8-16-2024 10:00 AM

Early Detection of Oral Potentially Malignant Disorders

Xinru Li, *Western University*

Supervisor: Darling, Mark, *The University of Western Ontario*
Co-Supervisor: Olea-Popelka, Francisco, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Pathology and Laboratory Medicine
© Xinru Li 2024

Abstract

Oral potentially malignant disorders (OPMDs) represent a diverse array of conditions with an elevated propensity for progressing into oral squamous cell carcinoma (OSCC). My study aims to enhance the early detection of OPMDs by investigating the expression levels of specific biomarkers, their correlation with disease severity, and the concordance between clinical and pathological diagnoses. Using nanoString gene expression analysis and immunohistochemical (IHC) staining, I identified significantly elevated levels of S100A7, Ki67, and Vimentin in OPMD tissues compared to normal controls, suggesting their potential as biomarkers for early detection and monitoring of disease progression. Conversely, E-cadherin showed reduced expression in OPMDs, indicating disruptions in cellular adhesion and the MAPK signaling pathway. My findings on moderate concordance between clinical and histopathological diagnoses, highlighting the complexity of diagnosing OPMDs based solely on clinical oral examination and the importance of histopathological confirmation. This underscores the need for improved diagnostic accuracy through enhanced clinical training and the use of molecular diagnostic tools. Despite limitations such as small sample size and geographic constraints, my research underscores the critical role of integrating molecular data with clinical diagnostics to improve the early detection and risk stratification of OPMDs. Future research should focus on developing comprehensive predictive models by integrating multiple biomarkers and leveraging digital pathology and artificial intelligence to refine these models. This approach holds promise for early intervention and better management of patients at risk of malignant transformation, ultimately enhancing patient care and outcomes.

Keywords: Oral potential malignant disorders, S100A7, immunohistochemistry, molecular evaluation, oral squamous cell carcinoma, digital pathology

Summary for Lay Audience

In my study, I explored how molecular methods can help identify early signs of disease in certain oral lesions known as oral potentially malignant disorders (OPMDs). These disorders, which include various conditions in the mouth, have a higher risk of developing into a serious form of oral cancer called oral squamous cell carcinoma (OSCC). Early detection of these conditions is crucial, as they have a higher likelihood to become cancerous. To achieve this, I collected biopsy samples from patients diagnosed with OPMDs and analyzed the genetic material (RNA) extracted from these samples. I also used a technique called Immunohistochemistry (IHC) to visualize specific proteins within the cells, which could indicate the disease's progression. My study found that a protein called S100A7 was present in higher amounts in the precancerous tissues compared to normal mouth tissue. The levels of S100A7 were also higher in more severe cases, suggesting it could be used to identify early cancer risk. Additionally, proteins called Ki67 and Vimentin were also higher in OPMD tissues, which are linked to cell growth and inflammation. On the other hand, proteins called E-cadherin was lower in OPMD patient group, indicating problems with cell connections and signaling. The study also highlighted the importance of combining clinical diagnoses with molecular analysis to improve the accuracy of detecting OPMDs. While clinical examinations are important, the integration of molecular data can provide a deeper understanding of the disease and help identify high-risk patients. I believe this approach could lead to early detection and intervention of OPMDs, and better care for patients.

Acknowledgments

Frist, I would like to express my deepest gratitude to my supervisors, Dr. Mark Darling and Dr. Francisco Olea-Popelka, for their unwavering support, guidance, and encouragement throughout the course of my research. Their insightful feedback, patience, and expertise have been invaluable to my growth as a researcher and have significantly contributed to the completion of this thesis. I am also immensely thankful to the members of my advisory committee, Dr. Zia Khan, and Dr. Noha Gomaa, for their valuable time, critiques, and suggestions. Their diverse perspectives and constructive feedback have greatly enriched my research, helping me to refine my ideas and methodologies.

I would like to thank David Carter, the facility manager at Robarts Research Institute for providing valuable training and access to lab equipment needed for my project. I would also like to express my gratitude to Linda Jackson for her unequivocal support in the oral pathology lab department. They have supported me in my research by offering me advice and helping me in troubleshooting methodological problems in my experiments.

I would like to acknowledge the people in the Pathology department, especially the administrative staff, Ms. Tracey Koning, Ms. Susan Underhill, a Ms. Elizabeth Goldhawk. They have helped me every step of the way since I started my MSc program. I would like to thank Ms. Kathilyn Allewell for helping me create presentable posters for conferences.

Finally, I would like to express my deep and sincere gratitude to my peers at the pathology department for helping me throughout my journey and sharing their expertise with me. I would like to thank Carolyn Twible and Rober Abdo for taking the time out to help me, and I am grateful for their contributions to my academic development.

This journey would not have been possible without the collective support, encouragement, and wisdom of these remarkable individuals. I am honored to have had the opportunity to learn from and work with such dedicated professionals. Thank you for believing in me, challenging me, and guiding me towards achieving this significant milestone.

Table of Contents

List of Tables

List of Figures

List of Appendices

<div align="center">Chapter 1</div>

1   Introduction and Background

Oral potentially malignant disorders (OPMD) encompass a variety of oral lesions and conditions that have a higher risk of progressing to oral cancer, particularly oral squamous cell carcinoma (OSCC). Early diagnosis of OPMDs is crucial in preventing the development of OSCC, since early detection of OPMDs allows for timely intervention and effective monitoring and follow-up of patients. The progression from OPMD to OSCC involves various molecular pathways and mechanisms that contribute to cellular changes, including alterations in cell cycle regulation, cellular signaling pathways, and phenotypic changes in affected cells.

1.1   Oral Potential Malignant Disorders

1.1.1   Overview

The word "precancer" was first introduced in 1805; and in 1978, The World Health Organization (WHO) classified "precancer" into "lesions" and "conditions". The worldwide prevalence rate of OPMDs ranges from 1% to 5% (Warnakulasuriya et al., 2021). Till today, the working group of WHO has defined OPMDs as "*any oral mucosal abnormality that is associated with a statistically increased risk of developing oral cancer.*" (Warnakulasuriya et al., 2021). According to the Journal of Oral and Maxillofacial Pathology, an oral lesion is recognized as "precancerous" is if fits the following criteria:

   i.   In longitudinal studies, areas of tissue with certain alterations in clinical appearances identified at the first assessment as 'precancerous' have undergone malignant change during follow-up.

  ii.   Some of these alterations, particularly red and white patches, are seen to co-exist at the margins of overt OSCCs.

 iii.   A proportion of these may share morphological and cytological changes observed in epithelial malignancies, but without frank invasion.

iv.   Some of the chromosomal, genomic and molecular alterations found in clearly invasive oral cancers are detected in these presumptive 'precancer' or 'premalignant' phases (Sarode et al., 2014).

Oral potentially malignant disorder is a collective term used to describe asymptomatic clinical lesions in the oral cavity. Patients diagnosed with OPMDs are at higher risks to develop oral cancer in their mouth during their lifetime, but not all lesions will undergo transformation since cancer is a multistep process. The most common types of OPMD include leukoplakia, which presents as white patches on the oral mucosa and is the most frequent OPMD. Erythroplakia is less common, appears as a well-defined, red, velvety patch on the oral mucosa. It is often asymptomatic but is more concerning than leukoplakia due to its higher risk of malignant transformation. Oral submucous fibrosis (OSMF), prevalent in South Asian populations and linked to betel quid use, causes stiffening of the oral mucosa and the mucosa may appear pale and fibrotic, with a loss of elasticity. Oral lichen planus, particularly its erosive form, presents with painful ulcers and erythematous areas. It is a chronic inflammatory condition associated with increased cancer risk. Actinic cheilitis is often due to chronic sun exposure of the lips, clinically presents as persistent dryness, scaling, crusting, and ulceration; with an increased chance of developing squamous cell carcinoma. Proliferative verrucous leukoplakia (PVL) is a more aggressive form of leukoplakia that is multifocal and persistent. Lastly, dysplasia of the oral mucosa, often graded as mild, moderate, or severe, represents a spectrum of premalignant changes, with the risk of cancer increasing with the severity. These OPMDs require regular monitoring and appropriate intervention to prevent progression to oral cancer (Ranganathan & Kavitha, 2019) (**Appx. Table 1**). These are chronic conditions with a high risk of transformation to OSCC, whose prevalences are influenced by environmental factors. Leukoplakia and erythroplakia are associated with alcohol and tobacco smoking, and are common in Western countries (Mello et al., 2018). Oral submucous fibrosis is more commonly reported in areca nut users in Asian countries (Mello et al., 2018). OPMDs are clinical diagnoses made based on the lesions' clinical appearances; the histological diagnoses of OPMDs may be hyperplasia, hyperkeratosis, or oral epithelial dysplasia (Ho et al., 2009).

A meta-analysis done by Iocca summarized the malignant transformation rate of the different OPMD subtypes using 92 studies with over 37,000 patients. The overall malignant transformation rate was 7.9% with the transformation rates for specific subgroups ranging from less than 2% to almost 50% (Iocca et al., 2020). Oral erythroplakia and proliferative verrucous leukoplakia had the highest malignant transformation value of 33.1% and 49.5%, respectively (Iocca et al., 2020). Given these high risks, it is suggested that these lesions should be completely removed and that the patients have intensive follow-up.

1.1.2　Oral Epithelial Hyperplasia and Dysplasia

The NCI dictionary defines hyperplasia as "an increase in the number of cells in an organ or tissue. These cells appear normal under a microscope. They are not cancer but may become cancer". Oral epithelial hyperplasia comprises a broad spectrum of histological changes which are characterized by cellular and structural abnormalities and preservation of the basement membrane (Žerdoner, 2003). Often OSCC begins as a simple epithelial hyperplasia which then progresses to dysplasia, followed by more severe dysplastic changes and genetic alterations until malignant transformation occurs. Oral epithelial dysplasia (OED) is considered a potential histologic precursor of subsequent squamous cell cancer (Pritzker et al., 2021). OED is characterized by cytological and architectural alterations reflecting the loss of normal maturation and stratification of the surface epithelium. Dysplasia is a term used to describe abnormal cell growth; it can range from mild to severe depending on the extent of involvement of the epithelium by dysplastic cells using the WHO three-tier grading system (**Table 1.1**). There is also a 2-tier grading system that separates dysplasia into high and low grades. This system makes it easier for clinicians and non-specialists to stratify patients to treatment regimes. Additionally, this system better predicts for malignant transformation in moderate dysplasia lesions (Odell et al., 2021). Regardless of the grading system, the overall evidence indicates a positive correlation between the likelihood and time to malignant transformation with increasing severity of dysplasia (Lorini et al., 2021). In the oral cavity, dysplasia is identified by cellular atypia and loss of normal cellular maturation and stratification (**Table 1.2**).

However, dysplasia is not a permanent event and it can revert to normal under the ideal environment (Ranganathan & Kavitha, 2019).

Table 1.1: The 3-tier grading includes, mild, moderate, and severe; indicated by cytological changes in the epithelial layers.

Grades	Epithelial Layers
MILD	basal one-third
MODERATE	middle one-third
SEVERE	Upper one-third

Table 1.2: World Health Organization criteria for diagnosing epithelial dysplasia (2017).

Architectural changes	Cellular changes
Irregular epithelial stratification	Abnormal variation in nuclear size
Loss of polarity of basal cells	Abnormal variation in nuclear shape
Drop-shaped rete ridges	Abnormal variation in cell size
Increased number of mitotic figures	Abnormal variation in cell shape
Abnormally superficial mitotic figures	Increased nuclear-cytoplasmic ratio
Premature keratinization in single cells	Atypical mitotic figures
Keratin pearls within rete ridges	Increased number and size of nucleoli

Different subtypes of OPMD lesions are associated with a higher chance of developing OED. For example, the prevalence of severe oral epithelial dysplasia and oral carcinoma in situ of red lesions is much higher than white lesions (Warnakulasuriya et al., 2021). Severe dysplasia is considered the "gold standard" predictor for OPMD progression. Areas with severe dysplasia are thought to have the highest probability of subsequent cancer development (Dionne et al., 2014). It is possible for OSCC to arise from clinically normal-looking oral mucosa. These cases often have a high level of genetic aberrations including activation of oncogenes and deletion of tumor suppressor genes.

1.1.3   Progression to Cancer

Oral potentially malignant disorders are characterized by tissue alterations associated with increased oncogenic potential compared to normal oral mucosa. The rate of

malignant transformation varies between OPMD subtypes (Mello et al., 2020). The most common cancer arising from OPMD is oral squamous cell carcinoma. The rate of cancer progression has been documented to be as high as 49.5% over a follow-up period ranging from 1 to 20 years (Iocca et al., 2020). Predicting the risk of transformation accurately remains a significant challenge in the field of oral healthcare. The lesion appearance along with the presence and grade of epithelial dysplasia have been widely used to estimate the risk of cancer development. A meta-analysis in 2009 completed by Mehannah et al. showed that high grade dysplastic lesions are more likely to progress to cancer (Mehanna et al., 2009). Four risk factors were evaluated, including sex, cigarette smoking, alcohol, and site of lesions. The relative risk (RR=1.87) for malignant transformation appears to be higher for lesions of the tongue. Smoking and alcohol consumption after the initial diagnosis of dysplasia did not have a significant effect on the progression to cancer in this particular study (Mehanna et al., 2009). Male patients aged between 50 and 69 tended to have the highest transformation rate for all OPMD subtypes (Chuang et al., 2018). It is not surprising that progression to cancer from OPMD is strongly associated with specific lifestyle behaviours. For example, the risk of malignant transformation of OPMD is elevated in patients with oral submucous fibrosis and verrucous hyperplasia secondary to alcohol consumption and betel quid chewing (Chuang et al., 2018).

1.1.4   Treatment and Management

The method of treatment for OPMD depends on the estimated risk for potential malignant transformation. Lesions with a lower risk of transformation are often left untouched and patients are given close clinical follow-up (Mello et al., 2020). Although the likelihood of malignant transformation is the greatest within the first two years, patients should remain on regular follow-up and any other clinically suspicious lesions should be biopsied (Dionne et al., 2014). It is advised that patients with low-risk lesions reduce smoking and their exposure to other environmental risk factors. Treatment for high-risk lesions is generally more invasive in order to minimize the risk of future potential cancer development, including complete surgical excision or laser ablation (Mello et al., 2020). Although there are no specific guidelines for the best management of oral epithelial

dysplasia, surgery and surveillance remain to be the most popular methods. However, the effectiveness of preventing transformation following treatment is still unknown.

One study examined the clinical outcome and follow-up of 100 patients who all underwent standardised interventional laser surgery to remove dysplastic lesion. The laser surgery was carried out by the same surgeon with a standardized protocol, and patients were reviewed between 1 to 12-month intervals based on the severity of the disease (Diajil et al., 2013). Patients were reviewed with a mean follow-up time of 5 years following the laser excision procedure and no patients in this study exhibited persistent disease following surgery. Out of the 100 patients in this cohort, 62 patients were completely disease-free following laser surgery; 17 developed disease at the same site of the excision, while 14 developed disease at a distant site; only 7 patients developed OSCC following laser excision within the 5-year follow-up period (Diajil et al., 2013). The rate of transformation was less than 10%, suggesting that interventional laser excision of dysplastic lesions is a promising method to minimize the risk of malignant transformation.

Although surgical excision is frequently carried out to reduce the risk of malignant transformation, there are a few criteria that need to be met prior to surgical intervention. The following Figure 1.1 explains the Liverpool algorithm developed by Field et al. and published in *Oral Oncology* (Field et al., 2015). The principles underlying the Liverpool Multidisciplinary approach to oral epithelial dysplasia have formed the basis of patient management and have resulted in excellent patient survival outcomes (Field et al., 2015).
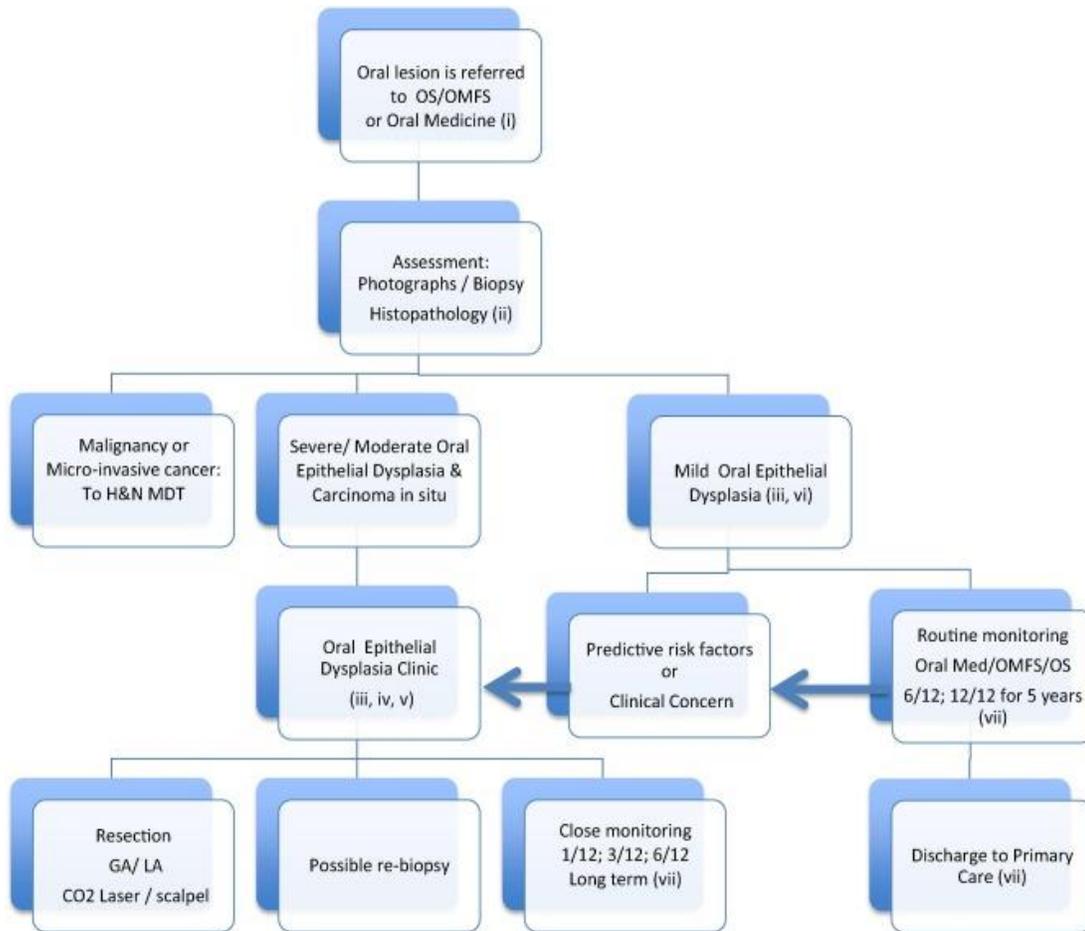
Figure 1.1: The Liverpool management algorithm for oral epithelial dysplasia.
According to the Liverpool protocol, incisional biopsy is preformed, and multiple
biopsies are taken for lesions larger than 200 mm^2. All patients with moderate and severe
dysplasia, carcinoma in situ (CIS) and mild dysplasia that have predictive risk factors are
offered surgical excision with the aim of achieving clear margins. In cases of extensive
lesions where complete excision cannot be arranged, patients are offered active specialist
surveillance with immediate re-biopsy in the event of clinical change. Moderate or severe
OED patients are offered long-term surveillance on the multidisciplinary clinic,
regardless of any surgical intervention; while mild OED patients are initially followed up
on the multidisciplinary clinic and then offered long-term review in the routine Oral
Medicine or Oral and Maxillofacial Surgery clinic (Field et al., 2015).

During the COVID19 pandemic, many in-person clinics were suspended, and clinicians

were forced to find alternative ways to provide patient support; the dramatic reduction in

face-to-face non-urgent services in the health care system was implemented to reduce the

risk of cross-infections among both health care providers and patients. Remote

consultations were offered during pandemic outbreaks, while face-to-face assessments were provided on a case-by-case basis (McCarthy et al., 2021). A study done in UK collected expert-opinion for the management of OED during the current and future pandemic outbreaks. Although the experts in this study were all clinicians in Oral and Maxillofacial Surgery and Oral Medicine in the UK, the consensus statements regarding the monitoring of patients with OED during pandemic outbreaks could be used on a global level. Patients with OED have a significant higher risk of developing oral cancer compared to the general population. Thus, close monitoring and appropriate surgical interventions should be considered in a timely manner. Urgent face-to-face consultations are only offered to patients reporting signs or symptoms of oral cancer, or if there are clinical changes suspicious for disease progression (McCarthy et al., 2021).

There have been several studies investigating possible alternative approaches to treat OPMDs without surgery. Retinoid treatment restores the expression of retinoic acid receptors and regulates the growth and squamous differentiation of the epithelium. However, relapse is common following treatment discontinuance suggesting that retinoids only delay carcinogenesis in the oral cavity (Dionne et al., 2014). Retinoids are the most extensively studied agents for preventing oral cancer; administration of retinoic acid for only 3 months yields a clinical response rate of 67% versus 10% for placebo, but toxicities remain a concern (Papadimitrakopoulou et al., 1997). Some natural agents were also considered as potential treatment for OPMDs including Bowman-Birk inhibitor concentrate, green tea extract, and curcumin; these agents exhibit high antioxidant activity and can protect cells from DNA damage from reactive oxygen species (Dionne et al., 2014). However, there is no evidence these natural agents can prevent oral cancer

development. One study investigated the efficacy of adding oncolytic adenovirus into mouthwash to target cancer cells with defects in the p53-dependent signaling pathways. Patients were divided into three treatment cohorts with the mouthwash containing various levels of viral load. Biopsies were taken after the 12th cycle (48 weeks mark). No significant difference of recurrence rates of severe dysplasia weas detected between patient cohorts (Chau et al., 2017).

1.2   Clinical Diagnosis of OPMDs

The diagnostic accuracy of clinicians based on visual disease presentation is critical in the early detection and management of patients with oral OPMDs. Oral lesions such as proliferative verrucous leukoplakia has been identified as the clinical condition that have the highest risk of transforming into OSCC (Iocca et al., 2020). Accurate and timely visual diagnosis of OPMD by clinicians can significantly impact the prognosis by facilitating early intervention and treatment. The ability to correctly identify and monitor these potentially malignant lesions is essential to prevent the progression to carcinoma. Ensuring high diagnostic accuracy of the visual assessment of OPMDs not only enhances patient outcomes but also reduces the need for more invasive diagnostic procedures. Therefore, evaluating and improving the visual diagnostic skills of clinicians is of the utmost importance in fighting against OSCC, as it should improve early detection and management of this serious and life-threatening disease.

1.2.1   Diagnostic Methods

Diagnostic tests for oral cancer and oral potentially malignant disorders encompass a comprehensive array of techniques to ensure precise diagnosis and effective management. The initial assessment involves clinical examination, where visual inspection and palpation are used to identify suspicious lesions (Jäwert et al., 2021). Vital staining with toluidine blue and additional tumoral markers may be employed to indentify areas of dysplasia or malignancy. Advanced imaging techniques, such as FDG-PET

scans, offer high-resolution images of soft tissues and detailed assessments of tumor spread and involvement of surrounding structures (Carreras-Torras & Gay-Escoda, 2015).

Biopsy techniques are the golden standard for achieving a definitive diagnosis. The following list include the common histological techniques to remove tissue from the oral cavity. Incisional biopsies involve removing a section of the lesion for histopathological examination and is particularly useful for larger or more complex lesions. Excisional biopsies involve the removal of the entire lesion and are often employed for smaller lesions, providing both diagnostic and therapeutic benefits. Fine needle aspiration biopsy (FNAB) offers a minimally invasive option to aspirate cells from a lesion or lymph node for cytological analysis and is particularly valuable for assessing deep-seated lesions and lymph nodes. However, screening by taking biopsies of clinically suspect oral tissue is often avoided by many clinicians since this often causes serious discomfort to the patient and is not suitable for repeated sampling at multiple sites (Carreras-Torras & Gay-Escoda, 2015).

1.2.2   Visual diagnostic Accuracy

The diagnostic process for detecting oral cancers and oral potentially malignant disorders in patients with visible lesions typically begins at that their dentist's office with a comprehensive clinical history and conventional oral examination (COE). The examination involves a thorough head and neck evaluation, visual inspection of the patient's oral mucosa under office lighting, and palpation. Features concerning for the presence of oral cancer or dysplasia include a non-homogeneous appearance of the oral mucosa, changes in oral mucosa texture and color, the presence of non-healing ulcer or tissue tethering (Essat et al., 2022). However, COE heavily relies on the clinician's expertise, and it can be difficult to visually identify dysplastic lesions. Despite the development of various adjuncts to COE, there is no consensus on their reliability, with COE followed by biopsy remaining the gold standard for diagnosis (Epstein et al., 2008).

A recent meta-analysis assessed the diagnostic accuracy of COE compared to incisional or excisional biopsy for detecting malignant and dysplastic lesions in patients with visible

lesions. Analysis across 14 studies revealed COE's sensitivity and specificity for diagnosing dysplastic or malignant oral lesions were 71% and 85%, respectively (Essat et al., 2022). One of the studies included in this meta-analysis performed by Brocklehurst et al. evaluated the diagnostic accuracy of various dental professionals, including primary care dentists, dental hygienists, hospital-based dentists, and dental nurses. They reported median sensitivity and specificity for each group. Hospital-based dentists achieved the highest median sensitivity (90%) and specificity (76%), followed by primary care dentists with 81% sensitivity and 73% specificity. Hygienists/therapists showed 77% sensitivity and 69% specificity, while dental nurses had the lowest performance with 68% sensitivity and 59% specificity. This indicated the overall accuracy of identifying oral lesions was greater for clinicians with more experience and medical training. Furthermore, significant variability within each group underscores the need for continuous training to enhance clinical diagnostic consistency across all dental professionals (Brocklehurst et al., 2015). A clinician's skill in forming an accurate provisional diagnosis for a specific oral mucosal lesion by integrating all relevant clinical data is crucial. This accuracy enables the clinician to take appropriate steps, such as conducting a biopsy, for further analysis when needed. Conventional oral examination for the identification of oral mucosal lesions can be sufficient in most non-high-risk patients with no significant clinical history. However, utilizing adjunctive tools such as chemiluminescent light and toluidine blue staining can further enhance the identification of suspicious lesions by improving the visualization of lesion margins and reduce false positives (Epstein et al., 2008).

1.2.3   Challenges in Diagnosis

The clinical oral examination serves as the gold standard for the initial identification of oral lesions that may contain dysplasia or early-stage oral squamous cell carcinoma. This screening method significantly reduces patients' discomfort since it is less invasive than taking biopsy samples from the oral cavity but there are limitations to the standard COE. First, early-stage OSCC and dysplastic lesions may manifest clinically as OPMDs (erythroplakia, leukoplakia, lichen planus, etc.). All may present as a mixture of red and white lesions, making it difficult to distinguish and evaluate by visual inspection (Epstein et al., 2012). According to Cancer Statistics, more than 30% of patients with OSCC and

oropharyngeal cancer had undergone oral cancer screening in the previous 3 years before receiving OSCC diagnosis (Jemal et al., 2003). Often the lesions may go unnoticed for years before a diagnosis is made. To add to the challenges in diagnosing oral lesions visually, oral carcinomas are often asymptomatic in the early stages of their development or appear in areas where lesions are not readily visible (González-Moles et al., 2022). Another challenge is the knowledge and experience of primary healthcare providers. Diagnostic delays often stem from local healthcare communities' lack of awareness about oral cancer symptoms. Although the primary healthcare provider (primarily dentists and family physicians) must know the symptoms and signs of the early onset of oral cancer; many articles have reported a lack of knowledge in primary providers regarding the symptoms of oral cancer (Guggenheimer et al., 1989). Additionally, patient social and cultural circumstances may play a role in these diagnostic challenges. For instance, New York City has one of the highest incidence and mortality rates of oral and pharyngeal cancer for Hispanics (Cruz et al., 2007). Cigarette smoking and alcohol consumption are the main behavioral risk factors for oral malignancies within the Hispanic community. The initial smoking age reported in this study was as early as 7 years old (Cruz et al., 2007). The lack of public awareness regarding the signs, symptoms, and risk factors associated with oral cancers poses another significant barrier to early detection and prevention. From this public survey, very few individuals have practiced preventive medical or dental care. Most admitted delaying a visit to the doctor or dentist due to lack of health insurance, and communication difficulties with the providers who don't speak their language (Cruz et al., 2007). This study highlights the importance of increasing public awareness about these risk factors in order to improve early detection and disease outcomes.

