ORAL POTENTIALLY MALIGNANT DISORDERS CLINICAL DATABASE AT THE ORAL AND MAXILLOFACIAL SURGERY DEPARTMENT, MONTREAL GENERAL HOSPITAL - MCGILL UNIVERSITY HEALTH CENTRE

Lojain Bassyoni
MSc Dental Sciences

Faculty of Dentistry
McGill University, Montreal

June 2018

A thesis submitted to McGill University in partial fulfillment of the requirement for a Master of Science degree

DEDICATION

To my mother Haifa, my husband Khuzam, my son Khaled.

ACNOWLEDGMENTS

"One who doesn't thank people, doesn't thank god" Prophet Mohamed

I would like to express my gratitude to my supervisor, Dr. Belinda Nicolau and my co-supervisor Dr. Nicholas Makhoul for their support, understanding and immense knowledge. They always steered me in the right direction and provided me with insightful comments and advice whenever I needed it. Despite their busy schedules, their doors were always open, and their guidance helped me all the way during the research project and the writing of this thesis.

I would like to extend my appreciation to Dr. Genevieve Castonguay for her support and help. I have learned a lot from her constructive feedback and comments. Also, I would like to thank Mrs. Huwaida Makhoul for her experience, help and constructive comments during customising the database software.

I am thankful to King Abdulaziz University and the Saudi Cultural Bureau in Canada for granting me a full scholarship that enabled me to be where I am today.

I am so grateful and thankful to all my friends at the Division of Oral Health and Society for the stimulating discussions we had and the fun times we spent together; my journey would not have been the same without them. Especially, I must thank my lab mate Dr. Sreenath Arekunnath Madathil for the tremendous valuable help he provided me with data exportation and analysis. I am always amazed by his knowledge, humbleness and continuous willingness to help and support others.

I feel so lucky to have the mother I have, she has been always a source of infinite unconditional love, always proud of me and offering support and help. She would cross the ocean all the way from Saudi Arabia to visit me in Montreal each time she felt I needed her and provided me with love and care, so I would be able to focus on work. I learned from her how to be strong, caring and resourceful. Words will never be enough to express my gratitude toward her.

My beloved husband Khuzam has been a major source of support. He always believed in me and encouraged me to follow my dreams. He has always been a great listener, charging me with positive energy and comforting me during my postgrad journey and in life in general. My precious son Khaled who in his laughter I find joy and hope, to me it is the sound of everything that is right in this world. This little family of mine was and will always be instrumental to my success.

Finally, I would like to thank my siblings Lojan, Aws and Yousef, and my friends Mashail and Ghada for their care, support and for being the wonderful people they are.

Praise is to Allah by whose grace good deeds are accomplished, thank you Allah for guiding me and giving me all the help I needed to complete this task.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

OPMD: Oral potentially malignant disorders

REDCap: Research electronic data capture

OMFS: Oral and maxillofacial surgery

MUHC: McGill university health center

OC: Oral cancer

OPL: Oral premalignant lesions

PVL: Proliferative verrucous leukoplakia

OSF: Oral submucous fibrosis

OED: Oral epithelial dysplasia

OSCC: Oral squamous cell carcinoma

OLP: Oral lichen planus

SES: Socioeconomic status

HPV: Human papilloma virus

OPC: Oropharyngeal cancer

BMI: Body metabolic index

CONTRIBUTION OF AUTHORS

Lojain Bassyoni, MSc candidate: conceived the idea of the project presented in this thesis, performed all the steps involved in building the database and writing this thesis including the manuscript, data collection and statistical analysis.

Belinda Nicolau, Associate professor, Division of Oral Health and Society, Faculty of Dentistry, McGill University, Montreal, Quebec, Canada: Supervised all the steps of the database development and thesis writing, supervised and contributed to the manuscript writing and design of statistical analysis.

Sreenath Madathil, PhD Candidate, Division of Oral Health and Society, Faculty of Dentistry, McGill University: facilitated data exportation process and contributed to manuscript writing and data analysis design.

Nicolas Makhoul, Assistant Professor, Faculty of Dentistry, and Chief Department of Dentistry and Oral and Maxillofacial Surgery, MUHC, McGill University, Montreal, Quebec, Canada: Owner of the OPMD database, supervised all the steps of the database development and contributed to manuscript writing.

ABSTRACT

Aim:

This thesis was set up to develop and implement an oral potentially malignant disorders (OPMD) database at the Oral Maxillofacial Surgery (OMFS) Division at the Montreal General Hospital (MGH), McGill University Health Centre (MUHC). In addition, we aim to describe the clinical, biological and socio-demographic factors among people with OPMD. Our ultimate goal for establishing this database is to facilitate future research in OPMD epidemiology as well as monitor quality of care.

Methodology:

Using an HTML based data collection system (REDCap), we created an OPMD clinical database at the OMFS clinic at the MGH-MUHC. The process of creating the database included identifying the purpose of the database, choosing database elements and data sources, establishing a database team, developing the extraction forms and questionnaires customising the REDCap software to the database needs and to enter the data. The database included two arms of patients: retrospective and prospective. Retrospective patients were OPMD patients receiving care at the OMFS clinic from 2008 to 2017. Data were collected on retrospective subjects and included patients'demographics, type of treatment provided, risk factors for OPMD and progression to oral cancer. Descriptive analysis was conducted to show the distribution of demographics, clinical findings, habits and progression to oral cancer among the population.

Results:

We conducted pilot testing procedures and amended several technical issues. For the retrospective arm of the study, we identified 155 patients with OPMD, the majority were males (57%), and have been exposed to tobacco and alcohol at a certain point in their life. The most common location was the tongue (39%), and the most common diagnosis was leukoplakia (52.9%). Dysplasia was present in 56% of patients. Lesions were mostly smaller than 200 mm^2 (45.8%) and white in color (63.23%). Sixteen patients in the sample progressed into OC with a proportion of 10.32% and an incidence rate of 4.65 per 100-person years (CI: 2.85 – 7.59).

Conclusion:

The development of the OPMD clinical database in an academic clinical setting provides a platform for research and education which is a huge advantage and opportunity for researchers, clinicians and students.

RÉSUMÉ

Objectif :

Cette thèse a été conçue pour développer et mettre en place une banque de données sur les troubles oraux potentiellement malins (TOPM) à la Division de chirurgie maxillo-faciale buccale (*Oral Maxillofacial Surgery*; OMFS) de l'Hôpital général de Montréal (HGM), Centre universitaire de santé McGill (CUSM). En outre, nous visons à décrire les facteurs cliniques, biologiques et sociodémographiques chez les personnes atteintes de TOMP. Notre objectif ultime en établissant cette banque de données est de faciliter la recherche future en épidémiologie des TOPM ainsi que de surveiller la qualité des soins.

Méthodologie :

À l'aide d'un système de collecte de données basé sur HTML (REDCap), nous avons créé une banque de données cliniques de TOPM à la clinique OMFS de l'HGM-CUSM. Le processus de création de la banque de données comprenait l'identification de l'objectif de la banque de données, le choix des éléments de la banque de données et des sources de données, l'établissement d'une équipe pour construire la banque de données, l'élaboration de formulaires d'extraction et de questionnaires pour adapter le logiciel REDCap aux besoins de la banque de données et pour entrer les données. La banque de données comprenait deux groupes de patients: rétrospectif et prospectif. Les patients rétrospectifs étaient des patients atteints de TOPM ayant reçu un traitement à la clinique OMFS de 2008 à 2017. Les données qui ont été recueillies sur les sujets rétrospectifs comprenaient des données démographiques concernant les patients, le type de traitement fourni, les facteurs de risque de TOPM et la progression vers le cancer buccal. Une analyse descriptive a été menée pour montrer la

distribution des données démographiques, les résultats cliniques, les habitudes de vie et la progression vers le cancer buccal dans la population.

Résultats :

Nous avons mené des essais pilotes et avons corrigé plusieurs problèmes techniques. Pour le volet rétrospectif de l'étude, nous avons identifié 155 patients atteints de TOPM, la majorité était des hommes (57%) et ont été exposés au tabac et à l'alcool à un certain moment de leur vie. L'emplacement le plus fréquent était la langue (39%) et le diagnostic le plus commun était la leucoplasie (52,9%). La dysplasie était présente chez 56% des patients. Les lésions étaient généralement inférieures à 200 mm^2 (45,8%) et de couleur blanche (63,23%). Seize patients dans l'échantillon ont progressé au cancer buccal avec une proportion de 10,32% et un taux d'incidence de 4,65 par 100 personnes-années (IC: 2,85 - 7,59).

Conclusion :

Le développement de la banque de données clinique de TOPM dans un cadre clinique universitaire fournit une plateforme pour la recherche et l'éducation, ce qui constitue un énorme avantage et une opportunité pour les chercheurs, les cliniciens et les étudiants.

1. INTRODUCTION

The natural history of oral cancer (OC) can be conceptualized as a multistep progressive model, resulting from carcinogen exposure induced accumulation of genetic alterations. Histopathologically, this is presented by the progression from normal oral mucosa to hyperplasia, to dysplasia, to oral cancer [1-4].

Oral cancer has an incidence of 4.2 % in Canada; approximately 4,700 new cases were diagnosed and 1,250 Canadians died from it in 2017, of which 1,070 new cases and 300 deaths were in the province of Quebec [5]. The 5-year survival rates for this disease has remained unchanged at less than 50%, with no significant improvement over the past decades despite advances in different therapy modalities [6, 7]. This low survival is mainly attributable to late detection, which means an advanced stage disease at time of diagnosis, mostly because of lack of public education [8, 9]. The stage at diagnosis of OC remains an important prognostic indicator to predict patient survival; while the survival rates for early stage OC are approximately 93%, these rates drop significantly to 38% for advanced stages, and 20% for metastatic disease [10]. Therefore, it is wise to direct efforts toward prevention and early detection.

Sometimes OC arises from Oral Potentially Malignant Disorders (OPMD). Depending on location in the oral cavity and clinical presentation of OPMD, up to 51% of these lesions have a degree of dysplasia at time of presentation [8]. In a SEER based case-cohort study, 7% of oral squamous cell carcinoma cases were preceded by OPMD (leukoplakia); these cancers had 64% less odds to be diagnosed at a regional/distant stage (stage III, IV) compared to cancer cases with no preceding OPMD [11].

OPMD is an important diagnosis, as approximately 40%, and in certain lesions (erythroplakia) up to 50%, will undergo malignant transformation [12-14]. The surveillance of patients with OPMD has been shown to improve oral cancer patients' outcomes [15]. Some clinical and demographic parameters that predispose patients to a higher risk of malignant transformation have been identified. However, it is likely that further histopathologic features and molecular biomarkers exist and may further help risk-stratify patients. Therefore, identifying these lesions, their risk factors, and providing proper treatment and care to patients, will help to decrease the incidence of OC, detect cases at an early stage and improve patients' survival. This, in turn, is likely to improve treatment outcomes, patients' quality of life, healthcare planning, and decrease the economic burden of OC on the healthcare system. At present, however, specific data are sparse when looking at this patient population (OPMD population), with a low level of evidence for treatment decisions. This is likely due to relatively low incidence of the disease; leukoplakia one of the most common OPMD has a global prevalence of 2.6%.[16] Databases and cancer registries offer an excellent way to deal with this problem. Indeed, premalignant lesions are usually documented within cancer registries, for example the Cancer Registry of Norway (CRN), [17, 18] or as a part of another database (non-disease specific/ patient registries), such as the nationwide histopathology registry in the Netherlands (PALGA) [19] or U.S. Medicare claims, [11].

In this project, we aim to develop a database for OPMD to assist clinicians in the treatment and management of patients with these lesions. Building such a database will enable the identification of factors allowing clinicians to predict which patients are at risk for developing OPMD and what is the optimal treatment of these lesions. In addition, we will use this infrastructure to monitor the prevalence and incidence of this disease and to study the risk factors for OPMD and its risk of progression into cancer.

2. LITERATURE REVIEW

2.1 Definition of Oral Premalignant Lesions (OPL)

OPL or oral epithelial precursor lesions are defined as altered epithelium with an increased likelihood for progression to squamous cell carcinoma; the word altered here refers to dysplasia.[20] These lesions can occur anywhere in the oral cavity and carry the potential of transforming into OC.

In 1978, a WHO working group proposed that the clinical presentations of the oral cavity referred to as precancerous should be classified into “precancerous lesions” and “precancerous conditions”, where precancerous lesions are defined as clinically altered mucosa in which cancer is more likely to happen compared to clinically normal appearing mucosa. On the other hand, precancerous conditions would carry the definition of a generalized state associated with a significantly increased risk of cancer.

In other words, these definitions imply that in cases of premalignant lesions the mucosa at increased risk of developing cancer would be the clinically morphologically altered area only, whereas in cases of precancerous conditions the cancerous lesion could arise anywhere in the oral mucosa as it is a generalized state that puts the whole oral cavity at risk. However, nowadays we know that patients with premalignant lesions in one area could harbor dysplastic mucosa or an epithelium with molecular changes on the contralateral side with no necessarily evident clinical changes. Therefore, in such cases even oral mucosal areas with no clinically appearing lesions or even histological changes are at increased risk of developing oral cancer. [21, 22]

As a way to overcome the aforementioned shortcoming of the current nomenclature and classification in describing precancerous oral presentations, the WHO Collaborating Centre for Oral Cancer and Pre-cancer reached a consensus in 2005 to refer to both premalignant or precancerous lesions and conditions as Oral Potentially Malignant Disorders (OPMD).

Yet, to date no definition has been given to this new nomenclature by the WHO. However, Sarodi et al. proposed in 2012 the following definition for the term OPMD: "*It is a group of disorders of varying aetiologies, usually tobacco; characterized by mutagen-associated, spontaneous or hereditary alterations or mutations in the genetic material of oral epithelial cells with or without clinical and histomorphological alterations that may lead to oral squamous cell carcinoma transformation*" [23]. Based on this definition, OPMD would incorporate all oral mucosal lesions, medical conditions or genetic and molecular alterations that would place one at increased risk of developing oral cancer.

The most commonly seen OPMD are leukoplakia, erythroplakia, proliferative verrucous leukoplakia (PVL) and oral submucous fibrosis (OSF). However, other lesions and conditions are included under OPMD nomenclature such as actinic keratosis, palatal lesions in reverse smoking, epidermolysis bullosa, discoid lupus erythematosus and prolonged immunosuppression state.

2.2 Common oral potentially malignant disorders

2.2.1 Leukoplakia: (definition, prevalence and transformation rate)

Schwimmer was the first to use the term leukoplakia in 1877 to describe a white lesion on the tongue [8]. Oral leukoplakia (OL) is defined as a white plaque of questionable risk having excluded

other known diseases or disorders that carry no increased risk for cancer. It does not have any specific histopathological appearance. It is a clinical diagnosis based on the exclusion of any other entity that would present itself similarly to oral leukoplakia, yet is diagnosed otherwise based on distinguishing clinical or pathological findings. Examples of other lesions and conditions that should not be noted as oral leukoplakia are: hairy leukoplakia, leukoedema, frictional lesions, and white sponge nevus.

Oral leukoplakia has a global prevalence of 2.6% (95% confidence interval: 1.72%-2.74%). [16] The condition has a higher prevalence among South East Asian populations compared to those living in the western world; this is mainly due to betel quid and areca nut chewing habits, which are popular in that geographical area. Few studies have reported the incidence of oral leukoplakia. Nevertheless, findings from a study conducted in Japan in 2005 showed that the age adjusted incidence rate per 100,000 person-years for leukoplakia was 409.2 (95% CI: 90.6-727.9) in men and 70.0 (95% CI: 17.9-121.8) in women. [24] In addition, we found two studies from India. One followed up a cohort of Bombay policemen and reported a 10-year incidence of 2.9%. [25] The other one was a 10-year follow up of a house-to-house survey in 3 rural areas in India, which reported an incidence rate between 1.4 and 2.6 per 1,000 for males, and ranging from 1.5 to negligible per 1,000 for females. [26]

Consuming tobacco and alcohol, and chewing betel quid/ areca nut causes the majority of oral leukoplakia cases. However, some leukoplakia are idiopathic and no known risk factors could be identified. Although leukoplakia can occur at any age, it is more common in people above 40 years old. [27] Also, it is more common in smokers than non-smokers by approximately six folds [28]. However, malignant transformation is consistently reported to be higher in non-smokers.

The rate of malignant transformation of oral leukoplakia varies considerably in the literature, from 0.1% to 36%. [29-33] This disparity could be related to the heterogeneity in the studies in terms of design, population, definition of cases and length of follow up. Some studies considered only biopsied lesions in their population [29] or only leukoplakia with OED, while in others an individual's clinical diagnosis of leukoplakia mentioned in a Medicare claim was sufficient to be included in the sample. [11]

A recent U.S. population-based study published in 2016 calculated a total 5-year cumulative incidence of oral cancer arising from a precursor leukoplakia of 2.54%. [11] In another study from the Netherlands, researchers estimated a 2.9% annual malignant transformation rate, and also noted that the longer the follow up period is, the higher is the number of lesions transforming into oral squamous cell carcinoma (OSCC). Moreover, from a Kaplan-Meier survival curve based on 166 patients, they proposed a 50% transformation rate within 200 months. [34] A Japanese study recently reported a cumulative malignant transformation rate of 11.6% in 10 years with no surgical intervention. [35] This corresponds to an annual transformation rate of 1.16%, which is lower than the numbers reported in the Netherlands. This is probably because in the Japanese study there was only one case of severe dysplasia, the number of non-homogenous leukoplakia was lower than in the Netherlands study, and 32.5% of the lesions either disappeared or regressed due to smoking cessation and elimination of possible irritants.

Proliferative verrucous leukoplakia (PVL) is a rare form of oral leukoplakia. It was first described by Hansen et al in 1985. PVL is part of a progressive condition that starts as a hyperkeratoic lesion progressing over time to a multifocal exophytic proliferative lesion called PVL, and eventually to verrucous carcinoma or squamous cell carcinoma. Compared to other oral

leukoplakia, PVL has a more aggressive biological behaviour with a higher risk of recurrence and a malignant transformation rate that reaches up to 100%.[36-38]

2.2.2 Erythroplakia

Erythroplakia is defined as a fiery red patch that cannot be characterized clinically or pathologically as any other definable disease. It is rare and considerably less common than leukoplakia, with a prevalence ranging between .02% and .83%. Erythroplakia is more commonly seen in middle aged and elderly males and mostly affects the areas of the soft palate and floor of the mouth. [39] When lesions are mixed and have changes of both red and white colors, they are called erythroleukoplakias. Erythroplakia is considered the most dangerous of OPMD; it has a high rate of malignant transformation reaching up to 50%, [40] and when examined histopathologically at the time of initial diagnosis, 51% of the cases come back as OSCC with the remainder having some grade of dysplasia, mostly severe dysplasia or carcinoma in situ. [41]

2.2.3 Oral Submucous Fibrosis (OSF)

OSMF is a chronic and potentially malignant disorder characterized by juxtaepitelial fibrosis of the oral cavity. It is considered an OPMD. [42] In early stages, the patient experiences a burning sensation and increased sensitivity to spicy food and the mucosa appears blanching and leathery. Eventually in later stages, fibrosis bands form and woody changes occur to the oral mucosa and tongue, leading to stiffness and trismus.[43] The strongest risk factor for OSMF is betel quid and areca nut chewing. It usually affects both sexes in the second and third decades and is more common in Asian populations. [44] Reported malignant transformation rates range between 7% and 30% depending on the sample and study design used. [42, 45-47]

2.2.4 Oral Lichen Planus (OLP)

Lichen planus is a chronic inflammatory mucocutaneous disease. Oral lichen planus (OLP) is a subtype of lichen planus and is considered an OPMD. The prevalence of OLP is 1-4%, [48] with a higher incidence in middle-aged women. The risk of malignant transformation varies between 0.9% and 3.2%. [49-52]

2.2.5 Actinic Cheilitis

Actinic cheilitis is an OPMD of the lip, caused mainly by excessive solar exposure with other factors such as smoking and irritation also playing a role. It predominantly occurs in men, seldom in women. The rate of malignant transformation ranges between 1.4 % and 36% over a period of 1 to 30 years. [53]

2.3 Epidemiology of OPMD

2.3.1 Incidence and prevalence

The incidence and prevalence of OPMD vary widely geographically depending on the amount and manner of exposure to tobacco and other carcinogens and socio-demographic characteristics in that area. In addition, reported numbers vary according to the clinical definition of OPMD used in the study.

The reported global prevalence of OPMD ranges between 1 and 5%. [54-57] However, higher numbers are described in some South East Asian countries such as Taiwan (12.7%), [58] India (13.7%) [59] Sri Lanka (11.3%), [60] and Papua New Guinea (11.7%). [61] Petti [16] conducted a meta-analysis in 2003 to estimate the pooled global prevalence of leukoplakia. Twenty-three

studies focusing primarily on leukoplakia from 17 countries, published between 1986 and 2002, were included in this work. The calculated point prevalence using random effects was 2.6% (95% confidence interval (CI): 1.72% - 2.74%).

Few studies have measured the incidence of OPMD. Some have reported the incidence based on individual diagnosis, mainly the more common ones such as leukoplakia and oral submucous fibrosis. In a study conducted in Taiwan, the age-standardised incidence rates for quid lesions and oral submucous fibrosis were 267.0 and 374.1 per 100,000 person-years, respectively, for areca/betel quid chewers.[62] In contrast, the incidence of leukoplakia has been reported in a number of studies, which we have described in the previous section.

2.3.2 Age, sex, ethnicity and intraoral site distribution

In general, OPMD are seen in middle aged adults, yet no age is immune as leukoplakia has been diagnosed in those aged 15 - 24 years old. [63] The differences in age distribution are probably due to geographical location, ethnicity and lifestyle variations. For example, in developed countries, leukoplakia is mainly found between the fourth and seventh decades of life, whereas in developing countries it is seen 5-10 years earlier. [64] As mentioned before, this could be attributed to earlier exposure to carcinogens and lifestyle differences.

Moreover, a retrospective review conducted in South Africa in 2012 has shown that black South Africans are less affected by leukoplakia than white South Africans. However, the number of black people in their sample was much lower than that of white, which could be explained by limited access to care for black South Africans during apartheid time. [65] A population-based prevalence survey from Malaysia found that among the different ethnic groups living in that geographical

region, OPMD prevalence was the highest among Indians and lowest among Chinese.[66] Ethnic diversity in the Southeast Asian countries usually caries along a diversity of lifestyle practices (e.g., smoking and chewing habits), which could explain the findings of this survey.

Findings from the previously mentioned meta-analysis [16] show that males are predominantly affected. However, the male to female ratio varies depending on the geographical area and the smoking habits of both sexes. In addition, some OPMD such as lichen planus are usually more common in women.

2.4 Risk factors for OPMD

2.4.1 Race and demographic risk factors

Higher socioeconomic status (SES) index, education and income are associated with a decreased risk of oral premalignant lesions [67, 68]. This could be due to limited health care access for individuals with low SES. Moreover, low SES is associated with a higher prevalence of tobacco and alcohol use and food practices that are less healthy, which might increase the risk of developing OPMD in this population. Usually OPMD are associated with adults older than 40 years of age, with an average of 58 years old, and has a male predominance. Some recent studies have shown equal prevalence in both sexes and a higher number of cases in younger adults, which could be attributed to the increased popularity of smoking among women and younger adults [69].

2.4.2 Tobacco smoking and alcohol drinking

Smoking tobacco and drinking alcohol are well-established risk factors for OC. [70-75]. Several studies show positive associations between smoking tobacco and both leukoplakia and oral epithelial dysplasia (OED) but not with oral submucous fibrosis [76-79]. However, the association between alcohol consumption and OPMD is not as strong [79]. Nonetheless, consuming alcohol is still considered a potential risk factor for OED [77, 78]. Although smoking tobacco and alcohol consumption are usually practiced together and have an additive synergistic effect, each of them is an independent risk factor for both OC and OPMD [72, 77].

2.4.3 Tobacco smoking

Tobacco smoking is practiced worldwide and, although the numbers of smokers are declining in the developed world (e.g., Canada), it is still high in some developing countries. Tobacco smoke contains around 4,000 chemical compounds, of which 69 have been classified as carcinogens to date; of those carcinogens, 11 are considered group 1 human carcinogens by IARC. [80]

In Canada, the highest prevalence of smoking based on the 2012 Canadian Tobacco Use Monitoring Survey (CTUMS) falls in the 20–24 years old age group. Although this group remains the highest in prevalence along the years compared with the other age groups, the ratio of people smoking in that category has declined over time, with the percentage being 32% and 20% in 2001 and 2012, respectively. With increasing awareness of the negative effects of smoking and the availability of tobacco cessation programs, tobacco smoking as a risk factor for OPC and OPMD could be modifiable even though tobacco is addictive. [81]

Although modifiable, tobacco smoking remains the strongest risk factor in the developed world, and one of the significant risk factors associated with OC and OPMD in the developing world [69,

77, 82-86]. Moreover, betel quid/ areca nut chewing habits, which usually include tobacco, also play a significant role. The risk of OPC and OPMD increases with the amount of tobacco used, the duration of smoking, and age at which the smoking habit developed and stopped, where cessation in younger ages (< 30 years old) is associated with a 90% decreased risk of developing OC in the future [68, 73, 87]. This shows a possible dose-response relationship between exposure levels and risk of developing OPMD, yet this relationship is not very well understood and needs to be further investigated. Moreover, the risk of OPMD, OC and second cancer decrease with cessation of smoking and continue to decrease with time until it is comparable to that of never smokers after 10–15 years of stopping the habit [87, 88].

There is an agreement in the literature that OPMD risk increases with smoking, [68, 86] yet non-smokers are at significantly greater risk of malignant transformation of precancerous oral lesions. [89] [34, 90]. It could be that smokers have a modifiable risk factor they can act on and control by cessation, whereas non-smokers do not. However, a recent meta-analysis that reviewed the data of 19,676 patients with lichen planus showed that malignant transformation was higher among smokers when compared to non-smokers, but they did not explain if these results are confounded by clinical subtype of OLP.[91] Also, it could be that different biological behaviour of different OPMD tower smoking when it comes to malignant transformation. Indeed, studies reporting a higher risk of transformation in non-smokers have mostly focused on leukoplakia or had higher numbers of leukoplakia cases compared to other lesions. This is evident in a systematic review of observational studies of leukoplakia that did not identify smoking as a factor increasing the risk of malignant transformation. [92]

In addition to the synergistic effect between tobacco and alcohol consumption, cigarette smoking adds to the effect of betel quid chewing on oral leukoplakia development based on the additive interaction model [68].

2.4.4 Alcohol

Similar to tobacco, IARC has classified alcohol as a Group 1 human carcinogen.[93] Several alcohol exposure characteristics such as amount, duration and cessation are strongly associated with cancer risk. [94] The two major components in alcoholic beverages are ethanol and water, in addition to a multitude of other compounds obtained from fermentation, contamination and the use of food additives and flavours. Ethanol and its metabolic by-products like acetaldehyde have carcinogenic and damaging effects on the human body.[93] There is a differential OC risk depending on the concentration of alcohol, where increased ethanol concentrations are associated with increased risk. [95] The amount and concentration of ethanol depends on the type of alcoholic beverage, the lowest being in beer (2.3% - 5%) and the highest in spirits (40%).

