopsies only	28	No statistical significance between groups
Biopsies and wide local excisions/ resections of OED and/or OSCC	72	

Note: Please refer to Table S3, survey questions.

*$p<0.05$.

the thought processes used by pathologists when grading and diagnosing OED and, in turn, deduce further reasons for the existing intra- and interobserver variability in the diagnosis of OED.

One of the significant findings revealed within our results was related to the experience and frequency of reporting OED by the pathologist. In our study, most pathologists worked full-time, were trained in Oral and Maxillofacial Histopathology and reported OED cases on a weekly basis. Specifically, pathologists who reported OED cases on a weekly basis noted increased satisfaction when using the 3-tier grading system and in grading architectural histological features. They also noted an increased ease of recognition of architectural changes, including irregular epithelial stratification, abnormal superficial mitosis and cytological changes, including anisocytosis. A study by Geetha et al. (2015) found that pathologists who were accustomed over the years to the use of a grading system routinely in their practice had low interobserver variability/high agreeability, potentially suggesting that how experienced and satisfied the pathologists are with the grading system may be associated with interobserver agreement. Additionally, our results demonstrated that most pathologists had taken a continuing education course in OED within the past 12 months. This was significantly associated with the ease of recognising loss of epithelial cell cohesion and nuclear pleomorphism, suggesting that regular training improves the ease of recognising histological features of dysplasia. A recent study by Kallarakkal et al. (2024) found significant improvement in interobserver agreement of the WHO 2017 OED grading system after calibration of oral pathologists in Malaysia, improving the reproducibility of OED grading. The study also noted that interobserver agreement reduced variability amongst pathologists after calibration, irrespective of the educational background and experience of oral pathologists. Continuous calibration and consensus meetings have been further promoted as contributors to improving consistency between pathologists (Sathasivam et al., 2022). These findings further add value to the ongoing training of pathologists and provide an excellent avenue for calibration.

The list of features for grading OED is not regarded as definitive and has evolved over the years. Additional features have been included in the latest edition of the WHO Blue Book (Lingen et al., 2023). This is mainly due to pathologists generally accepting certain features as relevant in diagnosing dysplasia but currently lacking a formal evidence base (Odell et al., 2021). Again, variability in interobserver agreement may be impacted by the ease of recognising these features by various pathologists. Generally, it is considered that oral pathologists in academia are faster to adopt these updates, while those in health services may lag behind (Kallarakkal et al., 2024). In our study, pathologists who were primarily based in academia/university positions found that drop-shaped rete ridges and abnormal superficial mitosis were the most easily recognised architectural features, while atypical mitotic figures and hyperchromasia were the most easily recognised cytological features. These findings are similar to Kujan et al., who noted the highest level of agreement amongst pathologists in an increased number of mitotic figures and drop-shaped rete ridges as architectural features (Kujan et al., 2007). However, atypical mitotic figures and hyperchromatism were cytological features that corresponded with the highest disagreement between pathologists, whereas increased nuclear size and abnormal variation in cell shape corresponded with the highest level of agreement (Kujan et al., 2007). In our study, however, increased nuclear size and abnormal variation in cell shape corresponded to the third and fourth highest features in ease of recognition. Additionally, the ease of recognising irregular epithelial stratification and loss of polarity of basal cells was dependent on the grading system used at statistically significant levels. Specifically, using the 2017 WHO grading criteria was statistically significant for ease of recognising abnormal superficial mitosis. It should be noted that Kujan et al. (2007) evaluated the 2005 WHO classification for OED, while our survey used the WHO 2017 grading criteria. Again, the evolution of grading classifications may partly account for some of these discrepancies with previous studies (De Vet et al., 1992). These findings suggest that architectural features, particularly drop-shaped rete ridges and abnormal superficial mitosis, should be considered relevant features to retain in future classifications to improve interobserver agreement.