1.3   Oral Squamous Cell Carcinoma

1.3.1   Epidemiology

Oral squamous cell carcinoma is the most common epithelial malignancy affecting the oral cavity in the head and neck region. Worldwide, head and neck cancer accounts for up to 4% of all cancer cases and it is estimated that over 90% of all oral neoplasms are

OSCC (Markopoulos, 2012). These tumors can arise from various locations in the oral cavity, including the tongue, floor of mouth, buccal surface, alveolar surface, and hard palate. The most common site is the tongue region and it is often associated with aggressive tumors with the worst prognosis (Ng et al., 2017). As with many other types of cancers, the majority of OSCC cases are detected at an advanced stage, with a low 5-year survival rate ranging from 40-50% (Ali, 2022). Despite the advances in therapeutic approaches and early detection technologies, morbidity and mortality rates of OSCC have not improved over the past 30 years (Markopoulos, 2012). Data from the Global Cancer Observatory shows that the annual incidence of OSCC in 2020 was 377,713 cases worldwide, while the five-year prevalence (2015 – 2020) of OSCC approached nearly one million cases (Sung et al., 2021).

Oral squamous cell carcinoma more frequently affects older men than women, with over 60% cases diagnosed in the male population (Feller & Lemmer, 2012). This trend is largely attributed to the participation in high-risk habits, including alcohol and tobacco consumption. The probability of developing OSCC increases with the period of exposure to risk factors, and increasing in age further contributes to cancer development due to mutagenic and epigenetic changes within the genome (Feller & Lemmer, 2012). As society has become more aware of the detrimental effects of tobacco use, there has been a decline in the number of tobacco-associated cancers in countries where anti-tobacco campaigns have been implemented. However, recent publications have reported that there is an increase in the number OSCC in younger patients, women, and individuals with no prior exposure to the risk factors mentioned above but are at higher risk for human papillomavirus (HPV) infection (Ng et al., 2017).

1.3.2 Risk Factors

Oral squamous cell carcinoma is a significant global health concern. In addition to OPMDs, various other risk factors also contribute to the development of this cancer. Tobacco (chewing and smoking) and alcohol consumption are two of the most common risk factors for oral cancer and work synergistically to increase the risk by up to 35% (Pulte & Brenner, 2010). Tobacco is responsible for over eight million deaths annually, and there are approximately 1.3 billion tobacco users worldwide and over 80% of these

are based in low- and middle-income countries. Cigarette smoking is a well-established risk factor for periodontal disease and oral cancer. It alters the host immune response by increasing the release of inflammatory mediators leading to oxidative stress and cellular damage in the oral cavity. (Zhang et al., 2019). Smokers are 7 to 10 times more likely to develop oral cancer and 3 times more likely to develop a second primary cancer compared to non-smokers (Warnakulasuriya et al., 2005). Tobacco smoke contains several chemical carcinogens (nitrosamines, polycyclic aromatic hydrocarbons, volatile aldehydes, and aromatic amines) to humans (Khariwala et al., 2012). Not only active smokers are in contact with tobacco smoke, but countless individuals are exposed to tobacco smoke indirectly. According to data from 192 countries, 33% of male, 35% of female and 40% of children are exposed second-hand tobacco smoke (Öberg et al., 2011). Exposure to second-hand smoke (SHS) increases the risk of oral cancer in humans, and there is a positive correlation between the length of exposure and the overall risk of oral cancer development (Mariano et al., 2022). Lastly, chewing tobacco can also contribute to the development of OSCC since smokeless tobacco is predominantly taken orally, resulting in prolonged exposure and absorption of chemicals by the oral mucosa (Mello et al., 2019). Like tobacco use, alcohol consumption is also responsible for millions of deaths annually. Alcohol plays a direct causal role in the development of oral cancer by activating oncogenes and promotes the initiation of oral cancer (Madani et al., 2014). Reactive oxygen species are one of the major by-products of alcohol, which causes DNA damage and the suppression of DNA damage repair mechanisms within the nucleus of rapidly dividing cells (Mello et al., 2019). A synergistic effect of tobacco and alcohol consumption has been shown in several studies regarding the developing and progression of oral cancer. Smoking and alcohol Sumption are the most important known risks for developing oral cancer and are the cause of 30% of all oral cancers. Alcohol increases the permeability of the oral mucosa and allows more carcinogens in the tobacco smoke to enter the epithelium causing oncogenic effects (Madani et al., 2014).

Cancer is a multistep process at the molecular level, resulting from the sequential accumulation of the appropriate genetic modifications. In OSCC, the genetic alterations include oncogene activation and tumor suppressor gene (TSG) inactivation, leading to unregulated cell cycle progression and uncontrolled cell proliferation (Mehrotra et al.,

2004). For example, p53 is a TSG that regulates apoptosis. The deregulation of apoptosis-related genes allows cells with DNA repair defects to continue to proliferate, causing further genomic instability and aids in successful carcinogenesis (Mehrotra & Yadav, 2006). Studies have shown that tobacco increases oxidative stress and causes abnormal expression of p53 (TSG) and other genes in oral epithelial cells (Jiang et al., 2019). The MAPK pathway is an important signaling pathway involved in cell proliferation and survival. Activation of RAS leads to the phosphorylation and activation of RAF kinase. BRAF is strongly associated with mutations in cancer (Davies et al., 2002). When the pathway is triggered with all the required signals, transcription protein ERK1/2 will enter the cell nucleus and push transcriptional programs related to cell proliferation forward (Pouysségur et al., 2002).

1.3.3   Prognosis and Management

According to the American Cancer Society, the 5-year survival rate for localized OSCC (no lymph node involvement) is over 75%. However, lymph node metastasis decreases survival rate by about half, and later stage metastasis to the lungs decreases survival rate even more. Despite numerous advances in management and treatment, the 5-year survival of OSCC has remined close to 50% for the last 50 years (Lingen et al., 2008). Oral cancer is usually diagnosed at a late or advanced stage, even often at initial clinical presentation. The death rate from oral cancer and the negative consequences of the disease are also not declining compared to other more common cancers (Brocklehurst et al., 2013). Over 60% of patients present with advanced cancers, and it is worth noting that the stage at diagnosis and additional therapeutic procedures strongly impacts survival and quality of life of the patient (Brocklehurst et al., 2013). The TNM system classifies stage I and II oral cancer as early stage, while stage III and IV as advanced stage. Several studies have demonstrated that the 5-year survival rate is better for patient with stage I disease and that the survival rate for patients with stage IV disease is almost 50% lower that of patients with stage I disease (Le Campion et al., 2017).

The treatment of oral cancer is determined according to the stage of the disease at diagnosis. Therefore, staging and grading based on physical and histological examination need to be accurate (Omura, 2014). Survival rates and the quality of life are considered

when selecting treatment, which should be tailored individually to the patient's needs. Early stage cancer is primarily managed with surgery alone with low morbidity; whereas advanced-stage cancer requires multidisciplinary treatment, including radiation, surgery, and possibly other pharmaceutical therapies (Omura, 2014). The goal for surgical resection is to completely remove tumor tissue and limit tumor growth. In many cases, complete surgical removal of OSCC with adequate margins would result in unacceptable aesthetic and functional morbidities (McMahon et al., 2003). Often skin grafting is required following the primary tumor resection to restore oral function and an acceptable cosmetic appearance (Omura, 2014). Although the course of OSCC is unpredictable, the primary tumor TNM stage is a good predictor for survival rate. Small primary tumor without any regional lymph node involvement or distant metastasis have a five-year survival rate close to 90% (Feller & Lemmer, 2012). Patients with OSCC have a higher chance of developing additional malignancies, since most carcinogens to the oral cavity also affect the entire upper respiratory tract. The risk of developing a second carcinoma is significantly higher 10 years following the initial OSCC onset. As such patients are advised to undergo frequent examination for up to fifteen years (Saikawa et al., 1991). Recently, targeted therapeutic drugs have become increasingly used for the treatment of malignant tumors that express specific biomarkers. These highly selective drugs have low toxicities and high therapeutic indexes. By targeting molecules such as, EGFR, VEGF, and various kinases, these drugs significantly improve the survival rate of cancer patients (L. Liu et al., 2019). Two immunotherapy drugs have been applied clinically for the treatment of patients with head and neck squamous cell carcinoma (HNSCC). These include monoclonal antibodies that interfere with the receptor action on the tumor cell surface, and tyrosine kinase inhibitors that enter the tumor cells and bind to cytoplasmic receptors and block growth signal transduction (L. Liu et al., 2019). For example, Cetuximab, an anti-EGFR antibody that binds to the ligand-binding domain of EGFR and prevents dimerization, internalisation and autophosphorylation of the receptor (Goerner et al., 2010, p.). EGFR-targeting approaches have already been approved for treatment in advanced HNSCC, but many molecular targeted drugs are still under evaluation (Goerner et al., 2010, p.).

1.4 Molecular Changes in OPMDs

1.4.1 MAPK Pathway

Mitogen activated protein kinases (MAPK) pathway is a central signaling pathway that regulate important cellular processes. This is an extremely complex pathway involving many components. Each signaling cascade typically consists of 3-5 tiers of kinases, and each cascade is initiated by specific extracellular cues leading to the successive activation of the target proteins (**Figure 1.2**) (Morrison, 2012). Out of the four pathways, the ERK1/2 pathway is the most well studied in the field of oncology. The pathway involves three key elements: RAS, RAF, and MEK. Upon the attachment of growth factors and hormones to cell surface receptors, there is an increase in the cellular levels of RAS-GTP, which facilitates the activation of the downstream kinases. This activation allows RAS-GTP to attract RAF kinases from the cytosol to the plasma membrane leading to activation of RAF, which in turn, activates MEK. This activation chain continues as MEK activates ERK, leading to the phosphorylation of various proteins within the cell. Activated ERK1/2 drives changes in cell movement and gene expression that support cell proliferation, differentiation, survival, and the formation of new blood vessels (Q. Peng et al., 2017).
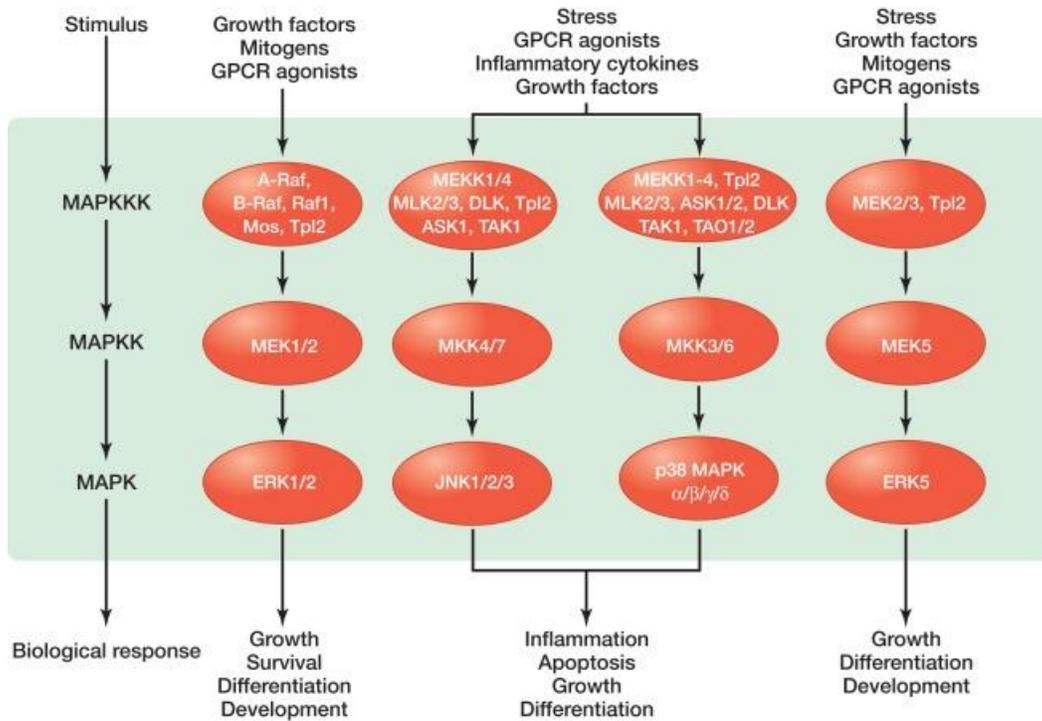
Figure 1.2: MAPK Pathway Overview.
MAPK modules containing three sequentially activated protein kinases are key components of a series of vital signal transduction pathways that regulate processes such as cell proliferation, differentiation, and cell death. The MAPK signaling pathway is comprised of four sub-pathways: the ERK1/2 sub-pathway, the JNK sub-pathway, the p38 sub-pathway and the ERK5 sub-pathway (Morrison, 2012).

1.4.2   Wnt/ β-catenin Pathway

The Wnt/β-catenin pathway comprises a family of proteins that play critical roles in embryonic development and adult tissue homeostasis. The canonical pathway involves Wnt proteins binding to their membrane components followed by β-catenin translocating to the nucleus where it associates with T-cell factor (TCF) to form a transcription factor. This functional transcription factor transcribes target genes involved in tumor progression, invasion, and metastasis (**Figure 1.4**). The canonical Wnt pathway mainly controls cell proliferation; whereas the noncanonical Wnt pathways regulates cell polarity and migration, and it is independent of β-catenin-T-cell factor/lymphoid enhancer-binding factor (J. Liu et al., 2022). Previous studies have examined the inactivation of APC or Axin, along with other mutations and epigenetic changes in the pathway. In this study we are focusing on the expression pattern of β-catenin and the involvement of the Wnt/β-catenin pathway at the different stages of oral carcinogenesis.

1.4.2.1   β-catenin

β-catenin is a highly evolutionary conserved molecule that serves critical roles in developmental and homeostatic processes. More specifically, β-catenin is an integral structural component of cadherin-based adherens junctions and the key molecule of the canonical Wnt signaling pathway in the nucleus. Modification of the molecular structure or change in expression of β-catenin is often associated with disease, deregulated growth, and cancer (He et al., 2004). β-catenin is located at the cell membrane, and in the cytoplasm and/or nucleus. At the cell membrane, it is bound to the cytoplasmic domain of E-cadherin and is essential for the structural organization and function of cadherins by linking to the actin cytoskeleton. Cytosolic β-catenin is subsequently degraded during the inaction of Wnt pathway signaling or translocated to the nucleus when the Wnt pathway is active (López-Knowles et al., 2010). Investigators have shown that, in oral dysplasia, β-catenin is detected at the nucleus while, in oral carcinoma, this protein is mostly accumulated in the cytoplasm with minimal detection in the nucleus (Reyes et al., 2020). In this study, cytosolic β-catenin expression was examined using IHC.
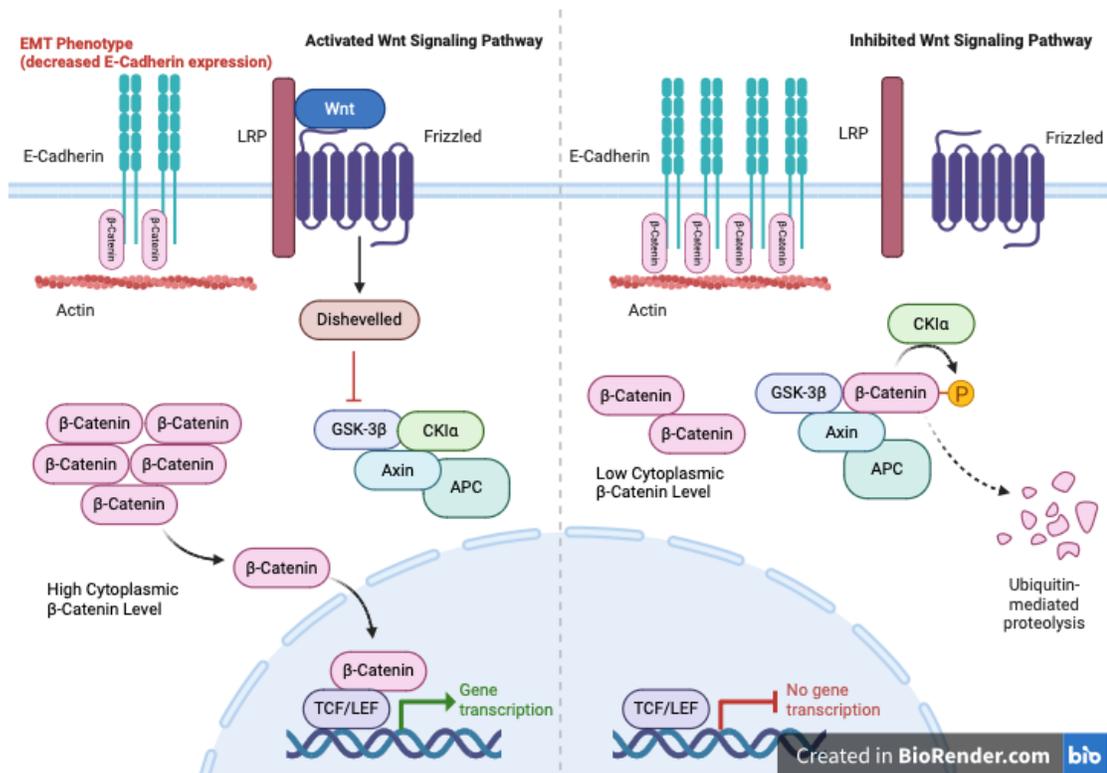
Figure 1.3: E-cadherin Sequesters β-catenin and Prevent Downstream Gene Activation and Proteolysis of β-catenin Due to Inactive Wnt Signaling Pathway. The Wnt β-catenin Signaling Pathway is a cellular pathway essential for signal transduction. Wnt proteins and β-catenin transmit signals from the cell surface, through frizzled and LRP receptors, to the nucleus, modulating gene expression. When the Wnt signaling pathway is inactive, free β-catenin is degraded by a complex including glycogen synthase kinase (GSK), adenomatous polyposis coli (APC), and Axin, which phosphorylate β-catenin (PP). The binding of Wnt to Frizzled (Frz) receptors activates Wnt signaling, and disheveled (Dsh) inhibits β-catenin phosphorylation by GSK. This results in β-catenin accumulation in the nucleus, where it complexes with T cell factor (TCF) and transactivates target genes such as *Cyclin D1* and *Myc*, genes that play crucial roles in development, cell proliferation, and tissue homeostasis. EMT phenotype results in a decrease in the expression of the transmembrane protein E-cadherin, which affects the formation of β-catenin–cadherin complexes and the disruption of adherens junctions, that directly contributes to oncogenesis.

1.4.2.2　E-cadherin

The Cadherin family proteins (Transmembrane glycoproteins) are cell surface proteins of adherens junctions consisting of multiple subtypes. These transmembrane proteins participate in Ca^{2+}-dependent cell adhesion that is necessary to form solid tissues (X. Tian et al., 2011). E-cadherin is the one subtype that is mostly expressed in epithelial

cells. Other subclasses include neural (N-cadherin), placental (P-cadherin) and vascular endothelial cadherin (VE-cadherin). E-cadherin links the actin cytoskeleton of adjacent cells for the formation of epithelial tissue (Borcherding et al., 2018). The extracellular portion binds to proteins on the surface of adjacent cells, whereas the intracellular region interacts with regulatory proteins (X. Tian et al., 2011). The protein expression closely correlates with tumor metastasis. E-cadherin sequesters β-catenin to the plasma membrane at a 1:1 ratio (**Figure 1.3**) preventing β-catenin from participating in the Wnt signaling pathway and reduces the level of nuclear transcription (Borcherding et al., 2018).

In OSCC, loss of epithelial cell polarity is associated with tumor aggressiveness, metastasis, and poor prognosis. Many research studies have focused on signaling factors on the cell membrane, cell adhesion proteins, transcription factors that interact with DNA, and microRNA which control gene expression at a post-transcriptional level (**Figure 1.4**) (Vallina et al., 2021). The loss of E-cadherin expression in epithelial cells supports cell migration and invasion due to loss of cell-to-cell integrations and is observed in many types of cancer (Borcherding et al., 2018). In OPMDs, which are potentially malignant lesions, there may be reduced expression of E-cadherin rather than a complete loss. This disruption could indicate early signs of loss in cell adhesion and polarity, signaling a potential progression towards malignancy.
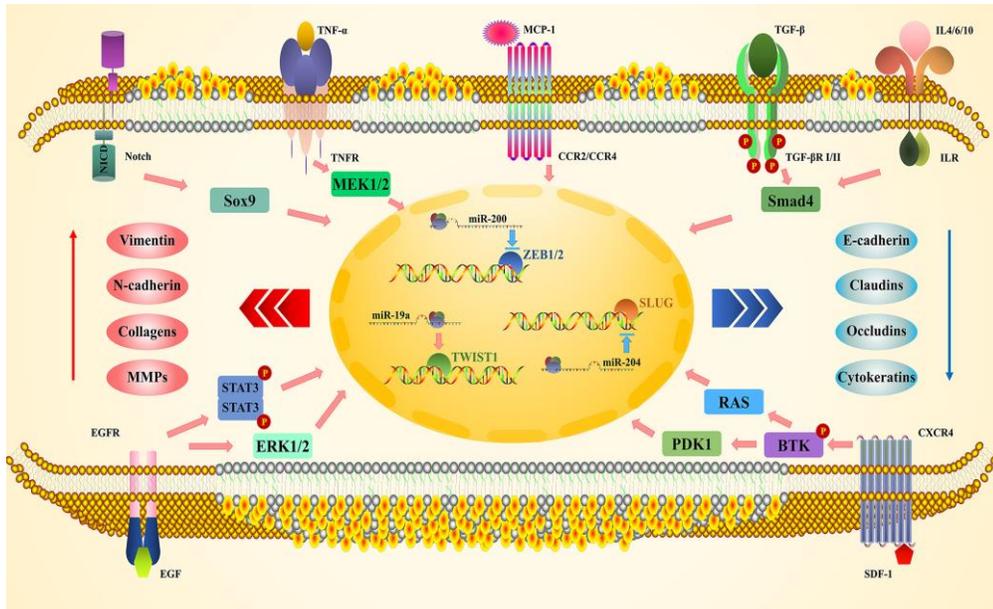
Figure 1.4: Cellular pathways contribute to the disruption of epithelial function.
Cytokines (TGF-β, TNF-α, etc.) bind to their corresponding cell membrane receptors and send stimulus signals to activate downstream intracellular pathways, such as MEK/ERK, STAT and other signaling cascades. These pathways collectively promote the loss of epithelial phenotype. As a result, transcription factors expression is altered inside the nucleus, leading to upregulation of mesenchymal molecules and a downregulation of cellular adhesion molecules (Vallina et al., 2021).

1.4.2.3   Vimentin

Vimentin is a cytoskeletal filamentous protein aiding in structural and functional support of the cell (Usman et al., 2021). Vimentin expression in cancers with epithelial origin is strongly associated with increased lymph node metastases, and poor overall survival rate (S. Liu et al., 2016). Figure 1.4 illustrates that increased vimentin expression is an element in the disruption of epithelial cell function, leading to the initial onset of metastasis and migration in many types of cancers. The levels of expression of both Vimentin and E-cadherin in colorectal carcinoma was investigated in relation to invasion and metastasis. Along with cancer progression level of Vimentin expression varied inversely with E-cadherin expression. Tumors with increased metastasis and invasion phenotypes had increased Vimentin levels (Niknami et al., 2020). Since vimentin is expressed in mesenchymal cells, examining its expression in the epithelial tissue of OPMDs could provide valuable insights.

1.4.3 Cell Cycle Progression

Cell cycle control is separated into two main events: the replication of genomic DNA and the segregation between daughter cells. The two events can be further separated into four main phases: G1, S, G2, and M (**Figure 1.5**). The control of the cell cycle is critical for the healthy functioning of organisms, and dysregulation of this process can lead to cancer. Cancer cells are characterized by their ability to proliferate uncontrollably, and it is assumed that cell cycle checkpoints need to be defective for a cell to become cancerous. For example, the DNA damage checkpoint is often compromised in cancer cells to allow continuous cell division in the presence of genetic errors (Matthews et al., 2022). Normal cells have several checkpoints (G1/S, G2/M, and the spindle assembly checkpoint during M phase) that ensure DNA is accurately replicated and divided. These checkpoints are regulated by various proteins, including cyclins and cyclin-dependent kinases (CDKs) which ensure that cells only proceed to the next phase of the cycle if conditions are right. In cancer cells, mutations often disrupt the function of these regulatory proteins, leading to the loss of checkpoint control and unregulated progression through the cell cycle (Matthews et al., 2022). Understanding the molecular underpinnings of the cell cycle is crucial for understanding cancer progression. Among the myriads of proteins involved in cell cycle control, Ki67, Minichromosome Maintenance Complex Component 2 (MCM2), and Geminin have emerged as pivotal players and will be examined in this study. These proteins not only serve as markers for cell proliferation but also play functional roles in the regulation of the cell cycle, rendering them significant in both the diagnosis and potential treatment of cancer.
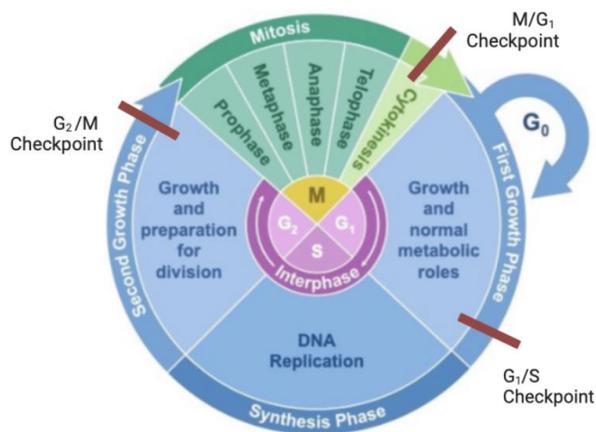
Figure 1.5: Cell Cycle Progression.
A complete cell cycle has G1 (growth 1), S (synthesis), G2 (growth 2), and M (mitosis). G1, S, and G2 are collective known as the interphase, while mitosis consists of prophase, metaphase, anaphase, telophase, and cytokinesis. G1 phase is a decision-making window for the cell where it can commit to initiate DNA replication and enter the cell cycle or stay in G1. During this time the cell can also exist the cell cycle and enter a non-proliferative state known as quiescence (G0 phase). DNA replication occurs during S phase. Following the completion of DNA replication, G2 is another decision window where the cell can commit to enter M phase by initiating chromatin condensation and the central alignment of chromosomes. During M phase, the mother cell is segregated into two daughter cells with the same genetic materials in both cells and reset the cell cycle to return to interphase.