Other components that might be present in alcohol and considered carcinogenic are Aflatoxins, N-Nitrosamines, lead and others. However, these contaminants are considered to be present at low concentrations and nowadays are even further reduced. [93] Another mechanism by which alcohol exposure could contribute to cancer risk is nutrition; heavy alcohol drinking limits and impairs the body's ability to absorb different nutrients and lowers the intake of healthy food elements leading to malnutrition, which could play a role in cancer risk. [96, 97]

Although alcohol is a well-established risk factor for OC and oropharyngeal cancer, [70-72] it is only considered a potential risk factor for OPMD. Numerous studies have found alcohol

significantly associated with OPMD diagnosis, [77, 84, 85] even after controlling for tobacco use [98]. However, it is not considered as strong a risk factor as smoking [78, 99]. Furthermore, some studies failed to identify it as an independent factor contributing to OPMD, [24, 68, 86, 100] while others reported only an association with heavy drinking. [101] One study showed an association of heavy drinking with presence of dysplasia at initial presentation and recurrence after controlling for smoking. [102] A dose-response relationship between alcohol consumption and risk of OPMD has been observed [77, 98]; the adjusted odds ratio increased with increasing levels of alcohol consumption, frequency, and duration, even in non-smokers and non-chewers of tobacco. Also, a synergistic effect of alcohol on betel quid chewing has been suggested.[100]

Varying results regarding the effect of alcohol drinking on the risk of malignant transformation of OPMD have been observed. While numerous studies did not identify any statistically significant relationship between alcohol intake and OPMD transformation, [69, 92, 103-105] others reported a positive association.[91] For example, a recent retrospective observational study reported that heavy alcohol drinking was significantly lower in oral cancer patients who had OLP and oral lichenoid lesions as precursors.[106] On the other hand, results from a meta-analysis on OLP showed that drinking alcohol was significantly associated with increased odds of malignant transformation.[91]

2.4.5 Betel quid

The WHO has classified areca nut and betel quid commonly chewed in South East Asia as human carcinogens. Betel quid, sometimes called paan, is composed of more than one element; generally it is constituted of betel leaf, areca nut, slaked lime and possibly tobacco. [107] The two major components of betel quid, tobacco and areca nut, can also be chewed on their own without the betel

leaf. However most of the time when either is mentioned as a chewing habit, the product being chewed is a composite, and it is more popular to chew areca nut in the form of a betel quid than on its own.[108]

Much epidemiological research including case-control and cohort studies have reported increased OC risk for betel quid chewing with and without tobacco, even after adjusting for alcohol and smoking as confounders. All of these studies were conducted in South East Asia countries where the habit of chewing is part of the culture and very popular.[107]

In India and South East Asia, betel quid chewing has been strongly associated with OPMD and could be considered the principal cause for oral leukoplakia and oral submucous fibrosis in this geographical area [68]. In addition, studies have shown a dose-response relationship, that is, the risk of precancerous lesions increases with the frequency and the duration of the chewing habit. [61, 108-111] Although there is a synergistic effect between betel quid and smoking, [68] betel quid and areca nut chewing without smoking or additional tobacco is an independent risk factor for the development of OPMD. [61, 108] Other researchers have even found that betel quid chewing carried the highest risk of developing OPMD compared to alcohol and smoking tobacco among their population.[100]

2.4.6 Immunosuppression

Consistent results from clinical epidemiological studies show an increased risk for head and neck cancer among transplant patients, graft versus host disease patients, and HIV patients. The risk factors among these groups of patients include treatment regimens and level of immune suppression. [112, 113] Based on the OPMD definition proposed by Sarodi, [23] the previously

mentioned diseases would fall under the same umbrella given the prolonged immune suppression state accompanying them.

Lip squamous cell carcinoma and leukoplakia were observed to be significantly associated with renal transplant patient status. The incidence of lower lip carcinoma and lip leukoplakia were 29 and 22 times higher among kidney transplant recipients compared to healthy controls, respectively. [114]-[115] However, the risk for head and neck cancer was not increased among liver transplant patients. [116]

Although there are a number of observed cases of oral cancer (tongue in particular) in HIV patients especially in the era before HAART therapy, the two main types of cancer associated with HIV patients are Kaposi sarcoma and lymphoma. However, it is noteworthy that HIV patients with tongue cancer were significantly younger when they were diagnosed than their matched controls.[117] The earlier onset could be attributed to the effect of HIV infection on specific aspects of the immune system. In addition, OC cases have been reported in patients with chronic graft versus host disease after bone marrow and stem cell transplant. [118, 119]

Another condition that could compromise the patient's immune response and that has been related to oral lesions is diabetes. A recent meta-analysis examining the relationship of diabetes type 2 with OPC and OPMD concluded that patients with type 2 diabetes have a higher incidence and mortality of OPC as well as OPMD compared to non-diabetics.[120]

Accumulating evidence is suggesting that the progression from OPMD to OC is strongly informed by an abnormal immunoenvironment. [121] A study investigating the genetic susceptibility to oropharyngeal squamous cell carcinoma and OPMD reported an association between these

diseases and the alteration of immune system genes and genes with metastatic potential (TNF-alpha, TGFbeta-1, MMP-1).[122, 123] Moreover, it has been suggested that Galectins, which are a group of proteins involved in immune responses, are associated with OED progression [124].

Chronic inflammation and the associated immunosuppressive microenvironment, iatrogenic immunosuppression, immunosuppressive state, and altered immune response as a result of a viral infection, should all be further studied to reach a more solid conclusion about the role of immunosuppression as a risk factor for OPMD and its progression to OC.

2.4.7 Sun exposure

Chronic solar exposure and artificial ultraviolet radiation have been well documented as causes for actinic keratosis.[39, 125, 126] There is no evidence that these exposures are associated with other OPMD.

2.4.8 Diet

The role of diet in oral carcinogenesis has been documented. Numerous observational studies have suggested that the diet of different populations may play a role in determining their rate of cancer. [127, 128] A low intake of antioxidants and fibres (fruits and vegetables), micronutrients (vitamin C, folate, zinc), and a high intake of fried food and red meat are strongly negatively correlated to oral cancer risk. [127] A daily intake of fruits and vegetables is associated with 49% and 50% reduced risk of cancer, respectively. Citrus fruits, orange and yellow vegetables were found to be particularly protective. [129] A similar protective effect of fruits and vegetables has been reported by others.[70, 130, 131] Therefore, diet is considered a modifiable risk factor of OC and OPC that could have either a carcinogenic or protective effect.

Unlike OC and OPC, the association between dietary habits and OPMD has not been as thoroughly investigated. However, a number of studies, mostly from India and South East Asia, have shown that fruits, vegetables, micronutrients (beta-carotene, ascorbic acid, calcium and fibres) play a protective role against OPMD and their progression to OC. [132-136] In conclusion, diet plays a role in oral carcinogenesis and further epidemiological studies are needed to determine its effect on OPMD risk and progression to cancer.

2.4.9 HPV

Although it is accepted worldwide that HPV is a primary and necessary cause of cervical cancer [137], and a well-established risk factor for a subset of oropharyngeal cancers [138], not much is known about its role in OC and OPMD and their progression to OC [139]. The rate of detection of HPV in OPC varies widely in the literature (20% - 90%) [140]. Around 30% to 70% of OPC could be attributed to HPV infection [141, 142]. HPV is less strongly associated with OC and OPMD. A systematic review conducted in 2005 reported a prevalence of HPV in OC cases of 23.5%[143] and an earlier one reported this figure to be as high as 46.5%[144]. In both reviews, HPV was shown to be an independent risk factor for oral squamous cell carcinoma.

Similar to oral and oropharyngeal cancer, the rates of HPV detection in OPMD are highly variable, ranging from 0% - 85% [145, 146]. This high variability could be attributed to geographical differences and heterogeneity in the definition and selection of cases, but mostly to inconsistency in the methods used for virus detection. A recent review and meta-analysis by Jayaprakash et al [141] estimated the prevalence of HPV16/18 in oral and oropharyngeal dysplasia to be 24.5%, and when oral dysplasia was considered separately, the prevalence of HPV was 25.3%. The prevalence did not significantly differ in relation to oral subsites, geographical location of the study, age group

or grade of dysplasia. However, males had twice the odds of having HPV detected in their lesions compared to females. The odds of detecting HPV in dysplastic lesions and invasive carcinoma were over 3 times higher compared to benign lesions and normal mucosa. An earlier systematic review has reported similar results; HPV was detected 3 times more in OPMD than in normal mucosa. Yet, the authors did not describe the relationship of HPV prevalence with age, sex, geographical location or method of detection [144].

Thus far, it is not clear whether HPV plays a similar role in OPMD progression to OSCC as in cervical cancer, because the evidence we have is limited and the association of HPV with OPMD presented in the literature does not imply causality. Stronger prospective larger scale epidemiological studies need to be conducted to be able to make such an inference.

2.4.10 Other potential risk factors: BMI, poor oral hygiene and oral health

A number of studies have found an association between low BMI and OPMD [59, 98, 147]; this association may be explained by malnutrition and low consumption of fruits and vegetables, which deprive the individual of their protective effect against OPMD and their progression to cancer. [132, 133, 135] Similarly, good oral hygiene seems to protect against OPMD. [148]

2.5 Transformation of OPMD to OC

To date, we cannot precisely predict the clinical behaviour of OPMD, including which lesions will progress to cancer and which ones will not or will even regress, and if progression occurs when the transformation will take place and who are the high-risk individuals. Reported rates of malignant transformation of OPMD vary greatly in the literature, from 0.1% to 40% [104]. Different geographical locations, definitions of cases, lengths of follow up and study designs could

have contributed to this variability. Some investigators only looked at histologically confirmed dysplastic cases, whereas others have used clinical diagnoses as an inclusion criterion. Some studies were prospective, and others were retrospective. In addition, different environmental factors and lifestyle habits related to different geographical locations may have played a role in producing such a large range of malignant transformation rates.

A meta-analysis published in 2009 that only included dysplastic lesions reported a malignant transformation of 12.1%. This work also showed that lesions treated surgically had lower transformation rates compared to lesions that were not excised, even after adjusting for grade of dysplasia. However, the transformation rate is only lowered after surgical treatment but not eliminated, given that some lesions recur and continue to progress. [33] The aforementioned meta-analysis and two recent systematic reviews did not identify alcohol drinking or smoking as factors related to transformation. [105, 120] On the contrary, higher rates of transformation are observed among non-smokers. [34, 89, 90] In addition, idiopathic leukoplakia lesions, the ones for which clinicians fail to identify possible causative factors, are known to carry an increased malignant potential. [64]

2.6 Risk factors for OPMD transformation to OC

2.6.1 Clinical characteristics

Different locations of the lesion in the oral cavity appear to be associated with varying risks of malignant transformation. Lesions on the tongue have an 87% increased risk compared to other oral subsites [33, 92, 149]. Similarly, the floor of the mouth seems to carry a higher risk [150, 151]. Large lesions exceeding 2 cm^2 in size have more than 5 times increased risk of developing a

malignancy [92, 152], and non-homogenous lesions, also referred to as erythroleukoplakia or speckled, are at increased risk [92].

2.6.2 Demographics

Individuals above 45 years old at the time of initial diagnosis are considered at higher risk [92]. Conflicting results about sex as a risk factor are being reported; both males and females seem to be at higher risk when compared to each other in different studies [92, 149]. However, two recent systematic reviews have concluded that females have a higher risk of transformation [92, 105].

2.6.3 Dysplasia

It is largely accepted that dysplastic lesions are at higher risk of malignant transformation than non-dysplastic ones [29, 149, 153]. The presence of dysplasia in the precancerous lesion is considered one of the most important predictors of transformation. Indeed, in most clinical settings the grade of dysplasia dictates the management strategy, that is, higher grade lesions receive more aggressive treatments (e.g., surgical excision), compared to lower grade lesions that are only followed up periodically.

Although the presence and degree of dysplasia are the golden standard to stratify lesions into high risk and low risk, there is no consensus that more severe grades are at higher risk of OC. This inconsistency is probably due to the subjectivity of the oral dysplasia grading system; there is a high variability in the pathologists' interpretation of the presence and degree of dysplasia and which morphological criteria are more important in diagnosing and grading dysplasia [154]. However, two recent studies, a meta-analysis[33] and a systematic review [92], have concluded that the grade of dysplasia is a significant determinant of malignant transformation, with more severe grades being at considerably higher risk.

Overall, the dysplasia histopathologic grading system is an insufficient tool on its own to determine which lesions have an increased potential of transformation, which highlights the need for new adjunctive diagnostic techniques to identify high risk OPMD.

2.6.4 Molecular biology

Given the subjectivity and shortcomings of clinical and histopathological predictive factors in determining the transformation of OPMD to OPC with high accuracy, genetic biomarkers have been investigated to find an objective measure that would act as a prognosticator for disease behaviour. Different genetic and epigenetic alterations have been associated with oral carcinogenesis.[155-160] Genetic alterations (mutations, deletions) are irreversible changes that affect the DNA itself, whereas epigenetic changes are heritable changes in gene expression that do not affect the underlying DNA sequence. In other words, epigenetic alterations are phenotypic rather than genotypic and are potentially reversible and transient unlike genetic changes. In cancer, epigenetic changes influence the DNA by silencing tumor suppressor genes and activating oncogenes mainly using three mechanisms: DNA methylation, histone modification and noncoding RNA regulation. [161]

Research has suggested numerous biomarkers that may aid in the prognosis and diagnosis of OPMD.[162] One of the promising markers that are consistently recognised as potentially independent risk predictors for malignant transformation is loss of heterozygosity at certain loci. [156, 158, 163, 164] A recent randomized clinical trial has validated loss of heterozygosity as a marker for OC risk in patients with OPMD, and suggested its use in routine clinical practice.[165] A prospective observational study showed that loss of heterozygosity in a specific group of loci (9p, 17p, 4q) is associated with a 52 folds increase in OPMD progression to OC. [163] The

accumulation of aberrant gene methylation is another possible marker associated with the malignant transformation of oral leukoplakia. [166] P16 and MGMT are two genes that are hypermethylated in OC, and they are also aberrantly methylated in oral leukoplakia cases.[167] P16 methylation, which leads to P16 gene silencing, is an independent risk factor for malignant transformation in oral epithelial dysplasia and has been proposed as a prognostic biomarker for progression. [168, 169] Moreover, OPC and oral leukoplakia seem to have the same pattern of microRNA dysregulation. [170]

Biomarker analysis tools (e.g., quantitative DNA methylation analysis and serum/saliva miRNA analysis tools) are available and the majority do not require invasive procedures to obtain a sufficient sample for the analysis. While tissue biopsy, dysplasia diagnosis and histopathological grading remain the golden standard to diagnose OPMD, stratify their risk of progressing to malignancy and dictate the treatment modality (surgical vs non-surgical), biomarkers could be used as adjuncts especially for low grade OPMD.

2.7 Management

Treatment for OPMD includes surgical and non-surgical options. A recent Cochrane review assessing interventions to treat leukoplakia reported that the available evidence on medical and complementary treatments for leukoplakia is limited. Although vitamin A and beta carotene might be effective in healing oral lesions, recurrence and systematic adverse effects are high. [171] To date, no randomised controlled trials have compared surgical treatment to observation of OPMD only; this explains why the Cochrane review did not include any articles that discussed surgical treatment. However, a number of observational studies have shown that surgically treated lesions had a significantly reduced risk of progressing to cancer compared to lesions that had not received

surgical treatment. [172, 173] Also, a longer interval to progression was noted in the treated lesions that developed cancer compared to untreated lesions. [172] Usually, if the patient is medically fit, the presence of dysplasia and grade of dysplasia would be the first factors influencing the clinical decision. In the case of no or only mild dysplasia without other concerning clinical features [e.g., size of the lesion >200 mm^2, leukoplakia subtype (speckled or verrucous), location (e.g., tongue is a high-risk)] [152], usually regular follow up is the path chosen. Regardless of the treatment chosen, clinicians must counsel patients with OPMD about the elimination of modifiable risk factors such as smoking and alcohol intake, among others.

2.8 Premalignant lesions registries

Usually premalignant lesions are documented within cancer registries, for example the Cancer Registry of Norway (CRN) [17, 18], or as a part of another database (non-disease specific/ patient registries), such as the nationwide histopathology registry in the Netherlands (PALGA) [19] or U.S. Medicare claims [11]. However, the literature is scarce on methods of registration and documentation of premalignant lesions in general and oral premalignant lesions specifically. Most of the epidemiological studies concerning OPMD were conducted based on histopathology databases, clinical records reviews, medical insurance claims, or retrospective analyses of the history of cancer cases. [11, 69, 174]

2.9 What is a clinical database (disease registry, patient registry)?

A registry in the field of health is defined as "a file of documents containing uniform information about individual persons, collected in a systematic and comprehensive way, in order to serve a predetermined purpose". [175] Sometimes such files are called a patient registry, clinical registry, clinical data registry or disease registry. All those terms are used to describe the same entity, that

is, a registry in a health care system.[176, 177] The US National Committee on Vital and Health Statistics defines a registry as "an organized system for the collection, storage, retrieval, analysis, and dissemination of information on individual persons who have either a particular disease, a condition (e.g., a risk factor) that predisposes (them) to the occurrence of a health-related event, or prior exposure to substances (or circumstances) known or suspected to cause adverse health effects". [178] Another definition by the Agency for Healthcare Research and Quality is: "an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves a predetermined scientific, clinical, or policy purpose(s)".[178]

Clinical registries can be classified based on a variety of factors; the target population (population based, hospital based, single provider), geography (provincial, national, international), type of exposure (disease or condition registries, product registries, health services registries) and other features.

It is important to distinguish between population-based and clinical data registries. Population-based registries aim to identify and include all cases among people living in a defined geographical area, whereas clinical registries include all patients carrying a certain disease or risk factor who are being treated in one or multiple centers. Also, some registries are based on a single provider.

Clinical registries can organise clinical data in an electronic form (digitized) with web-based access; studies have demonstrated that data captured in an electronic database can be aggregated, studied and analyzed to improve patient care and outcomes. [179] Because the use of electronic registries can improve healthcare delivery significantly, healthcare providers now have to familiarise themselves with health information technology including electronic medical records

(EMR), and web-based registries and databases. [179] Moreover, patients can contribute to the data collection process by answering questionnaires directly into the registry.[178, 180]

2.10 Significance and uses of clinical databases

Clinical registries incorporate a huge volume of clinical data on real life clinical cases and events (real world data). They are powerful tools that can act as a research and educational platform by providing a good repository for different types of studies, and allowing the investigation of various research questions starting from the effectiveness and safety of treatments, to the assessment of quality of care. The usefulness and value of clinical registries can be appreciated by different stakeholders including clinicians, professional organisations, regulatory agencies, patients, patient advocacy organisations, drug companies and device manufacturers, among others.

The research conducted based on clinical databases translates the available everyday clinical data into meaningful information that is critical to help policy makers take evidence-based decisions to change policy, as shown by real life examples [181, 182].

Moreover, by providing prospective and retrospective observational data, clinical registries allow researchers to complement RCT and fill gaps in the literature. For example, they may enable the investigation of questions that RCT cannot examine due to ethical, time, financial, or generalizability issues. [183] The generalizability of the data is supported by the huge number of patients included in registries, as multiple centers usually participate. This attracts the interest of researchers in conducting studies using registry data; it is cost-effective because the data collection system is already in place and working, and there is a possibility to study rare diseases and interventions.[180]

However, evidence derived from patient registries needs to be interpreted carefully; one must keep in mind the source of data, validity issues and different types of biases that may be encountered. Nevertheless, given their timeliness, comprehensive data gathering and low restrictions for inclusion, clinical registries can: 1) provide complementary information that extends RCT results to patient populations not included in those trials, 2) demonstrate real-world effects of treatments outside the research setting (efficacy vs. effectiveness), and 3) provide long-term follow up. [178] Additional uses of clinical data registries include evaluating the natural history of a disease and monitoring safety and harm by serving as an active surveillance system for unexpected side effects and complications of an intervention.

2.11 Planning and designing a clinical database

2.11.1 Purpose - why build it?

It is of paramount importance to plan ahead all the details of a registry before going into the design process. Good planning involves a clear description of the purpose of the registry, how the information collected will be used, and how this information will help to address the objectives of the registry. A clear purpose is fundamental to develop a good design, including determining aspects such as inclusion and exclusion criteria, the population, data to collect and the duration of the registry. Also, without a clear purpose, it would be difficult to know how to utilise the registry properly and whether it would be useful or not. Different clinical registries have different purposes, and some are multipurpose. The objective or the goal of the registry dictates its focus. For example, registries with objectives such as monitoring incidence and prevalence and describing the natural history of a disease would be disease-focussed registries, and registries targeting product and device users are built with the intention to measure and assess clinical effectiveness, cost-

effectiveness and safety. Regardless of whether a registry is single- or multi-purpose, all of its goals should be clear and translated into well-defined objectives and questions to be fulfilled by the registry.[178]

2.11.2 Database Team (personnel)

The availability of human resources is important when building a registry or database. To establish a team, one should think about the number of staff needed and the type of technical and professional expertise required. For example, to develop a registry focusing on OPMD, one needs to have access to or guidance from oral pathologists/ pathologists, maxillofacial / oral / head and neck surgeons, epidemiologists, statisticians, IT and data processing experts. In addition, technical staff including clerks, typists and administrators will be needed. [184]

2.11.3 Scope of the registry

The capacity and extent of the clinical registry in terms of size, duration, geography and cost must be determined and evaluated. Size refers to the number of data points, patients and participating sites. Duration is mainly governed by the purpose and questions the registry aims to answer. A registry looking into the cancer transformation potential of OPMD must run for at least five years, as cancers arising from a precursor lesion usually occur at that time period.[33] Also, the setting and the means by which the subjects are to be recruited and whether the registry is local or global all play a role in characterising the scope of the registry. [178]

2.11.4 Defining outcomes, data elements, target population

Setting the primary and secondary outcomes or end points for the clinical database is one of the most important early steps. It will help prioritise and select what information to collect and could

determine the duration of the registry.

Potential data elements to be included in the registry should be considered carefully. The elements chosen should properly and adequately address the registry questions. It is advisable to seek experts' counsel to ensure that the suitable and relevant data elements are selected. Also, planners and designers would benefit from consulting data standards for e.g., Clinical Data Interchange Standards Consortium (CDISC), other registries (e.g., SEER) and previous work by others.[185] Furthermore, the source of the data will have an effect on the feasibility of collecting certain information. For example, if the source of data are medical records, then they rarely contain detailed information about smoking, alcohol drinking and diet behaviours. If those data elements are important to address the registry questions, then another source of data should be sought to gather this information.

The target population for a clinical registry is the population to which the results derived from the registry apply and could be generalised, such as all patients with a given disease or risk factor. Some registries include the entire target population, but most include only a representative sample or use multiple phase sampling in which basic data are collected on the larger sample and more comprehensive data are collected on a subset of that target population (e.g., patients attending major centers contributing to the registry).[178, 185]

2.11.5 Internal and external validity

External validity is another name for generalizability and refers to whether the findings derived from the clinical database can be generalized to other subjects not included in the registry. The patient recruitment strategy used and differential loss to follow up could affect the generalizability

of the registry findings and therefore external validity. Internal validity, on the other hand, concerns the degree to which the resulting inference is actual, free of bias, confounding and errors. The definition of data elements is an important part in registry development and should conform to standard definitions used by well-established registries. In addition to facilitating data sharing and linking between different registries, this ensures consistency. Having a precise definition for each data element or variable to be collected, including ranges and acceptable values when applicable, helps standardisation and reproducibility of the data collection process and decreases data entry errors, thus increasing internal validity. [178]

2.11.6 Data sources

Depending on the type of clinical registry, its purpose and the elements chosen for data collection, appropriate data sources must be identified. Data sources are classified into primary and secondary based on the relationship of the data to the registry purpose. Primary data sources are created for the registry to ensure completeness, validity and reliability. They are prospectively planned, and data are collected following a protocol and a common procedure across all registry sites and patients. They are usually used when there is a need to collect data that are not available elsewhere, or there is a problem with the completeness and accuracy of the available source.[178] An example of this situation is a patient-filled questionnaire to collect information on lifestyle factors (e.g., diet, smoking, sexual behaviour). By contrast, secondary data sources are those readily available and established to serve purposes other than the registry, for example, standard medical care, insurance claims, other institutional databases, electronic medical records, and existing registries, among others.

2.11.7 Ethical and legal considerations

2.11.7.1 Database transparency and ownership

Registry transparency involves making the clinical registry operation information public. This could be achieved by making the registry objectives, inclusion criteria, enrolment/ registration processes, protocol and procedures, data sources and funding information available to whoever is interested. Registry transparency helps to educate the public and professionals about the scientific process and to build their confidence in the integrity and accuracy of the registry process and the results of data analysis. Such clarity is vital to appreciating the possible benefits of research using health information.

The concept of data ownership does not really fit with health information as it does not acknowledge individual patients' interest in the privacy of their health information. It is better to discuss registries in terms of custodianship or guardianship, where the guardian (usually the principal investigator) has legal rights and obligations to protect registry participants' privacy and dignity.[178]

2.11.7.2 Consent requirement

Permission should be obtained for registry participation from patients, and they should consent to any intended future use of the collected information. The consent form should include a paragraph describing the registry, its purpose, the duration of patient participation, the types of data collected, and the procedures followed. It should clearly state how confidentiality and patient identity will be handled, and if patients will be contacted in the future for further data collection. The description

of any anticipatable risks or discomforts the patients might feel are one of the core elements to include.

Moreover, ethically the consent form should clearly state that subjects' participation is completely voluntary, with no penalties or reason to fear that it would influence their care if they decided not to participate, and that they have the right to withdraw at any time, with the withdrawal procedure explained in the consent form. In addition, contact information of persons to whom questions can be addressed regarding the registry, subjects' rights, or to file a complaint should be listed.

2.11.7.3 Protection of patient privacy

During the design of the clinical registry, patient privacy and confidentiality must be established. This involves proper handling of patients' identifiers. There are multiple technical ways to de-identify patients and ensure that the data cannot be traced back to the individuals; data collection could be anonymized, or data could be made only indirectly identifiable by the use of encryption and pseudoanonymization [178].

Anonymization consists of permanently removing identifiers, changing the data into a form that does not recognise individuals. Anonymized data collection is not suitable if data linkage or health information exchange is planned or the data are used for studies requiring patient follow up. Encryption involves translating plain text data into codes (cypher text), readable only by a group of people who hold the decryption key. However, encryption requires an advanced and careful data management policy.[186]

Pseudoanonymization is considered a more sophisticated approach than anonymization, and involves two steps: depersonalization, in which identifiable data are separated from other clinical

data and stored in a separate location, and pseudonymization, where a unique identifier is given to each de-identified record. This unique identifier should not include any characters that could identify the patient or point toward his or her identity. The process could be reversible or irreversible. [178] Decryption or decoding keys for both encrypted or pseudanonymized data should remain solely with the health care provider who is the source of the patient information and no one else should have access to it.

2.11.8 Funding

A realistic budget covering every aspect of the development of the registry should be estimated, taking into account the intended size of the registry, pilot testing, continuous quality control and data analysis. Depending on the duration of the registry, long term funding might be crucial to ensure sustainability. Also, it is wise to expect registry costs to rise with time as the load of cases to follow up increases.[187]

2.11.9 Pilot testing

Pilot testing is a very important step in the implementation plan; it allows insight into issues that may be encountered in the future, gives an opportunity to develop solutions and refine the registry before official launching. Aspects to consider during pilot testing are: the ability of the registry procedures and protocol to correctly and comprehensively identify participants qualifying for inclusion, the time needed for recruitment/ enrolment, the time required to complete forms, the practicality of the data collection tool, the identification of issues in the data collection instruments, the accuracy and reproducibility of the data exported. In addition, pilot testing is useful to test the adequacy, relevance and completeness of personnel training.

2.12 Quality control

Four main aspects of quality assurance are used to indicate quality level in any registry: comparability, completeness, validity and timeliness. The sections below describe these indicators and quality assurance procedures.

2.12.1 Data quality

In addition to thoughtful consideration of what data elements to include, and ensuring that all important variables are being collected, it is important to make sure that the data are being collected accurately, completely and in a uniform standardized way to maximize their validity. The integrity and quality of data being entered must be monitored and audited while auditing must be done continuously to ensure the reliability of the database. This will influence whatever inference is obtained from the collected data, including the results of studies, and will impact decision making based on that source of information.

Almost all software systems used for clinical databases have a built-in validation tool or could be customised to include one. Those tools can identify missing data and out of range entries, and specify that some fields can only accept alphanumeric values or must be filled before advancing forward. Such intrinsic quality control methods ensure data completeness and minimise the amount of missing data. Validation tests should be done periodically, and results should be reviewed and documented. Issues identified through such processes should be resolved, and all unsolved matters must be followed-up and tracked.

Additional dimensions to consider in data quality are comparability, completeness and timeliness. Comparability is defined as the extent to which coding and classification procedures of a registry, together with the definitions for recording and reporting specific data items, adhere to agreed international guidelines. This includes the definitions of disease, incidence and outcome. Ensuring

comparability eases pooling and comparison of data from one registry with those from other registries following the same standardized guidelines.