Another source of interobserver variability has been attributed to the level of importance given to morphological characteristics in grading dysplasia. In a sense, this relates to the prognostic value given to each characteristic when grading epithelial dysplasia. Older studies suggested features of 'basal cell hyperplasia', nuclear hyperchromatism, enlarged nucleoli, loss of cohesion and basal polarity, and abnormal mitosis as features of particular significance. These studies, however, have now been disregarded due to limitations, including obsolete definitions and small sample size. In our study, irregular epithelial stratification was regarded as holding the highest prognostic value in grading epithelial dysplasia. Cytologically, increased nuclear/cytoplasmic ratio held the highest perceived prognostic value. Interestingly, these findings do not align with the recent suggestions by Odell et al., for features likely to have relative specificity for OED (Odell et al., 2021). In fact, loss of epithelial cell cohesion and increased number and size of nucleoli were regarded to have the least prognostic value in our study and were considered to have high specificity with OED based on underlying pathological processes (Odell et al., 2021). Another study by Mahmood et al. found loss of epithelial cell cohesion and bulbous rete pegs to demonstrate the highest correlation between individual OED histological features and prognosis (Mahmood et al., 2022). These discrepancies may provide some insight into the variability that exists in interobserver agreement, as reporting pathologists appear to have differing opinions regarding prognostic value when grading epithelial dysplasia.

The adoption of clinical information when diagnosing suspicious OED lesions was proposed by Krutchkoff et al. (1991). Following this, weighing available clinical information with histopathological diagnosis was also recommended by Abbey et al. (1995). In contrast to these suggestions, clinical information

provided during the diagnostic process was shown to induce bias between pathologists (Karabulut et al., 1995). Although not perfect, OED grading in conjunction with clinical findings is still the most frequently used method to stratify and manage patients with OPMD (Sathasivam et al., 2022). Our study noted that almost 95% of pathologists used clinical information when grading OED. This could be attributed to most of our respondents being oral and maxillofacial pathologists whose training involves access to clinical oral pathology. Additionally, it was noted that clinical information was statistically significant in increasing the ease of recognition of hyperchromasia and provided greater satisfaction when reporting architectural features. Furthermore, the previous history of oral cancer was regarded as the most important clinical demographic factor for clinicopathological correlation with OED. This is an interesting finding, given that studies evaluating interobserver variability in grading epithelial dysplasia generally blind the participating pathologists from any clinical information (Geetha et al., 2015; Krishnan et al., 2016; Kujan et al., 2006; Manchanda & Shetty, 2012). The lack of clinical information in these studies may, in part, influence some of the reported variability noted. Future studies investigating interobserver variability may consider including clinical information as part of their tested variable in influencing agreement when grading OED.

Many studies now recommend and support the use of an adjudicator or achieving consensus, when there is disagreement between pathologists during diagnosis of OED (Kallarakkal et al., 2024; Sathasivam et al., 2022; Speight et al., 2015). Our study supported these findings, demonstrating that 77% of pathologists sought second opinions for OED grading, and that this was also significant with requesting additional clinical photographs and information for diagnosis. These findings demonstrate a promising insight into the practicing patterns of pathologists, especially as there is now significant data to show that reproducibility in OED grading systems can be improved by consensus reporting (Sathasivam et al., 2022). Consideration should be given for future classifications to incorporate consensus reporting as a requirement when reporting OED.

The current study has some limitations, such as the relatively small sample size despite the efforts to invite all relevant stakeholders and the imbalance in the distribution of the samples, as most respondents are oral and maxillofacial pathologists whose training and scope of practice are different from those of head and neck pathologists. In addition, the impact of CME/CPD activities on OED grading could be more informative if further information was sought.

Ultimately, the current histological grading system for reporting OED has demonstrated a need for improvement in its consistency and reproducibility. To the best of our knowledge, this is the first study to look at pathologists' practice patterns and thought processes in grading OED. While the results from our study have provided some significant insights into pathologists' reporting practices, future studies may be able to build on these findings by using more open-ended questions, which may provide further details about the variability that exists in reporting OED amongst pathologists. Combined with the increasing demands for 'evidence-based medicine', there is an urgent need to refine and scientifically validate the diagnostic criteria used when examining histopathological evidence.