1.4.3.1 Ki67

Ki-67 has been shown to be an excellent biomarker for the estimation of the growth fraction in both normal and malignant tissue (Birajdar et al., 2014). It was first identified as an antigen in Hodgkin lymphoma cell nuclei that was highly expressed in cycling (mitotically active) cells but down-regulated in resting quiescent cells (X. Sun & Kaufman, 2018). The localization of Ki67 is associated with different roles. For instance, throughout interphase, Ki67 plays a crucial role in the proper distribution of heterochromatin antigens within the cell and in linking heterochromatin to the nucleolus. During the process of mitosis, Ki67 is involved in the development of the perichromosomal layer (PCL), a ribonucleoprotein covering that envelops the compacted chromosomes that prevents the clumping of chromosomes during mitosis (X. Sun & Kaufman, 2018). Nuclear expression of Ki67 is detected in cells with mitotic activity. However, with its short half-life, IHC staining of cells that have gone through the

proliferative stage is limited and makes it an ideal biomarker to only detect actively dividing cells. A recent study by Vieira et. al showed that Ki67 gene suffers "over expression" in epithelial cells of pre-malignant and malignant oral lesions (Vieira et al., 2008).

1.4.3.2   MCM2

MCM2 is one of six members of the minichromosome maintenance (MCM) protein family and is a vital regulator in DNA replication. MCMs are a group of proteins involved in the initiation of DNA replication. The six conserved proteins (MCM2, MCM3, MCM4, MCM5, MCM6, and MCM7) form a hexameric ring-shaped complex which acts as a DNA helicase (Y. Sun et al., 2022). The hexameric complex is activated by cyclin-dependent kinases (CDK) and Cdc7/ASK kinase leading to recruitment of elongation factors, Cdc45, DNA polymerases and RPA, to the origin of replication. Recruitment of these factors to replication forks results in unwinding of the DNA helix and initiation of DNA synthesis (Dudderidge et al., 2005). Some studies have indicated that insufficient MCM level cause genomic instability and impaired cell cycle progression, leading to early-onset cancer (Y. Sun et al., 2022). In the absent of DNA helicase, attempted DNA replication can lead to double stranded breaks in the genome. The cell will try to repair these breaks via non-homologous end joining (NHEJ) which will increase the chance of possible oncogenic mutations. Other research suggests that MCM2 is a possible prognostic marker and therapeutic target since MCM2 is highly expressed in solid tumors and silenced in normal tissue (Yuan et al., 2022). Increased levels of MCM2 causes proliferation, migration, and invasion of tumor cells. Overall, MCM2 is a sensitive biomarker for cancer cell proliferation like Ki67. Suppressing MCM2 expression through miRNA and siRNA resulted in less proliferation of tumor cells in cancers, making it a potential target for chemotherapy (Y. Sun et al., 2022).

1.4.3.3   Geminin

Geminin was originally identified as an inhibitor of DNA replication and substrate of the anaphase-promoting complex (APC). Geminin expression is restricted to the S-G2-M phases of the cell cycle; it acts by binding Cdt1 and blocking MCM loading onto

chromatin to prevent re-replication (Dudderidge et al., 2005). Geminin expression is triggered at the G_1-S transition and the protein level rises through the rest of the cell cycle to reach a maximum during M phase (**Figure 1.6**). During mitosis, an ubiquitin ligase called the APC, is activated leading to the proteolysis of geminin (Wohlschlegel et al., 2002). If geminin is a tumor suppressor, then overexpression of geminin might suppress cancer cell proliferation. However, this is not the case. Geminin expression has been found to be overexpressed in several cancer cell lines, and it is linked with a poor clinical outcome in patients with renal cell carcinomas (Torres-Rendon et al., 2009). It is possible, since Geminin modulates the stability and activity of Cdt1 throughout the cell cycle, imbalance in Cdt1 and Geminin levels can lead to genomic instability and the progression of cancer. An increase in Geminin level has been observed throughout various types of cancer and is associated with adverse prognosis, poor overall survival rate and the development of distant metastases (Kushwaha et al., 2016).
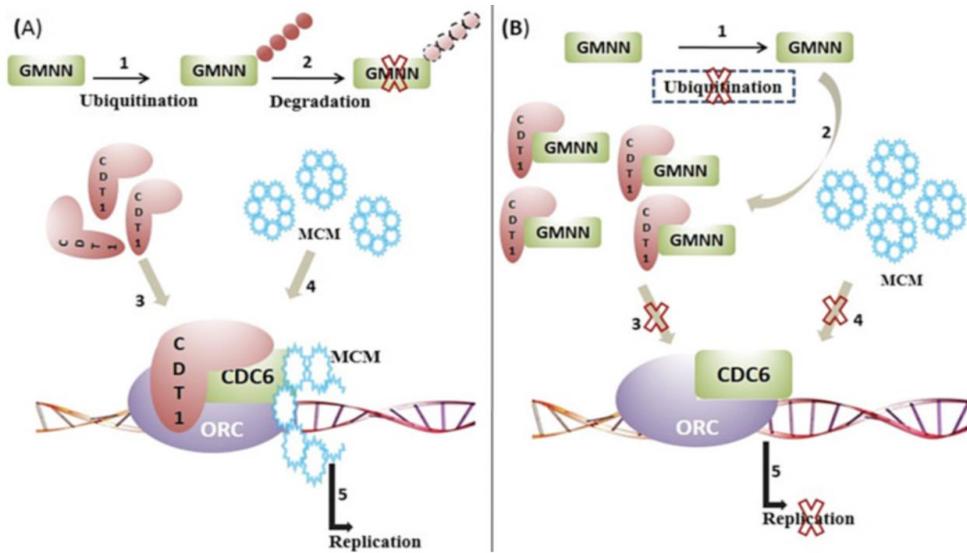
Figure 1.6: Licensing mechanism of replication initiation in absence and presence of GMNN protein.
During the late M phase and G1 phase of the cell cycle, Cdt1 is active and plays a crucial role in the formation of the pre-replicative complex (pre-RC) by recruiting the MCM complex to replication origins. The MCM complex is essential for the unwinding of DNA, which is a prerequisite for the initiation of DNA replication during the S phase. Geminin functions by directly binding to Cdt1, inhibiting its activity. When Geminin binds to Cdt1, it prevents Cdt1 from binding to CDC6 and recruiting the MCM complex to the replication origins, thereby blocking the assembly of new pre-RCs and ensuring that DNA replication does not occur more than once in a single cell cycle. Geminin levels are low during the M and early G1 phases, allowing Cdt1 to function and the MCM complex to be loaded onto DNA. As the cell progresses towards the S phase, Geminin levels increase, inhibiting Cdt1 and thus preventing re-replication. After DNA replication, Geminin is degraded, allowing the cell cycle to progress into the next phase (Kushwaha et al., 2016).

1.4.4   S100A7

The S100 protein family serves as calcium sensor proteins that regulate the function and distribution of specific target proteins (Eckert et al., 2004). Within the cell, S100A7 influences calcium homeostasis, energy metabolism, the regulation of cell proliferation, differentiation and apoptosis, production of reactive oxygen species and cytokines, as well as the improvement of skin barrier function (G. Peng et al., 2022). The calcium binding protein is mostly found in the human epidermis and is distributed in the cytoplasm and at the cell periphery in terminally differentiated keratinocytes. The expression of S100A7 is elevated in many epidermal inflammatory diseases as well as in

invasive skin cancers (Eckert et al., 2004). Recently, the first pancancer study conducted by Peng et. al. indicated that S100A7 expression was associated with the expression of DNA methyltransferase and mismatch repair genes in head and neck squamous cell carcinoma (G. Peng et al., 2022). Some evidence suggest that S100A7 may act as an inducer of Epithelial–mesenchymal transition (EMT) since S100A7 is an important ligand of RAGE, and RAGE increased MEK-EMT signaling and promoted migration, invasion and metastasis in cancer cells (T. Tian et al., 2017).

The change in S100A7 protein expression can be visualized using immunohistochemtry as has been determined in studies looking at OSCC, ovarian cancer, and cervival cancer (Das & Deshmukh, 2022). Using image analysis, S100A7 can better predict the clinical outcome for oral dysplasia cases as compared to using current histopathological techniques. This risk stratification is based on the size of the stained area and the intensity of staining. It is a promising prognostic and predictive biomarker for oral malignant transformation and cancer progression. Furthermore, S100A7 is also expressed by altered keratinocyte differentiation (Dey et al., 2016). Studies have shown that S100A7 expression is elevated in the early stage of oncogenesis, and the protein expression can be influenced by many environmental factors (Eckert et al., 2004). Straticyte™ is a test offered by Proteocyte AI that looks at biomarkers related to oral lesions. It measures the expression of S100A7 and provide a quantitative model for predicting the risk of transformation from premalignant lesions to cancer (Hwang et al., 2017). The Straticyte risk predicting process yields three groups: low, intermediate, and high-risk groups with minimal overlap. This method selects 2 cut-offs based on the Straticyte risk scores in comparison to the 3 grading histological method (**Table 1.3**).

Table 1.3: Straticyte Risk Groups

Cancer progression probability in 5 years	Risk group
Probability < 21%	Low
21% ≤ Probability < 55%	Intermediate
Probability ≥ 55%	High

1.5 Overall Research Goal

My overall research goal is to enhance the early detection of oral potentially malignant disorders through a comprehensive evaluation of clinical, histopathological, and

molecular diagnostic methods. This hypothesis-generating study aims to bridge the gap between clinical detection and molecular diagnostics, thereby enhancing the accuracy and reliability of early OPMD diagnosis.

1.6 Objectives of Thesis

Aim 1: Evaluate the concordance between clinical and histopathological detection of OPMDs.

- Determine the level of concordance between the clinical diagnoses made by clinicians based on visual examination and/or other diagnostic tools, and the definitive pathological diagnoses made through histopathological examination of biopsy samples.

Aim 2: Identify molecular biomarkers that are associated with clinical and pathological changes in OPMDs.

- Establish correlations between the presence of specific biomarkers and disease progression.

Chapter 2

2   Methods and Materials

My research involves the identification, quantification, and functional analysis of OPMD biomarkers, which can be genes, proteins, or other molecules that are involved in the progression of disease. The selection of appropriate methods and materials is critical for the reliability and validity of the study. We received 90 clinically diagnosed OPMD patient lesions (however due to the lack of tissue on some blocks, the final OPMD case samples were 82), along with 28 normal oral epithelium tissues from the Oral Pathology Diagnostic Service; all were FFPE tissue blocks ranging from 2015-2019. In this study, we started with RNA extraction from the FFPE blocks to conduct gene profiling using the nCounter Pancancer Immune Profiling Panel from nanoString. After looking at the overall gene expression data on the 770 genes using the nSolver software, we investigated deeper into the protein expression in these tissues. Immunohistochemistry staining was done on the protein of interest, which allowed us to see the amount of staining, the staining pattern, and the location of staining (cytoplasm, nucleus, or membrane). The stained areas were detected and measure using digital pathology software QuPath.

2.1   Tissue Selection and Patient Records

All patient biopsy reports showing an initial clinical diagnosis of OPMDs with clinical and histopathological features consistent with OPMDs were selected from the University of Western Ontario Oral Pathology database. Biopsy samples were collected from the time between years 2015 to 2019. Cases were excluded if the pathological diagnosis concluded that the lesion had already transformed to carcinoma.

From the patient reports, demographic data, such as sex and age at the time of biopsy, was obtained. Known risk factors for the development of OSCC (tobacco smoking and alcohol consumption) were asked during the initial assessment by physicians, however, not all patients were willing to answer the questions. As a result of this, the information

was taken from the report if provided and considered as environmental factors in this study.

2.2   RNA Extraction from FFPE

Three cases were selected from each subset group (control, mild dysplasia, moderate dysplasia, and severe dysplasia). The RNeasy FFPE Mini Kit from QIAGEN (Lot#: 175019577, Catalog#: 73504, QIAGEN) was used in the process of RNA extraction for this study. Three FFPE cores of the epithelial region from each case block were initially transferred into a 1.5mL microcentrifuge tube. Following this, 1mL of xylene was added to each tube, which was then subjected to vigorous vortexing for 10 seconds and subsequently centrifuged at full speed for 2 minutes to sediment the sample at the bottom of the tube. The supernatant was carefully removed without disturbing the pellet. To the pellet, 1mL of 100% ethanol was added, mixed by vortexing, and again centrifuged at full speed for 2 minutes. The supernatant was then pipetted out without perturbing the pellet, and the tube was left open at room temperature (22°C) to allow the residual ethanol to evaporate completely.

Subsequently, 150µL of Buffer PKD was added to the pellet, and the mixture was homogenized by vortexing. This step was followed by the addition of 10µL Proteinase K, which was mixed thoroughly by pipetting up and down. The samples were then incubated at 56°C for 4 hours and subsequently at 80°C for 30 minutes to improve RNA yield, as prior attempts with recommended incubation times resulted in low-quality RNA. Post-incubation, samples were cooled on ice for 3 minutes and then centrifuged for 15 minutes at 13500 rpm. Care was taken to transfer only the supernatant to a new 1.5mL microcentrifuge tube, ensuring the pellet was not disturbed. To the transferred supernatant, 16µL of DNase Booster Buffer and 10µL DNase I stock solution (previously prepared with 1500 Kunitz units of lyophilized DNase I and 550µL RNase-free water, stored at -20°C) were added and mixed by inverting the tube, followed by a brief centrifugation to collect any residual liquid from the sides of the tube. The mixture was then incubated at room temperature for 15 minutes.

For RNA purification, I added 320µL of the Buffer RBC to adjust binding conditions, followed by the addition of 720µL of 100% ethanol to the mixture. The sample (700µL)

was then transferred to a RNeasy MinElute spin column placed in a 2mL collection tube, centrifuged for 15 seconds at 12000 rpm, and the flow-through was discarded. This step was repeated until all the sample had been processed through the spin column. The column was then washed with 500µL Buffer RPE (prepared by adding 44mL 100% ethanol to 11mL Buffer RPE concentrate) and centrifuged for 15 seconds at 12000 rpm, discarding the flow-through. A second wash with 500µL Buffer RPE was performed, followed by centrifugation for 2 minutes at 12000 rpm to wash the spin column membrane thoroughly, discarding the flow-through and collection tube. To dry the spin column membrane, the RNeasy MinElute spin column was placed into a new 2mL collection tube, centrifuged with the lid open for 5 minutes at full speed, and the flow-through was discarded. Finally, the spin column was transferred to a new 1.5mL collection tube, and 20µL RNase-free water was added directly to the membrane. The tube was centrifuged for 1 minute at full speed to elute the RNA, considering the dead volume of the spin column, which resulted in an 18µL eluate.

2.3   RNA Quality Assessment

Both tests were done at the Robarts Research Institute with instruments at David Carter's Lab (**Appx. Table 2**).

2.3.1   NanoDrop

NanoDrop spectrophotometer was used for quantifying nucleic acids (DNA and RNA). It allows the users to determine the RNA concentration and purity with minimal sample volume. Prior to starting, the upper and lower pedestal were cleaned with lint-free Kimwipes to avoid contamination. A Blank run was conducted first using 1µL of RNase-free water to set the base line, click "Blank" to zero the instrument with the blank solution while making sure the NanoDrop software is on Nucleic Acid RNA mode. Following the Blank sample, 1µL of RNA extracted from the previous step was pipetted on to the same spot on the lower pedestal, click Measure to obtain the sample measurement. I repeated the previous steps for all the RNA samples.

2.3.2   Bioanalyzer High Sensitivity RNA Analysis

The Agilent 2100 Bioanalyzer system was used to separate, size, and quantify the RNA samples obtained from RNA extraction process. For this test, we used the Agilent RNA 6000 Pico Kit (Lot#: 2240, Catalog#: 5067-1513, Agilent Technologies) with the RNA Pico Chips (Lot#: BQ09BK30, Catalog#: 5067-1513, Agilent Technologies). The RNA filtered gel was prepared by David Carter in advance and stored at 4°C. The RNA dye concentrate was removed from the fridge and equilibrated to room temperate, then 1µL of the dye was added to 65µL of the RNA filtered gel to create the gel-dye mix. The mixture was vortexed at 13000g for 10 minutes and 9µL of the mix was pipetted into the appropriate well. The chip was placed under the plunger and the plunger pressed to evenly distribute the mix, then 9µL of the mix was added to 2 more wells and pressed again. After the gel-dye mix has been loaded, 9µL of the RNA conditioning solution was transferred into the well labeled with "CS", and 5µL of RNA marker was transferred into all 11 sample wells. For reference, 1µL of the heat denatured and aliquoted ladder is transferred into the ladder well; before 1µL of diluted RNA sample is loaded into the rest of the 11 wells. The chip was horizontally placed into a vortexer for 1 minute at 2400rpm and placed into the Agilent 2100 Bioanalyzer to obtain data.

2.4   Gene Profiling using NanoString Technologies

The NanoString nCounter Pancancer Immune Profiling Panel we ordered came with all the required instruments to run the nCounter Prep Station including a USB with the panel RLF file, a Reporter CodeSet, and a Capture ProbeSet (Lot#: 281123, Item#: 100052, NanoString Technologies).

2.4.1   Hybridization

The Reporter CodeSet and Capture ProbeSet both need to hybridize to the target of interest for the target to be detected in the downstream analysis. The Thermocycler was set to 65°C with the lid heated to 70°C. The Reporter CodeSet and Capture ProbeSet were stored at -80°C and had to be taken out of the freezer and thawed to room temperature. A hybridization master mix was made by adding the hybridization buffer to

the Reporter CodeSet tube (**Table 2.1**). The mixture was mixed and spun down using a microcentrifuge for 10 seconds. I added 8μL of hybridization master mix to each tube of the strip tube. Following this, 5μL of the sample was added to each tube containing the hybridization master mix (the total unamplified RNA should be 300ng for each sample). In instances where less than 5μL of RNA sample was added, I supplemented the volume with nuclease-free water to achieve a total reaction volume of 13μL. I then proceeded to mix the Capture ProbeSet tube by either inverting or flicking it and briefly spun down the contents., then added 2μL of the Capture ProbeSet to each reaction tube. To ensure a thorough mix, the strip tubes were tightly capped inverted several times. Finally, the tubes were spined and immediately placed the in a pre-heated thermal cycler at 65°C for 20 hours. After the incubation period, I programmed the thermal cycler to incubate at 4°C until the samples were collected the following morning (in this study the samples were left at 4°C for 2 hours).

Table 2.1: Hybridization master mix for one nCounter assay (12 reactions).

Component	Hybridization Master Mix (μL)	Per Reaction (μL)
Reporter CodeSet	42 (in tube)	3
Hybridization Buffer	70	5
Final Volume	112	8

2.4.2   nCounter Prep Station

Prior to the run, I removed the prep plates and nCounter cartridge from the fridge to allow them to thaw to room temperature. The prep plate was spun for 2 minutes at 2000g in a centrifuge to allow the magnetic beads to move to the bottom row. The nCounter Prep Station is a fully automated system with step-by-step instruction on the touch screen. This study was run using the "high sensitivity" option. I followed the instruction and loaded all the required pieces into the Prep Station. Finally, the 12 samples were collected from the thermal cycler and spun briefly using a microcentrifuge. The 12 sample tubes were loaded into the Prep Station and the run was initiated. Following the run, the nCounter cartridge was sealed immediately and stored away from light. Wastes were discarded following recommended manufacturer instructions.

2.4.3   nCounter Digital Analyzer and nSolver Analysis Software

The barcodes on the imaging surface for each sample in the cartridge were counted by the nCounter Digital Analyzer. The sealed sample cartridge was placed in one of the six slots in the Digital Analyzer. I uploaded the RLF file for the Pancancer Immune Profiling Panel, then created a CDF file directly using the machine with the FOV (field of view) count at 280. The data was collected using a USB drive after the run was completed.

The nSolver Analysis Software 4.0 (NanoString Technologies) was used for quality control assessment and data normalization, as per the manufacturer's recommendations. Expression levels of target genes were normalized to positive and negative controls and reference genes. The six positive controls were used to measure the efficiency of the hybridization reaction and to check the linearity performance of the assay. The eight negative controls were made to not hybridize to any targets, and this was used to set the background threshold. I selected the epithelial reference genes in the list of the housekeeping genes given within the panel to normalize expression levels of target genes (ALAS1, GUSB, HPRT1, PPIA, and TBP). The Advance Analysis function was used to obtain more information from the data, including Differential Expression, Gene Set Analysis, etc.

2.5   Immunohistochemistry on Protein Expression

Immunohistochemistry (IHC) was employed as a follow-up to gene expression analysis conducted using the NanoString nCounter system. IHC was utilized to visualize and validate the expression of the corresponding proteins within the tissue samples using specific antibodies against the target protein, followed by detection using a chromogenic marker. IHC provides insight into the proteins' distribution and relative expression levels in different areas of the tissue. It not only confirms the relevance of the gene expression findings at the protein level but also contributes to the mapping of molecular alterations within the structural context of the tissue, offering valuable insights into the biological mechanisms driving the transformation process within the oral epithelial cells. The IHC for this study was done at two different locations, the University Hospital Lab as well as the Oral Pathology Lab.

2.5.1   IHC Slide Preparation

Formalin fixed paraffin embedded tissue blocks were placed on ice for 20 minutes. The microtome (Microm HM 325; GMI Inc., Ramsey, MN) was set to 4μm, and each block was sectioned to expose a full surface. The fresh tissue surface was then put back onto the ice block. Tissue sections were cut from the block and dropped into a 45°C warm water bath. Positively charged glass slides were used to collect each thin tissue section. Once the tissue was on the slide, it was placed into a slide rack which was then placed into a 37°C oven for at least 24 hours before removal to allow fixation.

2.5.2   IHC Process for β-catenin, Ki67, E-cadherin, Vimentin, and S100A7 (University Hospital)

Automated staining utilizing the complimentary antibodies to the protein of interest was performed, according to the manufacturer's instructions (**Table 2.2**). Heat-induced epitope retrieval (HIER) was performed using diluted EnVision FLEX Target Retrieval Solution, High pH (50x) (Dako Omnis, Code GV805) for 20 minutes at 97°C followed by 5 minutes in EnVision FLEX Wash Buffer (20x) (Code: K8007). The staining steps and incubation times were pre-programmed using the Dako Omnis Basic User Guide at 32°C. Reagents were applied directly to the slide with tissue with a volume of 200 µL per slide. All incubation periods were completed at room temperature. The visualization system used was EnVision FLEX, High pH (Link) (Dako Omnis, Code GV800). Counterstaining was performed using EnVision FLEX Hematoxylin (Link) (Code: GC808). Positive and negative controls were mounted on the same slide and were run simultaneously using the same protocol as for the case specimens.

Table 2.2: List of antibodies used in this study.

Protein Target	Vendor	Species	Dilution	Product Code
β-catenin	Dako	Monoclonal Mouse	Ready-to-use	GA702
Ki67	Dako	Monoclonal Mouse	Ready-to-use	GA626
E-cadherin	Dako	Monoclonal Mouse	Ready-to-use	GA059
Vimentin	Dako	Monoclonal Mouse	Ready-to-use	GA630
S100A7	Cedarlane	Monoclonal Mouse	1:8000	NB100-56559SS

2.5.3   IHC Process for MCM2 and Geminin (Oral Pathology Department)

The rehydration technique was performed in the following order: the slides were immersed in 100% Xylene solution for 5 minutes. This was repeated for another 5 minutes, then 3 minutes, in two separate containers. Slides were then placed into various concentrations of ethanol solution, 100% ethanol two times (2 minutes and 1 minute), 95% ethanol two times (2 minutes and 1 minute), 70% ethanol for 2 minutes. Finally, the slides were placed into distilled water for 2 minutes. The slides were quenched with fresh 3% Hydrogen Peroxide in methanol for 5 minutes in order to block endogenous peroxidase activity, then rinsed in distilled water for 5 minutes, then placed in phosphate buffered saline (PBS) for 5 minutes on a shaker prior to the antigen retrieval process. Antigen retrieval was performed using a decloaking chamber (Serial #: DC1507, Biocare Medical) to establish optimal staining conditions. The decloaking chamber setting was set to 120°C for 90 seconds then 90°C for 10 seconds. Slides were emerged in buffer solution containing Tris – EDTA with 0.05% Tween 20 at pH=9.0 and put in the decloaking chamber until the time had reached as the setting indicated. Following antigen retrieval, slides were cooled by running them under cold tap water and washed with PBS for 5 minutes on a shaker. Once washed, each slide was blocked with 120μL of 2.5% normal horse serum (Lot #: ZH1118, Item #: 30022,

Vector Laboratories) for 30 minutes at room temperature to minimize non-specific antigen binding and reduce background staining. Positive and negative controls were mounted on the same slide and were run simultaneously using the same protocol as for the case specimens.

MCM2 Specific Steps

After the horse serum was removed (excluding the negative control slides), MCM2 rabbit polyclonal antibody (Cat. #: ab4461, ABCAM) was diluted to 1:2000 with 2.5% horse serum and 120μL of the diluted antibody was added onto the slides (excluding the negative control slides).

Geminin Specific Steps

After the horse serum was removed (excluding the negative control slides), recombinant rabbit anti-Gemini monoclonal [EPR14637] antibody (Cat. #: 195047, ABCAM) was diluted to 1:250 with 2.5% horse serum and 120μL of the diluted antibody was added onto the slides (excluding the negative control slides).

Slides were incubated at 4°C overnight (around 16 hours) in a humidified chamber filled with distilled water. After incubation, slides were rinsed in PBS for 5 minutes on a shaker twice and ImmPRESS HRP horse anti-rabbit IgG polymer reagent (Lot #: ZK0418, Item #: 30026, Vector Laboratories) was added to the slides as the secondary antibody for 30 minutes at room temperature and then rinsed with PBS two more times for 5 minutes on a shaker. The DAB solution was prepared by adding 2 drops of buffer, 4 drops of DAB, then 2 drops of H_2O_2 into a microtube containing 5 mL dH_2O (Lot #: ZG1119, Item#: SK4100, Vector Laboratories). Slides were covered with 120μL of DAB mixture for 10 minutes, then rinsed with distilled water to stop the reaction.

Following DAB staining, the slides were counterstained with Harris Hematoxylin for 1 minute and then rinsed under tap water. Once the Hematoxylin was removed, the slides were dipped 2 times in Acid Alcohol (1% Hydrochloric Acid in 70 % Alcohol) and rinsed in running tap water, then dipped 2 times in Ammonium Alcohol (2% Ammonium Hydroxide in 70% Alcohol) and rinsed in running tap water one last time. Dehydration of

the slides was then carried out in the opposite order of rehydration: submerged in 70%
Ethanol for one minute, 95% Ethanol for one minute twice, 100% Ethanol for one minute
three times, and finally Xylene for five minutes twice. Cover slips were applied to the slides
using Epredia Cytoseal 60 (Item #: 83104, Thermo Fisher Scientific). Finally, the slides
were left flat for at least 24 hours for the cover slips to adhere completely.

2.6   Digital Pathology Analysis

All IHC slides were scanned at 20x magnification using the Aperio Scanner (Leica
Biosystems Inc, Wetzler Hesse, Germany) at the University Hospital Pathology
Department. The digital images were saved as *.svs* files and imported into QuPath for
analysis.