Completeness refers to the successful inclusion of all incidental cases in the target population in the registry. Finally, timeliness is the rapidity with which a registry can collect, process and report sufficiently reliable and complete data. It could also refer to time from the diagnosis of a disease (e.g., OPMD) until registration in the database. [188]

2.12.2 Quality assurance of registry procedure

External audits could be used to maintain the quality of registry procedures. Maintaining registry quality could be done through the monitoring of sites with high enrolment numbers, or sites with a history of prior dissatisfactory audit results. In addition, keeping a detailed registry manual explaining policies, structure and procedures is very important to maintain quality control. This document will act as a reference for participating sites and for the auditing process. Audit results should be communicated to stakeholders and documented. [178]

2.13 Education and training

All personnel involved in the registry data and procedure management should receive continuous and up to date training in the domains related to their role. Proper training on how to meticulously perform data and procedural management and handle issues will first help to minimize the amount of error, and second to identify and share information on common mistakes, and therefore improve quality control.

2.14 Flexibility

Given the fast advancement and turnover of scientific research, it is important to choose a database design that is flexible toward future additions and changes. Examples of such modifications are adding new data elements or tools.

3. RATIONALE

It is well recognised that clinical data play an essential role in research and are an asset when making clinical decisions. Clinical databases or registries are fundamental to track clinical findings, uncover longitudinal relationships between exposures and outcomes, recognize complications early, and monitor disease prevalence, incidence and progression. In the case of OPMD, OC progressing from a pre-existing oral premalignant lesion tend to be smaller in size and to have better survival because they are detected early on in their course of development. In one series of patients with OPMD followed in a multidisciplinary clinic, 100% of patients who experienced malignant transformations were diagnosed early with stage I disease and the disease-specific survival was 100% at 24 months.[15] The aforementioned highly valuable benefits, combined with the considerable risk of malignant transformation in OPMD, and the high burden of OC and other cancers on the Canadian health system, as 1 in 2 Canadians is expected to develop cancer in his/her lifetime, [5] all warrant the initiation of a clinical registry for OPMD, and premalignant lesions in general. This need was considered and appreciated by our team at the Oral and Maxillofacial Surgery (OMFS) department at the McGill University Health Centre (MUHC), and the decision was made to build a clinical departmental database concerning OPMD. The OMFS clinic at the MGH-MUHC manages a wide scope of conditions (e.g., head and neck pathology, facial trauma, TMJ, dento-alveolar conditions), including OPMD. The clinic receives a large number of referrals that are seen and treated by specialists such as maxillofacial surgeons and oral pathologists. The department, therefore, provides a fertile ground for teaching and future research, research that would be facilitated by the robustness of clinical data and large number of subjects offered by an OPMD database.

4. OBJECTIVES

The overarching objective of this thesis was to develop and implement an OPMD database at the OMFS division at the MGH-MUHC. The specific objectives are to:

1- Describe the steps of the development and implementation of the OPMD database.

2- Describe clinical, biological and socio-demographic factors among people with OPMD attending the OMFS clinic at the MGH-MUHC.

5. METHODOLOGY

5.1 Research Ethics Board Approval (REB)

After the initial planning and before the actual implementation of the database, the project protocol was submitted to the MUHC REB for review. The REB committee approved the clinical OPMD database protocol and granted a waiver of consent for all retrospective patients who are not currently active (whose follow up at the clinic has stopped).

5.2 Database purpose and objectives

The objective of the OPMD database is to facilitate future epidemiological research, with the aim to identify possible clinical and biological determinants of premalignant lesions transformation. This objective includes the implementation of measures of quality assessment and quality assurance research.

5.3 Choosing data elements and defining outcomes

A review of literature was conducted to identify the variables to include in the database. We used the Medline electronic database to search for articles; and we limited our results to articles published in English. References of relevant articles were also explored. A list of variables that support and fulfil the database objectives were identified, and the data end points were established. Also, our previous experience in establishing the Maxillofacial Oncology and Reconstruction Surgery (MORS) registry helped us in identifying pertinent data elements. Chosen variables were reviewed and thereafter used to identify suitable primary and secondary data sources.

Having a clear research objective in mind was very helpful in determining the database outcomes or endpoints. We identified a total of four outcomes: the primary outcome was the development of oral squamous cell carcinoma, and three secondary outcomes or endpoints were: progressing on the dysplasia scale from mild to moderate to severe, developing a new premalignant lesion, and being disease free or stable at 5 years of follow-up.

5.4 Data sources

We used both primary and secondary data source; the latter includes: (i) patients' hospital electronic medical records (EMR) accessed through OACIS, which is the MUHC information system, and (ii) OMFS clinical paper charts (2008 – 2016) and OMFS clinical EMR (2016-2017) accessed through Medesync. However, some details regarding patients' lifestyle and clinical information would not have been satisfactorily obtained from the existing data sources. For that, the team felt the need to develop data sources for the purpose of the registry, for example, a self-complete questionnaire for the patients. In addition, we created standardised clinical forms to facilitate the accuracy and validity of the clinical data collection. Those standardised clinical forms will account for individual differences (different individuals filling them), while maintaining the capacity to collect important variables. The standardised clinical forms will not merely act as data collection tools, but also as clinical notes corresponding to patients' visits.

In summary we established three primary data sources: a patient self-completed questionnaire and two clinical forms. The former was developed to collect details on patients' sociodemographic and lifestyle factors that are too comprehensive to be recorded in clinical records. This includes extensive details on tobacco use, alcohol drinking, marijuana use, sexual behaviour and detailed demographic information about occupation and income.

The two clinical forms are: (i) a consult form that is used only once, either at the patient's initial visit to the OMFS clinic, or whenever a patient known to the clinic who is receiving treatment for other concerns develops a lesion that is clinically or histopathologically considered to be an OPMD; and (ii) a second form, the follow up form, for periodic subsequent visits of the patient to the OMFS clinic after the initial treatment has been provided. The clinical forms and patient filled questionnaire are displayed in appendix I, II.

5.5 Choosing the software

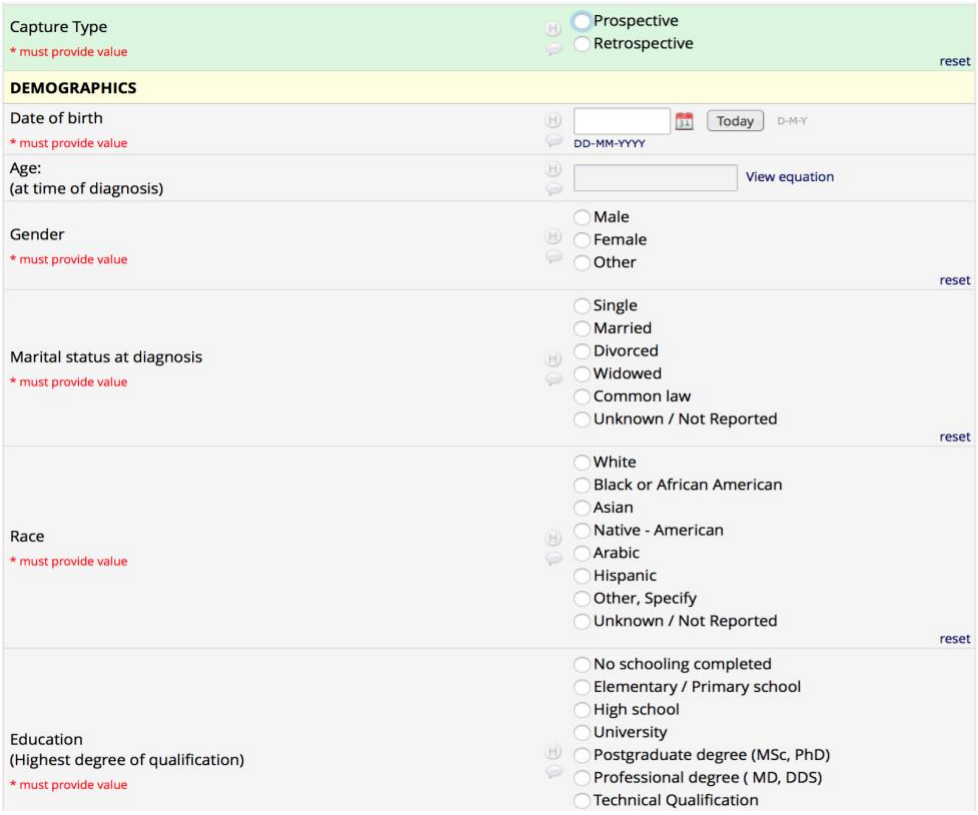
REDCap (Research Electronic Data Capture) was the software chosen to design the database. It is a web-based application for building and managing online databases and surveys that can be accessed from anywhere in the world over a secure web connection. It is supported by the MUHC research centre, met our team's needs and provided several advantages. REDCap is fully customisable and has built-in data quality measures for internal validity. It permits the designer to condition variables to be of a certain form (numbers, dates, letters only), specify ranges, create or customise data quality rules to check for discrepancies, and label some variables as required so all subjects within the database have a minimum data set. Moreover, the application offers the use of branching logic, which hides fields that are not applicable for a certain patient and condenses very long forms into very simple and short ones by showing only relevant fields. Finally, one of the great advantages of REDCap is that building a database using this software does not require any programing experience.

5.6 Database setup

The established data endpoints were: developing OC, progressing on the dysplasia scale, developing a new lesion and being disease free or stable. To record these endpoints accurately, the

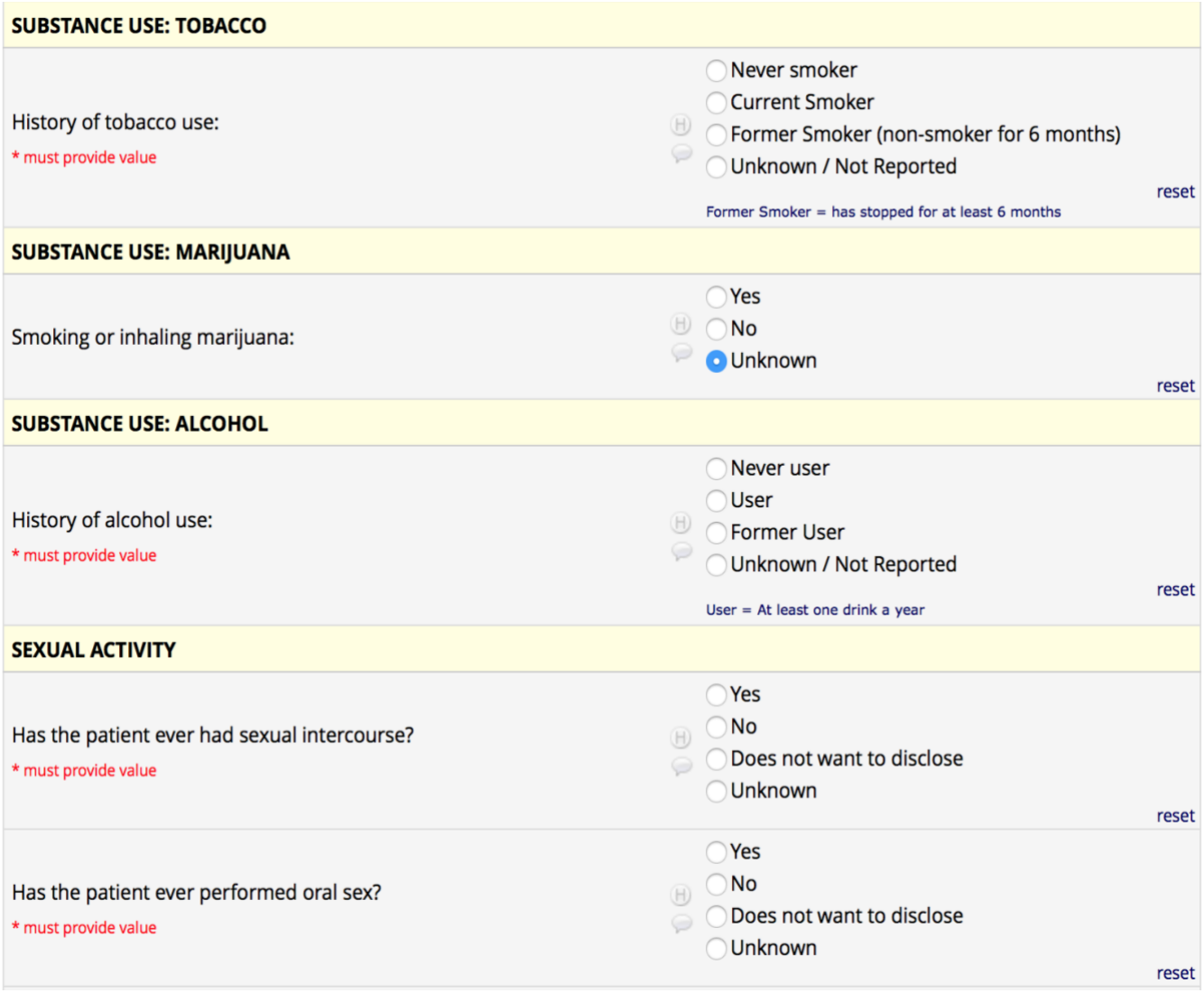
data capture must be done at different time points (consult event and subsequent follow ups). Therefore, the platform was customised to support a longitudinal data collection model. Afterwards, the database instruments (forms) and fields (variables) were created. The final database is composed of 5 instruments, and 900 different fields divided among those instruments. The first instrument comprises the patient's study ID and demographics section, titled "Patient's demographics" It contains demographic variables such as age, gender, marital and employment status.

Figure 1: Patient's demographics form



The second instrument is lifestyle history, which involves comprehensive details on relevant habits including tobacco, alcohol and marijuana consumption as well as sexual behaviour details.

Figure 2: Habits and lifestyle form



SUBSTANCE USE: TOBACCO

History of tobacco use:
* must provide value

- Never smoker
- Current Smoker
- Former Smoker (non-smoker for 6 months)
- Unknown / Not Reported

reset

Former Smoker = has stopped for at least 6 months

SUBSTANCE USE: MARIJUANA

Smoking or inhaling marijuana:

- Yes
- No
- Unknown

reset

SUBSTANCE USE: ALCOHOL

History of alcohol use:
* must provide value

- Never user
- User
- Former User
- Unknown / Not Reported

reset

User = At least one drink a year

SEXUAL ACTIVITY

Has the patient ever had sexual intercourse?
* must provide value

- Yes
- No
- Does not want to disclose
- Unknown

reset

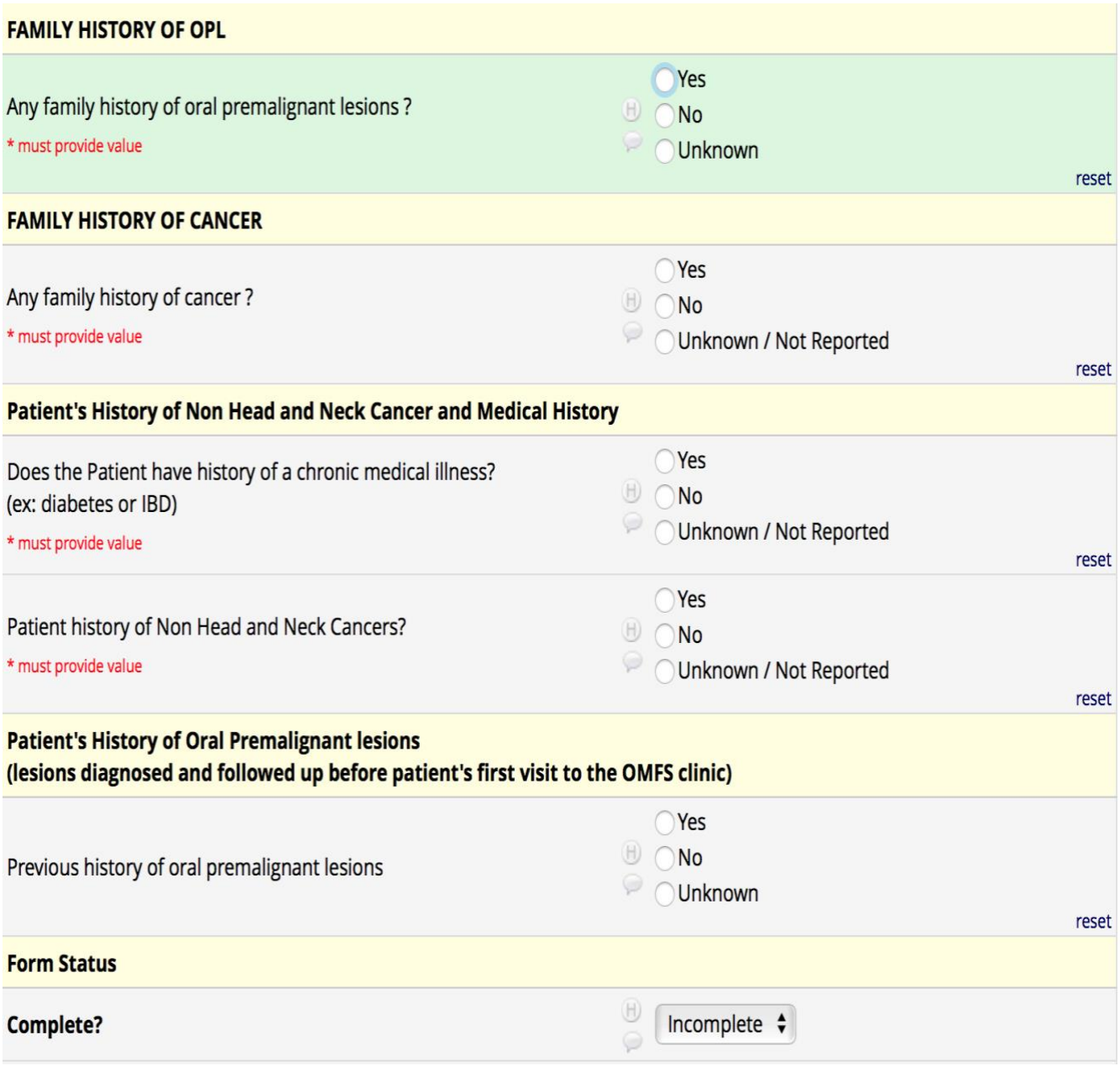
Has the patient ever performed oral sex?
* must provide value

- Yes
- No
- Does not want to disclose
- Unknown

reset

The third instrument is Past Medical History and Family History. It collects details about the patient's past history of OPMD diagnosis, non-head and neck cancer (HNC) history, and chronic medical conditions, in addition to family history of HNC, non-HNC and OPMD.

Figure 3: Past medical history form



FAMILY HISTORY OF OPL

Any family history of oral premalignant lesions ?
* must provide value
Yes
No
Unknown
reset

FAMILY HISTORY OF CANCER

Any family history of cancer ?
* must provide value
Yes
No
Unknown / Not Reported
reset

Patient's History of Non Head and Neck Cancer and Medical History

Does the Patient have history of a chronic medical illness?
(ex: diabetes or IBD)
* must provide value
Yes
No
Unknown / Not Reported
reset

Patient history of Non Head and Neck Cancers?
* must provide value
Yes
No
Unknown / Not Reported
reset

Patient's History of Oral Premalignant lesions
(lesions diagnosed and followed up before patient's first visit to the OMFS clinic)

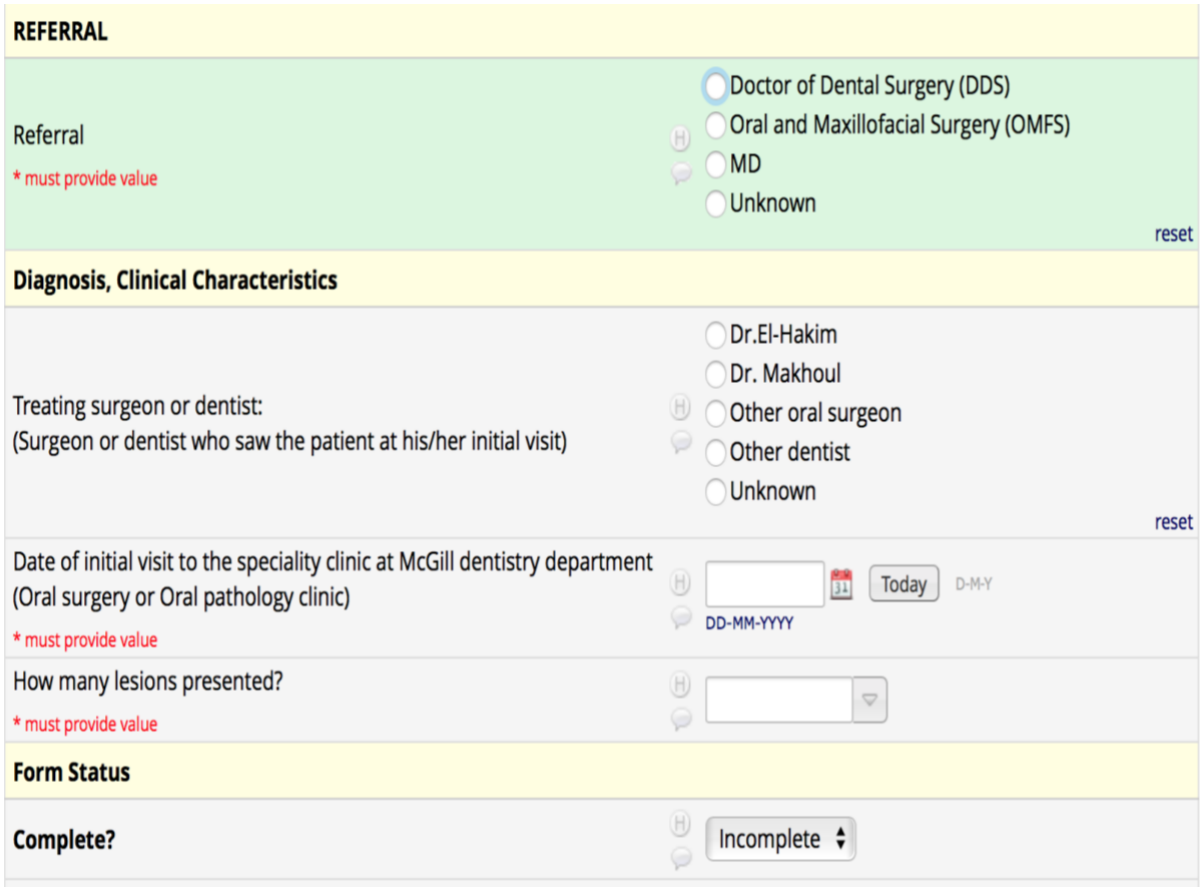
Previous history of oral premalignant lesions
Yes
No
Unknown
reset

Form Status

Complete? Incomplete

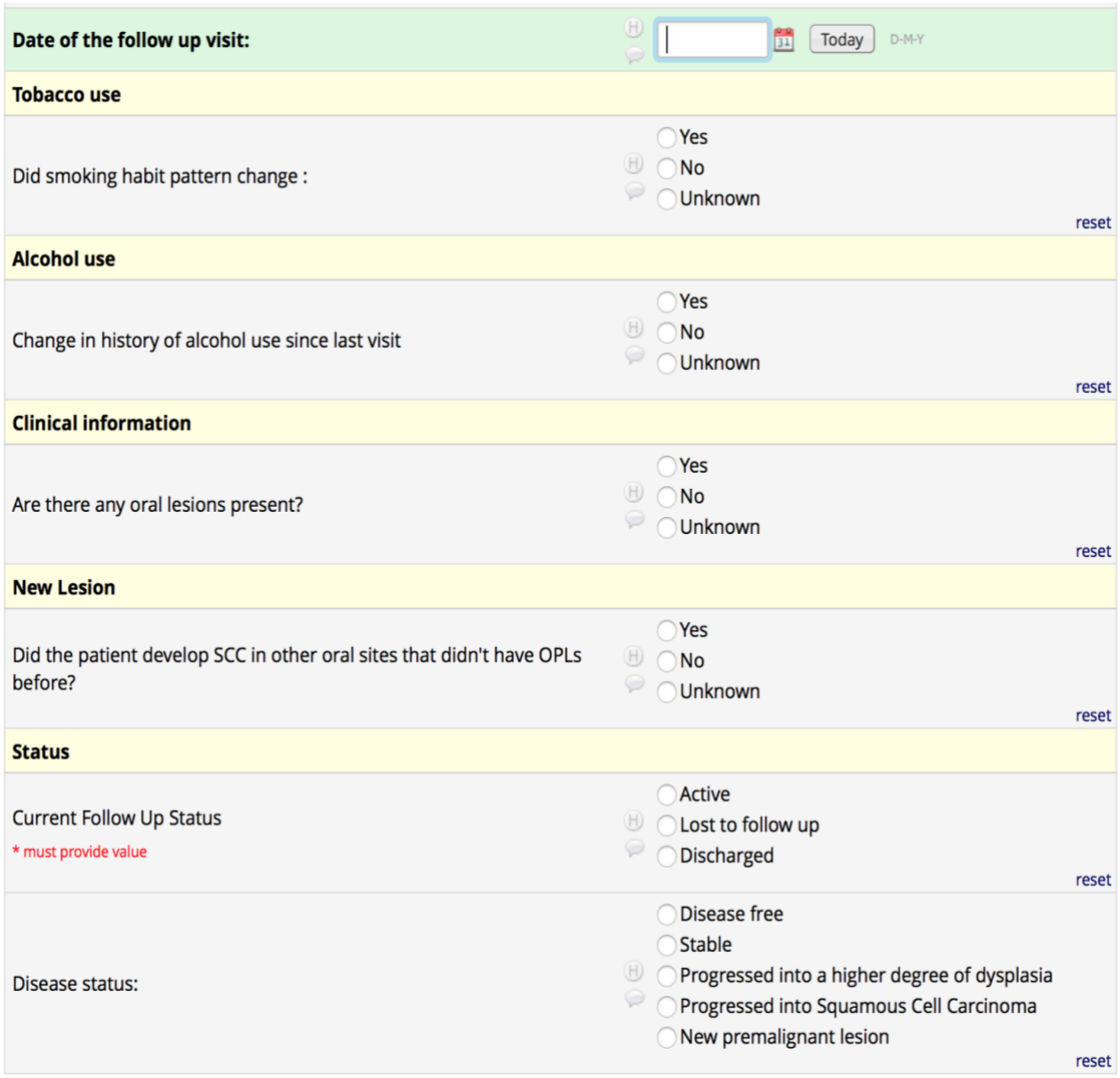
The fourth instrument is Referral Event Details, which comprises data elements that describe the present lesion's clinical features, symptoms, histopathological diagnosis and treatment provided.

Figure 4: Referral event details and current presentation form



The abovementioned four instruments are filled during the patient's initial visit (consultation visit). In addition, the patient will complete one questionnaire on the subsequent follow up visits (Figure 5). This instrument, called follow up visits details, is filled for each regular OPMD follow up visit which are usually scheduled biannually (each 6 months). Follow up visits' details section collects information on the patient's disease status and lifestyle factors (e.g., smoking, alcohol) at the time of follow up.

Figure 5: Follow up events' details form



5.7 Patients' identification, recruitment and de-identification

The registry includes two types of patients: retrospective and prospective, and therefore two sets of data. Prospective patients are identified through weekly screening of new consults seen in the OMFS clinic by the Chief of the Dentistry and Maxillofacial Surgery Department. Once the patients are identified as eligible for the OPMD database, the research coordinator contacts them

for recruitment either via email or phone call. If the patient agrees to participate, the research coordinator sends them a registration package that includes a detailed consent form and a questionnaire to fill.

For retrospective patients, it was challenging to find a way to identify eligible participants because clinic patients were not flagged or classified based on diagnosis. Therefore, we needed to identify those patients through another database. The pathology department database proved helpful. First we created queries to identify patients with OPMD who were treated at the OMFS department using clinical and histopathological keywords such as dysplasia and hyperkeratosis. The results of these queries were filtered and cleaned, and only subjects meeting the following inclusion criteria were included in the database: 18 years old and above at the time of initial consult, clinical or histopathological diagnosis of OPMD, and being treated/followed up or received treatment/ follow up at the OMFS clinic at the MGH-MUHC. Patients who had a history of OC, and patients who developed OC in less than 6 months of initial OPMD diagnosis were excluded.