5 | CONCLUSION

Findings from this study provide significant insights into various factors contributing to the interobserver variability in grading OED. Specifically, the frequency of OED reporting and attendance of CME/CPD events played a significant role in recognising histological features. Furthermore, specific histological features, including drop-shaped rete ridges and abnormal superficial mitosis, were the most easily recognised architectural features and could be considered relevant findings for formulating future OED classifications. The study also noted variations in the perceived prognostic value of individual histological features, which may further contribute towards interobserver variability. Clinical information was frequently used in the diagnosis of OED and should be considered as a variable to assess in future studies. Finally, consideration should be given for future OED classifications to incorporate consensus reporting as a requirement, given that current practices already reflect this, and can further improve reproducibility in OED grading.

AUTHOR CONTRIBUTIONS

Grace Tze Ern Ng: Writing – original draft; writing – review and editing; investigation; data curation. **Sarah Carmen Phang:** Investigation; data curation; writing – original draft; writing – review and editing. **Kae Shyang Yu:** Investigation; writing – original draft; writing – review and editing; data curation. **Lalima Tiwari:** Formal analysis; writing – review and editing; data curation. **Syed Ali Khurram:** Methodology; writing – review and editing. **Philip Sloan:** Methodology; writing – review and editing. **Omar Kujan:** Conceptualization; methodology; investigation; validation; supervision; project administration; writing – review and editing; resources.

ACKNOWLEDGEMENTS

The authors thank the respondents for their help in completing the questionnaire. Open access publishing facilitated by The University of Western Australia, as part of the Wiley - The University of Western Australia agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Lalima Tiwari https://orcid.org/0000-0001-5305-2953
Syed Ali Khurram https://orcid.org/0000-0002-0378-9380
Omar Kujan https://orcid.org/0000-0002-5951-8280

REFERENCES

Abbey, L. M., Kaugars, G. E., Gunsolley, J. C., Burns, J. C., Page, D. G., Svirsky, J. A., Eisenberg, E., Krutchkoff, D. J., & Cushing, M. (1995). Intraexaminer and interexaminer reliability in the diagnosis of oral epithelial dysplasia. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *80*(2), 188–191.

De Vet, H. C. W., Knipschild, P. G., Schouten, H. J. A., Koudstaal, J., Kwee, W.-S., Willebrand, D., Sturmans, F., & Arends, J. W. (1992). Sources of interobserver variation in histopathological grading of cervical dysplasia. *Journal of Clinical Epidemiology*, *45*(7), 785–790. https://doi.org/10.1016/0895-4356(92)90056-S

Dost, F., Lê Cao, K., Ford, P. J., Ades, C., & Farah, C. S. (2014). Malignant transformation of oral epithelial dysplasia: A real-world evaluation of histopathologic grading. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology*, *117*(3), 343–352. https://doi.org/10.1016/j.oooo.2013.09.017

Elmore, J. G., Barnhill, R. L., Elder, D. E., Longton, G. M., Pepe, M. S., Reisch, L. M., Carney, P. A., Titus, L. J., Nelson, H. D., Onega, T., Tosteson, A. N. A., Weinstock, M. A., Knezevich, S. R., & Piepkorn, M. W. (2017). Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: Observer accuracy and reproducibility study. *BMJ*, *357*, j2813. https://doi.org/10.1136/bmj.j2813

El-Sakka, H., Kujan, O., & Farah, C. S. (2018). Assessing miRNAs profile expression as a risk stratification biomarker in oral potentially malignant disorders: A systematic review. *Oral Oncology*, *77*, 57–82. https://doi.org/10.1016/j.oraloncology.2017.11.021

Eysenbach, G. (2004). Improving the quality of web surveys: The checklist for reporting results of internet E-surveys (CHERRIES). *Journal of Medical Internet Research*, *6*(3), e34. https://doi.org/10.2196/jmir.6.3.e34

Geetha, K. M., Leeky, M., Narayan, T. V., Sadhana, S., & Saleha, J. (2015). Grading of oral epithelial dysplasia: Points to ponder. *Journal of Oral and Maxillofacial Pathology*, *19*(2), 198–204. https://doi.org/10.4103/0973-029X.164533

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, *144*(5), 646–674.