2.6.1   QuPath Scoring of S100A7

The immunohistochemistry was analyzed digitally using the scanned images on QuPath.
To calculate the percentage of the stained epithelium, the area of interest was selected on
QuPath manually using the "*Annotation*" function. In this study, I selected the entire
epithelium measuring from the basement membrane to the edge of the outer most cell
layer, excluding the dead epithelium portion (keratinized layer of the stratum corneum).
After the area of interest was selected, I created a threshold using the parameters in
Figure 2.1 to detect positively stained area, and the stain percentage was calculated by the
QuPath software automatically. Within the same epithelial region, I was also interested in
seeing the difference between nuclear staining and cytoplasmic staining. To detect the
staining separately, the "*positive cell detection*" function was used to detect both
"*Nucleus: DAB OD mean*" and "*Cytoplasm: DAB OD mean*" keeping all other
parameters the same (**Figure 2.1**). After the settings were saved, the software would
count every cell within the selected region of interest and detect the numbers of positive
cells within that region.
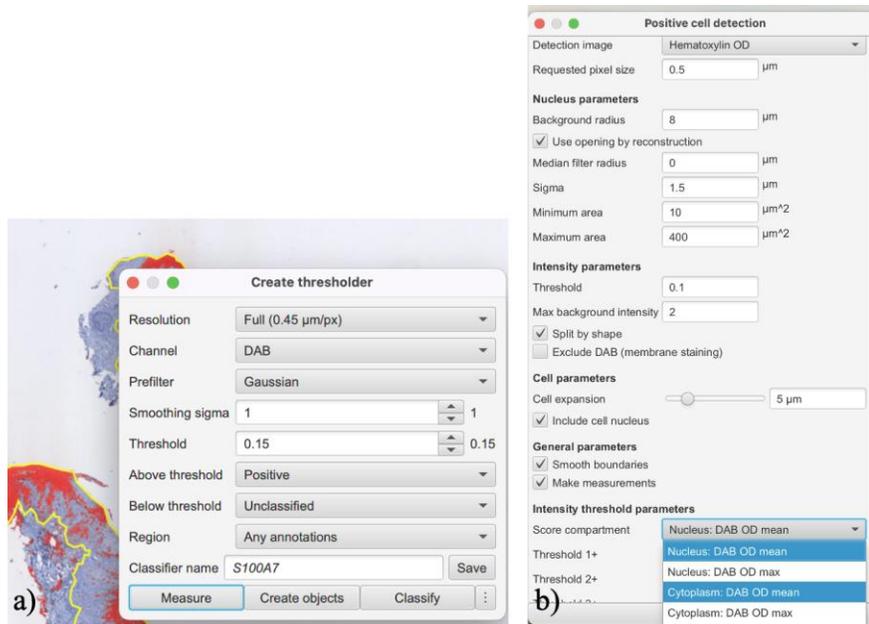
Figure 2.1: S100A7 detection settings in QuPath.
a) Positive threshold setting used to calculate the stain percentage. The resolution controls how 'blocky' the output is, and this should be the lowest resolution that is still considered accurate. The DAB channel was used to best suit the IHC stained images, and Gaussian is always used as the prefilter. Smoothing sigma and threshold values were chosen manually. b) Positive cell detection setting to detect nuclear and cytoplasmic stain. By using the same parameters only changing the score compartment, cells expressing S100A7 only in the nucleus or cytoplasm can be separated from the rest.

2.6.2   E-cadherin and Vimentin

The same detection method used for S100A7 was applied to E-cadherin. The stain intensity was used to quantify the immuno-staining around the cell borders;s a darker and thicker stain indicated the cell expresses an adequate amount of E-cadherin, while a lighter and thinner stain indicated there was down regulation of the protein. Using the digital image in QuPath, vimentin stain percentage was obtained by calculating the mean percentage of 4 random, non-overlapping rectangular fields selected within the epithelial layer. Each side of the rectangle is set to 100 units (μm) on QuPath, and the threshold was created using the same method as S100A7.

2.6.3   Nucleic expression of Ki67, Geminin, Mcm2

I started by using the annotation tools in QuPath to delineate the epithelial areas of interest on my tissue samples, making sure to include only the selected areas to maintain accuracy. Once the regions of interest were annotated, I set up the analysis protocol to detect the specific protein expression using the "*Positive Cell Detection*" function and configured the cell detection settings, adjusting parameters such as cell size and nuclear staining thresholds to match the specific characteristics of my stain and tissue type (**Figure 2.2**). QuPath processed the image and identified cells based on the criteria I had set in Figure 2.1. Once the process was completed, the cells were marked, and each cell's nucleus was outlined, which was crucial for assessing nuclear protein expression. To measure the protein expression specifically in the nuclei, I adjusted the measurement settings focused on the nuclei rather than the cytoplasm or the whole cell to specifically assess nuclear expression. Once the cells were classified, I reviewed the results visually to confirm the accuracy of the classification.
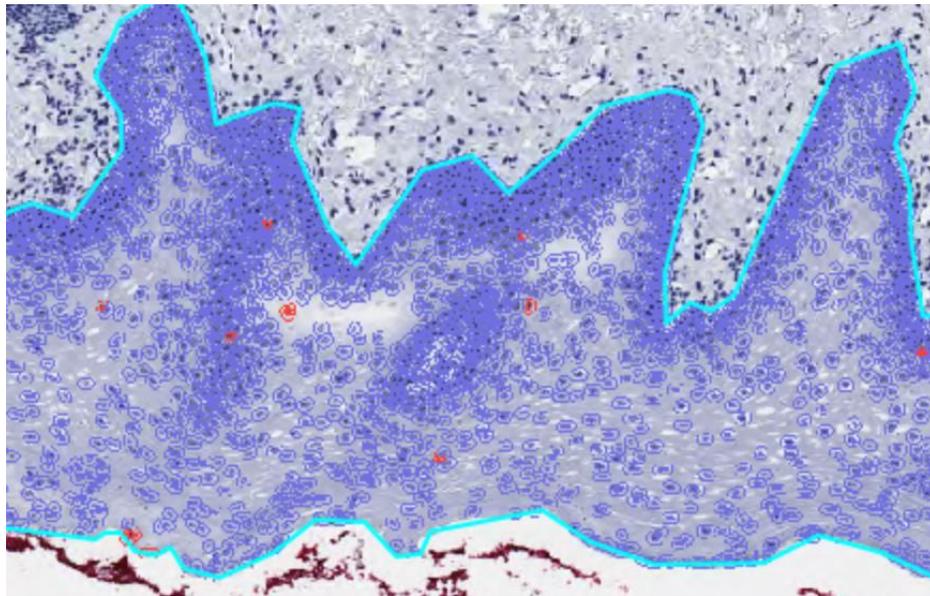
Figure 2.2: Positive Cell Detection in Oral Epithelium Using QuPath.
Positive cell detection performed on normal oral epithelium stained for Ki67. The image shows the oral epithelium outlined with cyan, individual cells are outlined in blue with positively stained cells highlighted in red. The detection was conducted using QuPath software.

2.7 Statistical analysis

2.7.1 Concordance of diagnoses between clinicians and pathologists

To assess the concordance of diagnoses between clinicians and pathologists, I used Cohen's Kappa (κ) test as the statistical measure of agreement. This process involved collecting diagnostic data for all patient samples, where each sample was independently classified by both a clinician, based on visual appearances and symptoms, and a pathologist, based on the microscopic examination of biopsy tissue samples. A contingency table was constructed, listing the frequencies of each category combination assigned by the clinicians and pathologists. The Observed Agreement (Po) was calculated as the proportion of cases where both raters agreed on the diagnosis. To account for chance agreement, the Expected Agreement (Pe) was computed by considering the marginal totals of each category. Cohen's Kappa was then derived using the formula κ = (Po - Pe) / (1 - Pe), providing a measure that adjusts for the agreement expected by chance. In this study, I separated the clinical diagnoses and pathological diagnoses into 3 groups: Dysplasia, Rule-out (R/O) dysplasia, and other OPMDs.

This method allowed for a quantitative evaluation of the inter-rater reliability, offering insights into the consistency of diagnostic practices between the two types of medical professionals. Similar to correlation coefficients, Cohen's Kappa can range from −1 to +1, where 0 represents the amount of agreement that can be expected from random chance, and 1 represents perfect agreement between the raters. Although it is possible to obtain a negative κ value, it is rare to see this in studies. The final κ value obtained was interpreted based on the suggested values in McHugh's study (**Table 2.3).**

Table 2.3: Interpretation of Cohen's kappa (McHugh, 2012).

Value of Kappa	Level of Agreement	% of Data that are Reliable
0 – 0.20	None	0–4%
0.21 – 0.39	Minimal	4–15%
0.40 – 0.59	Weak	15–35%
0.60 – 0.79	Moderate	35–63%
0.80 – 0.90	Strong	64–81%
Above 0.90	Almost Perfect	82–100%

Chapter 3

3   Results

In the results section of my thesis, I will present and analyze the data collected from my study on the early detection of oral potentially malignant disorders. The results will encompass a detailed examination of patient demographics (sex, age, smoking and alcohol consumption, etc.), clinical and histopathological diagnoses, and gene expression analysis obtained using the NanoString nCounter system. Additionally, I will discuss the IHC staining results of various protein expressions in the oral epithelium. This section will also include statistical analyses of the correlations between the severity of the disease and the expression levels of specific biomarkers.

3.1   Population Demographics

3.1.1   Age and Sex

A total of 110 tissue samples were selected in my study: 82 cases were identified as OPMDs by two separate oral surgeons, and 28 cases were identified as normal oral mucosa tissue (Controls) from the oral pathology department. The age at biopsy for the Controls group ranged from 21-82 years (average age = 59 years); and the age at biopsy for the OPMDs group ranged from 12-95 years (average age = 62 years) (**Table 3.1**). In the Controls group, there were 16 males and 12 females. The average age for females Controls group was 61 years and for males was 56 years (**Table 3.2**). In the OPMDs group, there were 42 males and 40 females. The average age for female OPMDs group was 60 years and 63 years for males (**Table 3.3**).

Table 3.1: Average age and median age of patients at initial biopsy between control group and OPMDs group.

Group	Average Age at Biopsy (SD)	Median Age at Biopsy
Controls (n=28)	59 (13.2)	57
OPMDs (n=82)	62 (12.6)	63

Table 3.2: Controls population (n=28) sex and average age at biopsy.

Sex	Amount	Percentage (%)	Average Age at Biopsy (SD)
Female	12	42.9	56 (11.5)
Male	16	57.1	61 (12.9)

Table 3.3: OPMDs population (n=82) sex and average age at biopsy.

Sex	Amount	Percentage (%)	Average Age at Biopsy (SD)
Female	40	48.8	60 (12.9)
Male	42	51.2	63 (12.5)

Since the Controls and OPMDs groups had difference sample size, I conducted Welch Two Sample t-test to determine whether there is a significant difference between the mean patient age of the two groups. The results indicated no statistically significant difference between the Control group (mean age = 59.68 years) and the OPMDs group (mean age = 61.80 years), with a t-value of -0.698, degrees of freedom (df) = 42.9, and a p-value of 0.49. The 95% confidence interval for the difference in means was [-8.24, 4.00], suggesting substantial overlap in the age distributions of the two groups (**Figure 3.1**). The same test was applied to see if there is a significant difference in age between male and female patients. In the Controls group, there is no significance difference between the patient age and patient sex (t = 1.70, df = 25.58, p-value = 0.10). In the OPMDs group, there is also no significance between the patient age and patient sex (t = -1.04, df = 74.48, p-value = 0.30).
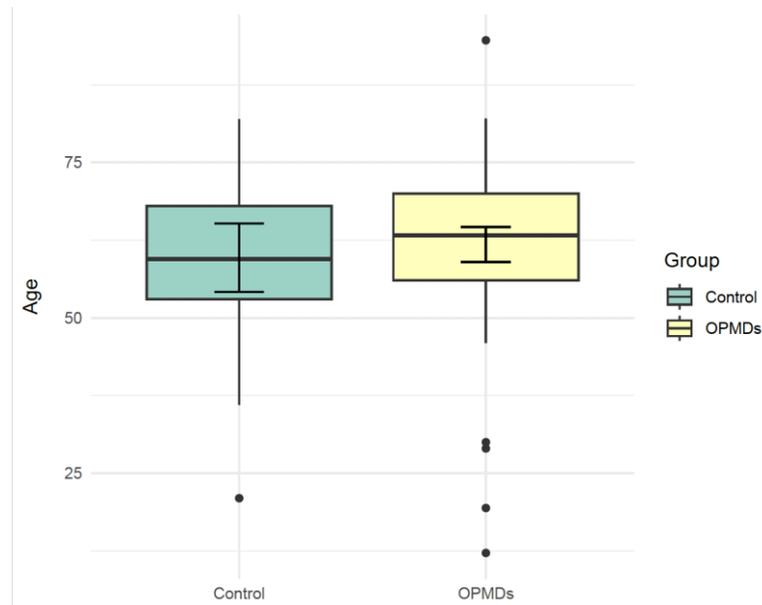
Figure 3.1: Boxplot comparison of age between Control and OPMDs groups.
Each box represents the interquartile range, with the median age indicated by a horizontal line within the box. Outliers are represented as individual points. Error bars represent the standard error of the mean.

After investigating the overall comparison of OPMDs to Controls, I examined whether patient age and sex would influence the severity of dysplasia in OPMD patients. Based on my ANOVA and Tukey's Honest Significant Difference (HSD) test results, there are no significant differences in age between the three dysplasia severity groups (**Figure 3.2**). ANOVA test was used to compare the ages across different dysplasia groups. The results indicated that there were no significant differences in age among the three groups (F= 0.16, p = 0.86). To further investigate the differences, a Tukey's HSD test was performed. The Tukey HSD test confirmed that there were no significant pairwise differences in age between the mild, moderate, and severe dysplasia groups. Specifically, the age difference between moderate and mild dysplasia was -0.43 years (p = 0.99), between severe and mild dysplasia was 2.34 years (p = 0.87), and between severe and moderate dysplasia was 2.77 years (p = 0.85). These results suggest that age is not significantly associated with the severity of OPMDs in this sample. Sex also had no significant effect on the severity of dysplasia, indicated by X^2 = 0.057 with 2 degrees of freedom and a p = 0.97 (**Figure 3.3**).
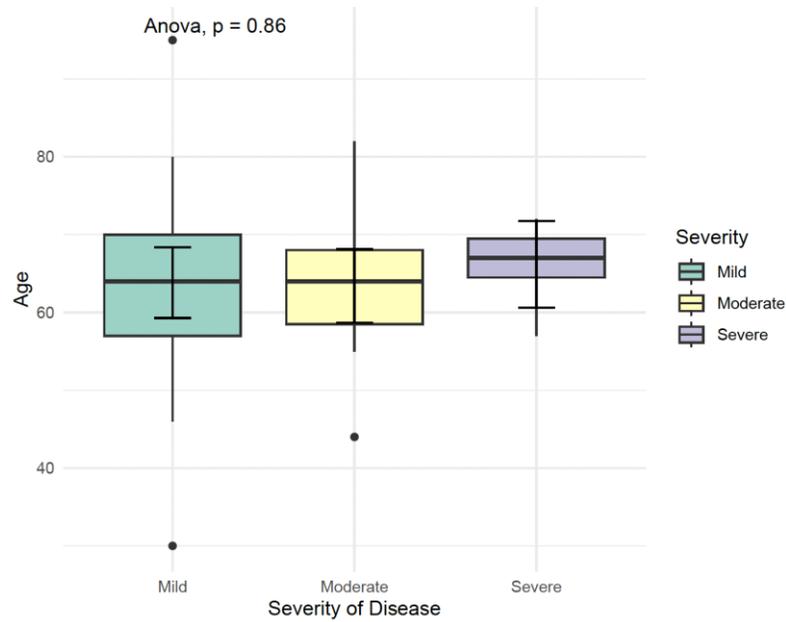
Figure 3.2: Boxplot comparison of age between different dysplasia Groups.
There was no significant difference between patient age and dysplasia severity. Outliers were indicated by black dots. Figure was generated using R.
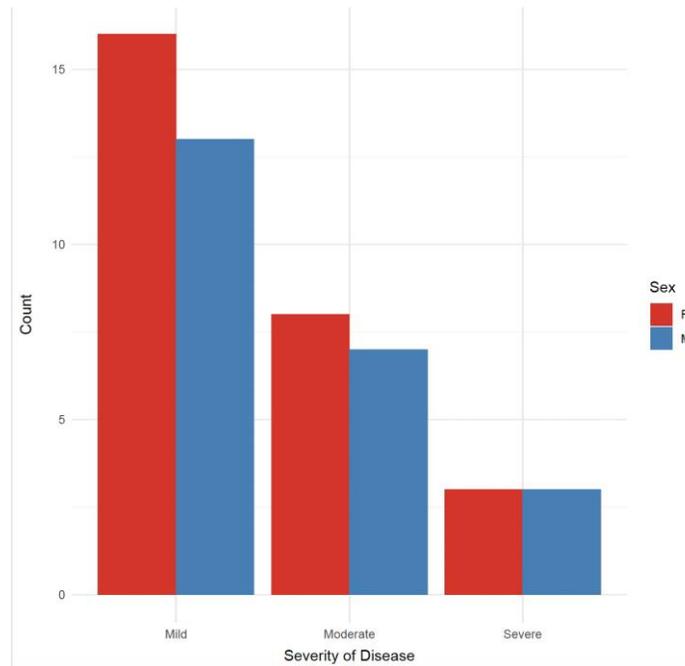
Figure 3.3: Distribution of dysplasia severity by sex.
The data suggests that dysplasia severity is similarly distributed across both male and female patients in this study. $X^2 = 0.057$, df = 2, and p-value = 0.9717. Bar graph was generated using R.

3.1.2   Tobacco and Alcohol Use

Additional demographic information reported in my study include smoking (tobacco) and alcohol use. In the Controls group: 39.3% (11/28) reported previous tobacco smoking, 60.7% (17/28) had no smoking history. The alcohol consumption data was no recorded in the Controls group. In the OPMDs group, 48.8% (40/82) had smoking history, and 51.2% (42/82) were non-smokers. Only 49 out of 82 OPMD cases reported alcohol history; 51% (25/49) were drinkers and 49% (24/49) were non-drinkers. Chi-square test for independence indicated that the prevalence of smoking is similar in both the Controls and OPMDs groups, suggesting that smoking alone is not significantly different between individuals with and without OPMDs (**Figure 3.4**). Finally, smoking also had no significant association with the severity of dysplasia, indicated by a p-value of 0.79 (**Figure 3.5**). Given the lack of data on alcohol usage in both the control and OPMDs groups, no statistical analysis was conducted on the drinking behavior.
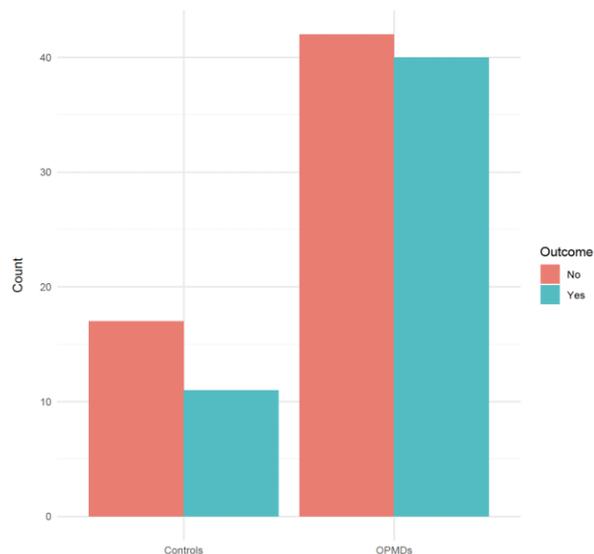
Figure 3.4: Distribution of Smoking Behavior in Controls vs. OPMDs Groups.
The bar graph illustrates the distribution of smoking behavior between Controls and OPMDs. The data was analyzed using a chi-square test for independence to determine if there is a significant association between the two groups and tobacco smoking. The chi-square test results indicate that X^2 = 0.42, df = 1, p-value = 0.52. Graph was generated using R.
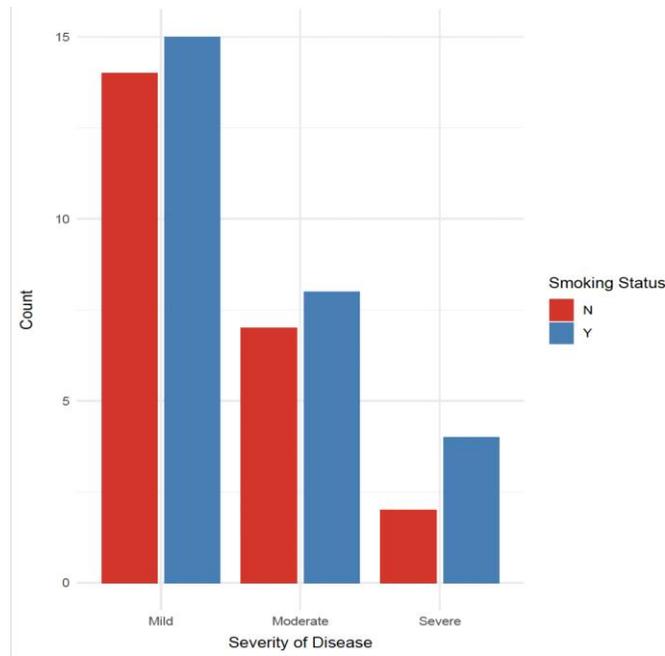
Figure 3.5: Distribution of dysplasia severity between smokers and non-smokers.
The counts of each severity level are displayed for both smokers and non-smokers. The chi-square test indicated no significant association between smoking status and disease severity (X^2 = 0.45, df = 2, p-value = 0.79).

3.2   Diagnostic Concordance

In this part of the result, I analyzed the data on the diagnostic concordance between clinical diagnoses made by oral surgeons and the final diagnoses provided by oral pathologists (**Appx. Table 3 & 4**). The clinical diagnoses (**Figure 3.6**) were categorized into five major disease groups based on the clinical reports received from the oral surgeons. Dysplasia was the most frequently diagnosed condition, accounting for 51 cases (75.6%). Leukoplakia was identified in 17 cases (20.7%), followed by chronic lichenoid mucositis and fibroma, each with 2 cases (2.5%). Additionally, 3 cases (3.6%) were diagnosed as lichen planus, and 7 cases (8.5%) were classified as N/A, indicating uncertainty or the absence of a definitive clinical diagnosis. The final diagnoses made by an oral pathologist listed in Figure 3.7, were based on histopathological examination of biopsy samples taken from the suspected area of concern within the oral region by the oral surgeons. Like the clinical diagnoses, the most commonly diagnosed condition was dysplasia. However, the dysplasia cases were further divided based on severity of the disease: mild dysplasia, moderate dysplasia, and severe dysplasia. Each were found

representing 29 cases (35.4%), 15 cases (18.3%), and 6 cases (7.3%), respectively. Chronic lichenoid mucositis was diagnosed in 13 cases (15.8%), while chronic candidiasis and fibroma were each identified in 5 cases (6.1%). Other conditions included lichen planus (3 cases, 3.7%), papilloma (4 cases, 4.9%), and traumatic ulcerative granuloma with stromal eosinophilia (TUGSE) (2 cases, 2.4%).
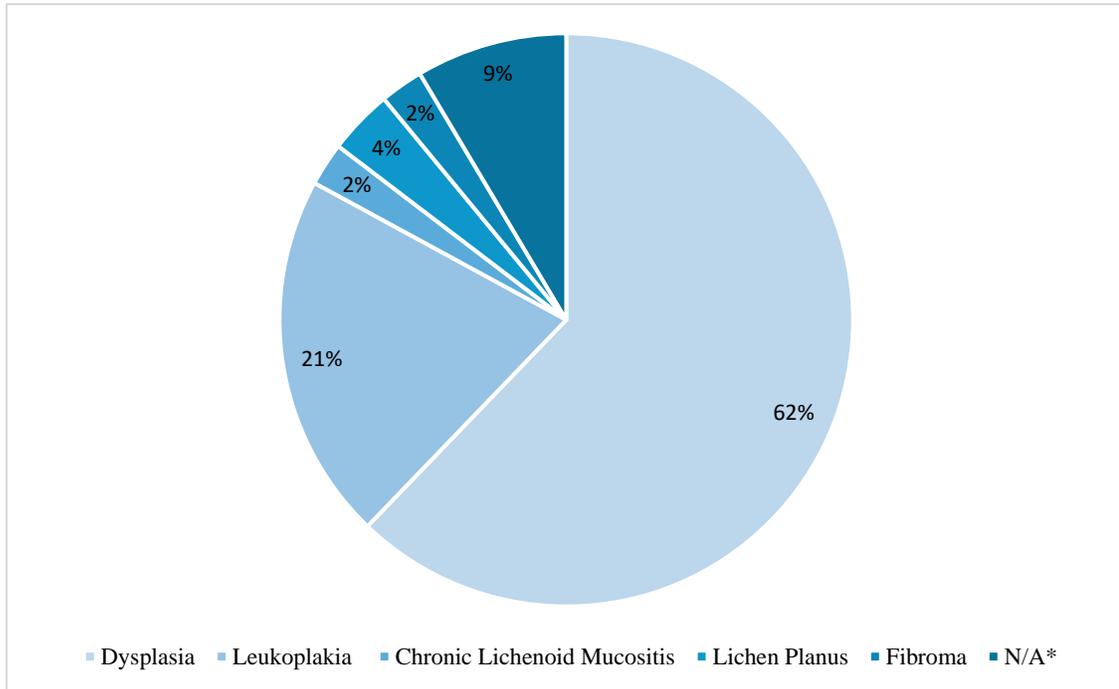
Figure 3.6: Clinical Diagnoses of OPMD Patient Cohort (n=82).
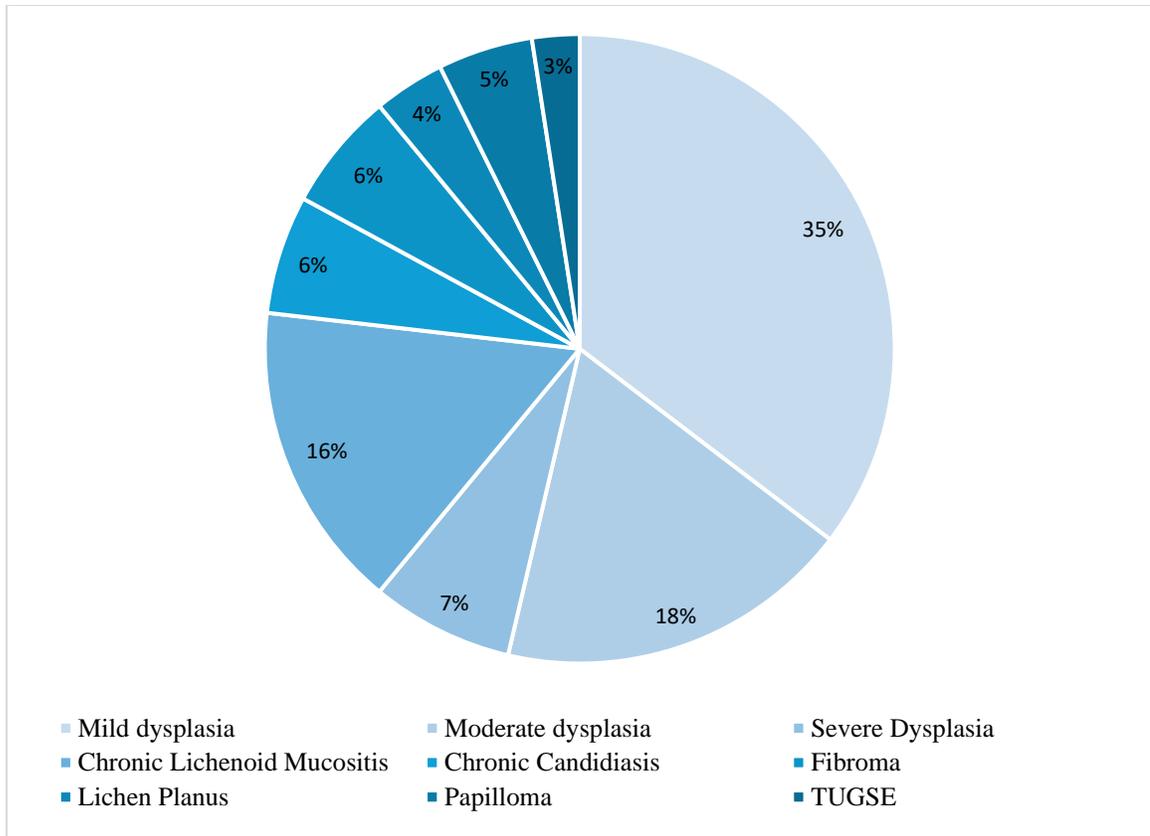
Figure 3.7: Final Diagnoses of OPMD Patient Cohort (n=82).