All patients' records found to be eligible for inclusion in the database undergo a process of de-identification. Whenever a new record is introduced into the database, a unique study number is generated for that record. This unique study number is kept in an excel worksheet where it is associated with the patient's identifying variables, such as medical record number (MRN) and name. This excel sheet acts as a decoding key for the unique study numbers used in the database, as there are no identifying variables entered into the database itself. The excel worksheet is saved on a password protected computer accessible only to the principal investigator and research coordinator.

5.8 Data entry trial, pilot testing and identified issues

After the database instruments were completed and before moving to the production phase, trials of data entry were conducted. For this trial, a small set of patients with OPMD (20 records) were added to the database and their data were exported and analysed. The pilot process allowed us to identify some technical malfunctions and implement solutions before moving to production. Among issues identified and addressed were branching logic and calculated fields syntax. Lastly, this process allowed us to make final adjustments to fine tune the instruments.

5.9 Production phase

After successful pilot testing, and all the technical issues identified had been amended, we moved to production mode. In this mode, patients' data were entered retrospectively and prospectively. After identifying all retrospective eligible patients (from 2008 until 2017), their names and medical record numbers were gathered and kept in an excel worksheet. Whenever a record was created in the database, the unique study number was generated and added to that sheet. Prospective data entry started in March 2018, and the patients' names corresponding to the prospective data entry are kept on a separate sheet. The analysis and results of this thesis use the retrospective data set.

5.10 Data analysis

Once the data entry phase of all eligible retrospective subjects was finalised, the data were meticulously cleaned using the REDCap data quality tool. The extensive cleaning process included checking for missing data, data entry errors, double checking unreasonable values and outliers. All detected errors were resolved before moving to analysis. Statistical analysis was conducted using Stata 15 statistical software. (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC)

The first step in our analysis was to carry out descriptive statistics. We examined the average values, dispersion, distribution shape and presence of outliers for all variables at each follow up. Second, we explored the associations between variables. X^2 test was used to check for dependence/independence between categorical variables. All variables were binary or categorical with the exception of age of patient at time of diagnosis and pack years of smoking, age was converted into a categorical variable. Associations were considered significant at $P < .05$.

Variables examined include demographics (age, sex) lifestyle (smoking and alcohol consumption) and clinic-pathological variables (presence of dysplasia at time of diagnosis, grade of dysplasia, lesion location, size and color, whether if the patient is under risk of immune suppression or not, and developing oral squamous cell carcinoma). Age, sex, alcohol history, presence of dysplasia, lesion size and developing oral SCC were binary variables whereas smoking history, lesion color and location were categorical. In addition, we created a dichotomous variable representing immunodeficiency by combining the following binary variables: presence of diabetes, HIV and transplant patients into one variable.

We first looked if there is an association between demographic variables and the presence of dysplasia at time of diagnosis, and progression into oral SCC. Then, we assessed for associations between clinicopathological variables and presence of dysplasia at time of diagnosis and progression into OC.

We also estimated the age standardised oral cancer incidence rate to facilitate accurate i) self-comparison over time, ii) comparison to other populations with different age distributions and, iii) in order to have the proper type of measure to use for comparison with the total Canadian

population as the measure published by the Canadian Cancer Statistics Advisory Committee for oral cancer incidence is the age standardised incidence rate. First, we created a new variable called "age group" in which the groups were defined at 5-year interval. Subsequently, we calculated the crude incidence per 100,000 for each age group with the denominator being people at risk in that age group. Afterward, we used weights for different age groups based on the 2011 Canadian population to project the standardised incidence rates according to age group. Finally, results for all age groups were added to yield the age standardised oral cancer incidence in our OPMD population.

6. RESULTS: MANUSCRIPT

The development and implementation of an Oral Potentially Malignant Disorders database: lessons learned

Bassyoni L[1,2], Nicolau B[1] , Madathil S[1], Makhoul N[1]

(1) Division of Oral and Maxillofacial Surgery, Faculty of Dentistry, McGill University, Montreal, Quebec, Canada
(2) Division of Oral Health and Society, Faculty of Dentistry, McGill University, Montreal, Quebec, Canada

***Corresponding author**
Nicholas M. Makhoul DMD. MD. FRCD(C). Dip. ABOMS. FACS.
Associate Professor and Director, Oral and Maxillofacial Surgery, McGill University
B3-149, 1650 avenue Cedar, Montreal General Hospital
Montreal, QC, H3G-1A4
Tel: (514) 934-1934 ext: 42492
Fax: (514) 934-8340
nicholas.makhoul@mcgill.ca

Manuscript word count : 4,508
Abstract word count : 353
Number of tables : 4

Keywords: registry, head and neck cancer, Montreal

Abstract

Background: Clinical databases can act as research platforms by converting everyday clinical data into meaningful information that help in evidence based decision making for both clinicians and policy makers. For a relatively rare and heterogeneous disease such as OPMD, a clinical database offers a source of recruitment for epidemiological research. Thus, it allows us to investigate the natural history of OPMD, its risk factors and the potential determinants of malignant transformation. In this work, we will share our experience in creating a web based OPMD clinical database at the OMFS clinic at the Montreal General Hospital (MGH), McGill University Health Centre (MUHC). In addition, we will describe the clinical, biological and socio-demographic factors among OPMD patients in our population.

Methods: We designed and customised a platform in REDCap software to support an OPMD web-based database, which comprised five forms and more than 900 fields. The database included two arms of patients: retrospective and prospective. Retrospective patients were OPMD patients seen at the OMFS clinic from 2008 to 2017 and a prospective arm for patients diagnosed after 2017. For the retrospective arm of the study, data were extracted from medical records and included patients' socio-demographics, treatment provided, OPMD risk factors, and progression to oral cancer. Data analysis involves descriptive statistics.

Results: A total of 155 retrospective patients were entered in the database and their data analysed. The majority of our sample were males (57%) and the mean age was 60.2 years old (SD: 13.27). In addition, more than of 80% of the patients were smokers or ex-smokers and 50% were either current or past alcohol users. The most common OPMD location was the tongue (39%), followed by the buccal mucosa (12%) and the maxillary gingiva (10%). The majority of the lesions were leukoplakia (52.9%) and dysplasia was present in 56% of all cases. Around 10.32% of the cases progressed to oral cancer with a mean follow up time of 3.3 years. Most of the oral cases were diagnosed at an early stage (87.5%).

Conclusion: The OPMD clinical database designed and implemented at the McGill University Health Centre (MUHC) has proven to be implementable clinically and proves to be a potentially valuable tool for clinical care and research.

Introduction

It is generally accepted that oral cancer is preceded by histopathological changes confined to the epithelium. These changes are precursors of malignancy, ranging from hyperplasia to oral epithelial dysplasia (OED). [1, 2] They are usually present as lesions on the oral mucosa known as Oral Potentially Malignant Disorders (OPMD) such as leukoplakia, erythroplakia, erythroleukoplakia, lichen planus, oral submucous fibrosis, actinic keratosis and others. Up to 51% of OPMD have dysplasia at time of diagnosis when assessed histopathologically, [3] and up to 40% progress to oral cancer (OC). [4, 5] These cancer cases are diagnosed at earlier stages compared those occurring in individuals without OPMD.[6] In a series of patients with OPMD followed in a multidisciplinary clinic, 100% of patients who experienced malignant transformations were diagnosed early with a localised stage and the disease-specific survival was 100% at 24 months.[7] While survival rates for early stage oral cancer are approximately 93%, these rates drop significantly to 38% for advanced stages, and 20% for metastatic disease [8]. According to Surveillance Epidemiology and End Results program (SEER), only 29.3% of oral cancer cases are diagnosed at an early localised stage [9], which could explain to a large extent the lack of improvement in survival outcomes in the past period.

Screening, diagnosing and following OPMD could potentially capture OC cases at earlier stages, and decrease its burden on public health resources. Although studies from countries with a high incidence of OC report an increase in survival rates for a simple visual screening of the oral cavity,[10] only opportunistic screening is recommended in Canada. [11] A solution for areas with a low incidence of OC may be to establish OPMD registries for efficient follow-up and early detection of malignant transformation among these high risk individuals.

OPMD is a relatively rare disease; leukoplakia, one of the most common OPMD, has a global prevalence of 2.6%.[12] However, the literature is scarce on methods of registration and documentation of premalignant lesions in general and oral premalignant lesions specifically. Premalignant lesions are usually documented within cancer registries, for example, the Cancer Registry of Norway (CRN), [13, 14] or as a part of another database (non-disease specific/ patient registries), such as the nationwide histopathology registry in the Netherlands (PALGA) [15] or U.S. Medicare claims [6]. To our knowledge, there is no database available of OPMD cases in Canada.

Most epidemiological studies concerning OPMD were conducted based on histopathology databases, clinical records reviews, medical insurance claims, or retrospective analyses of the history of cancer cases. [6, 16, 17] These studies have identified some demographic (e.g., advanced age, female gender), clinical and histological (non-homogenous lesions, lesions larger than 200 mm^2, and oral epithelial dysplasia) and molecular biomarkers (e.g., immunodeficiency) determinants of OPMD progression to OC. [18-19-24-25-28]

Numerous biomarkers have been suggested [29], among which genes and proteins involved in the immune response seem to be useful. [19-21] Immunodeficiency is considered a precancerous state [22-24], and is a risk factor for OPMD and their progression to OC. [25-28] Nevertheless, further longitudinal studies are needed to investigate the nature of these associations.

The majority of OC cases are preceded by OPMD, yet not all OPMD progress to OC, and the available evidence is not yet sufficient to accurately predict the progression of these lesions to malignancy. Therefore, a deeper understanding of the disease is needed and, given the rarity of

OPMD, a repository for cases is needed to recruit subjects for future epidemiological studies. Clinicians can contribute strongly to development of clinical databases, however, the process of implementation of them are rarely discussed in clinical literature. In this work, we describe the development and implementation of a web-based OPMD clinical database at MGH-MUHC. In addition, we describe socio-demographic, behavioural, clinical, biological characteristics among OPMD patients, and report the rate of malignant transformation in the study population.

Methodology and methods

i) OPMD database design

The initial and most important step in the planning process is determining the purpose and objective of the database so that its outcomes are clear and efficient. We started by identifying the main objective for the OPMD database, which is a disease focused database that will serve as a research platform to investigate OPMD. To achieve this objective, we determined the questions to be answered once the database is set up and running; examples of these questions include: what are the risk factors for OPMD? What are the clinical, histopathological and biological determinants for malignant progression in OPMD patients? Should surgical excision be the standard treatment for all OPMD lesions with dysplasia regardless of dysplasia grade?

Based on the above, we designed the database including the population, inclusion/ exclusion criteria, follow-ups and database duration. The main steps in the development process are discussed below.

a) Choosing the software

REDCap (Research Electronic Data Capture) was the software chosen to design the database, as it was supported by the MUHC research centre, met our team's needs and provided several

advantages. REDCap is secure a web-based application for building and managing online databases and surveys that can be accessed from anywhere in the world over a secure web connection. It was first introduced by Vanderbilt University in 2004 and since then, it has been adopted by more than 2800 institutions worldwide for research. It is fully customisable, fast and flexible, allowing the designer to (i) condition variables to be of a certain form (numbers, dates, letters only) specifying ranges, (ii) create data quality rules to check for discrepancies, and (iii) label some variables as required so all subjects within the database have a minimum data set. Moreover, the application uses branching logic, hiding fields that are not applicable for a certain patient, and condenses very long forms into simple and short ones showing only relevant fields. Finally, REDCap does not require any programing experience to build a database and has multi center access.

b) Database structure

Our research team, which included clinicians, epidemiologists and computer scientists, determined the data elements and outcomes that would fulfil the database objectives. We identified a total of four outcomes: main outcome - the development of Oral Squamous Cell Carcinoma, and three secondary outcomes: progressing from low grade dysplasia to high grade dysplasia, developing a new premalignant lesion, and being disease free or stable at 5-year follow up. Patients who are disease free or stable at five years are discharged from our clinic, they are sent back to receive care by their dentists who will continue to see them annually, their dentists would refer them back to our clinic if any significant clinical changes are observed.

Data were obtained from paper and electronic patients' records. The REDCap platform was customised to support a longitudinal data collection model with determined endpoints to capture the consult visit and subsequent follow ups. The final product was a comprehensive database that

has 5 sections: socio-demographics and lifestyle factors, past medical history, consult and follow up visits details (e.g., clinical, histopathological and treatment information).

c) Data sources

Data were obtained using both primary and secondary data sources and include: (i) patients' hospital electronic medical records (EMR), and (ii) OMFS clinical paper charts (2008 – 2016) and OMFS clinical EMR (2016-2017). In addition, we established three primary data sources: a patient self-completed questionnaire and two clinical forms. The former was developed to collect details on patients' sociodemographic and lifestyle factors that are too comprehensive to be recorded in clinical records (Figure 1 & Table 1). This includes extensive details on tobacco use, alcohol drinking, marijuana use, sexual behaviour and socio-demographic information. The two clinical forms comprise a consult form and a follow up form.

As explained below, the database is divided into two arms, prospective and retrospective, and the primary data sources are only used for patients in the prospective arm (Table 1).

d) Study population and inclusion/exclusion

OPMD patients treated at the OMFS clinic at the MGH-MUHC since 2008 were included in the database if they were of 18 years or older at the time of the initial consult. Patients were excluded if they had a history of OC or developed OC within the first 6 months of their initial OPMD diagnosis. The latter criteria was established to rule out any contamination of prevalent OC cases in OPMD patients. As mentioned above, the database has a retrospective arm and prospective arm. While the former includes patients who were diagnosed with OPMD from October 2008 to January 2017 and collected data through a retrospective chart review, the prospective arm comprises

patients who were diagnosed with OPMD after January 2017. The study protocol has been reviewed and approved by the MUHC Ethics Review Board.

e) Pilot testing and production phase

We conducted a data entry trial with a small set of patients with OPMD (20 records) before moving to the production phase. Information was added to the database and subsequently exported and analysed. The pilot process allowed us to identify some technical malfunctions and correct them before moving to production mode.

ii) Statistical analysis

We carried out descriptive statistics using data from the retrospective arm of the study. We examined the average values, dispersion, distribution shape and presence of outliers for all variables at each follow up. Continuous variables were presented as means $\pm$ SD. Pearson chi square test X^2 was used for univariate analysis and to estimate significant associations between categorical variables. Variables examined include demographics (age, sex), lifestyle (smoking and alcohol consumption) and clinical-pathological variables (presence of dysplasia at time of diagnosis, grade of dysplasia, lesion location, size and color, if the patient is under risk of immune suppression or not, and development of OC). Age, sex, alcohol consumption, presence of dysplasia, lesion size and developing OC were binary variables whereas smoking habits, lesion color and location were categorical. In addition, we created a dichotomous variable representing immunodeficiency by combining the following binary variables: presence of diabetes, HIV and transplant patients into one variable. Whenever detailed information was available about the subject's smoking habit (e.g. duration, age when smoking started, age when quit smoking, change of practice based on age, number of cigarettes/day in each smoking period), pack years were

calculated for each smoking period and then added together. All statistical analysis was done using Stata 15 statistical software. (StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC) and associations presenting a p value of $P < 0.05$ were considered statistically significant.

Results

The main issue we identified in the pilot study was the inability of the REDCap software to differentiate between true missing fields and not applicable fields. REDCap application uses branching logics to design parts of questionnaires, which are contingent on the response to a question. These branches of questionnaire will only be displayed if the answer to the filter question is appropriate. For example, details of the tobacco smoking behaviour will only be displayed for participants who reported they ever smoked tobacco. These hidden branches result in an interface that is efficient in data collection.

However, REDCap does not allow the imputation of values of hidden braches if not applicable. For example, non-smokers do not get any value for the corresponding branching questions. This results in a technical difficulty because the fields are exported as missing values in the dataset. It is important to distinguish these missing values from the true missing values occurring when the respondent missed a question or the information from the medical records is not available. This issue of how REDCap handles data fields associated with branching logic is well known among its users and the general consensus is to deal with it using external software once the dataset is exported.

To distinguish the 'true' missing values and impute the branching questions appropriate values denoting a 'Not Applicable' label, we adapted the 'redcapAPI' package [30] to our database to

implement an intermediate data processing step. This step will be baked into the process of data export from the server (Figure 1). This step is crucial to streamline the process and make the database sustainable for potential future end-users**.**

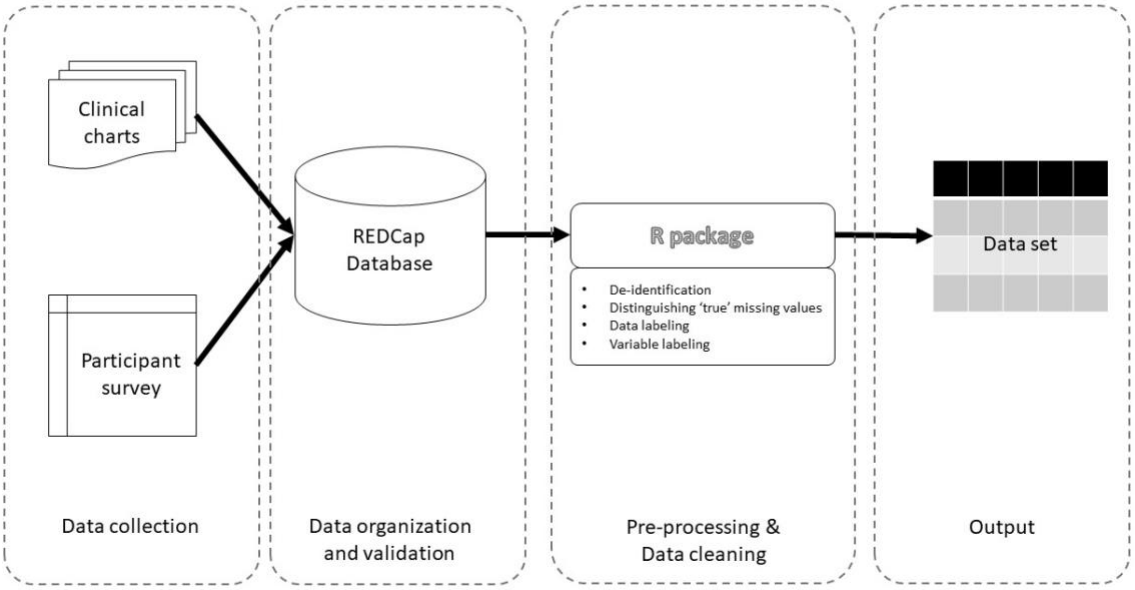


Figure 1: Data exportation process.

Given the retrospective chart review nature of the data collection process, extracting particular information from clinical charts was a challenging procedure for a number of reasons, of which are illegible handwriting and lack of standardised way of documentation of clinical information in the clinical charts. This lead to a non-organised way of collecting information and having incomplete data sets for some of the patients as some demographic and behavioral variables (e.g., education level, income, race, detailed smoking, alcohol drinking habits, and sexual behaviour) were not available. Having standardised clinical forms that once filled act as primary data sources for the database like the ones we established for the OPMD database will largely help to overcome this challenge.

Table 1. Description of type and availability of data

Variables	Retrospective Arm	Prospective Arm	Baseline visit	Follow-Up visits
Socio-demographic (sex, age, marital status, occupation, income)	■	■	■	
Habits and life style				
Tobacco smoking habit	■	■	■	
Yes or No	■	■	■	
Details (e.g., pack years, starting age)	■	■	■	□
Alcohol Drinking habit	■	■	■	
Yes or No	■	■	■	
Details (e.g., No. drinks/ day, duration)	■	■	■	□
Marijuana	■	■	■	
Yes or No	■	■	■	
Details (e.g., No. joint, grams / day, duration)	■	■	■	□
Sexual behavior	■	■	■	
Active or inactive	■	■	■	
Details (e.g., No. partners, oral sex)	■	■	■	
Past medical history				
Family history (cancer, oral cancer, OPMD)	■	■	■	
Medical history (e.g., medical conditions, cancer, OPMD history)	■	■	■	
Clinical and Histopathological data (lesion location, size, color, texture, date of diagnosis, histopathology, treatment)	■	■	■	□

■ Data collected

□ Change from previous visit collected

A total of 155 eligible subjects were recruited to the retrospective arm. Patients' age ranged from 22 to 84 years (mean: 60.2, SD: 13.27), and there were 66 (43%) females and 89 (57%) males (Table 2). Approximately 56% of the participant already had at least some grade of dysplasia at the baseline and 10% of participants experienced malignant transformation of their lesion.

Table.2: Socio-demographic and behavioral characteristics of patients attending the OMFS at MGH from 2008-2017 (n=155)

Variables	All subjects n=155 (%)	Dysplasia* n=87 (56.1)	Developed OC* n=16 (10.3)
Age, (Mean ± SD)	60 ± 13.27	61.91 ± 12.59	70.06 ± 11.84
< 45 years of age	18 (11.6)	7 (38.9)	1 (5.6)
≥ 45 years of age	137 (88.4)	80 (58.4)	15 (10.9)
Gender			
Male	89 (57.4)	47 (52.8)	5 (5.6)
Female	66 (42.6)	40 (60.6)	11 (16.7)
Smoking status			
Never Smoker	56 (41.5)	28 (50.0)	10 (17.9)
Former smoker	30 (22.2)	30 (100.0)	3 (10.0)
Current smoker	49 (36.3)	21 (42.9)	3 (6.12)
Pack-years(Mean ± SD)	10.53 ± 19.3	12.21 ± 20.76	11.4 ± 8.8
Missing	20 (12.9)	6 (30.0)	0 (0.0)
Alcohol drinking status			
Never	33 (21.3)	22 (66.7)	3 (23.1)
Ever	77 (49.7)	47 (61.0)	10 (76.9)
Missing	45 (29.0)	18 (40.0)	9 (20.0)

Table 2 shows the distribution of socio-demographic and behavioral characteristics of the total sample, patients with dysplasia at time of diagnosis and patients who developed OC during follow up. As expected, there are a greater proportion of men, and the majority of the patients use or used tobacco and alcohol in our population. However, proportion of participants who reported with

dysplasia at baseline and who experienced a malignant transformation were higher among females compared to males.

a) Clinicopathological characteristics (all patients)

Table 3 displays the clinical and histopathological characteristics of the total sample, patients with dysplasia at time of diagnosis, and patients who developed OC. The tongue, followed by the buccal mucosa, maxillary gingiva and lip were the most common oral lesion locations with the following proportions: 39%, 12%, 10% and 9%, respectively. Lesions were mostly smaller than 200 mm^2 (45.8%) and white in color (63.2%), the preponderance of patients presented with one lesion (65.1%), and 56% had dysplasia at time of diagnosis. Around 18% of the patients had a medical condition that causes secondary immunodeficiency, mainly diabetes mellitus type II.

The majority of the cases presented leukoplakia (82 cases, 52.9%); this is consistent with leukoplakia being the most common OPMD. The reminder was 14.8% (23 cases) lichen planus, 12.3% (19 cases) erythroleukoplakia, 9.7% (15 cases) erythroplakia, 5.81% (9 cases), and actinic keratosis and others (2.6%).

b) Characteristics of dysplasia cases

Patients with dysplasia were slightly older [mean: 62 years old (SD: 12.59)] than those without dysplasia, [58 years old (SD: 13.87)] and more than half of them who were 45 years old or more had dysplasia at time of diagnosis compared to 38% in those below 45 years old. Women (60.6%) and patients with medical conditions leading to secondary immunodeficiency (76%) had a higher proportion of dysplasia compared to men (52.8%) and to those who did not have secondary immunodeficiency (51.5%). As for alcohol and tobacco habits, the majority of patients were

current users or had used these substances in the past. The majority of the oral lesions locations had dysplasia except for the mandibular gingiva and the buccal mucosa; locations with the highest chance were the tonsillar pillar (100%) and soft palate (85.7%). The tongue location had the majority of lesions and the rarest location was the tonsillar pillar.

More than half of the lesions were low grade (mild and moderate) dysplasia (51 cases). Oral locations such as the floor of mouth, maxillary gingiva, hard palate, lip and tonsillar pillars had a high proportion of high grade dysplasia (severe dysplasia and CIS).

A lesser proportion of lesions smaller than 200 mm^2 (55%) had dysplasia compared to larger lesions (76%), while higher proportions of red and mixed lesions were dysplastic compared to white ones. (Table. 3)

c) *Characteristics of cases that transformed into squamous cell carcinoma*

Patients who later progressed to OC were older, with most aged 45 years or older. Sixteen patients in the sample (10.32%) progressed to OC with an incidence rate of 4.65 per 100-person years (CI: 2.85 – 7.59). The proportion of women who progressed to OC is 16.67% compared to only 5.6% of all men. If only subjects with dysplasia are considered, 22.5% of women progressed to OC while only 10.46% of men did. As for smoking status, 17.86% of non-smokers (never and former smokers) progressed to OC while only 10% of former smokers and 6.12% of smokers did. In the immunodeficient patients group, 76% (22 patients) had dysplasia at time of diagnosis; and 20.69% (6 patients) developed OC with a mean follow up of 3.4 years.

Table 2. Clinical and histopathological characteristics

Variable	All subjects N=155 (%)	Dysplasia N=87 (%)*	Developed OC N=16 (%)*
Lesion location			
Tongue	60 (38.7)	39 (65.0)	5 (8.3)
Buccal mucosa	19 (12.3)	6 (31.6)	1 (5.3)
Maxillary gingiva	16 (10.3)	9 (56.2)	5 (31.2)
Mandible gingiva	12 (7.7)	1 (8.3)	1 (8.3)
Lip	14 (9.0)	8 (57.1)	1 (7.1)
Hard palate	11 (7.0)	6 (54.5)	0 (0.0)
Floor of mouth	8 (5.2)	6 (75.0)	0 (0.0)
Soft palate	7 (4.5)	6 (85.7)	1 (14.3)
Retro molar trigone	6 (3.9)	4 (66.7)	1 (16.7)
Tonsillar pillars	2 (1.1)	2 (100.0)	1 (50.0)
Grade of dysplasia			
No dysplasia	68 (43.9)		2 (2.9)
Low grade	51 (32.9)		9 (17.7)
High grade	36 (23.2)		5 (13.9)
Size			
>= 200 mm^2	29 (18.7)	22 (75.9)	6 (20.7)
< 200mm^2	71 (45.8)	39 (54.9)	4 (5.6)
Missing	55 (35.4)	26 (47.3)	6 (10.9)
Color			
White	98 (63.2)	48 (48.9)	9 (9.2)
Mixed	30 (19.3)	17 (56.7)	5 (16.7)
Red	19 (12.3)	15 (78.9)	1 (5.3)
Missing	8 (5.2)	7 (87.5)	0 (0.0)
Number of lesions			
1	102 (65.8)	60 (58.8)	10 (9.8)
>1	53 (34.2)	27 (50.9)	6 (11.3)
Medical conditions			
Diabetes	26 (16.8)	19 (73.1)	5 (19.2)
GVH	1 (0.7)	0 (0.0)	0 (0.0)
Transplant	3 (1.9)	3 (100.0)	1 (33.3)
HIV	2 (1.3)	1 (50.0)	0 (0.0)
Other (HTN, Hypothyroidism, CAD)	83 (53.6)	51 (61.5)	12 (14.5)
No medical conditions	40 (25.8)	13 (32.5)	0 (0.0)
Immunodeficiency			
Yes	29 (18.7)	22 (75.9)	6 (20.7)
No	126 (81.3)	65 (51.6)	10 (7.9)

* Row percentages

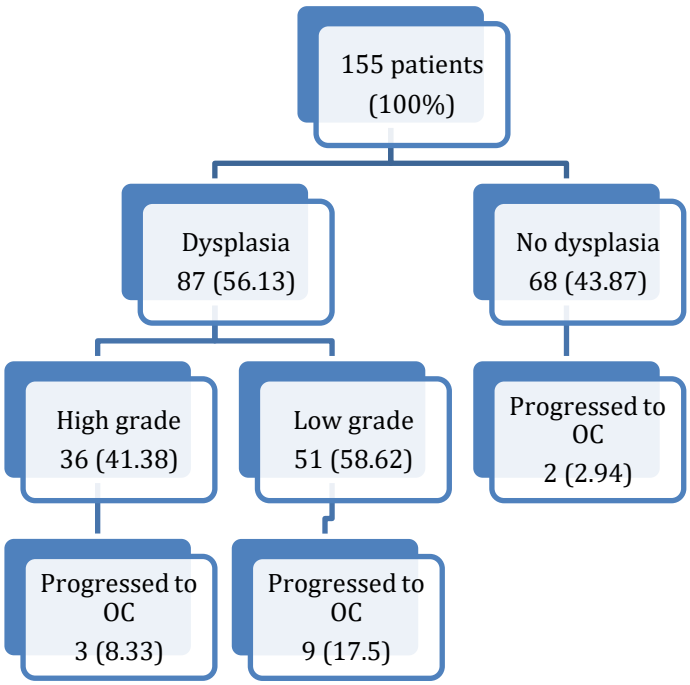
The rate of malignant transformation regarding dysplasia grading shows that 18% of low grade dysplasia, 14% of high grade dysplasia, and 3% of cases with no dysplasia have transformed into oral cancer (Table 3). The mean follow up time to malignant transformation was 3.3 years. More

than half of the lesions were equally divided between the tongue and the maxillary gingiva. However, the locations with the highest proportions of malignant transformation were the soft palate, tonsillar pillar, retro-molar trigone and maxillary gingiva. Around 40% arose from premalignant lesions larger than 200 mm^2, and almost all (87.5%) were diagnosed at an early stage (stage I, stage II).