Kallarakkal, T. G., Zaini, Z. M., Ghani, W. M. N., Karen-Ng, L. P., Siriwardena, B., Cheong, S. C., & Tilakaratne, W. M. (2024). Calibration improves the agreement in grading oral epithelial dysplasia-findings from a National Workshop in Malaysia. *Journal of Oral Pathology & Medicine*, *53*(1), 53–60. https://doi.org/10.1111/jop.13501

Karabulut, A., Reibel, J., Therkildsen, M. H., Praetorius, F., Nielsen, H. W., & Dabelsteen, E. (1995). Observer variability in the histologic assessment of oral premalignant lesions. *Journal of Oral Pathology & Medicine*, *24*(5), 198–200. https://doi.org/10.1111/j.1600-0714.1995.tb01166.x

Krishnan, L., Karpagaselvi, K., Kumarswamy, J., Sudheendra, U. S., Santosh, K. V., & Patil, A. (2016). Inter- and intra-observer variability in three grading systems for oral epithelial dysplasia. *Journal of Oral and Maxillofacial Pathology*, *20*(2), 261–268. https://doi.org/10.4103/0973-029X.185928

Krutchkoff, D. J., Eisenberg, E., & Anderson, C. (1991). Dysplasia of oral mucosa: A unified approach to proper evaluation. *Modern Pathology*, *4*(1), 113–119.

Kujan, O., Idrees, M., Anand, N., Soh, B., Wong, E., & Farah, C. S. (2021). Efficacy of oral brush cytology cell block immunocytochemistry in the diagnosis of oral leukoplakia and oral squamous cell carcinoma. *Journal of Oral Pathology & Medicine*, *50*(5), 451–458. https://doi.org/10.1111/jop.13153

Kujan, O., Khattab, A., Oliver, R. J., Roberts, S. A., Thakker, N., & Sloan, P. (2007). Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia: An attempt to understand the sources of variation. *Oral Oncology*, *43*(3), 224–231. https://doi.org/10.1016/j.oraloncology.2006.03.009

Kujan, O., Oliver, R. J., Khattab, A., Roberts, S. A., Thakker, N., & Sloan, P. (2006). Evaluation of a new binary system of grading oral epithelial dysplasia for prediction of malignant transformation. *Oral Oncology*, *42*(10), 987–993. https://doi.org/10.1016/j.oraloncology.2005.12.014

Kujan, O., Siddiqui, I., Lee, C., Idrees, M., Shearston, K., & Farah, C. S. (2023). Automated immunohistochemical quantification of hypoxia biomarkers shows correlation with dysplastic epithelial changes. *Journal of Oral Pathology & Medicine*, *52*(6), 504–513. https://doi.org/10.1111/jop.13427

Lingen, M., Vigneswaran, N., Kujan, O., Kurgo, Z. B., Poh, C., & Ranganathan, K. (Eds.). (2023). Chapter XX: Oral epithelial dysplasia. In *WHO classification of Tumours editorial board. Head and neck tumours. WHO classification of tumours series* (Vol. *9*, 5th ed.). International Agency for Research on Cancer. https://publications.iarc.who.int/629

Mahmood, H., Bradburn, M., Rajpoot, N., Islam, N. M., Kujan, O., & Khurram, S. A. (2022). Prediction of malignant transformation and recurrence of oral epithelial dysplasia using architectural and cytological feature specific prognostic models. *Modern Pathology*, *35*(9), 1151–1159. https://doi.org/10.1038/s41379-022-01067-x