The previous diagnoses made by oral surgeons and oral pathologists were further simplified into nominal variables (dysplasia and non-dysplasia) to assess the agreement between clinical and final histopathological diagnoses of OPMDs. Table 3.4 shows the Cohen's Kappa value was found to be $\kappa = 0.57$ with standard error of 0.093.

Table 3.4: Cohen's Kappa Test for Diagnostic Agreement.

Clinical Diagnosis	Pathological Diagnoses		Total Cases
	Dysplasia	**Non-Dysplasia**	**Total Cases**
Dysplasia	47	10	55
Non-Dysplasia	6	19	27
Total Cases	53	29	82

51

3.3 NanoString Gene Profiling

The NanoString nCounter technology was employed to assess the expression levels of mRNA of biomarkers in normal oral epithelium and three groups of oral dysplasia tissues. Using the nSolver software advance analysis function, I was able to analyze the gene expression patterns across all 12 samples. The analysis of gene expression patterns across the different stages of oral dysplasia revealed that both upregulation and downregulation of genes play critical roles in disease progression. These changes at the mRNA level likely represent the early molecular events that drive the transformation of normal tissue into dysplastic lesions.

Out of the 770 gene targets included in the panel, the identification of both upregulated and downregulated genes underscores the complex regulatory mechanisms at play, where activation of oncogenes and suppression of tumor suppressors might be involved. Additionally, the number of significant differentially expressed genes increases with disease severity is insightful. This trend suggests that as oral dysplasia progresses, there is a more extensive disruption of gene expression, which is often indicative of advancing pathology. In visualizing these findings, volcano plots were used to effectively distinguish between overexpressed and downregulated genes, enhancing the clarity of the results. The Benjamini-Yekutieli method was applied to obtain the adjusted p-value, assuming there may be some biological connection between genes to control the false discovery rate (FDR). These differential expression patterns not only serve as potential biomarkers for disease stratification but may also represent targets for therapeutic intervention in oral potentially malignant disorders (**Figure 3.8-10**).
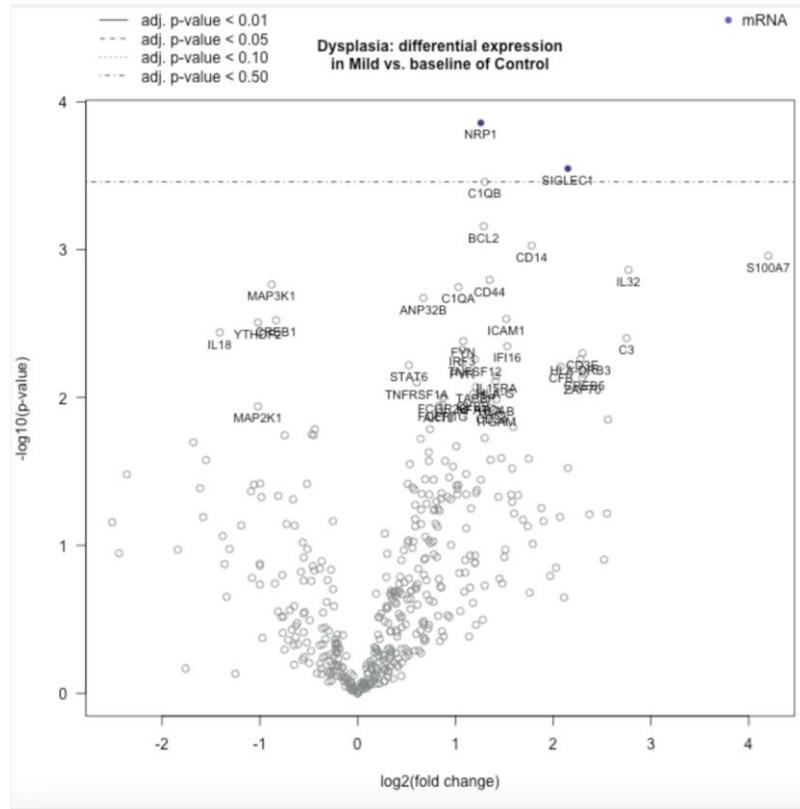
Figure 3.8: Volcano plot displaying the genes of significance between mild dysplasia cases and control tissues.

Highly statistically significant genes fall at the top of the plot above the horizontal lines (adj. p-value < 0.50), and highly differentially expressed genes fall to either side. Appx. Table 5 indicates the top 10 differentially expressed mRNA. Figure 3.8 illustrates a notable upregulation of NRP1-mRNA and SIGLEC1-mRNA expression in mild dysplasia tissue compared to normal tissue. The log2 fold change of 1.26 indicates that NRP1 expression is approximately 2.39 times higher in mild dysplasia tissue. The standard error is 0.185, with confidence intervals ranging from a lower limit of 0.893 to an upper limit of 1.62, reflecting the robustness of the observed difference. The statistical significance of this upregulation is supported by a low P-value (0.000139) and a Benjamini-Yekutieli adjusted P-value (0.349). Similarly, SIGLEC1-mRNA showed a log2 fold change of 2.15 corresponds to a linear fold change of approximately 4.44 times. The standard error for this measurement is 0.351, with confidence intervals ranging from 1.46 to 2.84, indicating a reliable result. The statistical significance of SIGLEC1

upregulation is further validated by a P-value of 0.000283 and a Benjamini-Yekutieli adjusted P-value (0.349).

Although S100A7 was above the adjusted p-value threshold, it had the highest log2 fold change of 4.2 indicates that S100A7 expression is approximately 18.4 times higher in mild dysplasia cases.
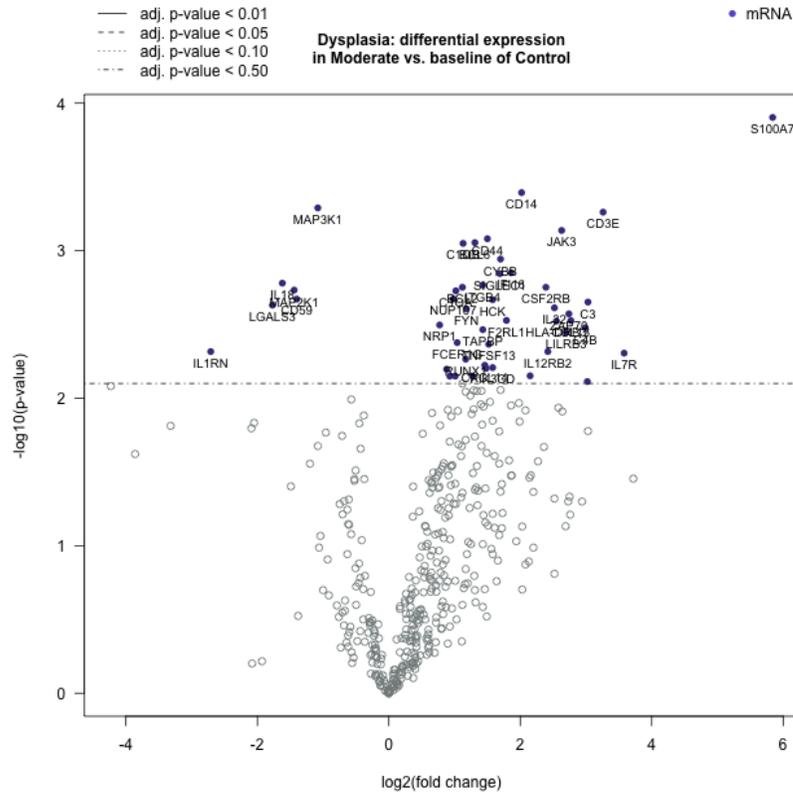
Figure 3.9: Volcano plot displaying the genes of significance between moderate dysplasia cases and control tissues.

The analysis between moderate dysplasia cases and control tissues revealed that the number of significantly differentially expressed genes was higher compared to the comparison between mild dysplasia cases and control tissues. Appx. Table 6 indicates the top 15 differentially expressed mRNA.

The volcano plot in Figure 3.9 illustrates a substantial upregulation of S100A7 mRNA expression in moderate dysplasia tissue compared to normal tissue. The log2 fold change of 5.84 indicates that S100A7 expression is approximately 57.4 times higher in moderate dysplasia tissue, with a narrow standard error (0.847) and tight confidence intervals

(lower limit: 4.18, upper limit: 7.5), suggesting robustness in the observed difference. The statistical significance of this upregulation is supported by a low P-value (0.0001) and Benjamini-Yekutieli adjusted P-value (0.311). In addition to S100A7, several other mRNAs exhibited significant differential expression. CD14-mRNA was significantly overexpressed in dysplasia tissues, exhibiting a log2 fold change of 2.02 (p-value = 0.0004), which corresponds to a linear fold change of about 4.04 times. Similarly, CD3E-mRNA showed a significant higher expression in dysplasia tissues, with a log2 fold change of 3.26 (p-value = 0.0005), resulting in a linear fold change of approximately 9.57 times. Lastly, JAK3-mRNA expression was significantly elevated in dysplasia tissues, with a log2 fold change of 2.63 (p-value = 0.0007), translating to a linear fold change of about 6.19 times. In contrast to these upregulations, one target, MAP3K1-mRNA, showed a significant decrease in expression levels. MAP3K1-mRNA expression was notably decreased in dysplasia tissues, with a log2 fold change of -1.08 (p-value = 0.0005) and a corresponding linear fold change of approximately 0.474 times.
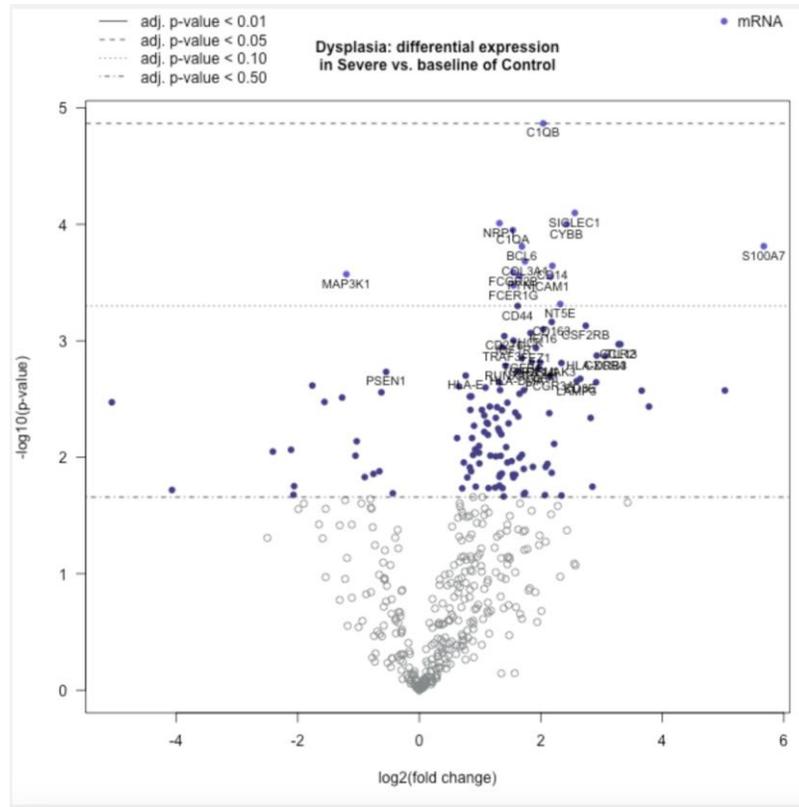
Figure 3.10: Volcano plot displaying the genes of significance between severe dysplasia cases and control tissues.

The analysis between severe dysplasia cases and control tissues revealed that the number of significantly differentially expressed genes was the highest. Appx. Table 7 indicates the top 20 differentially expressed mRNA. Similar to previous comparisons, S100A7 exhibited substantial upregulation. The log2 fold change of 5.67 indicates that S100A7 expression is approximately 50.8 times higher in severe dysplasia tissue, with a narrow standard error (0.847) and tight confidence intervals (lower limit: 4.01, upper limit: 7.33), suggesting robustness in the observed difference. The statistical significance of this upregulation is supported by a low P-value (0.000154) and a Benjamini-Yekutieli adjusted P-value of 0.065 indicating strong evidence for the differential expression of S100A7 in severe dysplasia.

3.4   Protein Levels in Oral Epithelium

3.4.1   S100A7 Levels

Visualization of S100A7 IHC Expression Using Digital Pathology
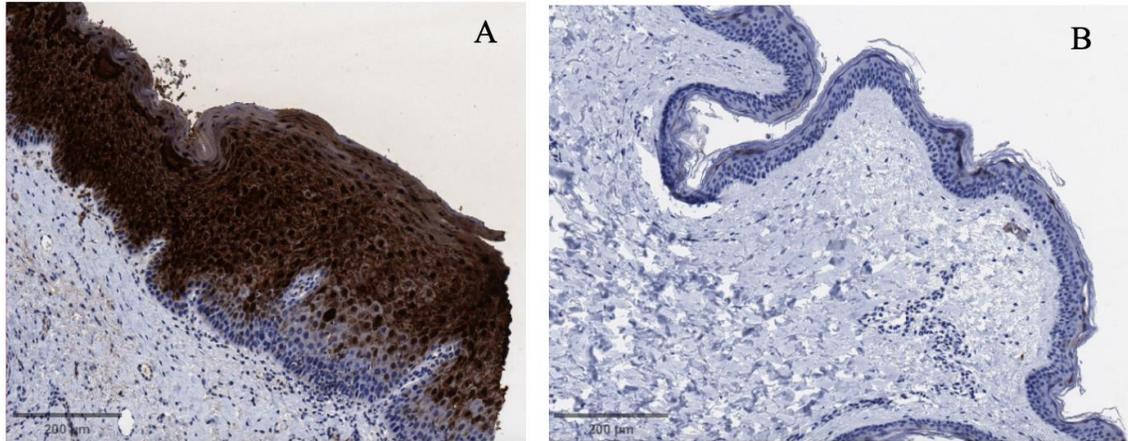
Figure 3.11: Positive and Negative Control of S100A7 Expression in Epithelium.
A) Positive control demonstrating strong and specific cytoplasmic and nuclear staining
for S100A7, validating the staining protocol. B) Negative control with no primary
antibody, displaying no specific staining, confirming the specificity of the
immunohistochemical procedure. Images were scanned at 20x magnification and
captured in QuPath.

Immunostaining of S100A7 in Control Group

Immunostaining of S100A7 level in the control cases revealed variability in staining
intensity across the group. The majority of control cases showed very little to no staining,
with only a few scattered epithelial cells displaying weak cytoplasmic and nuclear
positivity. However, a subset of control cases exhibited more pronounced staining on the
surface layer of the epithelium (**Figure 3.12**). This variability in S100A7 expression
among the control group indicates that while most cases exhibit minimal expression,
certain individuals may display higher levels of staining, suggesting intrinsic differences
in protein expression within normal tissue.
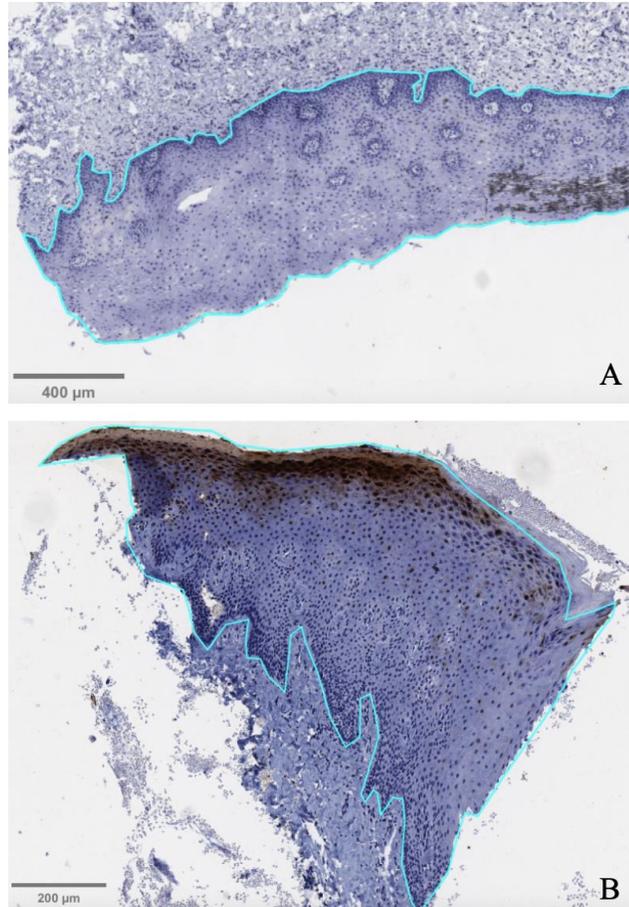
Figure 3.12: S100A7 Epithelial Expression in Control Cases.
A) Case 5 in Controls showing weak baseline cytoplasmic (1.33%) and nuclear (1.36%) stain for S100A7. B) Case 16 in Controls showing cytoplasmic (10.43%) and nuclear (10.2%) stain for S100A7. Images were scanned at 20x magnification and captured in QuPath.

Immunostaining of S100A7 in OPMD Group

Immunohistochemical staining for S100A7 protein expression across different levels of dysplasia in OPMD patient cases revealed a progressive increase in staining intensity with dysplasia severity (**Figure 3.13**). However, variability was evident within each dysplasia group. Figure 3.14 shows two cases belonging to the moderate dysplasia group. Both cases were biopsied by the same oral surgeon in 2019. One case showed intense staining of nearly 80%, while the other exhibited around 40%.
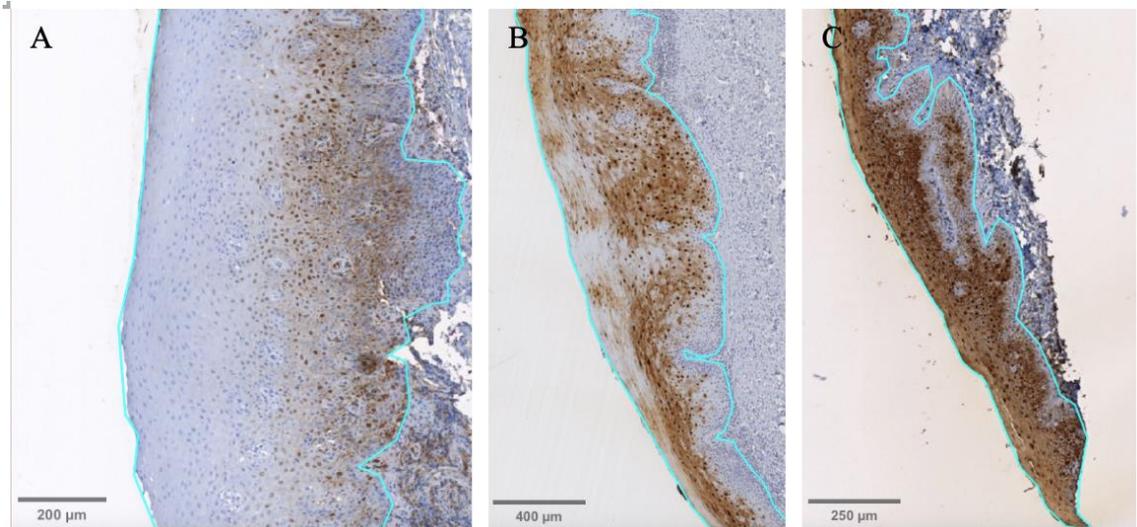
Figure 3.13: S100A7 Epithelial Expression in Dysplasia Cases.
Images showing IHC staining for S100A7 protein expression in epithelial tissues across different dysplasia severities. (A) Mild dysplasia showing moderate cytoplasmic staining (Case **2**, cytoplasmic staining 19%, positive nucleus staining 6.7%). (B) Moderate dysplasia (Case **3**) exhibiting more intense cytoplasmic (54.8%) and nuclear staining (30.8%). (C) Severe dysplasia (Case **1**) displaying extensive and intense cytoplasmic (91.6%) and nuclear (90%) staining. Images were scanned at 20x magnification and captured in QuPath.
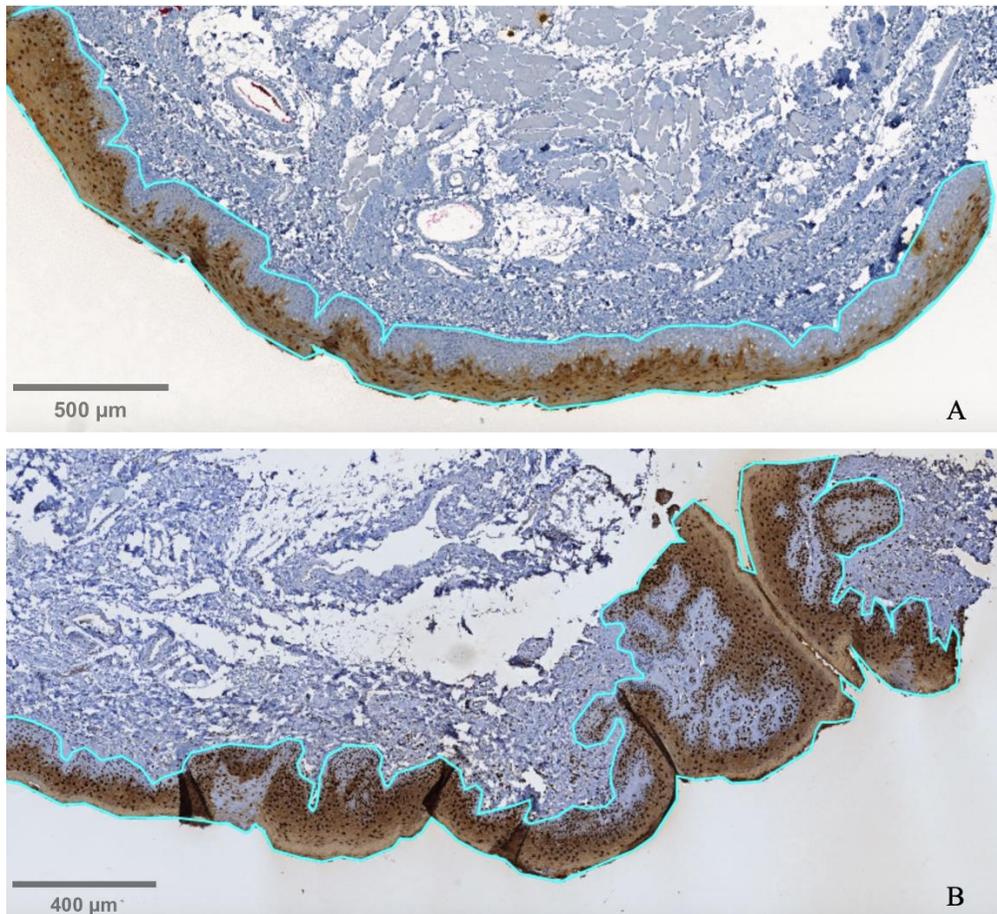
Figure 3.14: Variability in S100A7 Protein Expression Among Cases with Moderate Dysplasia.
A) Image moderate dysplasia with 39.14% S100A7 cytoplasmic staining intensity. B) Image of another case within the same moderate dysplasia group, showing higher S100A7 expression with 78.14% cytoplasmic staining intensity. Both images demonstrate the variability in S100A7 expression between cases with the same dysplasia level. Images were captured at 20x magnification using QuPath.

S100A7 Protein Level in Control and OPMDs Samples

Immunostaining of S100A7 protein in oral epithelium region were compared between Control and OPMDs groups using Welch's t-test. The analysis revealed a significant difference in the mean expression levels between the two groups. The mean protein expression in the Control group was 0.2123, while in the OPMDs group, the expression level was significantly higher at 0.5307. Welch's t-test indicated that this difference was statistically significant, with a t-value of -5.2492, degrees of freedom of 45.136, and a p-

value of 3.97x10^-6. The 95% confidence interval for the difference in means ranged from -0.4406 to -0.1963, suggesting that the true mean difference between the Control and OPMDs groups is not equal to zero. Figure 3.15 represents the distribution of S100A7 protein expression in the Control and OPMDs groups, with the OPMDs group showing a notably higher median expression level.
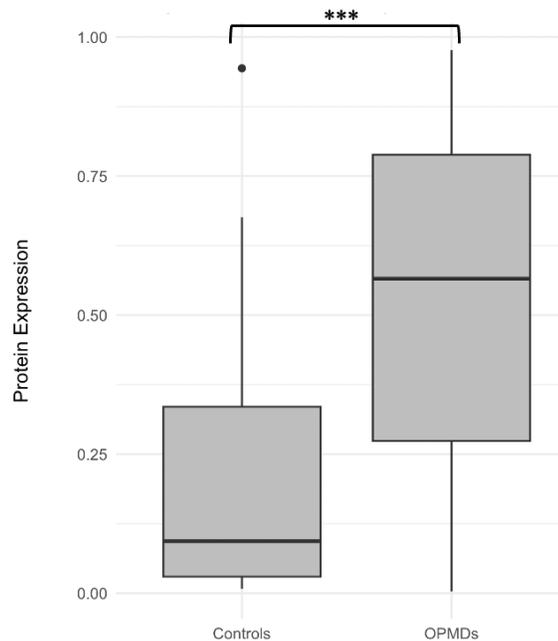
Figure 3.15: S100A7 protein expression in Control and OPMDs groups.
The grey boxes represent the interquartile range, and the horizontal lines indicate the median of expression levels. Black dot represents the outlier. *** indicates $p < 0.001$ based on Welch's t-test.

S100A7 Protein Expression Between Dysplasia Groups

I further investigated whether S100A7 protein expression correlates with the severity of dysplasia within the OPMDs group (**Figure 3.16**). ANOVA test did not show a statistically significant difference in S100A7 protein expression levels among the dysplasia groups ($p = 0.30$). The mean square error within groups (0.07589) was slightly lower than the mean square between groups (0.09381), indicating that dysplasia group does not significantly explain the variation in S100A7 protein expression. Post-hoc Tukey HSD tests were conducted to further examine pairwise differences in protein expression between dysplasia groups, and none of them had any significant difference

(mild vs. moderate p = 0.40; mild vs. severe p = 0.49; moderate vs. severe p = 0.99). The overall oral epithelial S100A7 expression had no significance between the three dysplasia groups, I further analyzed the positive nuclear expression by counting the positive stained nuclei within the oral epithelium. Like the results from the overall epithelial expression, there was no significant difference in S100A7 nuclear expression levels among the dysplasia groups (p = 0.55). My findings suggest that dysplasia severity does not influence S100A7 protein expression levels significantly in the studied population.
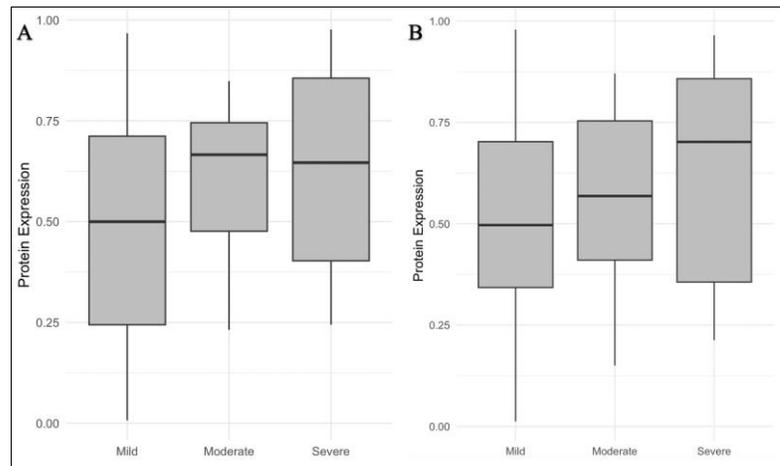
Figure 3.16: Positive S100A7 epithelial expression and nuclear expression measured across three dysplasia groups.
Epithelial expression (A) was measured based on positively stained area size in QuPath, and nuclear expression (B) was measured by counting individual cell nucleus in QuPath. Mild (n=29), moderate (n=15), and severe (n=6). No statistical significance found between the three-dysplasia stage in S100A7 expression.