Table. 4: Cancer stage among those who developed oral cancer

Stage	**n=16 (100%)**
I	12 (75.00)
II	2 (12.50)
III	0
IV	2 (12.50)

Figure 2: Cases which progressed to OC according to the presence and grade of dysplasia



Discussion

In this paper, we present our experience in developing a web-based clinical OPMD database and describe the characteristics of the study sample. A clinical OPMD database is a powerful tool that acts as a research and educational platform by providing a good repository of cases for different types of studies and answering various research questions; it allows for the identification of factors that will help clinicians to predict which patients are at risk for developing OPMD, which patients with OPMD are at higher risk of malignant transformation, and what is the optimal treatment for these disorders. In addition, it provides an infrastructure to monitor the prevalence and incidence of the disease and aids in the implementation of measures of quality assessment and assurance.

The Agency for Healthcare Research and Quality defines a clinical database as: "an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves a predetermined scientific, clinical, or policy purpose(s)". [31] Data organisation could be done in an electronic form (digitized) with web-based access; evidence has demonstrated that data captured in an electronic database could be aggregated, studied and analyzed to improve patient care and healthcare outcomes. [32] Moreover, if the platform is accessed through the web, patients can contribute to the data collection process by answering questionnaires directly into the database.[31, 33]

Clinical databases incorporate a huge volume of every day clinical data, coming from real life clinical cases and events (real world data). Clinical database research translates the available real world data into meaningful information that is critical to help policy makers take informed evidence-based decisions to change policy. [34, 35] Moreover, clinical databases given their

comprehensive data gathering and low restrictions for inclusion, could: 1) provide complementary information that extends RCT results to patient populations not included in those trials, 2) demonstrate real-world effects of treatments outside the research setting (efficacy vs. effectiveness), 3) provide long-term follow up, 4) evaluate the natural history of a disease, and 5) act as an active surveillance system for unexpected side effects and complications. [31, 36]

The clinical appearance of OPMD in terms of size, color, location, and homogeneity has been studied and identified as determinants for dysplasia and malignant potential. [18, 37-39] In our results, we observed that lesion color, location and size were significantly associated with presence of dysplasia at time of diagnosis. Indeed, these clinical characteristics have been associated with malignant transformation of OPMD, but except for lesion color they have not been investigated as potential risk factors for dysplasia [18]. Similarly, our findings showing that old age at diagnosis of OPMD, occurring predominantly in males and the most common OPMD (leukoplakia), are consistent with the literature. [40, 41]

Tobacco is one of the most significant risk factors associated with OPMD. [17, 42-47] The associated risk increases with the amount of tobacco used and duration of habit. [48-50] Although there is an agreement in the literature that smoking tobacco increases the risk for OPMD, the risk of malignant transformation is significantly higher in non-smokers [51-53] This is evident in our results as the majority of our OPMD study population were either smokers or former smokers. However, a higher percentage of never smokers (17.86 %) transformed into OC than former (10%) and current smokers (6.12%). Additionally, a higher percentage of the low grade dysplasia group progressed to OC in comparison to the high grade group. While it is largely accepted that dysplasia is an important predictor for OPMD malignant transformation, there is no consensus in the

literature on whether high grade dysplasia lesions have a higher risk of malignant progression. This may be due to the subjectivity of the dysplasia grading system [54], which includes contamination of the low-grade dysplasia sample, and different treatment modalities used to treat high and low-grade dysplasia. While the clinical guideline for treatment of high grade dysplasia is surgical excision, follow up is usually offered for low grade dysplasia. This in turn could lead to more OC cases arising from low-grade compared to high-grade dysplasia during the period of observation, as in our results. A possible explanation is that surgical excision either has cured the patient or at least increased the time to progression in high grade patients, while low grade subjects continued to progress along the dysplasia scale to OC with no intervention. However, in reality such relationships and interrelations are usually more complicated, and a group of factors (e.g., age, sex, immunodeficiency, treatment) interacting with each other instead of only one factor such as surgical excision, would play a role that could explain such an observation.

In this study, we report a higher age standardised OC incidence (4,987 per 100,000) (*adjustment was based on the 2011 Canadian population*)[55] than the Canadian general population (4,700 per 100,000). [56] This is probably attributable to the high-risk nature of the subjects involved in our analysis. However, the majority (87.5%) of OC incident cases were diagnosed at an early stage. Others have also demonstrated that OC cases arising from OPMD lesions that have been followed up periodically are usually smaller in size and diagnosed at an earlier stage.[6, 57]

The overall malignant transformation proportion in our cohort was 10.32%; this finding falls in the middle of the reported spectrum and divides the relevant literature into two groups; those who reported higher transformation rates [41, 51, 58-60], and others who reported lower transformation rates [6, 61, 62]. The relatively higher transformation rate in our analysis could be explained by the

nature of the population involved in our study. A clinic-based population is a high risk group to start with; our sample included OPMD patients, most of whom had dysplasia. Among studies reporting lower rates, some [6, 62] are population based including patients with and without OPMD diagnosis, and some [61] have smaller proportions of patients with oral epithelial dysplasia compared to their total population. On the other hand, in studies reporting higher rates of transformation [41, 59, 60], their samples comprised only OPMD dysplasia positive lesions. In addition, one study included patients with a previous history of OC [51]. The mean time to transformation was 3.3 years ranging between 1.5 years to 10.25 years, with the majority of cases transforming within the 5-year period (87%, 14 out of 16 patients). This coincides with the literature as most cases that undergo malignant transformation do within a 5-year period. [63]

In the present study, the proportion of women who progressed to OC is higher than men, the proportion remains higher in a sub-analysis of only subjects with dysplasia. These findings are in line with the results of two recent systematic reviews reporting that women have a higher risk of malignant transformation [18, 64]. However, conflicting results about gender as a risk factor for malignant transformation have been reported with some studies showing that men are at higher risk compared to women [65].

An increased risk for head and neck cancer, which include OC and OPMD, among transplant patients, graft versus host disease patients, type 2 diabetes patients, and HIV patients have been described. [26, 66-71] Accumulating evidence suggests that the progression from OPMD to OC is strongly informed by an abnormal immune environment. [72] A study investigating the genetic susceptibility to oropharyngeal squamous cell carcinoma (including OC) and OPMD reported an association between these diseases and the alteration of immune system genes and genes with

metastatic potential (TNF-alpha, TGFbeta-1, MMP-1).[20, 21] Moreover, galectins, which are a group of proteins involved in immune responses, have been associated with OED progression [19]. Our results show that OPMD patients who have secondary immunodeficiency had higher incidence rates of OC compared to non-immune deficient OPMD patients (8.29 vs 4.02 per 100-person years). However, our numbers are small and larger cohort studies are needed to investigate the effect of immunodeficiency and the degree of immunodeficiency in addition to other socio-demographic, behaviour and clinical characteristics on the progression potential of OPMD to OC.

In conclusion, patients with OPMD need long term clinical follow up as some lesions took more than 5 years to transform. Our sample included different OPMD, lesions with dysplasia and without dysplasia. Therefore, our results apply to a wide variety of OPMD patients. However, it should be interpreted carefully as the study is based on one centre, and retrospective in nature, which affected the availability of some data points for some subjects. Therefore, regardless of the associations presented it does not implement causality, but suggests avenues for future research. In addition, the strength of our conclusions is compromised by the small number of subjects. Larger cohorts of patients should be studied preferably in a randomised controlled approach. The OPMD clinical database could be used as an infrastructure for a prospective cohort study to investigate risk factors for OPMD (e.g., socio-demographic, clinical, histopathological, and biological) and their malignant transformation, in addition to future randomised controlled trial for a sub group of OPMD patients (low grade dysplasia and no dysplasia OPMD patients) to assess the effectiveness of surgical treatment in preventing future malignant transformation in these lesions.

List of References

1. Gimenez-Conti, I.B. and T.J. Slaga, *The hamster cheek pouch carcinogenesis model.* Journal of Cellular Biochemistry, 1993. **53**(17F): p. 83-90.
2. Todd, R., R.B. Donoff, and D.T.W. Wong, *The molecular biology of oral carcinogenesis: Toward a tumor progression model.* Journal of Oral and Maxillofacial Surgery, 1997. **55**(6): p. 613-623.
3. Neville, B.W. and T.A. Day, *Oral Cancer and Precancerous Lesions.* CA: A Cancer Journal for Clinicians, 2002. **52**(4): p. 195-215.
4. P.J, T., *Oral precancer – diagnosis and management of potentially malignant disorders*. 2012, Chichester: Wiley-Blackwell. 236
5. Goodson, M.L., et al., *Oral precursor lesions and malignant transformation--who, where, what, and when?* Br J Oral Maxillofac Surg, 2015. **53**(9): p. 831-5.
6. Yanik, E.L., et al., *Leukoplakia, Oral Cavity Cancer Risk, and Cancer Survival in the U.S. Elderly.* Cancer Prevention Research, 2015. **8**(9): p. 857-63.
7. Ho, M.W., et al., *Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: rationale for monitoring premalignant oral lesions in a multidisciplinary clinic.* Br J Oral Maxillofac Surg, 2013. **51**(7): p. 594-9.
8. *Survival statistics for oral cancer*

   5-year relative survival by stage and tumour site. [cited 2018 2018-04-29]; Available from: http://www.cancer.ca/en/cancer-information/cancer-type/oral/prognosis-and-survival/survival-statistics/?region=on.
9. *Survival by Stage*. Cancer Stat Facts: Oral Cavity and Pharynx Cancer [cited 2018 2018-05-11]; Available from: https://seer.cancer.gov/statfacts/html/oralcav.html.
10. Sankaranarayanan, R., et al., *Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India.* Oral Oncol, 2013. **49**(4): p. 314-21.
11. Examination, C.T.F.o.t.P.H., *The Canadian Guide to Clinical Preventive Health Care*, M.o.S.a.S. Canada, Editor. 1994: Ottawa. p. p. 1009.
12. Petti, S., *Pooled estimate of world leukoplakia prevalence: a systematic review.* Oral Oncol, 2003. **39**(8): p. 770-80.
13. Enerly, E., et al., *Quality assessment of the registration of vulvar and vaginal premalignant lesions at the Cancer Registry of Norway.* Acta Oncologica (Stockholm, Sweden), 2012. **51**(1): p. 45-50.
14. Norway, C.R.o., *Cancer in Norway 2015 - Cancer incidence, mortality, survival and prevalence in Norway.* Oslo.
15. Casparie, M., et al., *Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive.* Cell Oncol, 2007. **29**(1): p. 19-24.

16. Dost, F., et al., *Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading.* Oral Surg Oral Med Oral Pathol Oral Radiol, 2014. **117**(3): p. 343-52.

17. Starzynska, A., et al., *Oral premalignant lesions: epidemiological and clinical analysis in the northern Polish population.* Postepy Dermatologii I Alergologii, 2014. **31**(6): p. 341-50.

18. Warnakulasuriya, S. and A. Ariyawardana, *Malignant transformation of oral leukoplakia: a systematic review of observational studies.* Journal of Oral Pathology & Medicine, 2016. **45**(3): p. 155-66.

19. de Vasconcelos Carvalho, M., et al., *Alterations in the immunoexpression of galectins-1, -3 and -7 between different grades of oral epithelial dysplasia.* Journal of Oral Pathology & Medicine, 2013. **42**(2): p. 174-9.

20. Erdei, E., et al., *Cytokines and tumor metastasis gene variants in oral cancer and precancer in Puerto Rico.* PLoS ONE [Electronic Resource], 2013. **8**(11): p. e79187.

21. Hsu, H.J., et al., *Role of cytokine gene (interferon-gamma, transforming growth factor-beta1, tumor necrosis factor-alpha, interleukin-6, and interleukin-10) polymorphisms in the risk of oral precancerous lesions in Taiwanese.* Kaohsiung Journal of Medical Sciences, 2014. **30**(11): p. 551-8.

22. Katsanos, K.H., et al., *Oral Cancer and Oral Precancerous Lesions in Inflammatory Bowel Diseases: A Systematic Review.* Journal of Crohn's & colitis, 2015. **9**(11): p. 1043-52.

23. Piselli, P., et al., *Epidemiology of de novo malignancies after solid-organ transplantation: immunosuppression, infection and other risk factors.* Best Pract Res Clin Obstet Gynaecol, 2014. **28**(8): p. 1251-65.

24. Shah, A.T., E. Wu, and R.O. Wein, *Oral squamous cell carcinoma in post-transplant patients.* American Journal of Otolaryngology, 2013. **34**(2): p. 176-9.

25. Hernandez, G., et al., *Rapid progression from oral leukoplakia to carcinoma in an immunosuppressed liver transplant recipient.* Oral Oncol, 2003. **39**(1): p. 87-90.

26. King , G.N., et al., *Increased Prevalence of Dysplastic and Malignant Lip Lesions in Renal-Transplant Recipients.* New England Journal of Medicine, 1995. **332**(16): p. 1052-1057.

27. Meisel, P., et al., *Association Between Glycemia, Serum Lipoproteins, and the Risk of Oral Leukoplakia: The population-based Study of Health in Pomerania (SHIP).* Diabetes Care, 2010. **33**(6): p. 1230-1232.

28. P., D.R., et al., *Association between diabetes mellitus and pre-malignant oral diseases: A cross sectional study in Kerala, India.* International Journal of Cancer, 2006. **118**(2): p. 453-457.

29. Lingen, M.W., et al., *Genetics/epigenetics of oral premalignancy: current status and future research.* Oral Diseases, 2011. **17 Suppl 1**: p. 7-22.

30. S, N.B.a.L. *redcapAPI: Accessing data from REDCap projects using the API*. 2018; Available from: https://github.com/nutterb/redcapAPI/wiki.

31. *AHRQ Methods for Effective Health Care*, in *Registries for Evaluating Patient Outcomes: A User's Guide*, rd, et al., Editors. 2014, Agency for Healthcare Research and Quality (US): Rockville (MD).

32. Roberts, A.L. and J.P. Sewell, *Data Aggregation: A Case Study.* CIN: Computers, Informatics, Nursing, 2011. **29**(1): p. 3-7.

33. Black, N., *High-quality clinical databases: breaking down barriers.* The Lancet, 1999. **353**(9160): p. 1205-1206.

34. Tilson, H.H., et al., *The antiretrovirals in pregnancy registry: a fifteenth anniversary celebration.* Obstet Gynecol Surv, 2007. **62**(2): p. 137-48.

35. Eydelman, M.B., et al., *Ophthalmic Devices and Clinical Epidemiology*, in *Medical Device Epidemiology and Surveillance*. 2007, John Wiley & Sons, Ltd. p. 427-439.

36. Bravata, D.M., et al., *AHRQ Comparative Effectiveness Reviews*, in *Comparative Effectiveness of Percutaneous Coronary Interventions and Coronary Artery Bypass Grafting for Coronary Artery Disease*. 2007, Agency for Healthcare Research and Quality (US): Rockville (MD).

37. van der Waal, I., *Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management.* Oral Oncol, 2009. **45**(4-5): p. 317-23.

38. Reichart, P.A. and H.P. Philipsen, *Oral erythroplakia--a review.* Oral Oncol, 2005. **41**(6): p. 551-61.

39. G., A.P., et al., *Outcome of oral dysplasia: a retrospective hospital-based study of 207 patients with a long follow-up.* Journal of Oral Pathology & Medicine, 2009. **38**(6): p. 540-544.

40. Jaber, M.A., et al., *Oral epithelial dysplasia: clinical characteristics of western European residents.* Oral Oncology, 2003. **39**(6): p. 589-596.

41. Lumerman, H., P. Freedman, and S. Kerpel, *Oral epithelial dysplasia and the development of invasive squamous cell carcinoma.* Oral Surgery Oral Medicine Oral Pathology Oral Radiology & Endodontics, 1995. **79**(3): p. 321-9.

42. Shetty, P., et al., *Oral Leukoplakia: Clinicopathological Correlation and Its Relevance to Regional Tobacco-related Habit Index.* Journal of Contemporary Dental Practice [Electronic Resource], 2016. **17**(7): p. 601-8.

43. Casparis, S., et al., *Oral lichen planus (OLP), oral lichenoid lesions (OLL), oral dysplasia, and oral cancer: retrospective analysis of clinicopathological data from 2002-2011.* Oral & Maxillofacial Surgery, 2015. **19**(2): p. 149-56.

44. Hassona, Y., et al., *Oral potentially malignant disorders among dental patients: a pilot study in Jordan.* Asian Pacific Journal of Cancer Prevention: Apjcp, 2014. **15**(23): p. 10427-31.

45. Feng, J., et al., *Prevalence and distribution of oral mucosal lesions: a cross-sectional study in Shanghai, China.* Journal of Oral Pathology & Medicine, 2015. **44**(7): p. 490-4.

46. Morse, D.E., et al., *Smoking and drinking in relation to oral epithelial dysplasia.* Cancer Epidemiol Biomarkers Prev, 1996. **5**(10): p. 769-77.

47. Dietrich, T., P.A. Reichart, and C. Scheifele, *Clinical risk factors of oral leukoplakia in a representative sample of the US population.* Oral Oncol, 2004. **40**(2): p. 158-63.

48. Lee, C.H., et al., *The precancer risk of betel quid chewing, tobacco use and alcohol consumption in oral leukoplakia and oral submucous fibrosis in southern Taiwan.* Br J Cancer, 2003. **88**(3): p. 366-72.

49. Bosetti, C., et al., *Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers.* Am J Epidemiol, 2008. **167**(4): p. 468-73.

50. Reichart, P.A., *Identification of risk groups for oral precancer and cancer and preventive measures.* Clinical Oral Investigations, 2001. **5**(4): p. 207-13.

51. Ho, M.W., et al., *The clinical determinants of malignant transformation in oral epithelial dysplasia.* Oral Oncol, 2012. **48**(10): p. 969-76.

52. Silverman, S., Jr., M. Gorsky, and F. Lozada, *Oral leukoplakia and malignant transformation. A follow-up study of 257 patients.* Cancer, 1984. **53**(3): p. 563-8.

53. Schepman, K.P., et al., *Malignant transformation of oral leukoplakia: a follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands.* Oral Oncol, 1998. **34**(4): p. 270-5.

54. Warnakulasuriya, S., et al., *Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement.* Journal of Oral Pathology & Medicine, 2008. **37**(3): p. 127-33.

55. Canada, S., *Table 051-0001 - Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual (persons unless otherwise noted)*.

56. *Canadian Cancer Statistics 2017*, C.C. Society, Editor. 2017, Canadian Cancer Statistics Advisory Committee: Toronto, ON.

57. Ho, M.W., et al., *Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: rationale for monitoring premalignant oral lesions in a multidisciplinary clinic.* British Journal of Oral & Maxillofacial Surgery, 2013. **51**(7): p. 594-9.

58. Bánóczy, J. and Á. Csiba, *Occurrence of epithelial dysplasia in oral leukoplakia.* Oral Surgery, Oral Medicine, Oral Pathology, 1976. **42**(6): p. 766-774.

59. Cowan, C.G., et al., *Potentially malignant oral lesions in northern Ireland: a 20-year population-based perspective of malignant transformation.* Oral Dis, 2001. **7**(1): p. 18-24.

60. Wei, L., et al., *Malignant transformation of oral epithelial dysplasia: clinicopathological risk factors and outcome analysis in a retrospective cohort of 138 cases.* Histopathology, 2011. **59**(4): p. 733-740.

61. Hsue, S.S., et al., *Malignant transformation in 1458 patients with potentially malignant oral mucosal disorders: a follow-up study based in a Taiwanese hospital.* J Oral Pathol Med, 2007. **36**(1): p. 25-9.

62. Gupta, P.C., et al., *Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10-year follow-up study of Indian villagers.* Community Dentistry and Oral Epidemiology, 1980. **8**(6): p. 287-333.

63. Mehanna, H.M., et al., *Treatment and follow-up of oral dysplasia - a systematic review and meta-analysis.* Head & Neck, 2009. **31**(12): p. 1600-9.

64. Anderson, A. and N. Ishak, *Marked variation in malignant transformation rates of oral leukoplakia.* Evidence-Based Dentistry, 2015. **16**(4): p. 102-3.

65. Wang, Y.Y., et al., *Malignant transformation in 5071 southern Taiwanese patients with potentially malignant oral mucosal disorders.* BMC Oral Health, 2014. **14**: p. 99.

66. Gong, Y., et al., *Type 2 diabetes mellitus and risk of oral cancer and precancerous lesions: a meta-analysis of observational studies.* Oral Oncology, 2015. **51**(4): p. 332-40.

67. Vial, T. and J. Descotes, *Immunosuppressive drugs and cancer.* Toxicology, 2003. **185**(3): p. 229-40.

68. D'Souza, G., et al., *Epidemiology of head and neck squamous cell cancer among HIV-infected patients.* J Acquir Immune Defic Syndr, 2014. **65**(5): p. 603-10.

69. Scheifele, C., et al., *Incidence of oral, pharyngeal, and laryngeal squamous cell carcinomas among 1515 patients after liver transplantation.* Oral Oncology, 2005. **41**(7): p. 670-676.

70. Petti, S., et al., *Orofacial diseases in solid organ and hematopoietic stem cell transplant recipients.* Oral Diseases, 2013. **19**(1): p. 18-36.

71. Helenius-Hietala, J., et al., *Oral mucosal health in liver transplant recipients and controls.* Liver Transplantation, 2014. **20**(1): p. 72-80.

72. Ai, R., et al., *Microenvironmental regulation of the progression of oral potentially malignant disorders towards malignancy.* Oncotarget, 2017. **8**(46): p. 81617-81635.

7. DISCUSSION

The main objective of this MSc thesis is to create a clinical OPMD database at the OMFS department at the MGH, MUHC. This was driven by the lack of reported methods of registration of OPMD in the literature and motivated by the promising possible uses and outcomes of such project in terms of studying and understanding a relatively rare and heterogeneous disease such as OPMD that carries an increased risk for OC. An OPMD clinical database has numerous advantages and utilities of which acting as a research platform itself and as a repository for cases for research projects including cohort studies and RCTs.

In addition, in this work we described our experience in planning and building the OPMD database including the challenges and limitations we encountered and possible solutions to overcome these difficulties. We also describe the clinical, biological and socio-demographic parameters among people with OPMD attending the OMFS clinic at the MUHC between 2008 and 2017.

The main motive behind the new nomenclature OPMD was the imprecision of the previous subdivision of this entity as premalignant lesions and premalignant conditions; where premalignant lesions indicated a clinical morphological alteration of the mucosa and premalignant condition stands for any generalised condition that would put the normal appearing mucosa under increased risk of OC regardless if lesions are present or not. The dissatisfaction arose when patients with OPL were found to have dysplasia or molecular aberration in normally appearing oral mucosal sites. [210, 211] The new nomenclature is broader and would include the morphologically altered mucosa and normally appearing mucosa that possibly harbours dysplasia or molecular aberration due to external factors, chronic inflammation, inherited disorders or a condition/ disease that would place the oral

mucosa more susceptible to OC than usual. [212] For this reason, we felt more comfortable to use the term OPMD instead of OPL in naming our database.

The OPMD database was constructed to be a suppository of cases for future epidemiological research. Therefore, data elements to be collected included all the pertinent data elements that would help us achieve our objectives and answer the questions intended for the OPMD database.

The incidence and prevalence of OPMD varies significantly from one geographical location to another, which can be related to larger behavioural, sociodemographic and cultural differences. Given that the OMFS clinic at MGH- MUHC receives referrals from all the cities of greater Montreal area, our OPMD database will provide us with epidemiological data that represents the area of Montreal. Below I present a brief description of results and a discussion of the database strengths, challenges and limitations, future goals and research.

8.1 Summary of results

Our sample included 155 patients with OPMD who received care at the OMFS clinic at the MGH-MUHC between July 2008 and June 2017. There were more males in the sample (57%) and mean age was 60.2 years old (SE: 13.27), 85.5% were smokers or ex-smokers and 70% were either current or past alcohol users. The most common oral location was the tongue (39%), followed by the buccal mucosa (12%) and the maxillary gingiva (10%). The most common diagnosis was leukoplakia (52.9%) and dysplasia was present in 56% of all cases. Around 10.32% of the cases progressed to oral cancer with mean follow up time of 3.3 years, however, in some patients transformation to oral cancer occurred after more than 10 years. Most of the oral cases were diagnosed at an early stage (87.5%).

8.2 Strengths of the study

This work has several strengths. First, internal validity; REDCap, the software used to build the OPMD database, addresses internal validation by providing real time data verification to alert whether the type of information being entered (e.g., range, alphanumeric,) are compatible with the specifications defined by the user when the database was built. It has a tool to identify outliers, missing data, and sends real time alert error massage if a minimum dataset is not met for a subject. In addition, a number of patients are selected randomly and double entered to validate the accuracy and the completeness of the data entered.

Second, external validity concerns generalisability; in addition to self-referred patients our department receives referrals from all over greater Montreal area and from different healthcare providers including dentists, family doctors and other medical and dental specialists. This insures a wide variability and coverage of patients; however, the database covers only the cases the healthcare provider chose to refer which are more often categorised as a high-risk population. Therefore, the results might not be generalizable without limitations, for example, incidence of malignant transformation in the database might not reflect the incidence in the general population or a population from a non-speciality clinic.

Third, it is the solution provided to overcome data accuracy issues during data exportation from REDCap for statistical analysis purposes. The problem was the inability to differentiate true missing fields from not applicable fields as both would be coded the same in the exported data file; REDCap does not label those two conditions differently, therefore both entries are read as true missing data. This problem was overcome by using an R solution package that facilitates preparing the dataset for

the inducer by fixing all anticipated errors. The data is exported each time it is needed through the Application Programming Interface (API) tool provided by REDCap, during the exportation process the R solution package is programmed to identify certain errors and correct them, afterward prepare a data file that is compatible with Stata statistical software (or any other preferred statistical software). This solution is unique and important because it is not specific for this study and it is not a one-time use remedy for the issue. The R solution package created can be used to fix the above mentioned issue each time the data is exported and for all users, also it could be customised to different databases using the REDCap software.

Fourth, when building the database, we took into account the dynamicity of the literature and the production of new evidence everyday especially in the field of biomarkers. This is appreciated in the primary data sources developed as they are adjustable, and when the team chose a design and a software that allows addition of new data points and forms into the database whenever needed. However, most probably these new data points will be available only for future patients unless the team are able to implement a plan to collect these data pointe for current subjects, and to enter them into the database efficiently, for example, to collect the new information whenever current database subjects come for their biannual follow ups and update their information when the database is updated.