Manchanda, A., & Shetty, D. C. (2012). Reproducibility of grading systems in oral epithelial dysplasia. *Medicina Oral, Patología Oral y Cirugía Bucal*, *17*(6), e935–e942. https://doi.org/10.4317/medoral.17749

Morais, E. F., Pinheiro, J. C., Lira, J. A., Mafra, R. P., Barboza, C. A., Souza, L. B., & Freitas, R. D. (2020). Prognostic value of the immunohistochemical detection of epithelial-mesenchymal transition biomarkers in oral epithelial dysplasia: A systematic review. *Medicina Oral, Patología Oral y Cirugía Bucal*, *25*(2), e205–e216. https://doi.org/10.4317/medoral.23305

Muller, S., & Tilakaratne, W. M. (2022). Update from the 5th edition of the World Health Organization Classification of Head and Neck Tumors: Tumours of the oral cavity and mobile tongue. *Head and Neck Pathology*, *16*(1), 54–62. https://doi.org/10.1007/s12105-021-01402-9

Nankivell, P., Williams, H., Matthews, P., Suortamo, S., Snead, D., McConkey, C., & Mehanna, H. (2013). The binary oral dysplasia grading system: Validity testing and suggested improvement. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology*, *115*(1), 87–94. https://doi.org/10.1016/j.oooo.2012.10.015

Odell, E., Kujan, O., Warnakulasuriya, S., & Sloan, P. (2021). Oral epithelial dysplasia: Recognition, grading and clinical significance. *Oral Diseases*, *27*(8), 1947–1976. https://doi.org/10.1111/odi.13993

Ranganathan, K., & Kavitha, L. (2019). Oral epithelial dysplasia: Classifications and clinical relevance in risk assessment of oral potentially malignant disorders. *Journal of Oral and Maxillofacial Pathology*, *23*(1), 19–27. https://doi.org/10.4103/jomfp.JOMFP_13_19

Ranganathan, K., Kavitha, L., Sharada, P., Bavle, R. M., Rao, R. S., Pattanshetty, S. M., Hazarey, V. K., Madhura, M. G., Nagaraj, T., Lingappa, A., & Warnakulasuriya, S. (2020). 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*, *49*(9), 948–955. https://doi.org/10.1111/jop.13056

Reibel, J. (2003). Prognosis of oral pre-malignant lesions: Significance of clinical, histopathological, and molecular biological characteristics. *Critical Reviews in Oral Biology and Medicine*, *14*(1), 47–62.

Sathasivam, H. P., Sloan, P., Thomson, P. J., & Robinson, M. (2022). The clinical utility of contemporary oral epithelial dysplasia grading systems. *Journal of Oral Pathology & Medicine*, *51*(2), 180–187. https://doi.org/10.1111/jop.13262

Speight, P. M., Abram, T. J., Floriano, P. N., James, R., Vick, J., Thornhill, M. H., Murdoch, C., Freeman, C., Hegarty, A. M., D'Apice, K., Kerr, A. R., Phelan, J., Corby, P., Khouly, I., Vigneswaran, N., Bouquot, J., Demian, N. M., Weinstock, Y. E., Redding, S. W., … McDevitt, J. T. (2015). Interobserver agreement in dysplasia grading: Toward an enhanced gold standard for clinical pathology trials. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics*, *120*(4), 474–482.e2. https://doi.org/10.1016/j.oooo.2015.05.023

Speight, P. M., Khurram, S. A., & Kujan, O. (2018). Oral potentially malignant disorders: Risk of progression to malignancy. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology*, *125*(6), 612–627. https://doi.org/10.1016/j.oooo.2017.12.011

SUPPORTING INFORMATION

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

How to cite this article: Ng, G. T. E., Phang, S. C., Yu, K. S., Tiwari, L., Khurram, S. A., Sloan, P., & Kujan, O. (2024). Understanding interobserver variability of pathologists to improve oral epithelial dysplasia grading. *Oral Diseases*, *00*, 1–8. https://doi.org/10.1111/odi.15078