3.4.2   β -Catenin Levels

β-Catenin Expression between Control and OPMDs Groups

The average β-Catenin expression was 0.4214 in the Controls group and 0.4187 in the OPMDs group. These findings suggest that there is no significant difference β-Catenin expression between the two groups (**Figure 3.17**). Since here was no identifiable difference on β-Catenin staining between health and diseased individuals, quantitative evaluation was not performed on various grades of dysplasia cases. Figure 3.18 shows β-Catenin immunostaining in oral epithelium of control and dysplasia cases.
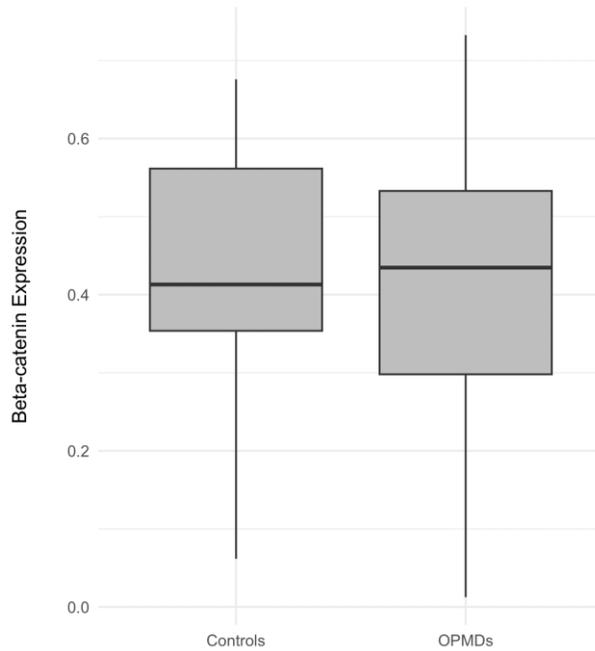
Figure 3.17: Boxplot of β-Catenin Expression.
The analysis yielded no significant difference between the two groups (p = 0.94). There were no outliers noted.
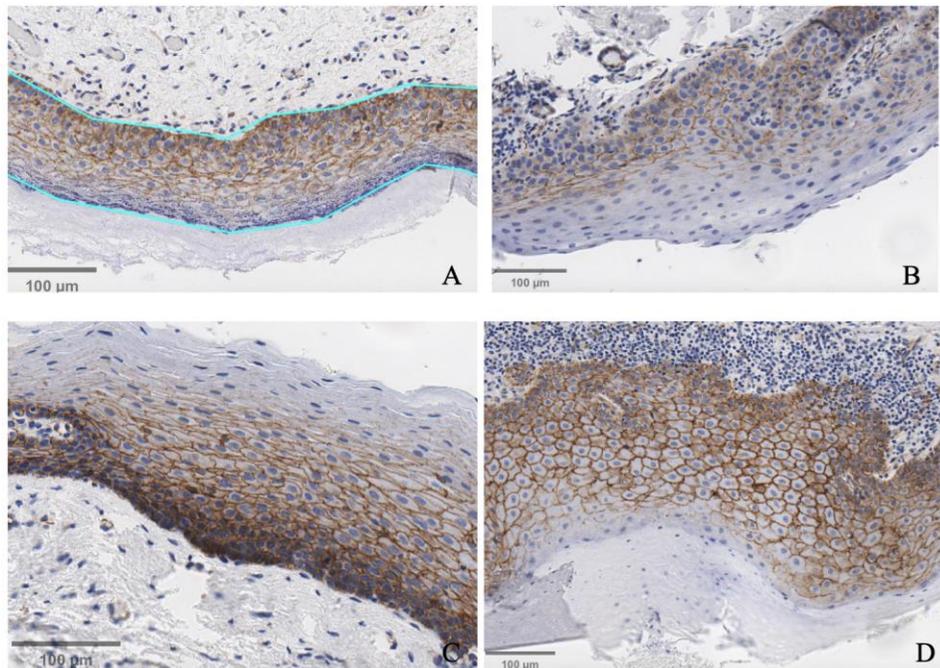
Figure 3.18: β-Catenin Immunohistochemical Staining in Oral Epithelium Across Different Dysplasia Levels.
Images showing β-catenin staining in the oral epithelium for different dysplasia levels. A) Control case. B) Mild dysplasia. C) Moderate dysplasia. D) Severe dysplasia. All images were captured at 20x magnification.

3.4.3   E-cadherin Levels

E-cadherin Expression between Control and OPMDs Groups

Based on the Welch's two-sample t-test, the mean E-cadherin expression in the Control population was significantly higher than OPMDs group (**Figure 3.19**). The mean E-cadherin expression as 0.754 (75.4%) in the Controls group, and 0.638 (63.8%) in the OPMDs group, with a 95% confidence interval for the difference in means ranging from 0.055 to 0.176. Figure 3.20 shows that the severe dysplasia case exhibits reduced E-cadherin staining around the cell boarders compared to the control case.

Figure 3.19: Boxplot of E-cadherin Expression.
The grey boxes represent the interquartile range, and the horizontal lines indicate the median of expression levels. Black dot represents the outlier. *** indicates $p < 0.001$ based on Welch's t-test.

Figure 3.20: Comparison of E-cadherin Staining Between a Control Case and a Severe Dysplasia Case.
A) Control case shows E-cadherin staining with an intensity of 89.2%, while B) severe dysplasia case exhibits reduced staining intensity at 65.6%. Images were scanned at 20x magnification.

E-cadherin Expression Between Dysplasia Groups

ANOVA test between the dysplasia groups had a F-value of 2.45 and p-value of 0.101.

The mean E-cadherin expression for severe dysplasia was lower compared to the other

two groups, but according to Tukey's HSD test it was not statistically different (**Figure**

3.21). This observation, although not statistically validated, suggests a potential trend that

could be worth exploring further with a larger sample size or additional biomarkers.

Figure 3.21: E-cadherin Expression Between Three Dysplasia Groups.
The results from the ANOVA and Tukey HSD test indicated that there were no significant differences in E-cadherin expression levels across the dysplasia groups. Black dots represent outliers. The p-values for all comparisons were higher than 0.05.

3.4.4   Vimentin Levels

Vimentin Expression between Control and OPMDs Groups

In oral potentially malignant disorders, elevated vimentin expression in the oral epithelial layer can suggest cellular changes since the protein is primary expressed in connective tissue. In Figure 3.22 Vimentin expression was observed within the epithelial layer, localized in clusters of cells. Based on the Welch's t-test, there was significant differences in Vimentin expression between the Controls and OPMDs groups (t = -3.8703, df = 94.741, p = 0.00019). Vimentin expression in the OPMDs group was almost double than the Controls group (mean = 0.071 and 0.036 respectively) (**Figure 3.23**).

Figure 3.22: Vimentin Expression Pattern.
Pockets of cells within the oral epithelium expressing Vimentin in a moderate dysplasia case. Image was scanned at 20x magnification and captured in QuPath.
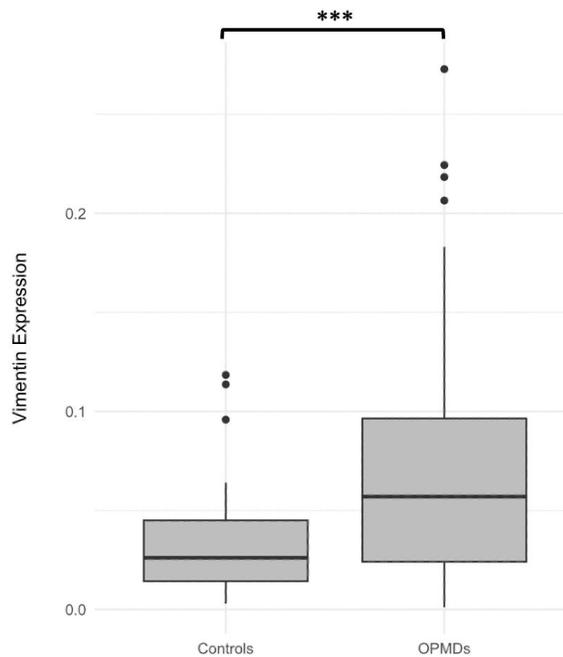
Figure 3.23: Boxplot of Vimentin Expression.
The grey boxes represent the interquartile range, and the horizontal lines indicate the median of expression levels. Black dots mean outliers. *** indicates $p < 0.001$ based on Welch's t-test.

Vimentin Expression Between Dysplasia Groups

ANOVA test did not show a statistically significant difference in Vimentin protein expression levels among the dysplasia groups (p = 0.65). Tukey HSD tests were conducted to further examine pairwise differences in protein expression between dysplasia groups, and none of them had any significant difference (mild vs. moderate p = 0.91; mild vs. severe p = 0.65; moderate vs. severe p = 086) (**Figure 3.24**).
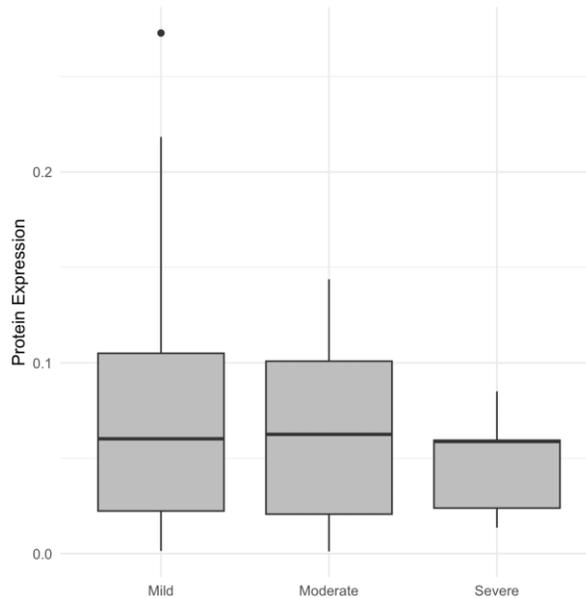
Figure 3.24: Vimentin Expression Between Three Dysplasia Groups.
The results from the ANOVA and Tukey HSD test indicated that there were no significant differences in Vimentin expression levels across the dysplasia groups. The black dot represents an outlier in the mild dysplasia population. The p-values for all comparisons were higher than 0.05.

3.4.5   Ki67, Geminin, MCM2 Levels

The three biomarkers related to cell cycle progression were compared between the Controls and OPMDs groups. Figure 3.25 indicates only Ki67 expression differed significantly between the two groups (Welch's t-test, t = -6.78, df = 77.64, p < 0.001). The mean Ki67 expression was 0.089 (8.9%) in the control group and 0.183 (18.3%) in the OPMDs group. Positive Ki67 cells were mostly located in the basal layer for the Control group, while positive cells were detected higher up in the oral epithelium for the OPMD group (**Figure 3.26**). Both Geminin and Mcm2 had elevated expression in the

OPMDs group compared to Controls, but the difference was not statically significant. The mean geminin expression was 0.045 (4.5%) in Controls and 0.061 (6.1%) in OPMDs (p = 0.23). For Mcm2, the mean expression was 0.056 (5.6%) in controls and 0.065 (6.5%) in OPMDs (p = 0.41). Further analysis was not performed on various grades of dysplasia cases for Geminin and Mcm2.
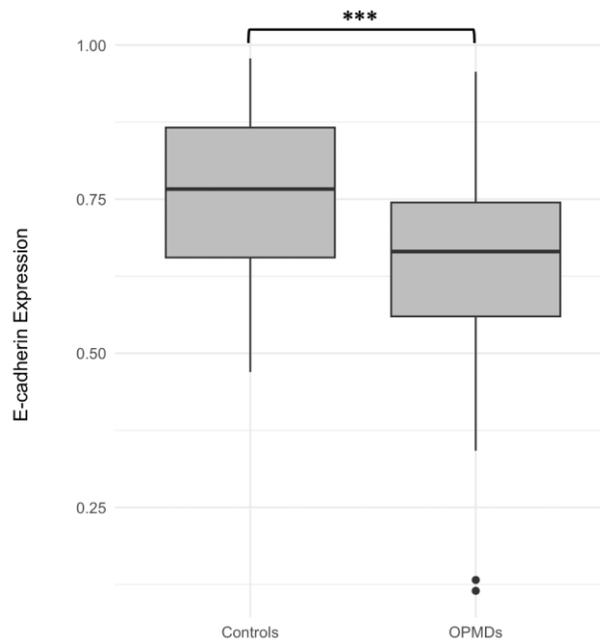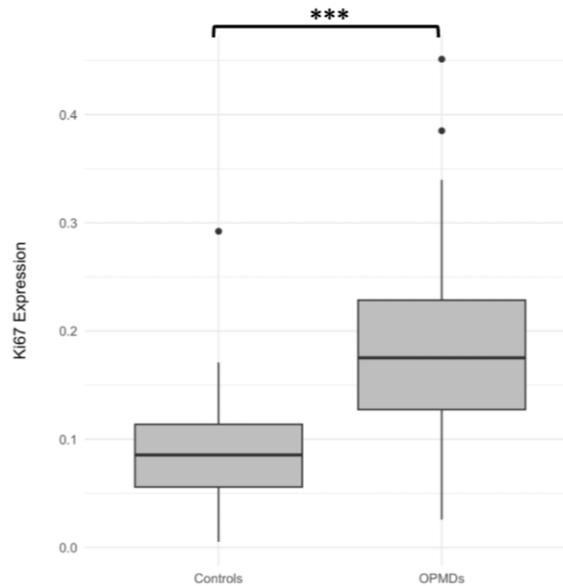

Figure 3.25: Boxplot of Ki67 Expression.
The grey boxes represent the interquartile range, and the horizontal lines indicate the median of expression levels. Black dots represent the outlier in the data. *** indicates p < 0.001 based on Welch's t-test.
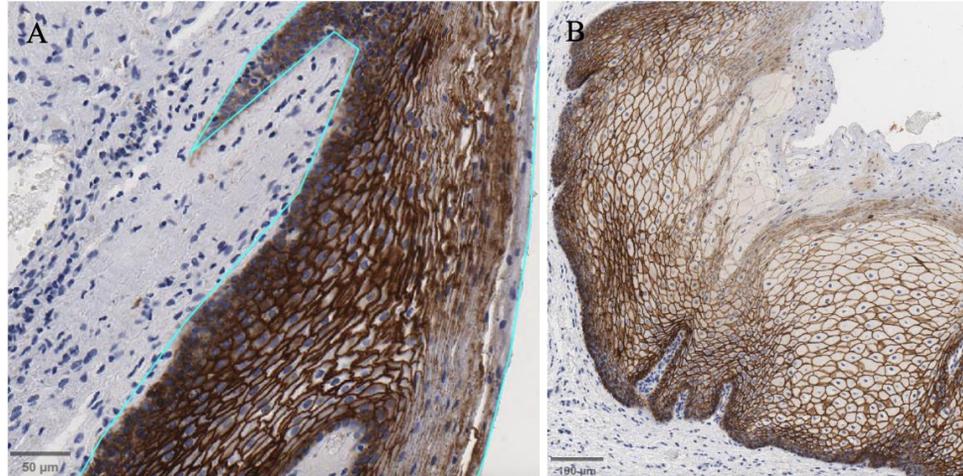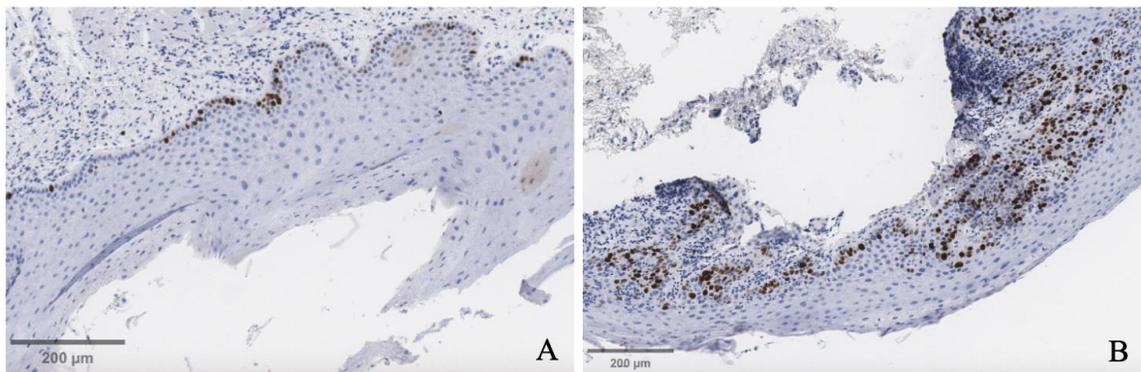

Figure 3.26: Comparison of Ki67 Staining Between a Control Case and a Severe Dysplasia Case.
A) Control case shows low (positive cell = 1.51%) Ki67 staining in the basal layer while B) severe dysplasia case exhibits increased positive cell percentage (21.85%) throughout the oral epithelium. Images were scanned at 20x magnification.

Ki67 Expression Between Dysplasia Groups

In my analysis of Ki67 expression across the dysplasia groups, I conducted an ANOVA to determine if there were any significant differences in mean expression levels; F value of 0.56 and a p-value of 0.57 suggesting that the variations in Ki67 expression across the groups were likely due to random chance. The Tukey HSD test for pairwise comparisons confirmed the ANOVA findings. Specifically, the comparisons between moderate and mild (diff: -0.031, p adj: 0.57), severe and mild (diff: -0.024, p adj: 0.85), and severe and moderate (diff: 0.0083, p adj: 0.98) showed no statistically significant differences, as all adjusted p-values were higher the 0.05 significance threshold. Interestingly, despite the lack of statistical significance, the boxplot (**Figure 3.27**) revealed that the severe dysplasia group had a visibly higher mean Ki67 expression compared to the mild and moderate groups. This observation, although not statistically validated, suggests a potential trend that could be worth exploring further with a larger sample size or additional biomarkers.
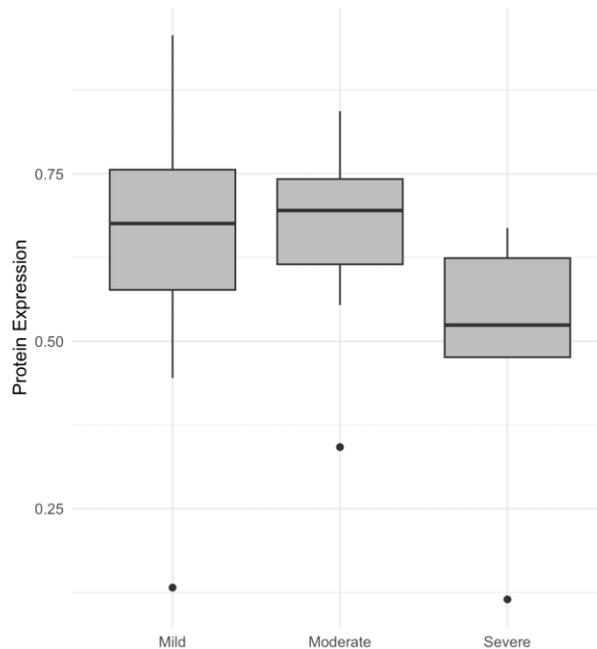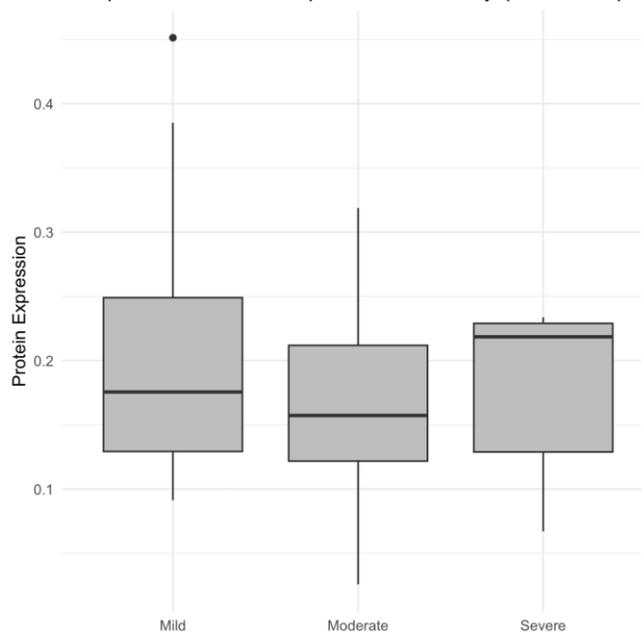

Figure 3.27: Ki67 Expression Between Three Dysplasia Groups.
The results from the ANOVA and Tukey HSD test indicated that there were no significant differences in the expression levels across the dysplasia groups. The black dot represents an outlier in the mild dysplasia population. The p-values for all comparisons were higher than 0.05.

4   Discussion

The identification of biomarkers is crucial for the early diagnosis of oral potentially malignant disorders, as they can provide specific and sensitive indicators of disease presence and progression. I will further discuss how the selected biomarkers in this study can aid in the early detection of OPMDs, and how the expression level of certain proteins correlated with the severity of disease. Additionally, clinical diagnosis through visual examination and clinical symptoms remains a cornerstone of early detection, indicating the importance of proper training for primary healthcare providers to accurately identify OPMDs. By incorporating gene analysis and IHC staining to detect protein levels, my research offers a more precise diagnostic approach that could complement the traditional clinical method. However, my study has limitations that should be considered, such as the small sample size for some OPMD subtypes and the limited geographic area where the samples were collected. Finally, future research can be directed using my data on biomarker expressions and further explore the integration of these biomarkers into routine clinical practice.

4.1   Population Demographics

4.1.1   Age

I calculated the average age at initial biopsy for each of the Control and OPMDs populations. The average age and age range of the Control group was 59 years with a range of 21-82 years, while the OPMDs group was 62 years, with a range of 12-92 years. My results align with a study done in Brazil reported that around 95% of OPMDs and OSCC occurs in individuals over the age of 40 (Martins-de-Barros et al., 2021). This underscores the importance of monitoring and screening middle age and older adults for these oral conditions. On the other hand, both groups had young patients that were identified as outliers based on the statistical test. This raises important questions about early-onset factors and the potential influence of genetic factors contributing to the development of OPMDs at a younger age.

4.1.2   Sex / Gender-based Risk Factors

Based on the sample population in my study, the number of male and female subjects were relatively equal in the OPMDs group, while the Controls had a slightly higher number of males. Within the dysplasia population, the number of females was slightly higher (no significance) than males in mild and moderate dysplasia, and equal to males in severe dysplasia. Although previous literature indicated that men are often more frequently affected by OPMDs compared to women, it is likely due to their higher rates of tobacco and alcohol consumption, and other gender-linked habits (Kumari et al., 2022). This study found no significant difference in the prevalence of OPMDs between male and females; and the number of female smokers was almost equal to the number of male smokers. This could be partly due to the similar social behaviors between men and women in the Ontario populations, since the OPMDs cases were received from two oral surgeons based in Ottawa and Newmarket Ontario. These findings highlight the importance of considering the prevalence of OPMDs and associated risk factors in diverse geographical regions and in in different populations. Many studies have also investigated the effect of alcohol consumption on OPMDs, however due to the lack of patient record data on alcohol history, I was unable to draw any conclusion on the subject. To address this gap and enhance our understanding of the relationship between alcohol consumption and OPMDs, primary healthcare professionals should implement comprehensive data collection protocols during patient's initial visit.

4.2   Diagnostic Concordance

In this study, I analyzed the concordance between clinical diagnoses made by oral surgeons and final diagnoses provided by oral pathologists based on histopathological examination of biopsy samples. This section discusses the findings and implications of this diagnostic concordance. The agreement was quantified using Cohen's Kappa statistic test, and the moderate level of agreement suggests that while clinical diagnoses are reasonably reliable, there is still room for improvement through enhanced training and use of additional diagnostic tools. This underlines the complexity of diagnosing OPMDs based solely on clinical examination and highlights the importance of histopathological

confirmation. The higher prevalence of dysplasia in histopathological findings compared to clinical diagnoses suggests that some lesions initially categorized as benign or uncertain clinically may contain dysplastic changes detected only through biopsy and histological analysis. Improving clinical diagnostic accuracy through enhanced clinical training could potentially reduce discrepancies between clinical and histopathological diagnoses. For this study, I simplified the diagnoses into two nominal variables: dysplasia and non-dysplasia. While this provided valuable insights to my study, future studies should expand diagnostic evaluations to encompass all possible outcomes in a larger sample population. Additionally, comparison among clinicians and pathologists can improve understanding about inter-observer variability. In conclusion, while clinical diagnoses provide valuable preliminary insights, histopathological examination remains essential for accurate diagnosis and management of OPMDs. Enhancing diagnostic protocols and interdisciplinary collaboration between primary physicians and pathologists are crucial for improving patient care and early detection of OPMDs.

4.3   Biomarkers

4.3.1   Gene Analysis

The differential expression of the mRNA targets highlights their potential roles in the pathogenesis and progression of oral potentially malignant disorders (OPMDs). Changes in expression levels of certain mRNAs indicate their involvement in inflammatory, immune, and signaling responses, which are crucial in the early stages of malignant transformation. These findings suggest that these mRNAs may be key players in the cellular mechanisms driving the transition from normal oral mucosa to dysplasia and, potentially, to oral squamous cell carcinoma (OSCC). From the NanoString data, there is a clear increase in significant changes in mRNA expression related to dysplasia stage, underscoring the progressive nature of the disease. The data reveal distinct patterns of upregulation and downregulation of specific mRNAs as dysplasia advances from mild to severe stages. The correlation between mRNA expression levels and dysplasia severity supports the hypothesis that these mRNAs contribute to the early events in malignant

transformation and may serve as potential biomarkers for assessing disease progression and identifying high-risk lesions.

S100A7 stands out due to its significantly elevated expression in dysplastic tissues compared to other targets. The nCounter analysis revealed that S100A7 mRNA expression is highly elevated across all dysplasia tissue compared to normal tissue. This substantial increase in S100A7 expression suggests its potential as a biomarker for distinguishing between normal and dysplastic lesions in OPMDs. Given the importance of early detection in improving patient outcomes, the identification of S100A7 as a biomarker is particularly relevant for early detection and risk stratification. High S100A7 expression could indicate a higher risk of malignant transformation, providing valuable information for clinicians in developing targeted intervention strategies. Interestingly, there have been no previous studies specifically examining the mRNA expression of S100A7 in OPMD lesions. Most existing research has focused on the detection of S100A7 protein expression using IHC. This study is among the first to explore the mRNA expression of S100A7 in OPMDs, providing a novel perspective on its role in the disease process.

In addition to S100A7, the elevated expression of CD14, CD3E, and JAK3 mRNAs further supports the role of immune and inflammatory responses in the progression of OPMDs. CD14 is a surface antigen preferentially expressed on monocytes/macrophages associated with the innate immune response regulation. A previous study on the increase in immune cell infiltration with disease progression reported a higher percentage of CD14 positive cells was associated with more advanced epithelial transformation, and this increase was also seen in epithelium underwent malignant transformation (Gannot et al., 2002). The over expression of CD14-mRNA observed in my study indicates an active immune environment within the epithelium. CD3E is a critical component of T-cell receptor signaling. A study in 2021 looked at the tumor immune microenvironment (TME) in head and neck squamous cell carcinomas (HNSCC). The study identified 407 immune-related genes and CD3E was one of the genes showing prominent differential expression (Chengcheng et al., 2021). Another study reported CD3E to be an indicator of bladder cancer TME regulation, since CD3E is closely related to immune infiltration (Y.