8.3 Limitations and challenges

We have faced some challenges and limitations and foresee others when planning and working on this project. One of the challenges we anticipate is recruiting enough numbers of patients to understand the epidemiology and the natural history of a relatively rare disease such as OPMD. It is a challenge for a single clinic even though it is a large center for referrals in a tertiary health care

facility in a big city. Expanding the database coverage to include other departments in the hospital, and private clinics who treats and follow up OPMD patients is a possible solution.

Another challenge was the size of the database and large number data points to be collected. The scope of the database is broad in terms of variables to collect, variables that could play a protective or a promoting role in causing the disease, and in malignant transformation because of the inconsistency and the variability in the literature regarding potential biological, clinical and socio-demographic risk factors of OPMD and their malignant transformation, as well as the heterogeneity of the disorders falling under the umbrella of OPMD. Nevertheless, the ability to create different, separate forms and utilising the branching logic design in the database helps in hiding all not applicable fields for each particular subject and makes each record as specific as possible to the pertinent information. However, it does not address the large number of variables and the time needed to enter them. In addition, most of the variables are entered only once corresponding to the patient's initial (consult) visit, only one form of the five database forms is filled when the patient record is updated corresponding to follow up visits.

Censoring is a limitation and a challenge that are usually faced in epidemiological research. Since the database is designed in longitudinal format, some patients will be lost to follow up and it will be challenging to minimise lost to follow up subjects and retain participation from both patients and healthcare provider (in case of database expansion). For that the database team should discuss possible solution to keep censoring at minimum. Proposed solution ideas include: for outside hospital providers, a low burden of participation and providing them with help and assistance by the database team when needed could aid in maintaining participation. Regarding loss of subjects, staff members should be aware and reinitiate contact with patients who lost the follow up visit. This would help in retention.

8.4 Future research

The absence of a standard of care for treatment at least for low grade dysplasia cases, the inadequacy of histopathology proven OED as the current golden standard for malignant transformation prediction, and the large variety and lack of clinical validation of most proposed biomarkers to complement the OED system, keeps ideas for future research numerous and variable.

Future tentative research ideas include a cohort study to observe the effect of immunodeficiency, and other clinical and biological elements as risks for malignant transformation among OPMD patients; a randomised controlled trial for OPMD patients with no and low grade dysplasia to assess the effectiveness of surgical excision as a treatment in preventing future malignant transformation in this subgroup of patients.

8.5 Future goals

One of our future goals for the OPMD clinical database is to expand its coverage to incorporate OPMD patients treated in departments other than OMFS at the MGH-MUHC and other hospitals in Montreal area, and further expand it to provide coverage for private sector OPMD patients treated by different healthcare providers including maxillofacial surgeons, dentists and other surgeons who treats and follow up OPMD patients. This could be approached by unifying the clinical data collection process among the participating clinics using the data collection forms that were initiated for the purpose of the OPMD database (clinical forms and the patient filled questionnaires).

Similarly, the OPMD database variables should be updated, so that the database reflects the current research developments especially for new biomarkers been discovered, and if possible to establish a biorepository for biological specimens. The more our knowledge is advanced about OPMD the

more equipped and capable we are to perform analysis on existing biological samples to validate and study newly discovered biomarkers. Indeed, while patient's rights and privacy are protected, and operating under clear guidelines with a planned approach, that covers regulations and details about what type of bio-sample to be collected? How often is the collection process? How long will the sample be stored? And including adequate information in the consent form about the process and the possible types of research the sample could be used in.

8. CONCLUSION

The development of the OPMD clinical database in an academic clinical setting provides a platform for research and education which is a huge advantage and opportunity for researchers, clinicians and students. It acts as a monitor for patients' status and facilitate reinitiating contact with lost to follow up patients. In addition, evidence derived from future research projects conducted based on the OPMD clinical database that can help policy markers decisions.

9. List of references

1. Gimenez‐Conti, I.B. and T.J. Slaga, *The hamster cheek pouch carcinogenesis model.* Journal of Cellular Biochemistry, 1993. **53**(17F): p. 83-90.

2. Vered, M., N. Yarom, and D. Dayan, *4NQO oral carcinogenesis: animal models, molecular markers and future expectations.* Oral Oncol, 2005. **41**(4): p. 337-9.

3. Tanaka, T. and R. Ishigamori, *Understanding Carcinogenesis for Fighting Oral Cancer.* Journal of Oncology, 2011. **2011**: p. 603740.

4. Todd, R., R.B. Donoff, and D.T.W. Wong, *The molecular biology of oral carcinogenesis: Toward a tumor progression model.* Journal of Oral and Maxillofacial Surgery, 1997. **55**(6): p. 613-623.

5. *Canadian Cancer Statistics 2017*, C.C. Society, Editor. 2017, Canadian Cancer Statistics Advisory Committee: Toronto, ON.

6. Jemal, A., et al., *Cancer statistics, 2003.* CA: a Cancer Journal for Clinicians, 2003. **53**(1): p. 5-26.

7. RB., L., *La survie reliée au cancer pour les nouveaux cas déclarés au Québec, de 1984 ŕ 1998: Survie observée et survie relative. Québec, 2003.*

8. Neville, B.W. and T.A. Day, *Oral Cancer and Precancerous Lesions.* CA: A Cancer Journal for Clinicians, 2002. **52**(4): p. 195-215.

9. Rogers, S.N., S.V. Vedpathak, and D. Lowe, *Reasons for delayed presentation in oral and oropharyngeal cancer: the patients perspective.* Br J Oral Maxillofac Surg, 2011. **49**(5): p. 349-53.

10. *Survival statistics for oral cancer 5-year relative survival by stage and tumour site*

    [cited 2018 2018-04-29]; Available from: http://www.cancer.ca/en/cancer-information/cancer-type/oral/prognosis-and-survival/survival-statistics/?region=on.

11. Yanik, E.L., et al., *Leukoplakia, Oral Cavity Cancer Risk, and Cancer Survival in the U.S. Elderly.* Cancer Prevention Research, 2015. **8**(9): p. 857-63.

12. Suter, V.G., et al., *[Oral erythroplakia and erythroleukoplakia: red and red-white dysplastic lesions of the oral mucosa--part 2: cytodiagnosis, pathogenesis, therapy, and prognostic aspects].* Schweiz Monatsschr Zahnmed, 2008. **118**(6): p. 510-8.

13. Dionne, K.R., et al., *Potentially malignant disorders of the oral cavity: Current practice and future directions in the clinic and laboratory.* International Journal of Cancer, 2015. **136**(3): p. 503-515.

14. P.J, T., *Oral precancer – diagnosis and management of potentially malignant disorders*. 2012, Chichester: Wiley-Blackwell. 236

15. Ho, M.W., et al., *Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: rationale for monitoring premalignant oral lesions in a multidisciplinary clinic.* Br J Oral Maxillofac Surg, 2013. **51**(7): p. 594-9.

16. Petti, S., *Pooled estimate of world leukoplakia prevalence: a systematic review.* Oral Oncol, 2003. **39**(8): p. 770-80.

17. Enerly, E., et al., *Quality assessment of the registration of vulvar and vaginal premalignant lesions at the Cancer Registry of Norway.* Acta Oncologica (Stockholm, Sweden), 2012. **51**(1): p. 45-50.

18. Norway, C.R.o., *Cancer in Norway 2015 - Cancer incidence, mortality, survival and prevalence in Norway.* Oslo.

19. Casparie, M., et al., *Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive.* Cell Oncol, 2007. **29**(1): p. 19-24.

20. L, B., *Pathology and Genetics of Head and Neck Tumours*, in *WHO/IARC Classification of Tumours*, E. JW, Editor. 2005, IARC Press: Lyon.

21. Thomson, P.J., *Field change and oral cancer: new evidence for widespread carcinogenesis?* International Journal of Oral and Maxillofacial Surgery, 2002. **31**(3): p. 262-266.

22. Bremmer, J.F., et al., *A noninvasive genetic screening test to detect oral preneoplastic lesions.* Lab Invest, 2005. **85**(12): p. 1481-8.

23. Sarode, S.C., G.S. Sarode, and J.V. Tupkari, *Oral potentially malignant disorders: precising the definition.* Oral Oncol, 2012. **48**(9): p. 759-60.

24. Nagao, T., et al., *Incidence rates for oral leukoplakia and lichen planus in a Japanese population.* J Oral Pathol Med, 2005. **34**(9): p. 532-9.

25. Mehta, F.S., et al., *Oral leukoplakia in relation to tobacco habits: A ten-year follow-up study of Bombay policemen.* Oral Surgery, Oral Medicine, Oral Pathology, 1972. **34**(3): p. 426-433.

26. Gupta, P.C., et al., *Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10 ‐ year follow ‐ up study of Indian villagers.* Community Dentistry and Oral Epidemiology, 1980. **8**(6): p. 287-333.

27. Greer, R.O., *Pathology of malignant and premalignant oral epithelial lesions.* Otolaryngol Clin North Am, 2006. **39**(2): p. 249-75, v.

28. van der Waal, I., *Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management.* Oral Oncology, 2010. **46**(6): p. 423-5.

29. Cowan, C.G., et al., *Potentially malignant oral lesions in northern Ireland: a 20-year population-based perspective of malignant transformation.* Oral Dis, 2001. **7**(1): p. 18-24.

30. Hsue, S.S., et al., *Malignant transformation in 1458 patients with potentially malignant oral mucosal disorders: a follow-up study based in a Taiwanese hospital.* J Oral Pathol Med, 2007. **36**(1): p. 25-9.

31. Saito, T., et al., *Development of squamous cell carcinoma from pre-existent oral leukoplakia: with respect to treatment modality.* Int J Oral Maxillofac Surg, 2001. **30**(1): p. 49-53.

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

33. Mehanna, H.M., et al., *Treatment and follow-up of oral dysplasia - a systematic review and meta-analysis.* Head & Neck, 2009. **31**(12): p. 1600-9.

34. Schepman, K.P., et al., *Malignant transformation of oral leukoplakia: a follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands.* Oral Oncol, 1998. **34**(4): p. 270-5.

35. Kuribayashi, Y., et al., *Long-term outcome of non-surgical treatment in patients with oral leukoplakia.* Oral Oncol, 2015. **51**(11): p. 1020-1025.

36. Hansen, L.S., J.A. Olson, and S. Silverman, Jr., *Proliferative verrucous leukoplakia. A long-term study of thirty patients.* Oral Surg Oral Med Oral Pathol, 1985. **60**(3): p. 285-98.

37. Silverman, S., Jr. and M. Gorsky, *Proliferative verrucous leukoplakia: a follow-up study of 54 cases.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 1997. **84**(2): p. 154-7.

38. Bagan, J.V., et al., *Proliferative verrucous leukoplakia: high incidence of gingival squamous cell carcinoma.* J Oral Pathol Med, 2003. **32**(7): p. 379-82.

39. van der Waal, I., *Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management.* Oral Oncol, 2009. **45**(4-5): p. 317-23.

40. Reichart, P.A. and H.P. Philipsen, *Oral erythroplakia--a review.* Oral Oncol, 2005. **41**(6): p. 551-61.

41. Hardy, H., S. Persac, and J.M. Péron, *Les érythroplasies des voies aérodigestives supérieures.* Revue de Stomatologie et de Chirurgie Maxillo-faciale, 2010. **111**(4): p. 213-215.

42. Pindborg, J.J., et al., *Oral submucous fibrosis as a precancerous condition.* European Journal of Oral Sciences, 1984. **92**(3): p. 224-229.

43. Neville, B.W. and T.A. Day, *Oral cancer and precancerous lesions.* CA Cancer J Clin, 2002. **52**(4): p. 195-215.

44. Angadi, P.V. and S.S. Rao, *Areca nut in pathogenesis of oral submucous fibrosis: revisited.* Oral and Maxillofacial Surgery, 2011. **15**(1): p. 1-9.

45. Arakeri, G. and P.A. Brennan, *Oral submucous fibrosis: an overview of the aetiology, pathogenesis, classification, and principles of management.* British Journal of Oral and Maxillofacial Surgery, 2013. **51**(7): p. 587-593.

46. Aziz, S.R., *Oral Submucous Fibrosis: Case Report and Review of Diagnosis and Treatment.* Journal of Oral and Maxillofacial Surgery, 2008. **66**(11): p. 2386-2389.

47. Pindborg, J.J. and S.M. Sirsat, *Oral submucous fibrosis.* Oral Surgery, Oral Medicine, Oral Pathology, 1966. **22**(6): p. 764-779.

48. McCartan, B.E. and C.M. Healy, *The reported prevalence of oral lichen planus: a review and critique.* J Oral Pathol Med, 2008. **37**(8): p. 447-53.

49. Carbone, M., et al., *Course of oral lichen planus: a retrospective study of 808 northern Italian patients.* Oral Dis, 2009. **15**(3): p. 235-43.

50. Bermejo-Fenoll, A., et al., *A retrospective clinicopathological study of 550 patients with oral lichen planus in south-eastern Spain.* J Oral Pathol Med, 2010. **39**(6): p. 491-6.

51. Fitzpatrick, S.G., S.A. Hirsch, and S.C. Gordon, *The malignant transformation of oral lichen planus and oral lichenoid lesions: a systematic review.* J Am Dent Assoc, 2014. **145**(1): p. 45-56.

52. Silverman, S., Jr., et al., *A prospective study of findings and management in 214 patients with oral lichen planus.* Oral Surg Oral Med Oral Pathol, 1991. **72**(6): p. 665-70.

53. Vieira, R.A., et al., *Actinic cheilitis and squamous cell carcinoma of the lip: clinical, histopathological and immunogenetic aspects.* An Bras Dermatol, 2012. **87**(1): p. 105-14.

54. Napier, S.S., et al., *Potentially malignant oral lesions in Northern Ireland: size (extent) matters.* Oral Diseases, 2003. **9**(3): p. 129-137.

55. Kleinman, D.V., et al., *Toward assessing trends in oral mucosal lesions: lessons learned from oral cancer.* Adv Dent Res, 1993. **7**(1): p. 32-41.

56. Bouquot, J.E., *Oral leukoplakia and erythroplakia: a review and update.* Pract Periodontics Aesthet Dent, 1994. **6**(6): p. 9-17; quiz 19.

57. Smith, L.W., et al., *Oral cancer and precancerous lesions in 57,518 industrial workers of Gujarat, India.* Indian J Cancer, 1975. **12**(2): p. 118-23.

58. Chung, C.H., et al., *Oral precancerous disorders associated with areca quid chewing, smoking, and alcohol drinking in southern Taiwan.* J Oral Pathol Med, 2005. **34**(8): p. 460-6.

59. Kumar, S., et al., *Prevalence and Risk Factors for Oral Potentially Malignant Disorders in Indian Population.* Advances in Preventive Medicine, 2015. **2015**: p. 208519.

60. Amarasinghe, H.K., et al., *Betel-quid chewing with or without tobacco is a major risk factor for oral potentially malignant disorders in Sri Lanka: a case-control study.* Oral Oncol, 2010. **46**(4): p. 297-301.

61. Thomas, S.J., et al., *Betel quid not containing tobacco and oral leukoplakia: a report on a cross-sectional study in Papua New Guinea and a meta-analysis of current evidence.* Int J Cancer, 2008. **123**(8): p. 1871-6.

62. Yang, Y.H., et al., *Incidence rates of oral cancer and oral pre-cancerous lesions in a 6-year follow-up study of a Taiwanese aboriginal community.* J Oral Pathol Med, 2005. **34**(10): p. 596-601.

63. Mehta, F.S., et al., *Epidemiologic and histologic study of oral cancer and leukoplakia among 50,915 villagers in India.* Cancer, 1969. **24**(4): p. 832-849.

64. Napier, S.S. and P.M. Speight, *Natural history of potentially malignant oral lesions and conditions: an overview of the literature.* Journal of Oral Pathology & Medicine, 2008. **37**(1): p. 1-10.

65. Chandran, R., S. Meer, and L. Feller, *Oral leukoplakia in a South African sample: a clinicopathological study.* Oral Dis, 2013. **19**(6): p. 592-7.

66. Zain, R.B., et al., *A national epidemiological survey of oral mucosal lesions in Malaysia.* Community Dent Oral Epidemiol, 1997. **25**(5): p. 377-83.

67. Hashibe, M., et al., *Socioeconomic status, lifestyle factors and oral premalignant lesions.* Oral Oncology, 2003. **39**(7): p. 664-671.

68. Lee, C.H., et al., *The precancer risk of betel quid chewing, tobacco use and alcohol consumption in oral leukoplakia and oral submucous fibrosis in southern Taiwan.* Br J Cancer, 2003. **88**(3): p. 366-72.

69. Starzynska, A., et al., *Oral premalignant lesions: epidemiological and clinical analysis in the northern Polish population.* Postepy Dermatologii I Alergologii, 2014. **31**(6): p. 341-50.

70. Ahrens, W., et al., *Oral health, dental care and mouthwash associated with upper aerodigestive tract cancer risk in Europe: the ARCAGE study.* Oral Oncol, 2014. **50**(6): p. 616-25.

71. Blot, W.J., et al., *Smoking and drinking in relation to oral and pharyngeal cancer.* Cancer Res, 1988. **48**(11): p. 3282-7.

72. Hashibe, M., et al., *Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium.* J Natl Cancer Inst, 2007. **99**(10): p. 777-89.

73. Bosetti, C., et al., *Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers.* Am J Epidemiol, 2008. **167**(4): p. 468-73.

74. Hashibe, M., et al., *Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(2): p. 541-50.

75. Garrote, L.F., et al., *Risk factors for cancer of the oral cavity and oro-pharynx in Cuba.* Br J Cancer, 2001. **85**(1): p. 46-54.

76. Kramer, I.R.H., *Oral cancer and precancer.* British Journal of Oral and Maxillofacial Surgery. **18**(3): p. 274.

77. Morse, D.E., et al., *Smoking and drinking in relation to oral epithelial dysplasia.* Cancer Epidemiol Biomarkers Prev, 1996. **5**(10): p. 769-77.

78. Li, L., et al., *Smoking and drinking in relation to oral potentially malignant disorders in Puerto Rico: a case-control study.* BMC Cancer, 2011. **11**: p. 324.

79. Mayne, S.T., Morse, D. E., & Winn, D. M., *Cancers of the Oral Cavity and Pharynx. In Cancer Epidemiology and Prevention Oxford University Press. .* (2009).

80. *Tobacco smoke and involuntary smoking.* IARC Monogr Eval Carcinog Risks Hum, 2004. **83**: p. 1-1438.

81. Health Canada. *Canadian Tobacco Use Monitoring Survey (CTUMS) 2012*. [Internet] 2012 [cited 2018 Feb 1]; Available from: https://www.canada.ca/en/health-canada/services/publications/healthy-living/canadian-tobacco-use-monitoring-survey-ctums-2012.html.

82. Shetty, P., et al., *Oral Leukoplakia: Clinicopathological Correlation and Its Relevance to Regional Tobacco-related Habit Index.* Journal of Contemporary Dental Practice [Electronic Resource], 2016. **17**(7): p. 601-8.

83. Casparis, S., et al., *Oral lichen planus (OLP), oral lichenoid lesions (OLL), oral dysplasia, and oral cancer: retrospective analysis of clinicopathological data from 2002-2011.* Oral & Maxillofacial Surgery, 2015. **19**(2): p. 149-56.

84. Hassona, Y., et al., *Oral potentially malignant disorders among dental patients: a pilot study in Jordan.* Asian Pacific Journal of Cancer Prevention: Apjcp, 2014. **15**(23): p. 10427-31.

85. Feng, J., et al., *Prevalence and distribution of oral mucosal lesions: a cross-sectional study in Shanghai, China.* Journal of Oral Pathology & Medicine, 2015. **44**(7): p. 490-4.

86. Dietrich, T., P.A. Reichart, and C. Scheifele, *Clinical risk factors of oral leukoplakia in a representative sample of the US population.* Oral Oncol, 2004. **40**(2): p. 158-63.

87. Reichart, P.A., *Identification of risk groups for oral precancer and cancer and preventive measures.* Clinical Oral Investigations, 2001. **5**(4): p. 207-13.

88. Roed-Petersen, B., *Effect on oral leukoplakia of reducing or ceasing tobacco smoking.* Acta Derm Venereol, 1982. **62**(2): p. 164-7.

89. Ho, M.W., et al., *The clinical determinants of malignant transformation in oral epithelial dysplasia.* Oral Oncol, 2012. **48**(10): p. 969-76.

90. Silverman, S., Jr., M. Gorsky, and F. Lozada, *Oral leukoplakia and malignant transformation. A follow-up study of 257 patients.* Cancer, 1984. **53**(3): p. 563-8.

91. Aghbari, S.M.H., et al., *Malignant transformation of oral lichen planus and oral lichenoid lesions: A meta-analysis of 20095 patient data.* Oral Oncol, 2017. **68**: p. 92-102.

92. Warnakulasuriya, S. and A. Ariyawardana, *Malignant transformation of oral leukoplakia: a systematic review of observational studies.* Journal of Oral Pathology & Medicine, 2016. **45**(3): p. 155-66.

93. *Alcohol consumption and ethyl carbamate.* IARC Monogr Eval Carcinog Risks Hum, 2010. **96**: p. 3-1383.

94. Castellsague, X., et al., *The role of type of tobacco and type of alcoholic beverage in oral carcinogenesis.* Int J Cancer, 2004. **108**(5): p. 741-9.

95. Huang, W.Y., et al., *Alcohol concentration and risk of oral cancer in Puerto Rico.* Am J Epidemiol, 2003. **157**(10): p. 881-7.

96. Blot, W.J., *Invited commentary: more evidence of increased risks of cancer among alcohol drinkers.* Am J Epidemiol, 1999. **150**(11): p. 1138-40; discussion 1141.

97. Enwonwu, C.O. and V.I. Meeks, *Bionutrition and oral cancer in humans.* Critical Reviews in Oral Biology & Medicine, 1995. **6**(1): p. 5-17.

98. Hashibe, M., et al., *Alcohol drinking, body mass index and the risk of oral leukoplakia in an Indian population.* International Journal of Cancer, 2000. **88**(1): p. 129-134.

99. Kulasegaram, R., et al., *Case-control study of oral dysplasia and risk habits among patients of a dental hospital.* European Journal of Cancer Part B: Oral Oncology, 1995. **31**(4): p. 227-231.

100. Chher, T., et al., *Prevalence of oral cancer, oral potentially malignant disorders and other oral mucosal lesions in Cambodia.* Ethn Health, 2018. **23**(1): p. 1-15.

101. Carrard, V., et al., *Prevalence and risk indicators of oral mucosal lesions in an urban population from South Brazil.* Oral Diseases, 2011. **17**(2): p. 171-9.

102. Goodson, M.L., O. Hamadah, and P.J. Thomson, *The role of alcohol in oral precancer: observations from a North-East England population.* Br J Oral Maxillofac Surg, 2010. **48**(7): p. 507-10.

103. Fitzpatrick, S.G., S.A. Hirsch, and S.C. Gordon, *The malignant transformation of oral lichen planus and oral lichenoid lesions: a systematic review.* Journal of the American Dental Association, 2014. **145**(1): p. 45-56.

104. Goodson, M.L., et al., *Oral precursor lesions and malignant transformation--who, where, what, and when?* Br J Oral Maxillofac Surg, 2015. **53**(9): p. 831-5.

105. Anderson, A. and N. Ishak, *Marked variation in malignant transformation rates of oral leukoplakia.* Evidence-Based Dentistry, 2015. **16**(4): p. 102-3.

106. Ruokonen, H.M.A., et al., *High percentage of oral lichen planus and lichenoid lesion in oral squamous cell carcinomas.* Acta Odontol Scand, 2017. **75**(6): p. 442-445.

107. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans / Betel-Quid and Areca-Nut Chewing and some Areca-Nut Related Nitrosamines.* IARC Scientific Publications 2004. **85**.

108. Chatterjee, R., B. Gupta, and S. Bose, *Oral Screening for Pre-cancerous Lesions Among Areca-nut Chewing Population from Rural India.* Oral Health & Preventive Dentistry, 2015. **13**(6): p. 509-14.

109. Shiu, M.N., et al., *Risk factors for leukoplakia and malignant transformation to oral carcinoma: a leukoplakia cohort in Taiwan.* Br J Cancer, 2000. **82**(11): p. 1871-4.

110. Tang, J.G., et al., *Epidemiological survey of oral submucous fibrosis in Xiangtan City, Hunan Province, China.* Community Dent Oral Epidemiol, 1997. **25**(2): p. 177-80.

111. Yen, A.M., S.C. Chen, and T.H. Chen, *Dose-response relationships of oral habits associated with the risk of oral pre-malignant lesions among men who chew betel quid.* Oral Oncol, 2007. **43**(7): p. 634-8.

112. Vial, T. and J. Descotes, *Immunosuppressive drugs and cancer.* Toxicology, 2003. **185**(3): p. 229-40.

113. D'Souza, G., et al., *Epidemiology of head and neck squamous cell cancer among HIV-infected patients.* J Acquir Immune Defic Syndr, 2014. **65**(5): p. 603-10.

114. Blohme, I. and H. Brynger, *Malignant disease in renal transplant patients.* Transplantation, 1985. **39**(1): p. 23-5.

115. King , G.N., et al., *Increased Prevalence of Dysplastic and Malignant Lip Lesions in Renal-Transplant Recipients.* New England Journal of Medicine, 1995. **332**(16): p. 1052-1057.

116. Scheifele, C., et al., *Incidence of oral, pharyngeal, and laryngeal squamous cell carcinomas among 1515 patients after liver transplantation.* Oral Oncology, 2005. **41**(7): p. 670-676.

117. Demopoulos, B.P., et al., *Non–Acquired Immunodeficiency Syndrome-Defining Malignancies in Patients Infected With Human Immunodeficiency Virus.* Archives of Pathology & Laboratory Medicine, 2003. **127**(5): p. 589-592.

118. Abdelsayed, R.A., et al., *Oral precancerous and malignant lesions associated with graft-versus-host disease: report of 2 cases.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2002. **93**(1): p. 75-80.

119. Szeto, C.H., et al., *Squamous cell carcinoma of the tongue complicating chronic oral mucosal graft-versus-host disease after allogeneic hematopoietic stem cell transplantation.* Am J Hematol, 2004. **77**(2): p. 200-2.

120. Gong, Y., et al., *Type 2 diabetes mellitus and risk of oral cancer and precancerous lesions: a meta-analysis of observational studies.* Oral Oncology, 2015. **51**(4): p. 332-40.

121. Ai, R., et al., *Microenvironmental regulation of the progression of oral potentially malignant disorders towards malignancy.* Oncotarget, 2017. **8**(46): p. 81617-81635.

122. Erdei, E., et al., *Cytokines and tumor metastasis gene variants in oral cancer and precancer in Puerto Rico.* PLoS ONE [Electronic Resource], 2013. **8**(11): p. e79187.

123. Hsu, H.J., et al., *Role of cytokine gene (interferon-gamma, transforming growth factor-beta1, tumor necrosis factor-alpha, interleukin-6, and interleukin-10) polymorphisms in the risk of oral precancerous lesions in Taiwanese.* Kaohsiung Journal of Medical Sciences, 2014. **30**(11): p. 551-8.

124. de Vasconcelos Carvalho, M., et al., *Alterations in the immunoexpression of galectins-1, -3 and -7 between different grades of oral epithelial dysplasia.* Journal of Oral Pathology & Medicine, 2013. **42**(2): p. 174-9.

125. Ron, E., *Ionizing radiation and cancer risk: evidence from epidemiology.* Radiat Res, 1998. **150**(5 Suppl): p. S30-41.

126. Picascia, D.D. and J.K. Robinson, *Actinic cheilitis: a review of the etiology, differential diagnosis, and treatment.* J Am Acad Dermatol, 1987. **17**(2 Pt 1): p. 255-64.

127. Taghavi, N. and I. Yazdi, *Type of food and risk of oral cancer.* Arch Iran Med, 2007. **10**(2): p. 227-32.

128. Doll, R. and R. Peto, *The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today.* J Natl Cancer Inst, 1981. **66**(6): p. 1191-308.

129. Pavia, M., et al., *Association between fruit and vegetable consumption and oral cancer: a meta-analysis of observational studies.* Am J Clin Nutr, 2006. **83**(5): p. 1126-34.

130. Lagiou, P., et al., *Diet and upper-aerodigestive tract cancer in Europe: the ARCAGE study.* International Journal of Cancer, 2009. **124**(11): p. 2671-6.

131. Lucenteforte, E., et al., *Dietary factors and oral and pharyngeal cancer risk.* Oral Oncol, 2009. **45**(6): p. 461-7.

132. Gupta, P.C., et al., *Influence of dietary factors on oral precancerous lesions in a population-based case–control study in Kerala, India.* Cancer, 1999. **85**(9): p. 1885-1893.

133. Hebert, J.R., et al., *Dietary exposures and oral precancerous lesions in Srikakulam District, Andhra Pradesh, India.* Public Health Nutr, 2002. **5**(2): p. 303-12.

134. Singh, V.N. and S.K. Gaby, *Premalignant lesions: role of antioxidant vitamins and beta-carotene in risk reduction and prevention of malignant transformation.* Am J Clin Nutr, 1991. **53**(1 Suppl): p. 386s-390s.

135. Gupta, P.C., et al., *Dietary factors in oral leukoplakia and submucous fibrosis in a population-based case control study in Gujarat, India.* Oral Dis, 1998. **4**(3): p. 200-6.

136. Nagao, T., et al., *Serum antioxidant micronutrients and the risk of oral leukoplakia among Japanese.* Oral Oncol, 2000. **36**(5): p. 466-70.

137. Walboomers, J.M., et al., *Human papillomavirus is a necessary cause of invasive cervical cancer worldwide.* J Pathol, 1999. **189**(1): p. 12-9.

138. Gillison, M.L., et al., *Evidence for a causal association between human papillomavirus and a subset of head and neck cancers.* J Natl Cancer Inst, 2000. **92**(9): p. 709-20.

139. Jayaprakash, V., et al., *Human papillomavirus types 16 and 18 in epithelial dysplasia of oral cavity and oropharynx: a meta-analysis, 1985-2010.* Oral Oncol, 2011. **47**(11): p. 1048-54.

140. Syrjänen, S., *Human papillomavirus (HPV) in head and neck cancer.* Journal of Clinical Virology, 2005. **32**: p. 59-66.

141. Jayaprakash, V., et al., *Human papillomavirus types 16 and 18 in epithelial dysplasia of oral cavity and oropharynx: A meta-analysis, 1985–2010.* Oral oncology, 2011. **47**(11): p. 1048-1054.

142. Plummer, M., et al., *Global burden of cancers attributable to infections in 2012: a synthetic analysis.* Lancet Glob Health, 2016. **4**(9): p. e609-16.

143. Kreimer, A.R., et al., *Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(2): p. 467-75.

144. Miller, C.S. and B.M. Johnstone, *Human papillomavirus as a risk factor for oral squamous cell carcinoma: A meta-analysis, 1982-1997.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, 2001. **91**(6): p. 622-635.

145. Campisi, G., et al., *Human papillomavirus: its identity and controversial role in oral oncogenesis, premalignant and malignant lesions (review).* Int J Oncol, 2007. **30**(4): p. 813-23.

146. Bouda, M., et al., *"High risk" HPV types are frequently detected in potentially malignant and malignant oral lesions, but not in normal oral mucosa.* Mod Pathol, 2000. **13**(6): p. 644-53.

147. Amarasinghe, H.K., et al., *Diet and risk of oral potentially malignant disorders in rural Sri Lanka.* Journal of Oral Pathology & Medicine, 2013. **42**(9): p. 656-62.

148. Meisel, P., et al., *Association of periodontitis with the risk of oral leukoplakia.* Oral Oncology, 2012. **48**(9): p. 859-63.

149. Wang, Y.Y., et al., *Malignant transformation in 5071 southern Taiwanese patients with potentially malignant oral mucosal disorders.* BMC Oral Health, 2014. **14**: p. 99.

150. Kramer, I.R., N. El-Labban, and K.W. Lee, *The clinical features and risk of malignant transformation in sublingual keratosis.* Br Dent J, 1978. **144**(6): p. 171-80.

151. Waldron, C.A. and W.G. Shafer, *Leukoplakia revisited. A clinicopathologic study 3256 oral leukoplakias.* Cancer, 1975. **36**(4): p. 1386-1392.

152. Holmstrup, P., et al., *Long-term treatment outcome of oral premalignant lesions.* Oral Oncol, 2006. **42**(5): p. 461-74.

153. Warnakulasuriya, S., et al., *Factors predicting malignant transformation in oral potentially malignant disorders among patients accrued over a 10-year period in South East England.* Journal of Oral Pathology & Medicine, 2011. **40**(9): p. 677-83.

154. Warnakulasuriya, S., et al., *Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement.* Journal of Oral Pathology & Medicine, 2008. **37**(3): p. 127-33.

155. Warnakulasuriya, K.A. and N.W. Johnson, *Expression of p53 mutant nuclear phosphoprotein in oral carcinoma and potentially malignant oral lesions.* Journal of Oral Pathology & Medicine, 1992. **21**(9): p. 404-8.

156. Mao, L., et al., *Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment.* Nat Med, 1996. **2**(6): p. 682-5.

157. Zhang, L. and M.P. Rosin, *Loss of heterozygosity: a potential tool in management of oral premalignant lesions?* Journal of Oral Pathology & Medicine, 2001. **30**(9): p. 513-20.

158. Lippman, S.M. and W.K. Hong, *Molecular markers of the risk of oral cancer.* N Engl J Med, 2001. **344**(17): p. 1323-6.

159. Yang, Q., et al., *Identification of the key genes implicated in the transformation of OLP to OSCC using RNA-sequencing.* Oncol Rep, 2017. **37**(4): p. 2355-2365.

160. Kil, T.J., et al., *Genetic Abnormalities in Oral Leukoplakia and Oral Cancer Progression.* Asian Pacific Journal of Cancer Prevention: Apjcp, 2016. **17**(6): p. 3001-6.

161. Dawson, Mark A. and T. Kouzarides, *Cancer Epigenetics: From Mechanism to Therapy.* Cell, 2012. **150**(1): p. 12-27.

162. Lingen, M.W., et al., *Genetics/epigenetics of oral premalignancy: current status and future research.* Oral Diseases, 2011. **17 Suppl 1**: p. 7-22.

163. Zhang, L., et al., *Loss of heterozygosity (LOH) profiles--validated risk predictors for progression to oral cancer.* Cancer Prevention Research, 2012. **5**(9): p. 1081-9.

164. Lee, J.J., et al., *Predicting Cancer Development in Oral Leukoplakia: Ten Years of Translational Research.* Clinical Cancer Research, 2000. **6**(5): p. 1702-1710.

165. William, W.N., Jr., et al., *Erlotinib and the Risk of Oral Cancer: The Erlotinib Prevention of Oral Cancer (EPOC) Randomized Clinical Trial.* JAMA Oncol, 2016. **2**(2): p. 209-16.

166. Abe, M., et al., *High-risk oral leukoplakia is associated with aberrant promoter methylation of multiple genes.* BMC Cancer, 2016. **16**: p. 350.

167. López, M., et al., *Gene promoter hypermethylation in oral rinses of leukoplakia patients—a diagnostic and/or prognostic tool?* European Journal of Cancer, 2003. **39**(16): p. 2306-2309.

168. Cao, J., et al., *Methylation of p16 CpG island associated with malignant progression of oral epithelial dysplasia: a prospective cohort study.* Clin Cancer Res, 2009. **15**(16): p. 5178-83.

169. Liu, H., et al., *P16 Methylation as an Early Predictor for Cancer Development From Oral Epithelial Dysplasia: A Double-blind Multicentre Prospective Study.* EBioMedicine, 2015. **2**(5): p. 432-7.

170. Yang, Y., et al., *Progress risk assessment of oral premalignant lesions with saliva miRNA analysis.* BMC Cancer, 2013. **13**: p. 129-129.

171. Lodi, G., et al., *Interventions for treating oral leukoplakia to prevent oral cancer.* Cochrane Database of Systematic Reviews, 2016. **7**: p. CD001829.

172. Zhang, L., et al., *Should severe epithelial dysplasia be treated?* Oral Oncology, 2016. **60**: p. 125-9.

173. Arnaoutakis, D., et al., *Recurrence patterns and management of oral cavity premalignant lesions.* Oral Oncology, 2013. **49**(8): p. 814-7.

174. Dost, F., et al., *Malignant transformation of oral epithelial dysplasia: a real-world evaluation of histopathologic grading.* Oral Surg Oral Med Oral Pathol Oral Radiol, 2014. **117**(3): p. 343-52.

175. Brooke and E. M, *The current and future use of registers in health information systems Eileen M. Brooke.* 1974, Geneva . World Health Organization.

176. Hammill, B.G., et al., *Linking inpatient clinical registry data to Medicare claims data using indirect identifiers.* Am Heart J, 2009. **157**(6): p. 995-1000.

177. Dokholyan, R.S., et al., *Regulatory and ethical considerations for linking clinical and administrative databases.* Am Heart J, 2009. **157**(6): p. 971-82.

178. *AHRQ Methods for Effective Health Care*, in *Registries for Evaluating Patient Outcomes: A User's Guide*, rd, et al., Editors. 2014, Agency for Healthcare Research and Quality (US): Rockville (MD).

179. Roberts, A.L. and J.P. Sewell, *Data Aggregation: A Case Study.* CIN: Computers, Informatics, Nursing, 2011. **29**(1): p. 3-7.

180. Black, N., *High-quality clinical databases: breaking down barriers.* The Lancet, 1999. **353**(9160): p. 1205-1206.

181. Tilson, H.H., et al., *The antiretrovirals in pregnancy registry: a fifteenth anniversary celebration.* Obstet Gynecol Surv, 2007. **62**(2): p. 137-48.

182. Eydelman, M.B., et al., *Ophthalmic Devices and Clinical Epidemiology*, in *Medical Device Epidemiology and Surveillance*. 2007, John Wiley & Sons, Ltd. p. 427-439.

183. Bravata, D.M., et al., *AHRQ Comparative Effectiveness Reviews*, in *Comparative Effectiveness of Percutaneous Coronary Interventions and Coronary Artery Bypass Grafting for Coronary Artery Disease*. 2007, Agency for Healthcare Research and Quality (US): Rockville (MD).

184. Jensen, O.M. and S. Whelan, *Planning a cancer registry.* IARC scientific publications, 1991(95): p. 22-28.

185. de Groot, S., et al., *Balancing the Optimal and the Feasible: A Practical Guide for Setting Up Patient Registries for the Collection of Real-World Data for Health Care Decision Making Based on Dutch Experiences.* Value in Health, 2017. **20**(4): p. 627-636.

186. Miller, A.R. and C.E. Tucker, *Encryption and the loss of patient data.* J Policy Anal Manage, 2011. **30**(3): p. 534-56.

187. *Cancer registration: principles and methods.* IARC Sci Publ, 1991(95): p. 1-288.

188. Bray, F. and D.M. Parkin, *Evaluation of data quality in the cancer registry: Principles and methods. Part I: Comparability, validity and timeliness.* European Journal of Cancer, 2009. **45**(5): p. 747-755.

189. *GLOBOCAN 2012: Estimated Cancer Incidance, Mortality and Prevalence Worldwide in 2012* [cited 2018 2018-05-11]; Available from: http://globocan.iarc.fr/Pages/fact_sheets_population.aspx?country=900.

190. Warnakulasuriya, S., *Global epidemiology of oral and oropharyngeal cancer.* Oral Oncol, 2009. **45**(4-5): p. 309-16.

191. *Survival by Stage*. Cancer Stat Facts: Oral Cavity and Pharynx Cancer [cited 2018 2018-05-11]; Available from: https://seer.cancer.gov/statfacts/html/oralcav.html.

192. Speight, P.M., *Update on oral epithelial dysplasia and progression to cancer.* Head Neck Pathol, 2007. **1**(1): p. 61-6.

193. Bouquot, J.E., P.M. Speight, and P.M. Farthing, *Epithelial dysplasia of the oral mucosa—Diagnostic problems and prognostic features.* Current Diagnostic Pathology, 2006. **12**(1): p. 11-21.

194. Katsanos, K.H., et al., *Oral Cancer and Oral Precancerous Lesions in Inflammatory Bowel Diseases: A Systematic Review.* Journal of Crohn's & colitis, 2015. **9**(11): p. 1043-52.

195. Piselli, P., et al., *Epidemiology of de novo malignancies after solid-organ transplantation: immunosuppression, infection and other risk factors.* Best Pract Res Clin Obstet Gynaecol, 2014. **28**(8): p. 1251-65.

196. Shah, A.T., E. Wu, and R.O. Wein, *Oral squamous cell carcinoma in post-transplant patients.* American Journal of Otolaryngology, 2013. **34**(2): p. 176-9.

197. Hernandez, G., et al., *Rapid progression from oral leukoplakia to carcinoma in an immunosuppressed liver transplant recipient.* Oral Oncol, 2003. **39**(1): p. 87-90.

198. Meisel, P., et al., *Association Between Glycemia, Serum Lipoproteins, and the Risk of Oral Leukoplakia: The population-based Study of Health in Pomerania (SHIP).* Diabetes Care, 2010. **33**(6): p. 1230-1232.

199. P., D.R., et al., *Association between diabetes mellitus and pre‐malignant oral diseases: A cross sectional study in Kerala, India.* International Journal of Cancer, 2006. **118**(2): p. 453-457.

200. S, N.B.a.L. *redcapAPI: Accessing data from REDCap projects using the API*. 2018; Available from: https://github.com/nutterb/redcapAPI/wiki.

201. G., A.P., et al., *Outcome of oral dysplasia: a retrospective hospital‐based study of 207 patients with a long follow‐up.* Journal of Oral Pathology & Medicine, 2009. **38**(6): p. 540-544.

202. Jaber, M.A., et al., *Oral epithelial dysplasia: clinical characteristics of western European residents.* Oral Oncology, 2003. **39**(6): p. 589-596.

203. Lumerman, H., P. Freedman, and S. Kerpel, *Oral epithelial dysplasia and the development of invasive squamous cell carcinoma.* Oral Surgery Oral Medicine Oral Pathology Oral Radiology & Endodontics, 1995. **79**(3): p. 321-9.

204. Canada, S., *Table 051-0001 - Estimates of population, by age group and sex for July 1, Canada, provinces and territories, annual (persons unless otherwise noted)*.

205. Ho, M.W., et al., *Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: rationale for monitoring premalignant oral lesions in a multidisciplinary clinic.* British Journal of Oral & Maxillofacial Surgery, 2013. **51**(7): p. 594-9.

206. Bánóczy, J. and Á. Csiba, *Occurrence of epithelial dysplasia in oral leukoplakia.* Oral Surgery, Oral Medicine, Oral Pathology, 1976. **42**(6): p. 766-774.

207. Wei, L., et al., *Malignant transformation of oral epithelial dysplasia: clinicopathological risk factors and outcome analysis in a retrospective cohort of 138 cases.* Histopathology, 2011. **59**(4): p. 733-740.

208. Petti, S., et al., *Orofacial diseases in solid organ and hematopoietic stem cell transplant recipients.* Oral Diseases, 2013. **19**(1): p. 18-36.

209. Helenius-Hietala, J., et al., *Oral mucosal health in liver transplant recipients and controls.* Liver Transplantation, 2014. **20**(1): p. 72-80.

210. Thomson, P.J., *Field change and oral cancer: new evidence for widespread carcinogenesis?* Int J Oral Maxillofac Surg, 2002. **31**(3): p. 262-6.

211. F Bremmer, J., et al., *A noninvasive genetic screening test to detect oral preneoplastic lesions*. Vol. 85. 2005. 1481-8.

212. Sarode, S.C., et al., *A new classification for potentially malignant disorders of the oral cavity.* Oral Oncology. **47**(9): p. 920-921.

10. APENDICIES

10.1 Appendix I – Clinical form

MGH ORAL AND MAXILLOFACIAL SURGERY
MAXILLOFACIAL OPL CONSULT FORM

Date: _____
Age: _____ Sex: ☐ M ☐ F
Referral: ☐ General Dentist: _____ ☐ OMFS: _____ ☐ Others: _____
Reason for Referral: _____
Patient's complaint and History of present illness:
Date symptoms started/Lesion noticed for first time**:** _____
Date of First time patient sought medical consultation: _____Consult done by**:** _____
If patient notified about the diagnosis, give date: _____
Other pertinent history: _____
_____
_____
Symptoms:
If yes, Pain: ☐ No ☐ Yes Sensitivity: ☐ No ☐ Yes
Paresthesia: ☐ No ☐ Yes Discomfort: ☐ No ☐ Yes
☐ Other: _____ _____
Medical History:
☐ CVS: _____ ☐ Gastro: _____
☐ Respiratory:_____ ☐ Neurology: _____
☐ Endocrinology: _____ ☐ Other: _____
Medications: ☐ None ☐ List Attached B-Blocker: ☐ No ☐ Yes
List: _____
Allergies: ☐ None (NKDA) ☐ Yes:_____
Surgical History: _____
Social History:
Smoking Hx: ☐ Never ☐ Former smoker (more than 6 months): year quitting _____, ☐ Former smoker (less than 6 months), ☐ Smoker: Number of years: ____
Type1(Cig, Cigars, Pipes, Hours of smoking water pipe): _____, amount/day____
Type 2: _____amount /day_____.
Alcohol use: ☐ Never, ☐ Former user, ☐ User (if User then): ☐ Social ☐ Habitual(>3glasses/day)
If Habitual: Type1 (beer, wine, hard liquor): _____, amount /day_____.
Type 2: _____amount /day_____.

Sexual activity: ☐ yes ☐ No ☐ Unknown/ doesn't want to disclose
History of Oral sex: ☐ No ☐ Yes

Marital status: ☐ Single ☐ Married ☐ Others: _____

History of OPL ☐ No ☐ Yes If yes, How many lesions? _____

Lesion#1: When: _____, Dx ☐ histopathology ☐ clinical, Type: _____

Location: _____Treatment_____

Lesion#2: When: _____, Dx ☐ histopathology ☐ clinical, Type: _____

Location: _____Treatment_____

History of Cancer (Head and Neck or Non-Head and Neck): ☐ No ☐ Yes

If yes, Details (Type, location, When, Treatment): _____

_____

Family History of ☐ **Cancer** ☐**OPL**: ☐ No (Relationship/Type):

Family member #1: _____Type_____ Family member #2: _____Type_____

Family member #3: _____Type_____

Clinical Examination:

Patient General Look: _____

Head: _____

Neck: _____

Intra-Oral (Lips, Tongue, FOM, Pharynx, Buccal/Alveolar mucosa, Dentition, Saliva): _____

_____

- Lesion #1: Site: _____Size(mm): _____Color: _____ Borders: _____

Texture: _____ Ulceration: ☐ Yes ☐ No Homogenous: ☐Yes ☐No

Symptoms: _____, Biopsy done ☐No ☐Yes: Dx: _____

- Lesion #2: Site: _____ Size(mm): _____ Color: _____ Borders: _____

Texture: _____ Ulceration: ☐ Yes ☐ No Homogenous: ☐ Yes ☐ No

Symptoms_____, Biopsy done ☐No ☐Yes: Dx: _____

- Lesion #3: Site: _____ Size(mm): _____ Color: _____ Borders: _____

Texture: _____ Ulceration: ☐ Yes ☐ No Homogenous: ☐ Yes ☐ No

Symptoms: _____, Biopsy done ☐No ☐Yes: Dx: _____

Impression (Summary includes: Age, short description of the lesion/problem):

_____

DDx:_____

Plan:

☐ Eligible for OPL registry

☐ Follow up in ☐___Weeks ☐ 6 Months ☐___Months

☐ Biopsy: ☐ Incisional ☐ Excisional

☐ Blood test(Specify):_____

☐ Medications prescribed ☐ Medical work up needed ☐ Booking for OR (pre-op sheets, consent)

_____

_____

Resident: Staff: Date:

MGH ORAL AND MAXILLOFACIAL SURGERY
MAXILLOFACIAL OPL <u>FOLLOW-UP FORM</u>

McGill
Maxillofacial Oncology and Reconstructive Surgery
Chirurgie Oncologique et Reconstruction Maxillo-Faciale

Date: _____

Routine follow up ☐Yes ☐No, Date of last follow up_____
Age: _____ Sex: ☐ M ☐ F
Referral: General Dentist: _____ OMFS:_____ Others:_____

Lesion	Diagnosis Hx/ histopathology	Date of Dx	Location	Treatment provided and margins (f/u, excisional or incisional biopsy, WLE)	Date of Treatment
Lesion ____					
Lesion ____					
Lesion ____					
Lesion ____					
Lesion ____					

New histopathology results to follow up on: ☐ No ☐ Yes Results: _____

Change in Medical History: ☐ No ☐ Yes, _____

Contributing factors:
Change in smoking History: ☐ No ☐ Yes, If yes or no:
☐Smoker ☐ Never smoked ☐ Former smoker (more than 6 months) ☐ Former smoker (less than 6 months)
Smoking Habit change details:
Type 1(cigarettes, cigar, pipe, water pipe): _____, Amount (number of Cig/Cigars/Pipes/ Hours of smoking water pipe) /day_____, Duration of the habit new (or not new) pattern (in months) _____

Type 2 _____, Amount (number of Cig/Cigars/Pipes/ Hours of smoking water pipe) /day_____
Duration of the new habit (in months) Type 2 _____
Change in Alcohol Drinking History: ☐ No ☐ Yes, If yes or no:
☐ User ☐ Never ☐ Former user (if user then): ☐ Social ☐ Habitual (>= 4glasses/day)
Habitual: Type1 ☐ Beer ☐ Wine ☐ Hard Liquor (select all that apply), Amount/day:
_____.
Duration of the habit new (or not new) pattern (in months) _____

Type 2: ☐ Beer ☐ Wine ☐ Hard Liquor (select all that apply), Amount/day: _____.
Duration of the habit new (or not new) pattern (in months) _____

Clinical Examination:
Subjective:
_____
_____

Patient General Look: _____
Head: _____
Neck: _____

Intra-Oral:
Any lesions present: ☐ No ☐ Yes, if yes: ☐ New lesion ☐ Has been followed up (known) ☐Both
If known , location:_____, Color:_____,
Size(mm)_____ Homogenous: ☐ No ☐Yes, Ulceration: ☐No ☐Yes, Symptomatic: ☐ No ☐Yes, if yes_____

If new, location: _____Color: _____,
Size(mm)_____
Homogenous: ☐ No ☐Yes, Ulceration: ☐No ☐Yes, Symptomatic: ☐ No ☐Yes, if yes_____
(New/Known) location: _____, Color: _____,
Size(mm)_____
Homogenous: ☐ No ☐Yes, Ulceration: ☐No ☐Yes, Symptomatic: ☐ No ☐Yes, if yes_____
Assessment:
☐ No lesions present (treated) ☐ Stable from previous exam ☐ Clinical progression ☐ New lesion

Plan:
New biopsy: ☐ No ☐Yes
If yes, type of biopsy: _____
Follow-up:
☐ __Week(s)
☐ __ Month(s)
☐ 6 Months (usual follow up for all grades of dysplasia)
☐ 1 year

Procedure note/Other:
_____
_____

Resident: Staff: Date:

10.2 Appendix II – Patient's questionnaire

Patient's questionnaire

Date:
Name:
Date of birth:
File number:

1. Gender: ☐Male ☐Female ☐Other

2. Age:

3. Date of birth:

4. You live in: ☐Urban (city) ☐Rural area (farm)

5. What city you live in?
   City name:

1. Ethnicity: ☐White ☐ Black ☐Asian ☐ Aboriginal ☐Mixed ☐Other :……

2. Highest degree of qualification you obtained:
   - ☐ None
   - ☐ Primary / elementary
   - ☐ High school
   - ☐ Technical qualification
   - ☐ CEGEP
   - ☐ University
   - ☐ Post – graduate (MSc – PhD)
   - ☐ Professional degree (MD – DMD)

9. Household income:
   - ☐ less than $10,000
   - ☐ $10,000 – $19,000
   - ☐ $20,000 - $29,000
   - ☐ $30,000 - $39,000
   - ☐ $40,000 - $49,000
   - ☐ $50,000 - $59,000
   - ☐ $60,000 - $69,000
   - ☐ $70,000 - $79,000
   - ☐ $80,000 - $89,000
   - ☐ $90,000 - $99,000
   - ☐ $100,000 - $120,000
   - ☐ above $120,000

10. Employment status:
    - ☐ Employed for wages
    - ☐ Self-employed
    - ☐ Out of work and looking
    - ☐ Military
    - ☐ Out of work but not looking currently
    - ☐ Student
    - ☐ Homemaker
    - ☐ Retired
    - ☐ Unable to work

11. If you choose any of the first four options in the previous question, what is your profession?

    _____
    _____
    _____
    _____

12. Have you ever smoked in your life? (or chewed, any product, any amount)
    - ☐ NO
    - ☐ Yes, in the past
    - ☐ Yes (I still do)

❖ **Think of the periods in your life during which you smoked cigarettes, cigars, pipe, chewed tobacco products and/or took drugs, the amount you smoked / chewed / took and other details about the products. Please try to summarise the most important changes in the amount and type of product**.

13. Do you smoke cigarettes?
    - ☐ No
    - ☐ Yes, in the past
    - ☐ yes (I still do)

From age	To age [A]	Type	Brand	#cigarettes/day [B]

[A] If still smoking, write age at time of interview, if less than one year, write same age From and To

[B] If less than daily Make average if not constant frequency

14. Do you smoke cigar?
    - ☐ No
    - ☐ Yes, in the past
    - ☐ yes (I still do)

From age	To age [A]	Brand	#cigars/day [B]

[A] If still smoking, write age at time of interview, if less than one year, write same age From and To
[B] If less than daily Make average if not constant frequency

15. Do you smoke pipe?
    - ☐ No
    - ☐ Yes, in the past
    - ☐ yes (I still do)

From age	To age [A]	Brand	Unit [B]	#/day [C]

[A] If still smoking, write age at time of interview, if less than one year, write same age From and To
[B] Grams or Pipes
[C] If less than daily Make average if not constant frequency

16. Do you smoke water pipe?
    - ☐ No
    - ☐ Yes, in the past
    - ☐ yes (I still do)

From age	To age [A]	hours/day [B]

[A] If still smoking, write age at time of interview, if less than one year, write same age From and To
[B] If less than daily Make average if not constant frequency

17. Do / did you smoke or inhale drugs (marijuana, grass, dope, joints) at least once a week for at least 6 months in your lifetime?