Liu et al., 2021). Overexpression of CD3E in dysplastic tissue may indicate the early development of TME within the oral epithelium. JAK3 is involved in the cytokine signaling JAK-STAT pathway, and its over expression also indicated inflammatory response in the oral epithelium. A study by Khan *et al*. utilized total RNA-sequencing and advanced computational analyses to uncover the mechanisms driving OSCC development from potentially malignant lesions. They also found significant early activation of IL6/JAK-STAT3 signaling pathway, confirming the important role played by the immune system in malignant transformation (Khan et al., 2023).

Conversely, the downregulation of MAP3K1-mRNA in dysplastic tissues points to potential disruptions in signaling pathways that may contribute to disease progression. MAP3K1 (mitogen-activated kinase kinase kinase 1) is involved in various cellular processes including apoptosis and stress responses, and has been used as a therapeutic target in breast cancer (C. Liu et al., 2016). MAP3K1 was reported to be regulating both the canonical and non-canonical Wnt pathways through distinct functional domains. Specifically, MAP3K1 acts as an upstream activator of β-catenin in the Wnt signaling cascade (Sue Ng et al., 2010). The observation of low MAP3K1 expression in OPMDs contrasts with findings in other cancer literature, suggesting a need for further in-depth investigation.

One of the biggest advantages of using NanoString nCounter technology was the ability to precisely measure the mRNA levels directly from FFPE tissue sections. This precision is particularly important when extracting RNA from FFPE cores, as it ensures that only the epithelial tissue is sampled, excluding any nearby connective tissue that could confound the results. However, it is important to note that while the panel included 770 gene targets, not all genes were equally informative for my specific research aims. Many genes within the panel did not exhibit significant differential expression or relevance to OPMD population in my study cohort. Many genes within the panel were specifically designed to target immune cells, and less for detecting pathway activations. The panel also did not detect significant differential expression to the epithelial changes observed in OPMDs. This limitation highlights the need for more targeted gene panels that are specifically tailored to the study of epithelial biomarkers.

4.3.2   Protein Expression in Oral Epithelium

In this study, I investigated the expression levels of various proteins in the oral epithelium using IHC technique to understand their potential roles as biomarkers for OPMDs. Significantly elevated expression levels were observed for S100A7, Ki67, and Vimentin in the OPMD group compared to the normal control group; however, further comparison between the expression levels in different dysplasia groups showed no significance. Conversely, E-cadherin exhibited significantly lowered expression in the OPMD group. β-catenin, Geminin, and MCM2 did not exhibit significant changes in expression between the normal control and OPMD groups.

The elevated protein expressions are consistent with existing literature, suggesting that these proteins are involved in cell proliferation and are upregulated in OPMD lesions as part of the disease progression. S100A7 is known for its role in inflammation and epithelial differentiation, and its overexpression in OPMDs may reflect the inflammatory microenvironment and altered differentiation status of the diseased oral lesions. In other studies, S100A7 has also been identified as a marker for invasion in premalignant oral epithelium by comparing the nuclear, cytoplasmic and membrane staining as well as the staining intensity had statistically significant different scoring patterns among the OSCC group and OPMD group (Sood et al., 2022). Recent research has demonstrated that the increased in S100A7 can also be detected in saliva, making it a promising biomarker in identifying OPMDs and potential malignant transformation (Raffat et al., 2018). Interestingly, there is variability in S100A7 staining within cases of a single dysplasia group. This variability could be attributed to differences in individual patient responses, such as variations in the local inflammatory environment, genetic factors, or differences in the stage and progression of dysplasia at the time of biopsy. Such variability highlights the complexity of the disease and underscores the need for further investigation to understand the factors contributing to these differences. Identifying and accounting for these variations could improve the accuracy of S100A7 as a biomarker and enhance its utility in clinical settings for monitoring disease progression and response to treatment.

Ki67 is a well-established marker of cell proliferation, and its increased levels lead to an increased state of proliferation. A study by Shailaja *et. al* on Ki67 expression in oral lichen planus (OLP) and oral epithelial dysplasia (OED) using immunohistochemistry reported that Ki67 expression was similar in male and female population, while the OED group had significantly higher expression than OLP group (Shailaja et al., 2015). The elevated Ki67 expression observed in dysplastic lesions highlights its potential role as a biomarker for identifying and monitoring OPMDs.

Vimentin was found to have a higher expression in the OPMD tissues in this study. Upon examining the IHC slides, it was noted that vimentin expression in the epithelium was primarily observed in clusters of cells. This pattern could potentially indicate the presence of inflammatory cells within the oral epithelium and/or connective tissue papillae. The significant inflammation in the underlying connective tissue may have contributed to the influx of these vimentin-positive inflammatory cells into the epithelial layer. A similar study on inflammatory reaction (IR) in oral epithelial dysplasia (OED) found that the IR positively correlated with the expression of Vimentin; previous research also reported significantly higher Vimentin expression in severe OED compared to mild or moderate OED (Miguel et al., 2021). Although my data did not show a significant difference in Vimentin expression among severe, moderate, and mild OED, this was likely due to the small sample size.

Contrary to the trend observed previously, the key protein in cell-cell adhesion, E-cadherin, showed reduced levels of expression. This observation indicated the loss of cellular junction and epithelial cell phenotype. Yuwanati *et al.* conducted an in vivo study to examine E-cadherin expression in OPMD and OSCC. They found that E-cadherin played a significant role in the progression from OPMD to OSCC. Following the comparison of E-cadherin expression across normal oral mucosa, OPMD, and OSCC, they concluded that decreased E-cadherin expression might be a useful indicator for the transformation of OPMD into OSCC (Yuwanati et al., 2011).

β-catenin, Geminin, and MCM2 did not exhibit significant changes in expression between the normal control and OPMD groups. Although my study did not yield

significant results regarding the expression of MCM2 and Geminin, these proteins are essential for the normal regulation of the cell cycle. Both MCM2 and Geminin have been previously studied in the context of OED and OSCC, highlighting their relevance in the pathogenesis and progression of these conditions. A study in England looked at the expression patterns of Mcm2, Ki67, and Geminin in normal oral mucosa, OED, and their subsequent progression to OSCC. The findings revealed that Mcm2 expression was significantly higher (P=0.04) in OED cases that progressed to OSCC compared to those that did not progress; but Geminin expression was not significantly different between the two groups (Torres-Rendon et al., 2009). Research also indicated increased levels of Geminin and MCM2 in other cancers. For instance, a study on renal cell carcinoma (RCC) found that increased tumor grade was associated with elevated expression of MCM2 and Geminin; also an increasing MCM2 − Ki67 ratio was associated with reduced disease-free survival, suggesting that the proportion of noncycling MCM2-expressing tumor cells strongly influences tumor growth potential (Dudderidge et al., 2005). A similar IHC study on β-catenin expression reveled that the expression of β-catenin increased progressively from mild to moderate to severe epithelial dysplasia, with localization shifting from the membrane to the cytoplasm and eventually to the nucleus(Chowdhury et al., 2021). The stable expression of these proteins observed in my study suggests that the pathways they are involved in may not be prominently altered in the early stages of OPMDs. Further studies can dive deeper into the role of β-catenin, Geminin, and MCM2 in OPMDs by conducting longitudinal studies to observe the expression levels of these proteins over time as the disease progress, which can provide insights into whether these pathways become altered at later stages.

4.4   Limitations

My study has several limitations that need to be addressed to provide a better understanding on early detection of OPMDs and potentially improve diagnostic and treatment strategies. One of the primary limitations of my study was the small sample size. The limited number of cases restricts the generalizability of the findings and may not accurately represent the patient population. Similarly, the results on the three dysplasia levels were constrained by the small and unequal sample sizes in each group;

especially, the severe dysplasia group had only 6 cases, which significantly impacts the statistical power and the ability to detect meaningful differences between groups. Geographic limitation was another issue since all cases were received from two locations in Ontario, Canada: Newmarket and Ottawa. This hinders the ability to observe trends and patterns that may be present in other regions, thus limiting the study's external validity. All OPMD cases were initially identified by oral surgeons as suspicious lesions and sent to the oral pathology division for confirmation. This raises questions about whether the surgeons can accurately identify OPMDs clinically; because the accuracy of initial clinical diagnoses may vary based on the training and experience of the individual and can potentially affect the consistency and reliability of the clinical identification of OPMDs. Finally, improving the collection of patient demographic information, such as age, gender, ethnicity, and lifestyle factors, would enhance the understanding of the risk factors of OPMDs.

During the experimentation step of my study, I encountered a few technical limitations with the NanoString nCounter system. First, the cost of consumables and reagents was relatively high compared to other gene analysis methods, making it very expensive to utilize; and I was only able to conduct gene analysis on a single panel (12 samples) due to the high cost. Another limitation of the nCounter system was the requirement for high-quality RNA samples. In my study, the RNA samples needed to be extracted from FFPE tissue samples. RNA degradation occurs slowly within the FFPE tissue, the older the tissue sample the harder is it to extract and purify the RNA sample. Future studies should consider using fresh tissue samples or employing advanced RNA extraction techniques specifically designed to handle degraded RNA. Lastly, while the nCounter pre-designed panel had 770 gene targets including S100A7, it did not cover some other biomarkers of interest in my research. Although it is possible to create custom panels containing all the protein of interest, this adds significant cost and complexity to the study, involving time-consuming and expensive processes for designing, validating, and optimizing new probes.

Working with digitalized tissue slides was the most time-consuming part of my research. The QuPath software had difficulties in accurately distinguishing between tissue

components such as the oral epithelium and connective tissue underneath using automated tools. To accurately outline the epithelium, I used the annotate function and manually outlined the region of interest, this was the most time-consuming step depending on the size of the epithelium of the tissue sample. Additionally, adjusting staining intensity manually for each protein during IHC analysis further added to the time and complexity of the analysis process. Enhancing the software's ability to precisely differentiate between tissue components would significantly reduce the need for manual adjustments making the image analysis step much more efficient.

4.5 Future Work

Identifying biomarkers that can sufficiently distinguish OPMD lesions from healthy tissues can have significant impact on disease management and increase surveillance for OPMD patients to monitor disease progression. Future research should focus on integrating multiple biomarkers, such as S100A7, E-cadherin, and vimentin, to create a more robust predictive model for OPMD detection and prevent malignant transformation. The application of digital pathology holds promising future in enhancing the early detection of subtle epithelial changes associated with OPMDs. Utilizing machine learning algorithms and artificial intelligence (AI) models to analyze large-scale molecular and clinical datasets could facilitate the development of predictive tools for OPMDs. By cross comparing these biomarker expressions across different OPMD samples, AI can identify subtle but significant correlations and patterns that may indicate varying degrees of malignant potential. Further integration of biomarker analysis with comprehensive clinical data, such as patient demographics, medical history, and follow-up outcomes, could ultimately result in early disease detection and improved patient care.

Chapter 5

5   Conclusion

The identification of biomarkers is crucial for the early diagnosis of oral potentially malignant disorders due to their role in providing specific and sensitive indicators of disease presence and progression. By integrating gene analysis and immunohistochemical staining into my study, it offers a more precise diagnostic approach that complements the traditional clinical method.

One of the key findings of my research is the elevated expression of S100A7 mRNA in dysplastic tissues, which highlights its potential as a biomarker for distinguishing between normal and dysplastic lesions in OPMDs. Similarly, the increased levels of Ki67 and Vimentin suggest their roles in cell proliferation and inflammation, respectively, further emphasizing their utility in identifying OPMDs. Conversely, the decreased expression of E-cadherin indicates disruptions in cellular adhesion and signaling pathways, further indicating their roles in disease development. My study's findings also emphasize the importance of proper training for primary healthcare providers to accurately identify OPMDs through visual examination and clinical symptoms. Enhanced training, combined with molecular and protein expression analyses, could greatly improve diagnostic accuracy and patient outcomes.

Despite the promising results, my study faced several limitations. The small sample size, particularly within specific dysplasia subtypes, and the geographic restriction to samples from Ontario, Canada, limit the generalizability of the findings. Additionally, the high cost and technical challenges associated with the NanoString nCounter system and the manual annotation required for digitalized tissue slides presented practical hurdles. Addressing these limitations in future research could enhance the robustness and applicability of biomarker-based diagnostic strategies.

Moving forward, the integration of multiple biomarkers, such as S100A7, E-cadherin, and Vimentin, could lead to the development of a comprehensive predictive model for

OPMD detection. The application of digital pathology and machine learning algorithms holds great promise in refining these models and facilitating the early detection of subtle epithelial changes. By incorporating extensive clinical data and leveraging artificial intelligence, future studies can build on my findings to improve the early diagnosis, monitoring, and management of OPMDs, ultimately enhancing patient care and outcomes.

References

Ali, K. (2022). Oral cancer—The fight must go on against all odds…. *Evidence-Based Dentistry*, *23*(1), 4–5. https://doi.org/10.1038/s41432-022-0243-1

Birajdar, S., Radhika, M., Paremala, K., Sudhakara, M., Soumya, M., & Gadivan, M. (2014). Expression of Ki-67 in normal oral epithelium, leukoplakic oral epithelium and oral squamous cell carcinoma. *Journal of Oral and Maxillofacial Pathology*, *18*(2), 169. https://doi.org/10.4103/0973-029X.140729

Borcherding, N., Cole, K., Kluz, P., Jorgensen, M., Kolb, R., Bellizzi, A., & Zhang, W. (2018). Re-Evaluating E-Cadherin and β-Catenin. *The American Journal of Pathology*, *188*(8), 1910–1920. https://doi.org/10.1016/j.ajpath.2018.05.003

Brocklehurst, P., Kujan, O., O'Malley, L., Ogden, G. R., Shepherd, S., & Glenny, A.-M. (2013). Screening programmes for the early detection and prevention of oral cancer. *Cochrane Database of Systematic Reviews*, *2021*(3). https://doi.org/10.1002/14651858.CD004150.pub4

Brocklehurst, P., Pemberton, M. N., Macey, R., Cotton, C., Walsh, T., & Lewis, M. (2015). Comparative accuracy of different members of the dental team in detecting malignant and non-malignant oral lesions. *British Dental Journal*, *218*(9), 525–529. https://doi.org/10.1038/sj.bdj.2015.344

Carreras-Torras, C., & Gay-Escoda, C. (2015). Techniques for early diagnosis of oral squamous cell carcinoma: Systematic review. *Medicina Oral Patología Oral y Cirugia Bucal*, e305–e315. https://doi.org/10.4317/medoral.20347

Chau, L., Jabara, J. T., Lai, W., Svider, P. F., Warner, B. M., Lin, H.-S., Raza, S. N., & Fribley, A. M. (2017). Topical agents for oral cancer chemoprevention: A systematic review of the literature. *Oral Oncology*, *67*, 153–159. https://doi.org/10.1016/j.oraloncology.2017.02.014

Chengcheng, L., Wenwen, Q., Ningyue, G., Fangyuan, Z., Runtong, X., Zhenxiao, T., Fenglei, X., Yiming, Q., Miaoqing, Z., Xiaoming, L., & Ming, X. (2021). Identification of the Immune-Related Genes in Tumor Microenvironment That Associated With the Recurrence of Head and Neck Squamous Cell Carcinoma. *Frontiers in Cell and Developmental Biology*, *9*, 723721. https://doi.org/10.3389/fcell.2021.723721

Chowdhury, P., Nagamalini, B., Singh, J., Ashwini, B., & Sharada. (2021). Expression of β-catenin in oral leukoplakia and oral submucous fibrosis: An immunohistochemical study. *Journal of Oral and Maxillofacial Pathology*, *25*(1), 124. https://doi.org/10.4103/jomfp.JOMFP_41_20

Chuang, S.-L., Wang, C.-P., Chen, M.-K., Su, W. W.-Y., Su, C.-W., Chen, S. L.-S., Chiu, S. Y.-H., Fann, J. C.-Y., & Yen, A. M.-F. (2018). Malignant transformation to

oral cancer by subtype of oral potentially malignant disorder: A prospective cohort study of Taiwanese nationwide oral cancer screening program. *Oral Oncology*, *87*, 58–63. https://doi.org/10.1016/j.oraloncology.2018.10.021

Cruz, G. D., Shulman, L. C., Kumar, J. V., & Salazar, C. R. (2007). The Cultural and Social Context of Oral and Pharyngeal Cancer Risk and Control among Hispanics in New York. *Journal of Health Care for the Poor and Underserved*, *18*(4), 833–846. https://doi.org/10.1353/hpu.2007.0092

Das, P., & Deshmukh, R. (2022). Expression of S100A7 in oral potentially malignant disorders: An immunohistochemical study. *Journal of Oral and Maxillofacial Pathology*, *26*(3), 419. https://doi.org/10.4103/jomfp.jomfp_151_22

Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., … Futreal, P. A. (2002). Mutations of the BRAF gene in human cancer. *Nature*, *417*(6892), 949–954. https://doi.org/10.1038/nature00766

Dey, K. K., Bharti, R., Dey, G., Pal, I., Rajesh, Y., Chavan, S., Das, S., Das, C. K., Jena, B. C., Halder, P., Ray, J. G., Kulavi, I., & Mandal, M. (2016). S100A7 has an oncogenic role in oral squamous cell carcinoma by activating p38/MAPK and RAB2A signaling pathway. *Cancer Gene Therapy*, *23*(11), 382–391. https://doi.org/10.1038/cgt.2016.43

Diajil, A., Robinson, C. M., Sloan, P., & Thomson, P. J. (2013). Clinical Outcome Following Oral Potentially Malignant Disorder Treatment: A 100 Patient Cohort Study. *International Journal of Dentistry*, *2013*, 1–8. https://doi.org/10.1155/2013/809248

Dionne, K. R., Warnakulasuriya, S., Binti Zain, R., & Cheong, S. C. (2014). Potentially malignant disorders of the oral cavity: Current practice and future directions in the clinic and laboratory: OPMD review. *International Journal of Cancer*, n/a-n/a. https://doi.org/10.1002/ijc.28754

Dudderidge, T. J., Stoeber, K., Loddo, M., Atkinson, G., Fanshawe, T., Griffiths, D. F., & Williams, G. H. (2005). Mcm2, Geminin, and KI67 Define Proliferative State and are Prognostic Markers in Renal Cell Carcinoma. *Clinical Cancer Research*, *11*(7), 2510–2517. https://doi.org/10.1158/1078-0432.CCR-04-1776

Eckert, R. L., Broome, A.-M., Ruse, M., Robinson, N., Ryan, D., & Lee, K. (2004). S100 Proteins in the Epidermis. *Journal of Investigative Dermatology*, *123*(1), 23–33. https://doi.org/10.1111/j.0022-202X.2004.22719.x

Epstein, J. B., Güneri, P., Boyacioglu, H., & Abt, E. (2012). The limitations of the clinical oral examination in detecting dysplastic oral lesions and oral squamous cell carcinoma. *The Journal of the American Dental Association*, *143*(12), 1332–1342. https://doi.org/10.14219/jada.archive.2012.0096

Epstein, J. B., Silverman, S., Epstein, J. D., Lonky, S. A., & Bride, M. A. (2008). Analysis of oral lesion biopsies identified and evaluated by visual examination, chemiluminescence and toluidine blue. *Oral Oncology*, *44*(6), 538–544. https://doi.org/10.1016/j.oraloncology.2007.08.011

Essat, M., Cooper, K., Bessey, A., Clowes, M., Chilcott, J. B., & Hunter, K. D. (2022). Diagnostic accuracy of conventional oral examination for detecting oral cavity cancer and potentially malignant disorders in patients with clinically evident oral lesions: Systematic review and meta-analysis. *Head & Neck*, *44*(4), 998–1013. https://doi.org/10.1002/hed.26992

Feller, L., & Lemmer, J. (2012). Oral Squamous Cell Carcinoma: Epidemiology, Clinical Presentation and Treatment. *Journal of Cancer Therapy*, *03*(04), 263–268. https://doi.org/10.4236/jct.2012.34037

Field, E. A., McCarthy, C. E., Ho, M. W., Rajlawat, B. P., Holt, D., Rogers, S. N., Triantafyllou, A., Field, J. K., & Shaw, R. J. (2015). The management of oral epithelial dysplasia: The Liverpool algorithm. *Oral Oncology*, *51*(10), 883–887. https://doi.org/10.1016/j.oraloncology.2015.06.015

Gannot, G., Gannot, I., Vered, H., Buchner, A., & Keisari, Y. (2002). Increase in immune cell infiltration with progression of oral epithelium from hyperkeratosis to dysplasia and carcinoma. *British Journal of Cancer*, *86*(9), 1444–1448. https://doi.org/10.1038/sj.bjc.6600282

Goerner, M., Seiwert, T. Y., & Sudhoff, H. (2010). Molecular targeted therapies in head and neck cancer—An update of recent developements -. *Head & Neck Oncology*, *2*(1), 8. https://doi.org/10.1186/1758-3284-2-8

González-Moles, M. Á., Aguilar-Ruiz, M., & Ramos-García, P. (2022). Challenges in the Early Diagnosis of Oral Cancer, Evidence Gaps and Strategies for Improvement: A Scoping Review of Systematic Reviews. *Cancers*, *14*(19), 4967. https://doi.org/10.3390/cancers14194967

Guggenheimer, J., Verbin, R. S., Johnson, J. T., Horkowitz, C. A., & Myers, E. N. (1989). Factors delaying the diagnosis of oral and oropharyngeal carcinomas. *Cancer*, *64*(4), 932–935. https://doi.org/10.1002/1097-0142(19890815)64:4<932::AID-CNCR2820640428>3.0.CO;2-Y

He, X., Semenov, M., Tamai, K., & Zeng, X. (2004). LDL receptor-related proteins 5 and 6 in Wnt/β-catenin signaling:Arrows point the way. *Development*, *131*(8), 1663–1677. https://doi.org/10.1242/dev.01117

Ho, P.-S., Chen, P.-L., Warnakulasuriya, S., Shieh, T.-Y., Chen, Y.-K., & Huang, I.-Y. (2009). Malignant transformation of oral potentially malignant disorders in males: A retrospective cohort study. *BMC Cancer*, *9*(1), 260. https://doi.org/10.1186/1471-2407-9-260

Hwang, J. T. K., Gu, Y. R., Shen, M., Ralhan, R., Walfish, P. G., Pritzker, K. P. H., & Mock, D. (2017). Individualized five-year risk assessment for oral premalignant lesion progression to cancer. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *123*(3), 374–381. https://doi.org/10.1016/j.oooo.2016.11.004

Iocca, O., Sollecito, T. P., Alawi, F., Weinstein, G. S., Newman, J. G., De Virgilio, A., Di Maio, P., Spriano, G., Pardiñas López, S., & Shanti, R. M. (2020). Potentially malignant disorders of the oral cavity and oral dysplasia: A systematic review and meta-analysis of malignant transformation rate by subtype. *Head & Neck*, *42*(3), 539–555. https://doi.org/10.1002/hed.26006

Jäwert, F., Nyman, J., Olsson, E., Adok, C., Helmersson, M., & Öhman, J. (2021). Regular clinical follow-up of oral potentially malignant disorders results in improved survival for patients who develop oral cancer. *Oral Oncology*, *121*, 105469. https://doi.org/10.1016/j.oraloncology.2021.105469

Jemal, A., Murray, T., Samuels, A., Ghafoor, A., Ward, E., & Thun, M. J. (2003). Cancer Statistics, 2003. *CA: A Cancer Journal for Clinicians*, *53*(1), 5–26. https://doi.org/10.3322/canjclin.53.1.5

Jiang, X., Wu, J., Wang, J., & Huang, R. (2019). Tobacco and oral squamous cell carcinoma: A review ofcarcinogenic pathways. *Tobacco Induced Diseases*, *17*(April). https://doi.org/10.18332/tid/105844

Khan, M. M., Frustino, J., Villa, A., Nguyen, B.-C., Woo, S.-B., Johnson, W. E., Varelas, X., Kukuruzinska, M., & Monti, S. (2023). Total RNA sequencing reveals gene expression and microbial alterations shared by oral pre-malignant lesions and cancer. *Human Genomics*, *17*(1), 72. https://doi.org/10.1186/s40246-023-00519-y

Khariwala, S. S., Hatsukami, D., & Hecht, S. S. (2012). Tobacco carcinogen metabolites and DNA adducts as biomarkers in Head and Neck cancer: Potential screening tools and prognostic indicators. *Head & Neck*, *34*(3), 441–447. https://doi.org/10.1002/hed.21705

Kumari, P., Debta, P., & Dixit, A. (2022). Oral Potentially Malignant Disorders: Etiology, Pathogenesis, and Transformation Into Oral Cancer. *Frontiers in Pharmacology*, *13*, 825266. https://doi.org/10.3389/fphar.2022.825266

Kushwaha, P. P., Rapalli, K. C., & Kumar, S. (2016). Geminin a multi task protein involved in cancer pathophysiology and developmental process: A review. *Biochimie*, *131*, 115–127. https://doi.org/10.1016/j.biochi.2016.09.022

Le Campion, A. C. O. V., Ribeiro, C. M. B., Luiz, R. R., da Silva Júnior, F. F., Barros, H. C. S., dos Santos, K. de C. B., Ferreira, S. J., Gonçalves, L. S., & Ferreira, S. M. S. (2017). Low Survival Rates of Oral and Oropharyngeal Squamous Cell Carcinoma. *International Journal of Dentistry*, *2017*, 1–7. https://doi.org/10.1155/2017/5815493

Lingen, M. W., Kalmar, J. R., Karrison, T., & Speight, P. M. (2008). Critical evaluation of diagnostic aids for the detection of oral cancer. *Oral Oncology*, *44*(1), 10–22. https://doi.org/10.1016/j.oraloncology.2007.06.011

Liu, C., Wang, S., Zhu, S., Wang, H., Gu, J., Gui, Z., Jing, J., Hou, X., & Shao, Y. (2016). MAP3K1-targeting therapeutic artificial miRNA suppresses the growth and invasion of breast cancer in vivo and in vitro. *SpringerPlus*, *5*(1), 11. https://doi.org/10.1186/s40064-015-1597-z

Liu, J., Xiao, Q., Xiao, J., Niu, C., Li, Y., Zhang, X., Zhou, Z., Shu, G., & Yin, G. (2022). Wnt/β-catenin signalling: Function, biological mechanisms, and therapeutic opportunities. *Signal Transduction and Targeted Therapy*, *7*(1), 3. https://doi.org/10.1038/s41392-021-00762-6

Liu, L., Chen, J., Cai, X., Yao, Z., & Huang, J. (2019). Progress in targeted therapeutic drugs for oral squamous cell carcinoma. *Surgical Oncology*, *31*, 90–97. https://doi.org/10.1016/j.suronc.2019.09.001

Liu, S., Liu, L., Ye, W., Ye, D., Wang, T., Guo, W., Liao, Y., Xu, D., Song, H., Zhang, L., Zhu, H., Deng, J., & Zhang, Z. (2016). High Vimentin Expression Associated with Lymph Node Metastasis and Predicated a Poor Prognosis in Oral Squamous Cell Carcinoma. *Scientific Reports*, *6*(1), 38834. https://doi.org/10.1038/srep38834

Liu, Y., Wu, Y., Zhang, P., Xu, C., Liu, Z., He, C., Liu, Y., & Kang, Z. (2021). CXCL12 and CD3E as Indicators for Tumor Microenvironment Modulation in Bladder Cancer and Their Correlations With Immune Infiltration and Molecular Subtypes. *Frontiers in Oncology*, *11*, 636870. https://doi.org/10.3389/fonc.2021.636870

López-Knowles, E., Zardawi, S. J., McNeil, C. M., Millar, E. K. A., Crea, P., Musgrove, E. A., Sutherland, R. L., & O'Toole, S. A. (2010). Cytoplasmic Localization of β-Catenin is a Marker of Poor Outcome in Breast Cancer Patients. *Cancer Epidemiology, Biomarkers & Prevention*, *19*(1), 301–309. https://doi.org/10.1158/1055-9965.EPI-09-0741

Lorini, L., Bescós Atín, C., Thavaraj, S., Müller-Richter, U., Alberola Ferranti, M., Pamias Romero, J., Sáez Barba, M., de Pablo García-Cuenca, A., Braña García, I., Bossi, P., Nuciforo, P., & Simonetti, S. (2021). Overview of Oral Potentially Malignant Disorders: From Risk Factors to Specific Therapies. *Cancers*, *13*(15), 3696. https://doi.org/10.3390/cancers13153696

Madani, A. H., Dikshit, M., Bhaduri, D., Aghamolaei, T., Moosavy, S. H., & Azarpaykan, A. (2014). Interaction of Alcohol Use and Specific Types of Smoking on the Development of Oral Cancer. *International Journal of High Risk Behaviors and Addiction*, *3*(1). https://doi.org/10.5812/ijhrba.12120

Mariano, L. C., Warnakulasuriya, S., Straif, K., & Montiero, L. (2022). Secondhand smoke exposure and oral cancer risk: A systematic review and meta-analysis.