From age	To age [A]	Type [B]	Unit [C]	#/Day [D]

[A] If still smoking, write age at time of interview, if less than one year, write same age From and To
[B] Marijuana, Grass, Crack, Hashish
[C] Grams, Joints
[D] If less than daily Make average if not constant frequency

18. Have you ever chewed tobacco or any other substance?
☐ No
☐ Yes, in the past
☐ yes (I still do)

From age	To age [A]	Type	Brand

[A] If still practicing the habit, write age at time of interview, if less than one year, write same age From and To

19. Have you ever drink alcoholic beverages at least once a month?
☐ No
☐ Yes, in the past
☐ yes (I still do)

❖ **Please describe the periods in your life during which you consumed alcoholic beverages. Please try to summarise the most important changes in your life regarding the amount and type of beverage.**

From age	To age	Beverage type [A]	Unit [B]	# Glasses/ Day

Beverage (A)	**Unit (B)**
- Wine - Beer / cider - Hard liquor (>35) (whisky, cognac, vodka, brandy, grappa, marc, gin, rum) - Aperitif (<35) (Martini, port, sherry, vermouth) - Other, specify: _____	- Small glass (50ml) (1-2oz) - Medium glass (100ml) (2-3oz) - Big glass (250ml) (7oz) (1/2 pint) - ½ small bottle (330ml) (1beer) - Bottle (700-750 ml) (21oz)

20. Marital status:
    - ☐ Married / Common low
    - ☐ Divorced
    - ☐ Widowed
    - ☐ Single
    - ☐ Unknown

21. Have you ever had sexual intercourse?
    - ☐ No
    - ☐ yes
    - ☐ prefer not to say

22. How old were you when you had your first sexual intercourse?

23. You consider yourself to be:
    - ☐ Heterosexual/straight
    - ☐ Bisexual
    - ☐ Lesbian/homosexual
    - ☐ Other, please specify_____

24. How many sexual partners have you had in total in your life?

25. have you ever performed oral sex?
    - ☐ No
    - ☐ yes

☐ prefer not to say

26. How old were you when you performed oral sex?

27. How many oral sex partners have you performed oral sex too in total in your life?

28. Have you ever had skin warts?
☐ No
☐ yes
☐ prefer not to say

23. If yes, where?
☐ Hands
☐ Feet
☐ Head & Neck
☐ Other, specify…….

24. How old were you?

25. Since you have started your sexual life have you ever had candida albicans?
☐ No
☐ yes
☐ prefer not to say

26. If yes, where?
☐ Genital
☐ Mouth
☐ Other, specify……

10.3 Appendix III – Consent

SUBJECT INFORMATION AND CONSENT FORM
STUDY TITLE: MCGILL MAXILLOFACIAL ORAL PREMALIGNANT LESIONS REGISTRY

Version 2: 28 - 06 - 2017

RESEARCH TEAM
CAOMS

Principal Investigators

Nicholas Makhoul, DMD, MD, FRCD(C), Dip.ABOMS
Maxillofacial surgeon
Maxillofacial oncology and reconstructive surgery

Mcgill University
Director and Assistant Professor
Oral and Maxillofacial surgery

McGill University Health Center
Chief, Department of Dentistry and Oral and Maxillofacial Surgery

1650 Cedar Ave, Room B3-194
Montreal, Quebec, H3G 1A4
Tel.: 514-934-1934 ext: 42468
Fax: 514-934-8340

Belinda Nicolau, DDS, PhD
Associate Professor
Division of Oral Health and Society Faculty of Dentistry, McGill University

2001 McGill College Ave, Suite 527
Montreal, Quebec, H3A 1G1
Tel.: 514-398-7203 ext: 094655
Fax: 514-398-7220

Co-investigators

Michael El-Hakim, DMD, MD, FRCD(C), Dip. ABOMS
Maxillofacial surgeon
Maxillofacial oncology and reconstructive surgery

Mcgill University
Assistant Professor
Oral and Maxillofacial surgery

McGill University Health Center
Maxillofacial Surgeon

1650 Cedar Ave, Room B3-194
Montreal, Quebec, H3G 1A4
Tel.: 514-934-1934 ext: 42468
Fax: 514-934-8340

Lojain Bassyoni, DDS, FRCD(C)
MSc Student
Division of Oral and Maxillofacial Surgery Faculty of Dentistry, McGill University

1650 Cedar Ave, Room B3---112.2
Montreal, Quebec, H3G 1A4
Tel.: 514-934-1934 ext: 45892
Fax:514-934-834

Additional key personnel

Huwaida Makhoul, BSc, MSHS Research Assistant

Division of Oral and Maxillofacial Surgery Faculty of Dentistry, McGill University

1650 Cedar Ave, Room B3---112.2
Montreal, Quebec, H3G 1A4
Tel.: 514-934-1934 ext: 42468
Fax: 514-934-4429

INTRODUCTION

We are asking you for permission to enter your personal, medical and treatment-related information in an electronic database. We are building a database to record information about patients with pre-cancerous lesions of the oral cavity, such as yourself, to conduct research that will help us to better understand different aspects of this disease as well as improve results of cancer treatment and patients' quality of life.

Before providing your answer, you should read and understand the content of this consent form. It contains a full explanation of the project including potential risks and benefits associated with it. If there is anything you do not understand; please ask questions so that you can make an informed decision. If you agree, you will be asked to sign and date this form, and a copy will be given to you.

PURPOSE OF THE STUDY

Oral cancer is potentially preventable, yet it is very prevalent. In Canada, three people die from oral cancer every day – making the five-year survival rate of this cancer lower than that of other cancers such as breast, cervical and prostate cancer. However, if the disease is detected early, the treatment is easier, less invasive and more than 90% curable. Therefore, early detection of this cancer significantly improves the outcome of the treatment.

The data about oral pre-cancer lesions are sparse; therefore using them to improve treatment strategies and evaluate the outcomes is difficult. Our goal is to build a databank that will allow us to study different variables (e.g., risk factors, symptoms and disease pathogenesis) this may lead to a better understanding of precancerous lesions, when and why precancerous lesion progress into cancer and improve treatments.

STUDY DESCRIPTION

We will collect information on several factors including your socio-demographic details (e.g., age, sex, education, income), medical health problems, behavioral factors (e.g., smoking, alcohol, sexual), family history of cancer, type of pre-cancer lesion, clinical features of the lesion (e.g., size, location, color, homogenous or non-homogenous) type of treatment provided, transformation time if occurred, number of lesions and duration, histology and grade of dysplasia (dysplasia is the presence of cells of abnormal type within a tissue, which may lead to the development of cancer), how many patients developed subsequent new sites of dysplasia during follow-up. All the information will be entered into a database using software called REDCap (servers are located at the MUHC research center), which has been specifically designed for this study.

STUDY PROCEDURES

If you agree for your information to be entered in this database, a clinical research coordinator will introduce you to a questionnaire during your first visit to the Oral and Maxillofacial surgery clinic. Completing the questionnaire, which contains the information mentioned above, will take about 15 minutes and will be done using an electronic tablet with the help of the research coordinator, while you are in the clinic. Subsequently, the doctor will do your clinical examination. All the other diagnostic, treatment and follow-up information will be retrieved from medical records entered by the research coordinator.
Occasionally, the research team may contact you again to ask some questions or verify some information.

POTENTIAL BENEFITS

You will not directly benefit from taking part in this project; however, your permission to enter your information in the database may help to further understand this disease. This may eventually lead to the development and improvements of new therapies that benefit future patients with a condition similar to yours.

POTENTIAL RISKS AND DISCOMFORTS

There are no known harms associated with your participation in the establishment of the registry. However, some questions related to your behavior may cause you discomfort.

NEW INFORMATION

You will be informed if any new information becomes available that would affect your willingness to continue participating in this project.

CONFIDENTIALITY

If you agree for your information to be entered in this database, we will remove any information that may identify you including your name, street address, e-mail address, telephone and fax numbers; full face photos and any other comparable images; medical record numbers, health plan beneficiary numbers and biometric identifiers, such as finger and voice prints. Our data set may include the following (potentially identifying) information: date of admission, discharge, and service dates; year of birth and, if applicable, death. Your information will

be coded before being entered into a database by assigning each subject a unique study number.
Only the study team members will have access to information that may potentially identify you and to your study number. This information will be stored separately from the coded data. Although direct identifiers are not retained in the study database, multiple data factors put together could result in the identification of individuals. Your confidentiality will be protected to the extent permitted by applicable laws and regulations.

The information entered in the database will be available for research purposes to the study investigators, co-investigators as well as their support staff and students.

Participation: the results from any research conducted using information from this database may be published in scientific journals in an anonymous form. Your confidentiality will be protected to the extent permitted by applicable laws and regulations. Your identity will never be revealed and no identifiers will leave the MUHC. All research studies that will use the data in the database will first receive an ethics approval from the researcher's designated ethics board and will reiterate the focus on oral maxillofacial cancer of the registry information.

VOLUNTARY PARTICIPATION AND WITHDRAWAL FROM THE STUDY

Your approval to enter your information in this database is entirely voluntary. You are perfectly free to withdraw at any time, even after filling out the questionnaires, without explanation and without penalty or loss of benefits to which you are otherwise entitled.

If you decide not to participate or to discontinue your participation, you will not suffer any prejudice regarding medical care or your participation in any other research study. You may refuse to answer any question you do not want to answer.
The study doctor may end your participation for administrative reasons unrelated to the purpose of the study. Additionally, the McGill University Health Center (MUHC) Research Ethics Board (REB) may terminate this study.

COST AND COMPENSATION

You will not be paid for contributing your information to this data base or for any research conducted using this information.

CONTROL OF THE ETHICAL ASPECTS OF THE RESEARCH PROJECT

The McGill University Health Center (MUHC) Research Ethics Board (REB) has conducted an ethics review and provided an ethics approval for the research database. Any subsequent changes to the consent form, database and/or procedures will require an ethics review and approval before the researchers can implement the changes.

FUNDING OF THE RESEARCH PROJECT

The investigators of this project received funds as a grant from CAOMS.

STUDY RECORDS RETENTION POLICY

Any information that links your identity to the database will be coded and kept in a secure location with very limited access. All the anonymized information will be kept in the database indefinitely.

QUESTIONS AND/OR CONTACT INFORMATION

If you have any questions regarding this research study, please contact: Nicholas Makhoul
Assistant Professor

Division of Oral and Maxillofacial Surgery Faculty of Dentistry, McGill University
1650 Cedar Ave, Room B3-119.1 Montreal, Quebec, H3G 1A4
Tel.: 514-934-1934 ext: 42468 | Fax: 514-934-8340

If you have any questions concerning your rights as a research participant, please contact the MUHC Ombudsman office at 514-934-1934 ext: 48306, and they will provide you with independent advice.

STUDY TITLE: MCGILL MAXILLOFACIAL ORAL PREMALIGNANT LESIONS REGISTRY

DECLARATION OF CONSENT

I have read the contents of this consent form, and I agree to participate in this research study.

☐ I allow access to my medical records for the purpose of the registry.

☐ I allow the research team to contact me again to ask some questions or verify some information if needed.

I have had the opportunity to ask questions and all of my questions have been answered to my satisfaction. I have been given sufficient time to consider the above information and to seek advice if I choose to do so. I understand that I will be given a copy of this consent form. By signing this consent form, I am not giving up any of my legal rights.

Participant Signature	Date	Participant Printed Name

Signature of Investigator/ Delegate obtaining the consent	Date	Name of Investigator/ Delegate obtaining the consent

10.4 Appendix IV - REDCap instruments explanatory tables

Instrumnet	Section	Field	Options
Patient's demographics			
		Study id	First patient in the database will have number 1, second 2, and so on…
		Capture Type	- Prospective; if the patient is new or still following up in the clinic (not lost to follow up or discharged) - Retrospective; patients who are not following up in the clinic anymore.
	DEMOGRAPHICS		
		Date of birth	
		Age	At time of diagnosis
		Gender	As documented on the consult form or on oasis
		Marital status at diagnosis	As documented on the consult form
		Race	As documented on the patient's filled questionnaire
		Education	As documented on the patient's filled questionnaire
		Household income	As documented on the patient's filled questionnaire
		Employment status	As documented on the patient's filled questionnaire
		Profession	As documented on the patient's filled questionnaire
		Lives in:	As documented on the patient's filled questionnaire
		City name:	As documented on the patient's filled questionnaire
Initial visit: Habits History			
	SUBSTANCE USE: TOBACCO		
		History of tobacco use	- Former smoker; non-smoker for at least 6 months

			- Information to be captured from patient filled questionnaire
		Year of quitting	- 999 if unknown - Information to be captured from consult form
		Types of tobacco use	- Choose all that apply - Information to be captured from consult form and/ or patient filled questionnaire
	SUBSTANCE USE: TOBACCO Cigarettes		
		How many times cigarettes smoking habit changed based on age?	- Information to be captured from patient filled questionnaire
	Cigarettes, period 1		
		Period 1, From age: (Age when first started smoking cigarettes)	- Information to be captured from patient filled questionnaire - 999 if unknown
		To age: (end of period 1)	- Information to be captured from patient filled questionnaire - 999 if unknown
		Number of Years Smoked (period1)	Calculated field
		Do you know the number of cigarettes smoked/ day during the above specified period?	- Information to be captured from patient filled questionnaire
		Number of cigarettes smoked /day during the	- Information to be captured from patient filled questionnaire - 999 if unknown

		above specified years (period 1)	
		Pack years for period 1	Calculated field
Guide on the section above applied to other periods and other types of substances smoked			
	SUBSTANCE USE: TOBACCO Water-pipe		
	Water-pipe, Period1:		
		Do you know the number of 20 minutes session of water pipe smoked/ day during the above specified period?	- Information to be captured from patient filled questionnaire - Convert hours reported by the patient to 20 min sessions (multiply x 3)
	SUBSTANCE USE: ALCOHOL		
		History of alcohol use: User = At least one drink a year	- Information to be captured from patient filled questionnaire and /or consult form
		Type of User	- Information to be captured from consult form - Social = no more than 2-3 drinks/day, Habitual = more than 2-3 drinks/day
		Select the type of Drinks	- Information to be captured from patient filled questionnaire and /or consult form
	ALCOHOL Type: Unknown		
		Period 1, From age:	- Information to be captured from patient filled questionnaire - 999 if unknown

		To age: (end of period 1)	- Information to be captured from patient filled questionnaire - 999 if unknown
		Number of Years Smoked (period1)	Calculated field
		Number of drinks/ day during the above specified period	- Information to be captured from patient filled questionnaire
Guide on the section above applied to other periods and other types of alcohol (wine, beer, hard liquor)			
	SEXUAL ACTIVITY		
		Has the patient ever had sexual intercourse?	- Information to be captured from patient filled questionnaire
		Age at first intercourse	- Information to be captured from patient filled questionnaire - 999 if unknown
		If Yes, total number of sexual partners	- Information to be captured from patient filled questionnaire - 999 if unknown
		Has the patient ever performed oral sex?	- Information to be captured from patient filled questionnaire
		Age when performed oral sex the first time:	- Information to be captured from patient filled questionnaire
		If Yes, total number of oral sex partners	- Information to be captured from patient filled questionnaire
		Patient considers him/herself	- Information to be captured from patient filled questionnaire
		Skin warts	- Information to be captured from patient filled questionnaire
		If yes, where?	- Information to be captured from patient filled questionnaire

		Age when warts first diagnosed:	- Information to be captured from patient filled questionnaire - 999 if unknown
		Have you ever been diagnosed with Candida albicans since you have been sexually active?	- Information to be captured from patient filled questionnaire
		If yes, where?	- Information to be captured from patient filled questionnaire
initial visit: Past Medical History & Family History			
	FAMILY HISTORY OF OPL		
		Any family history of oral premalignant lesions?	- Information to be captured from consult form
		How many family history of oral premalignant lesions?	- Information to be captured from consult form
		Please specify Relationship	- Information to be captured from consult form
		Type of oral premalignant lesion	- Information to be captured from consult form
	FAMILY HISTORY OF CANCER		
		Any family history of cancer?	- Information to be captured from consult form
		How many family history of cancer?	- Information to be captured from consult form

		Please specify Relationship	- Information to be captured from consult form
		Type of Cancer	- Information to be captured from consult form
	Patient's History of Non-Head and Neck Cancer and Medical History		
		Does the Patient have history of a chronic medical illness?	- Information to be captured from consult form - Choose all that apply - For other: 1. HTN for hypertention 2. HypoT4 for hypothyroidism 3. MI for myocardial infarction 4. CAD for coronary artery disease
		Patient history of Non Head and Neck Cancers?	- Information to be captured from consult form
		Type	- Information to be captured from consult form
		Year of diagnosis	- Information to be captured from consult form
	Patient's History of Oral Premalignant lesions (lesions diagnosed and followed up before patient's first visit to the OMFS clinic)		
		Previous history of oral	- Information to be captured from consult form

		premalignant lesions	
		If yes, how many lesions?	- Information to be captured from consult form
	Lesion 1		
		Specify location:	- Information to be captured from consult form
		Laterality	- Information to be captured from consult form
		Specify location:	- Information to be captured from consult form
	Diagnosis		
		Biopsy done:	- Information to be captured from consult form
		If biopsy, histologic diagnosis:	- Information to be captured from consult form - Always choose dysplasia if present
		If dysplasia, specify	- Information to be captured from consult form
		If yes, biopsy date:	- Information to be captured from consult form
		If NO, type of premalignant lesion or condition:	- Information to be captured from consult form
	Treatment		
		Type of treatment provided	- Information to be captured from consult form
- Above guide is applied to lesion 2, 3, etc.			
initial visit: Referral Event Details and Presentation of current OPMD			
	REFERRAL		

		Referral	- Information to be captured from consult form
		If DDS, specify	- Information to be captured from consult form
		If MD, specify	- Information to be captured from consult form -
	Diagnosis, Clinical Characteristics		
		Treating surgeon or dentist: (Surgeon or dentist who saw the patient at his/her initial visit	- Information to be captured from consult form
		Date of initial visit to the speciality clinic at McGill dentistry department (Oral surgery or Oral pathology clinic)	- Information to be captured from consult form -
		How many lesions presented?	- Information to be captured from consult form
	Lesion 1		Lesion 1 should be the primary lesion of interest
		Is this lesion:	- Information to be captured from consult form
	Clinical Presentation and Details		
		Site:	- Information to be captured from consult form
		Laterality:	- Information to be captured from consult form

		Presentation	- Information to be captured from consult form
		If Symptomatic	- Information to be captured from consult form - Choose all that apply
		Color	-
		Ulceration	-
		Texture:	- Choose all that apply
		Homogenous:	-
		Duration (in weeks): (How long since the lesion was first noticed)	- Information to be captured from consult form - How long since the lesion was first noticed - 999 if unknown
		Size Text box	- Information to be captured from consult form - As documented clinically - document in mm2 - 999 if unknown
		Size radio button	- Information to be captured from consult form - As documented clinically
	Diagnosis		
		Biopsy done at MUHC Dentistry and Maxillofacial Surgery department:	- Information to be captured from consult form
		If no biopsy done at initial visit or as part of the initial plan, how was the diagnosis reached?	- check for outside pathology report, usually scanned into the patient chart on medesync or present on oasis - if not choose clinically - clinical diagnosis usually is one of the following options (Lichen planus, Leukoplakia, Erythroplakia, Erythroleukoplakia, PVL) - If the diagnosis dysplasia, hyperkeratosis, hyperplasia there must be a biopsy and a report.

			- diagnosis like actinic keratosis, lichen planus, oral submucous fibrosis could be clinical only or confirmed with biopsy. always check for a report before choosing clinical diagnosis
		If No biopsy or recent outside pathology report, type of premalignant lesion or condition:	- Information to be captured from consult form
		If biopsy done at the MUHC OMFS and Dentistry department or recent pathology report, type of biopsy:	- Information to be captured from consult form or procedure note
		Histopathological diagnosis:	- Information to be captured from pathology report on oasis - If the patient received more than one initial biopsy, or have more than one diagnosis in the pathology report, choose the most sever, (ex: LP and dysplasia, always report dysplasia here)
		If dysplasia, specify:	- Information to be captured from pathology report on oasis - choose the most sever type, if the report has both mild and moderate, always report moderate

		HPV status	- Information to be captured from pathology report on oasis if available
		P16 status	- Information to be captured from pathology report on oasis if available
		Was S100a7 testing complete?	- Information to be captured from pathology report on oasis if available
		What was the risk reported from S100a7 testing?	- Information to be captured from pathology report on oasis if available
		If Biopsy or outside recent pathology report, do you know the procedure date?	- Information to be captured from pathology report on oasis or outside report scanned and saved in the patient chart on medesync
		If yes, Date of procedure	- Information to be captured from pathology report on oasis or outside report scanned and saved in the patient chart on medesync
	Treatment		
		If incisional biopsy results (MUHC or outside) came back as Dysplasia, did the patient receive WLE or excisional biopsy later on based on these results?	- Information to be captured from procedure note or OR note in the patient chart on medesync.
		If yes, procedure date	- Information to be captured from procedure note or OR note in the patient chart on medesync
		Histopathological diagnosis:	- Information to be captured from pathology report on oasis corresponding to the tissue obtained from this procedure

			- compare date of procedure with date of collection on pathology report
		If dysplasia, specify:	- Information to be captured from pathology report on oasis corresponding to the tissue obtained from from the wide local excision or excisional biopsy procedure - choose the most sever type, if the report has both mild and moderate, always report moderate
	Margins		
		If WLE done, was there any dysplasia left behind (positive margins):	- Information to be captured from pathology report on oasis corresponding to the tissue obtained from the wide local excision or excisional biopsy procedure - or information could be captured from follow up form
		Positive margins were:	- choose the most sever dysplasia grade left behind - Information to be captured from pathology report on oasis corresponding to the tissue obtained from the wide local excision or excisional biopsy procedure - or information could be captured from follow up form
		If positive margins, was re excision done?	- information to be captured from follow up form and/ or procedure notes and /or OR notes in patient's chart on medesync
		Date:	- information to be captured from follow up form and/ or procedure notes and /or OR notes in patient's chart on medesync
- Above guide is applied to lesion 2, 3, etc.			

Follow up visits' details			
		Date of follow up visit	As documented on follow up form
	Tobacco use		
		Did smoking habit pattern change?	Yes or no or unknown as reported in the clinical follow up form
		Specify status	- This field should be addressed wither the answer to the previous field is yes or no or unknown. - As reported in the clinical follow up form pertaining to the visit of interest
		Type of tobacco being used	- Choose all that apply - As recorded in the clinical follow up form
	Cigar		
		Date cigar smoking habit started / changed	- Document if reported - As recorded in the clinical follow up form
		Year cigar habit started / changed	- As recorded in the clinical follow up form - Document if reported - 999 if unknown
		How long have the patient been practicing this pattern of cigar smoking	- As recorded in the clinical follow up form - Duration in months - 999 if unknown
		Number of cigars smoked …	- As reported in the clinical follow up form - 999 if unknown
		Pack years	Calculated field
	Pipe, cigarettes, waterpipe	Same fields apply	
	Alcohol use		
		Change in alcohol use hx	Yes or no or unknown as reported in the clinical follow up form
		If change…, Specify:	- This field should be addressed wither the answer to the previous field is yes or no or unknown.

			- As reported in the clinical follow up form pertaining to the visit of interest
		Type of user	As reported in the follow up clinical form
		Select type of drinks	- Choose all that apply - As recorded in the clinical follow up form
	Beer		
		How long have the patient….	- Duration in months - As recorded in the clinical follow up form
		Number of drinks in the above specified period	- As recorded in the clinical follow up form
	Wine, unknown type, hard liquor	Same fields apply	
	Clinical information		
		Are there any oral lesions present?	- Yes: there are oral lesions present as reported in clinical note - No: there isn't oral lesions present as reported in clinical note - Unknown: it is not mentioned in clinical notes if oral lesions are present or not.
		If yes:	- Lesion has been followed up for a while: as reported in the clinical follow up form - New lesion: as reported in the clinical follow up form - Both: as reported in the clinical follow up form
	Progression		
		Is there a clinical change (progression) warranting biopsy or re biopsy for any of the lesions	- Was there a re-biopsy done for any of the lesions being followed up. - Yes: as reported in the clinical follow up form

		being followed up?	- No: as reported in the clinical follow up form
		Did any of the patient's OPL progressed into a higher-grade dysplasia or SCC? (as mentioned in histopathology report)	- Yes: one or more lesions which have been followed up in the clinic have progressed into higher-grade dysplasia or SCC - No: none of the lesions which have been followed up in the clinic have progressed into higher-grade dysplasia or SCC. - Unknown: progression status is unknown
		How many lesions Progressed to higher grade dysplasia or Squamous Cell Carcinoma (SCC) or from hyperkeratosis to dysplasia	Choose from drop down menu how many lesions progressed so far
	Lesion one		
		Lesion progressed to	- Always choose the most sever. Ex : if sever dysplasia and CIS are reported choose CIS
		P16 status	As reported in the histopathology report.
		HPV status	As reported in the histopathology report.
		Date of progression as documented in the new biopsy report:	- Date must be taken from the histopathology report reporting progression
	Lesion one: clinical presentation		
		Presentation	- Symptomatic: symptoms like pain or discomfort are reported by the patient

			- Asymptomatic: patient didn't report any symptoms - Unknown: no information in the chart about patient's symptoms
		If symptomatic	- Choose all that apply, as reported in the clinical follow up form for the lesion of interest.
		If other	- Other symptoms not mentioned above
		Site	- Choose site (location in the oral cavity)
		Size	Should be documented in mm, if reported otherwise in clinical note it should be converted.
		Size	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - multiple choice question, choices are in mm2, if reported otherwise in clinical note it should be converted.
		Homogenous	Yes or no or unknown, As reported in the clinical follow up form for the lesion of interest.
		Ulceration	Yes or no or unknown, As reported in the clinical follow up form for the lesion of interest.
		Texture	Choose all that apply, as reported in the clinical follow up form for the lesion of interest.
		Duration	Should be reported in weeks, if reported otherwise should be converted.
		Color	- As reported in the clinical follow up form for the lesion of interest.
	Lesion one: Treatment		

		Treatment	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - Follow up: no surgical or medical treatment was provided. Only regular follow up. - Surgical excision: also called wide local excision - Laser ablation: lesion removed using laser - Unknown: treatment not reported in the patient chart
		Date of treatment	Date on which the surgical treatment was done (excisional biopsy or wide local excision).
	Lesion 2	Same fields are repeated for lesions 2,3,4,5.	
	New lesion		
		Did the patient develop SCC in other oral sites that didn't have OPLs before?	- information obtained from oncology consult - Yes: patient developed SCC in an oral location that didn’t harbor an OPMD before. - No
		Patient developed a new pre-malignant lesion	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - Yes: patient developed an oral lesion that is considered premalignant that wasn’t present at the latest previous follow up.
	New lesion: clinical presentation and details		
		Date of diagnosis	- As reported in the clinical follow up form for the visit

			of interest, for the lesion of interest. - Date of clinical diagnosis, the date the lesion first noticed clinically
		Presentation	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - Symptomatic: symptoms like pain or discomfort reported by the patient - Asymptomatic: patient didn't report any symptoms - Unknown: no information in the chart about patient's symptoms
		Size	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - Should be documented in mm2, if reported otherwise in clinical note it should be converted.
		Size	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - multiple choice question, choices are in mm2, if reported otherwise in clinical note it should be converted.
		Duration	- As reported in the clinical follow up form for the visit of interest, for the lesion of interest. - Should be reported in weeks, if reported otherwise should be converted.
	New lesion: Biopsy and diagnosis		

		If biopsy, histopathological diagnosis:	- Diagnosis as reported in the histopathology report pertaining to the biopsy procedure for the lesion of interest. - Always choose dysplasia if present in the histopathology report. - Actinic chelitis is the same as actinic keratosis.
		If dysplasia, specify:	Always choose the most sever grade reported.
		Is there a second diagnosis in the histopathology report? ex: Actinic keratosis and dysplasia	- report any other diagnosis mentioned in the histopathology report. - Always other than dysplasia; dysplasia is always captured above. for example: if both lichen planus and dysplasia are mentioned in the report, lichen planus is chosen here,
		If yes, Date of biopsy	- As reported in the histopathology report under collection date
		P16 status	- As reported in the histopathology report.
		HPV status	As reported in the histopathology report.
		If no biopsy, specify:	Clinical diagnosis as reported in the clinical notes.
	Treatment		
		Treatment	- As documented in the clinical note. - Follow up: no surgical or medical treatment was provided. Only regular follow up.
		Date of treatment	Date on which the surgical procedure (excisional biopsy or wide local excision) was done.

		If Surgical excision, histopathological diagnosis:	- As reported in the histopathology report pertaining to the procedure of interest for the lesion of interest. - always choose dysplasia if present in the histopathology report
	Status		
		Current follow up status	- Active: patient is still following up in the dentistry and Oral and Maxillofacial surgery department. - Loss to follow up: patient stopped following up in the department - discharged: if it is mentioned in the follow up form
		Disease status	- Disease free: no clinically apparent disease (no lesion in the oral cavity) - Stable: oral presentation hasn't changed since last visit - Progressing into a higher grade of dysplasia: ex: progressing from mild dysplasia to moderate proven by biopsy. - Progressing into SCC: ex: progressing from dysplasia to SCC proven by biopsy. - New premalignant lesion: patient presenting with a new lesion at the time of follow up visit.
		If SCC, indicate stage:	- Information to be captured from pathology report on oasis corresponding to the tissue obtained from the wide local excision or resection procedure - or information could be captured from the new oncology consult

		Follow up visit in months	- Total period of follow up in months up until the date of the follow up visit. - follow up periods starts to be calculated after surgical excision if surgical excision was used as treatment
		time elapsed since last surgical treatment in Months (if applicable):	- time elapsed from latest surgical procedure done (WLE or excisional biopsy)