Tobacco Control, *31*(5), 597–607. https://doi.org/10.1136/tobaccocontrol-2020-056393

Markopoulos, A. K. (2012). Current Aspects on Oral Squamous Cell Carcinoma. *The Open Dentistry Journal*, *6*(1), 126–130. https://doi.org/10.2174/1874210601206010126

Martins-de-Barros, Av., Barros, Am., Silva, Cc., Ramos, Lf., Ferreira, Sj., Araújo, Fa., Silva, Ed., & Carvalho, Md. (2021). High prevalence of oral potentially malignant disorders and risk factors in a semi-urban brazilian city: A population-based cross-sectional study. *Medicina Oral Patología Oral y Cirugia Bucal*, e778–e785. https://doi.org/10.4317/medoral.24747

Matthews, H. K., Bertoli, C., & De Bruin, R. A. M. (2022). Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology*, *23*(1), 74–88. https://doi.org/10.1038/s41580-021-00404-3

McCarthy, C. E., Fedele, S., Ho, M., & Shaw, R. (2021). UK consensus recommendations on the management of oral epithelial dysplasia during COVID-19 pandemic outbreaks. *Oral Oncology*, *112*, 105110. https://doi.org/10.1016/j.oraloncology.2020.105110

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

McMahon, J., O'Brien, C. J., Pathak, I., Hamill, R., McNeil, E., Hammersley, N., Gardiner, S., & Junor, E. (2003). Influence of condition of surgical margins on local recurrence and disease-specific survival in oral and oropharyngeal cancer. *British Journal of Oral and Maxillofacial Surgery*, *41*(4), 224–231. https://doi.org/10.1016/S0266-4356(03)00119-0

Mehanna, H. M., Rattay, T., Smith, J., & McConkey, C. C. (2009). Treatment and follow-up of oral dysplasia - A systematic review and meta-analysis: Treatment and Follow-Up of Oral Dysplasia. *Head & Neck*, *31*(12), 1600–1609. https://doi.org/10.1002/hed.21131

Mehrotra, R., Vasstrand, E. N., & Ibrahim, S. O. (2004). Recent Advances in Understanding Carcinogenicity of Oral Squamous Cell Carcinoma: From Basic Molecular Biology to Latest Genomic and Proteomic Findings. *Cancer Genomics & Proteomics*, *1*(4), 283–294.

Mehrotra, R., & Yadav, S. (2006). Oral squamous cell carcinoma: Etiology, pathogenesis and prognostic value of genomic alterations. *Indian Journal of Cancer*, *43*(2), 60. https://doi.org/10.4103/0019-509X.25886

Mello, F. W., Melo, G., Guerra, E. N. S., Warnakulasuriya, S., Garnis, C., & Rivero, E. R. C. (2020). Oral potentially malignant disorders: A scoping review of

prognostic biomarkers. *Critical Reviews in Oncology/Hematology*, *153*, 102986. https://doi.org/10.1016/j.critrevonc.2020.102986

Mello, F. W., Melo, G., Pasetto, J. J., Silva, C. A. B., Warnakulasuriya, S., & Rivero, E. R. C. (2019). The synergistic effect of tobacco and alcohol consumption on oral squamous cell carcinoma: A systematic review and meta-analysis. *Clinical Oral Investigations*, *23*(7), 2849–2859. https://doi.org/10.1007/s00784-019-02958-1

Mello, F. W., Miguel, A. F. P., Dutra, K. L., Porporatti, A. L., Warnakulasuriya, S., Guerra, E. N. S., & Rivero, E. R. C. (2018). Prevalence of oral potentially malignant disorders: A systematic review and meta-analysis. *Journal of Oral Pathology & Medicine*, *47*(7), 633–640. https://doi.org/10.1111/jop.12726

Miguel, A. F. P., Embaló, B., Alves Dias, H. B., & Rivero, E. R. C. (2021). Immunohistochemical Expression of MMP-9, TIMP-1, and Vimentin and its Correlation With Inflammatory Reaction and Clinical Parameters in Oral Epithelial Dysplasia. *Applied Immunohistochemistry & Molecular Morphology*, *29*(5), 382–389. https://doi.org/10.1097/PAI.0000000000000910

Morrison, D. K. (2012). MAP Kinase Pathways. *Cold Spring Harbor Perspectives in Biology*, *4*(11), a011254–a011254. https://doi.org/10.1101/cshperspect.a011254

Ng, J. H., Iyer, N. G., Tan, M.-H., & Edgren, G. (2017). Changing epidemiology of oral squamous cell carcinoma of the tongue: A global study: Changing epidemiology of tongue cancer. *Head & Neck*, *39*(2), 297–304. https://doi.org/10.1002/hed.24589

Niknami, Z., Muhammadnejad, A., Ebrahimi, A., Harsani, Z., & Shirkoohi, R. (2020). Significance of E-cadherin and Vimentin as epithelial-mesenchymal transition markers in colorectal carcinoma prognosis. *EXCLI Journal; 19:Doc917; ISSN 1611-2156*. https://doi.org/10.17179/EXCLI2020-1946

Öberg, M., Jaakkola, M. S., Woodward, A., Peruga, A., & Prüss-Ustün, A. (2011). Worldwide burden of disease from exposure to second-hand smoke: A retrospective analysis of data from 192 countries. *The Lancet*, *377*(9760), 139–146. https://doi.org/10.1016/S0140-6736(10)61388-8

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

Omura, K. (2014). Current status of oral cancer treatment strategies: Surgical treatments for oral squamous cell carcinoma. *International Journal of Clinical Oncology*, *19*(3), 423–430. https://doi.org/10.1007/s10147-014-0689-z

Papadimitrakopoulou, V. A., Hong, W. K., Lee, J. S., Lippman, S. M., Martin, J. W., Lee, J. J., & Batsakis, J. G. (1997). Low-Dose Isotretinoin Versus -Carotene to Prevent

Oral Carcinogenesis: Long-term Follow-up. *JNCI Journal of the National Cancer Institute*, *89*(3), 257–258. https://doi.org/10.1093/jnci/89.3.257

Peng, G., Tsukamoto, S., Okumura, K., Ogawa, H., Ikeda, S., & Niyonsaba, F. (2022). A Pancancer Analysis of the Oncogenic Role of S100 Calcium Binding Protein A7 (S100A7) in Human Tumors. *Biology*, *11*(2), 284. https://doi.org/10.3390/biology11020284

Peng, Q., Deng, Z., Pan, H., Gu, L., Liu, O., & Tang, Z. (2017). Mitogen-activated protein kinase signaling pathway in oral cancer (Review). *Oncology Letters*. https://doi.org/10.3892/ol.2017.7491

Pouysségur, J., Volmat, V., & Lenormand, P. (2002). Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling. *Biochemical Pharmacology*, *64*(5–6), 755–763. https://doi.org/10.1016/S0006-2952(02)01135-8

Pritzker, K. P. H., Darling, M. R., Hwang, J. T.-K., & Mock, D. (2021). Oral Potentially Malignant Disorders (OPMD): What is the clinical utility of dysplasia grade? *Expert Review of Molecular Diagnostics*, *21*(3), 289–298. https://doi.org/10.1080/14737159.2021.1898949

Pulte, D., & Brenner, H. (2010). Changes in Survival in Head and Neck Cancers in the Late 20th and Early 21st Century: A Period Analysis. *The Oncologist*, *15*(9), 994–1001. https://doi.org/10.1634/theoncologist.2009-0289

Raffat, M. A., Hadi, N. I., Hosein, M., Mirza, S., Ikram, S., & Akram, Z. (2018). S100 proteins in oral squamous cell carcinoma. *Clinica Chimica Acta*, *480*, 143–149. https://doi.org/10.1016/j.cca.2018.02.013

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: JOMFP*, *23*(1), 19–27. https://doi.org/10.4103/jomfp.JOMFP_13_19

Reyes, M., Flores, T., Betancur, D., Peña-Oyarzún, D., & Torres, V. A. (2020). Wnt/β-Catenin Signaling in Oral Carcinogenesis. *International Journal of Molecular Sciences*, *21*(13), 4682. https://doi.org/10.3390/ijms21134682

Saikawa, M., Ebihara, S., Yoshizumi, T., & Ohyama, W. (1991). Multiple Primary Cancers in Patients with Squamous Cell Carcinoma of the Oral Cavity. *Japanese Journal of Cancer Research*, *82*(1), 40–45. https://doi.org/10.1111/j.1349-7006.1991.tb01743.x

Sarode, S., Sarode, G., & Tupkari, J. (2014). Oral potentially malignant disorders: A proposal for terminology and definition with review of literature. *Journal of Oral and Maxillofacial Pathology*, *18*(4), 77. https://doi.org/10.4103/0973-029X.141322

Shailaja, G., Kumar, J. V., Baghirath, P. V., Kumar, U., Ashalata, G., & Krishna, A. B. (2015). Estimation of malignant transformation rate in cases of oral epithelial dysplasia and lichen planus using immunohistochemical expression of Ki-67, p53, BCL-2, and BAX markers. *Dental Research Journal*, *12*(3), 235–242.

Sood, A., Mishra, D., Kharbanda, O., Chauhan, S., Gupta, S., V Deo, S., Yadav, R., Ralhan, R., Kumawat, R., & Kaur, H. (2022). Role of S100 A7 as a diagnostic biomarker in oral potentially malignant disorders and oral cancer. *Journal of Oral and Maxillofacial Pathology*, *26*(2), 166. https://doi.org/10.4103/jomfp.jomfp_402_20

Sue Ng, S., Mahmoudi, T., Li, V. S. W., Hatzis, P., Boersema, P. J., Mohammed, S., Heck, A. J., & Clevers, H. (2010). MAP3K1 functionally interacts with Axin1 in the canonical Wnt signalling pathway. *Biological Chemistry*, *391*(2–3), 171–180. https://doi.org/10.1515/bc.2010.028

Sun, X., & Kaufman, P. D. (2018). Ki-67: More than a proliferation marker. *Chromosoma*, *127*(2), 175–186. https://doi.org/10.1007/s00412-018-0659-8

Sun, Y., Cheng, Z., & Liu, S. (2022). MCM2 in human cancer: Functions, mechanisms, and clinical significance. *Molecular Medicine*, *28*(1), 128. https://doi.org/10.1186/s10020-022-00555-9

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Tetzlaff, M. T., Pattanaprichakul, P., Wargo, J., Fox, P. S., Patel, K. P., Estrella, J. S., Broaddus, R. R., Williams, M. D., Davies, M. A., Routbort, M. J., Lazar, A. J., Woodman, S. E., Hwu, W.-J., Gershenwald, J. E., Prieto, V. G., Torres-Cabala, C. A., & Curry, J. L. (2015). Utility of BRAF V600E Immunohistochemistry Expression Pattern as a Surrogate of BRAF Mutation Status in 154 Patients with Advanced Melanoma. *Human Pathology*, *46*(8), 1101–1110. https://doi.org/10.1016/j.humpath.2015.04.012

Tian, T., Li, X., Hua, Z., Ma, J., Wu, X., Liu, Z., Chen, H., & Cui, Z. (2017). S100A7 promotes the migration, invasion and metastasis of human cervical cancer cells through epithelial-mesenchymal transition. *Oncotarget*, *8*(15), 24964–24977. https://doi.org/10.18632/oncotarget.15329

Tian, X., Liu, Z., Niu, B., Zhang, J., Tan, T. K., Lee, S. R., Zhao, Y., Harris, D. C. H., & Zheng, G. (2011). E-Cadherin/ β -Catenin Complex and the Epithelial Barrier. *Journal of Biomedicine and Biotechnology*, *2011*, 1–6. https://doi.org/10.1155/2011/567305

Tian, Y., & Guo, W. (2020). A Review of the Molecular Pathways Involved in Resistance to BRAF Inhibitors in Patients with Advanced-Stage Melanoma. *Medical Science Monitor*, *26*. https://doi.org/10.12659/MSM.920957

Torres-Rendon, A., Roy, S., Craig, G. T., & Speight, P. M. (2009). Expression of Mcm2, geminin and Ki67 in normal oral mucosa, oral epithelial dysplasias and their corresponding squamous-cell carcinomas. *British Journal of Cancer*, *100*(7), 1128–1134. https://doi.org/10.1038/sj.bjc.6604967

Usman, S., Waseem, N. H., Nguyen, T. K. N., Mohsin, S., Jamal, A., Teh, M.-T., & Waseem, A. (2021). Vimentin Is at the Heart of Epithelial Mesenchymal Transition (EMT) Mediated Metastasis. *Cancers*, *13*(19), 4985. https://doi.org/10.3390/cancers13194985

Vallina, C., López-Pintor, R. M., González-Serrano, J., De Vicente, J. C., Hernández, G., & Lorz, C. (2021). Genes involved in the epithelial-mesenchymal transition in oral cancer: A systematic review. *Oral Oncology*, *117*, 105310. https://doi.org/10.1016/j.oraloncology.2021.105310

Vieira, F. L., Vieira, B. J., Guimaraes, M. A., & Aarestrup, F. M. (2008). Cellular profile of the peritumoral inflammatory infiltrate in squamous cells carcinoma of oral mucosa: Correlation with the expression of Ki67 and histologic grading. *BMC Oral Health*, *8*(1), 25. https://doi.org/10.1186/1472-6831-8-25

Warnakulasuriya, S., Kujan, O., Aguirre-Urizar, J. M., Bagan, J. V., González-Moles, M. Á., Kerr, A. R., Lodi, G., Mello, F. W., Montiero, L., Ogden, G. R., Sloan, P., & Johnson, N. W. (2021). 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 Diseases*, *27*(8), 1862–1880. https://doi.org/10.1111/odi.13704

Warnakulasuriya, S., Sutherland, G., & Scully, C. (2005). Tobacco, oral cancer, and treatment of dependence. *Oral Oncology*, *41*(3), 244–260. https://doi.org/10.1016/j.oraloncology.2004.08.010

Wohlschlegel, J. A., Kutok, J. L., Weng, A. P., & Dutta, A. (2002). Expression of Geminin as a Marker of Cell Proliferation in Normal Tissues and Malignancies. *The American Journal of Pathology*, *161*(1), 267–273. https://doi.org/10.1016/S0002-9440(10)64178-8

Yuan, J., Lan, H., Huang, D., Guo, X., Liu, C., Liu, S., Zhang, P., Cheng, Y., & Xiao, S. (2022). Multi-Omics Analysis of MCM2 as a Promising Biomarker in Pan-Cancer. *Frontiers in Cell and Developmental Biology*, *10*, 852135. https://doi.org/10.3389/fcell.2022.852135

Yuwanati, M. B., Tupkari, J. V., & Avadhani, A. (2011). Expression of E-cadherin in oral epithelial dysplasia and oral squamous cell carcinoma: An in vivo study.

Journal of Clinical and Experimental Investigations, *2*(4).
https://doi.org/10.5799/ahinjs.01.2011.04.0070

Žerdoner, D. (2003). The Ljubljana classification – its application to grading oral epithelial hyperplasia. *Journal of Cranio-Maxillofacial Surgery*, *31*(2), 75–79. https://doi.org/10.1016/S1010-5182(02)00186-5

Zhang, Y., He, J., He, B., Huang, R., & Li, M. (2019). Effect of tobacco on periodontal disease and oral cancer. *Tobacco Induced Diseases*, *17*(May). https://doi.org/10.18332/tid/106187

Appendix

Appx. Table 1: Clinical Appearance and definition of oral potentially malignant lesions.

Oral Potential Malignant Disorders	Clinical Appearance
Leukoplakia	A predominantly white plaque of questionable risk having excluded (other) known diseases or disorders that carry no increased risk for cancer
Proliferative Verrucous Leukoplakia	Progressive, persistent, and irreversible disorder characterized by the presence of multiple Leukoplakia that frequently become warty
Erythroplakia	A predominantly fiery red patch that cannot be characterized clinically or pathologically as any other definable disease
Oral Submucous Fibrosis	A chronic, insidious disease that affects the oral mucosa, initially resulting in loss of fibroelasticity of the lamina propria and as the disease advances, results in fibrosis of the lamina propria and the submucosa of the oral cavity along with epithelial atrophy
Oral Lichen Planus	A chronic inflammatory disorder of unknown etiology with characteristic relapses and remissions, displaying white reticular lesions, accompanied or not by atrophic, erosive, and ulcerative and/or plaque type areas. Lesions are frequently bilaterally symmetrical. Desquamative gingivitis may be a feature
Actinic Keratosis/ Cheilitis	A disorder that results from sun damage and affects exposed areas of the lips, most commonly the vermilion border of the lower lip with a variable presentation of atrophic and erosive areas and white plaques
Palatal Lesions in Reverse Smokers	White and/or red patches affecting the hard palate in reverse smokers, frequently stained with nicotine
Oral Lupus Erythematosus	An autoimmune connective tissue disease which may affect the lip and oral cavity, where it presents as an erythematous area surrounded by whitish striae, frequently with a target configuration
Dyskeratosis Congenita	A rare cancer-prone inherited bone marrow failure syndrome caused by aberrant telomere biology. It is characterized clinically by the presence of the diagnostic triad of dysplastic nails, lacy reticular skin pigmentation and oral leukoplakia
Oral Lichenoid Lesion	Oral lesions with lichenoid features but lacking the typical clinical or histopathological appearances of OLP, that is, may show asymmetry or are reactions to dental restorations or are drug-induced
Oral Graft versus Host Disease	Clinical and histopathological presentations like oral lichen planus in a patient developing an autoimmune, multi-organ complication after allogeneic hematopoietic cell transplantation

Table is taken from Oral Diseases Volume 27 Issue 8 (Warnakulasuriya et al., 2021).

Appx. Table 2: RNA Concentration Used for nanoString nCounter

Sample #	RNA Concentration *(ng/µl)*	DV200 *(%)*	RNA volume *(µl)*	RNA-free water volume *(µl)*	Total volume *(µl)*
15-12700 (Control)	48.63	25	5.0	0	5
16-190 (Control)	28.89	41	5.0	0	5
20-3820 (Control)	39.79	53	5.0	0	5
17-43990 (Mild dysplasia)	236.61	85	1.3	3.7	5
18-49201 (Mild dysplasia)	6.84	66	5.0	0	5
19-57445 (Mild dysplasia)	360.46	28	0.9	4.1	5
17-18204 (Moderate dysplasia)	392.67	37	0.8	4.2	5
17-47099 (Moderate dysplasia)	377.06	40	0.8	4.2	5
19-36499 (Moderate dysplasia)	281.66	35	1.1	3.9	5
15-29820 (Severe dysplasia)	161.16	66	1.9	3.1	5
16-16667 (Severe dysplasia)	105.85	41	2.9	2.1	5
16-45618 (Severe dysplasia)	42.53	33	5.0	0	5

Appx. Table 3: Clinical diagnosis categories of 82 OPMD cases.

Final Diagnosis	Total Cases	Percentage (%)
Dysplasia	51	75.6
Leukoplakia	17	20.7
Chronic Lichenoid Mucositis	2	2.5
Lichen Planus	3	3.6
Fibroma	2	2.5
N/A*	7	8.5

Appx. Table 4: Final diagnosis categories of 82 OPMD cases.

Final Diagnosis	Total Cases	Percentage (%)
Mild dysplasia	29	35.4
Moderate dysplasia	15	18.3
Severe Dysplasia	6	7.3
Chronic Lichenoid Mucositis	13	15.8
Chronic Candidiasis	5	6.1
Fibroma	5	6.1
Lichen Planus	3	3.7
Papilloma	4	4.9
TUGSE	2	2.4

Appx. Table 5: Top 10 Differentially Expressed mRNAs Between Mild Dysplasia Cases and Control Tissues.

	Log2 fold change	std error (log2)	Lower confidence limit (log2)	Upper confidence limit (log2)	Linear fold change	Lower confidence limit (linear)	Upper confidence limit (linear)	P-value
NRP1-mRNA	1.26	0.185	0.893	1.62	2.39	1.86	3.07	0.000139
SIGLEC1-mRNA	2.15	0.351	1.46	2.84	4.44	2.75	7.15	0.000283
C1QB-mRNA	1.3	0.22	0.874	1.74	2.47	1.83	3.33	0.000348
BCL2-mRNA	1.29	0.241	0.814	1.76	2.44	1.76	3.38	0.000695
CD14-mRNA	1.78	0.349	1.09	2.46	3.43	2.13	5.5	0.000938
S100A7-mRNA	4.2	0.847	2.54	5.86	18.4	5.83	58.2	0.0011
IL32-mRNA	2.77	0.579	1.64	3.91	6.84	3.11	15	0.00137
CD44-mRNA	1.35	0.288	0.781	1.91	2.54	1.72	3.76	0.0016
MAP3K1-mRNA	-0.879	0.19	-1.25	-0.506	0.544	0.42	0.704	0.00172
C1QA-mRNA	1.03	0.224	0.588	1.47	2.04	1.5	2.76	0.00179

Appx. Table 6: Top 15 Differentially Expressed mRNAs Between Moderate Dysplasia Cases and Control Tissues.

	Log2 fold change	std error (log2)	Lower confidence limit (log2)	Upper confidence limit (log2)	Linear fold change	Lower confidence limit (linear)	Upper confidence limit (linear)	P-value
S100A7-mRNA	5.84	0.847	4.18	7.5	57.4	18.2	181	0.000125
CD14-mRNA	2.02	0.347	1.33	2.7	4.04	2.52	6.48	0.000404
MAP3K1-mRNA	-1.08	0.192	-1.45	-0.699	0.474	0.365	0.616	0.000513
CD3E-mRNA	3.26	0.588	2.11	4.41	9.57	4.3	21.3	0.000548
JAK3-mRNA	2.63	0.496	1.66	3.6	6.19	3.15	12.2	0.000729
CD44-mRNA	1.5	0.288	0.932	2.06	2.82	1.91	4.18	0.00083
BCL6-mRNA	1.31	0.255	0.812	1.81	2.48	1.76	3.51	0.000882
C1QB-mRNA	1.13	0.221	0.7	1.56	2.19	1.62	2.96	0.000892
CYBB-mRNA	1.7	0.344	1.02	2.37	3.25	2.03	5.18	0.00114
IFI16-mRNA	1.86	0.391	1.1	2.63	3.64	2.14	6.19	0.00141
SIGLEC1-mRNA	1.69	0.355	0.994	2.39	3.23	1.99	5.23	0.00143
IL18-mRNA	-1.62	0.35	-2.31	-0.938	0.325	0.202	0.522	0.00166
ITGB4-mRNA	1.43	0.309	0.823	2.04	2.69	1.77	4.1	0.00171
BCL2-mRNA	1.12	0.243	0.64	1.59	2.17	1.56	3.01	0.00177
CSF2RB-mRNA	2.39	0.52	1.37	3.41	5.24	2.58	10.6	0.00177

Appx. Table 7: Top 20 Differentially Expressed mRNAs Between Severe Dysplasia Cases and Control Tissues.

	Log2 fold change	std error (log2)	Lower confidence limit (log2)	Upper confidence limit (log2)	Linear fold change	Lower confidence limit (linear)	Upper confidence limit (linear)	P-value
C1QB-mRNA	2.04	0.218	1.62	2.47	4.12	3.07	5.54	1.36E-05
SIGLEC1-mRNA	2.56	0.348	1.88	3.24	5.9	3.68	9.48	7.97E-05
NRP1-mRNA	1.32	0.185	0.958	1.68	2.5	1.94	3.21	9.77E-05
CYBB-mRNA	2.42	0.34	1.76	3.09	5.37	3.38	8.52	9.96E-05
C1QA-mRNA	1.54	0.22	1.11	1.97	2.91	2.16	3.93	0.000112
S100A7-mRNA	5.67	0.847	4.01	7.33	50.8	16.1	160	0.000154
BCL6-mRNA	1.69	0.252	1.19	2.18	3.22	2.29	4.53	0.000155
COL3A1-mRNA	1.74	0.272	1.21	2.27	3.34	2.31	4.83	0.000208
CD14-mRNA	2.19	0.346	1.51	2.87	4.56	2.85	7.31	0.000227
FCGR2B-mRNA	1.55	0.251	1.06	2.04	2.94	2.09	4.13	0.000259
MAP3K1-mRNA	-1.2	0.194	-1.57	-0.816	0.437	0.336	0.568	0.000268
FYN-mRNA	1.65	0.269	1.12	2.18	3.14	2.18	4.53	0.000278
ICAM1-mRNA	2.16	0.353	1.47	2.85	4.47	2.77	7.21	0.000282
FCER1G-mRNA	1.55	0.259	1.04	2.05	2.92	2.05	4.15	0.000335
NT5E-mRNA	2.32	0.411	1.51	3.12	4.99	2.86	8.71	0.000483
CD44-mRNA	1.62	0.288	1.05	2.18	3.07	2.08	4.54	0.000502
CD163-mRNA	2.18	0.407	1.38	2.97	4.52	2.6	7.85	0.000689
CSF2RB-mRNA	2.74	0.519	1.73	3.76	6.7	3.31	13.6	0.000742
IFI16-mRNA	2.04	0.39	1.27	2.8	4.11	2.42	6.99	0.000797
HCK-mRNA	1.83	0.353	1.13	2.52	3.54	2.19	5.73	0.000857

Curriculum Vitae

Name: Xinru Li

Post-secondary Western University
Education and London, Ontario, Canada
Degrees: 2017-2022 BMSc
      Western University
      London, Ontario, Canada
      2022-Present MSc

Honours and One Health Graduate Student Research Award
Awards: 2022
      Dr. Fredrick Winnett Luney Graduate Research Award
      2024
      Western Graduate Research Scholarship
      Western University
      2022- 2024

Related Work Graduate Teaching Assistant
Experience   Western University
      2022-2024
      Research Assistant
      Oral Pathology, Western University
      2021-2022

Conference Presentations:
Pathology and Laboratory Medicine Research Day
Western University
Oral Potentially Malignant Disorders and Oral Cancers in Humans and Dogs
March 2023

Pathology and Laboratory Medicine Research Day
Western University
Molecular Evaluation of Clinical Diagnosis of Oral Potentially Malignant Disorders
March 2024

London Health Research Day
Western University
Molecular Evaluation of Clinical Diagnosis of Oral Potentially Malignant Disorders
May 2024

The 22nd International Congress on Oral Pathology and Medicine (IAOP)
Cancun, Mexico
June 2024