Oral Health and Head and Neck Cancer Associations: Addressing Confounding through Negative Control and Quantitative Bias Analyses

Praveen Kumar Elango

Faculty of Dental Medicine and Oral Health Sciences

McGill University, Montreal

December 2024

A thesis submitted to McGill University in partial fulfilment of the requirements

Of the degree of Master of Science.

DEDICATION

To my family and friends, for always believing in me.

ACKNOWLEDGEMENTS

I am deeply grateful to the many people whose support, guidance, and encouragement have been instrumental in the completion of this thesis.

First and foremost, I express my deepest gratitude and appreciation to my supervisor, Dr. Belinda Nicolau. Her mentorship has been instrumental in shaping both this work and my development as a researcher. Her insightful guidance, unwavering support, and belief in my potential provided the foundation I needed to navigate the complexities of this academic journey. I am truly privileged to have learned from her expertise and wisdom.

I am also immensely thankful to my co-supervisor, Dr. Sreenath Madathil whose steadfast support and invaluable advice have been central to the success of this project. His dedication and generosity in sharing his knowledge have made a significant impact on my research and personal growth.

I would also like to extend my sincere appreciation to my committee members, Dr. Marie-Claude Rousseau and Dr. Audrey Grant. Their constructive feedback, thorough reviews, and thoughtful suggestions have refined and strengthened this thesis. I am grateful for the time and effort they invested in helping me bring this work to fruition.

This journey would not have been possible without the love and support of my family, cheering me on from the other side of the world and my friends with their daily dose of support and comfort. Their constant encouragement and motivation have sustained me throughout this process. I would also like to thank the Faculty of Dental Medicine and Oral Health Sciences, McGill University for the opportunity and the Graduate Dentistry team for their efforts and timely action.

Finally, I thank God for the strength, perseverance, and grace that have guided me throughout this endeavour. I am forever indebted to you all. Thank you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

HNC - Head and neck cancers

OPC - Oro-pharyngeal cancers

HPV - Human Papilloma Virus

SCC - Squamous cell carcinoma

HNSCC- Head and neck squamous cell carcinoma

OC - Oral cancer

EBV - Epstein - Barr virus

SEP- Socioeconomic position

PSA - Probabilistic sensitivity analysis

OR - Odds ratio

AOR - Adjusted Odds ratio

CI - Confidence intervals

PCR - Polymerase chain reaction

HeNCe - Head and neck cancer life study

DAG - Directed acyclic graph

ICD - International Classification of Diseases

HR-HPV - High risk Human papilloma virus

LR-HPV - Low risk Human Papilloma virus

STD – Sexually transmitted diseases

LIST OF TABLES AND FIGURES

ABSTRACT

Background: In the past two decades, there is a growing interest in the relationship between oral health indicators and head and neck cancers (HNC). Even though there is a strong biological plausibility to support these associations, the literature remains controversial, with some studies reporting strong positive associations, while others report no, or negative associations. Some authors contend that these links are influenced by underlying mediators, unmeasured risk factors, errors, and biases in the data, leading to these discrepancies. The use of negative controls, an epidemiological tool routinely used to detect bias in observational studies, offers an opportunity to clarify this issue. Although negative controls help identify and distinguish spurious associations from true ones, they have not been routinely used in epidemiological studies in oral health research.

Objective: To estimate the extent to which the association between oral health indicators and HNC risk is due to unmeasured confounders using data from the HeNCe life study. Additionally, we plan to estimate the magnitude of non-differential misclassification bias due to exposure using probabilistic sensitivity analysis (PSA).

Methods: The data for this investigation come from the HeNCe life study - Canadian site. This hospital-based case-control study recruited incident cases of HNC (n=389) frequency matched to controls (n=429) by sex and age within five years from four major referral hospitals in Montreal, Canada. In-person interviews using life-grid-based questionnaires collected information on a wide array of life course exposures. Oral rinse and oral brush specimens were analysed for HPV positivity and genotyping. The main exposure variables (oral health indicators) included self-reported number of missing teeth, denture use, and mouthwash use. We estimated odds ratios (OR) and 95% confidence intervals (CI) for the associations between oral health indicators and HNC using unconditional logistic regression models. The negative control exposure, sexually transmitted diseases (STD), was selected based on whether any of the following diseases—syphilis, gonorrhea, chlamydia, and herpes—were present or absent. Current literature does not provide substantial evidence connecting these diseases to head and neck cancer (HNC). We used unconditional logistic regression models and OR and 95% CI to estimate the association between STD and HNC risk, assuming a null hypothesis. Probabilistic sensitivity analysis (PSA) was done to estimate the magnitude and direction of misclassification bias. Corrected estimates were obtained by using priors selected from previous validation studies of self-reported oral health measures.

Results: The use of complete dentures and having more than 9 missing teeth suggested an increase in HNC risk [OR= 1.33, 95%CI (0.93-1.90) & OR=1.31, 95%CI (0.93-1.83)], respectively. Similar results were obtained when stratified by HNC subsite. Negative control analysis yielded a null finding indicating no detectable presence of bias due to unmeasured confounders. Bias-corrected estimates of the association between oral health indicators and HNC risk for the predicted sensitivity and specificity values showed stronger associations that shifted further from the null.

Conclusion: Our findings suggest that the associations between oral health indicators and HNC risk observed in the existing literature are potentially true and are not influenced by unmeasured confounders or biases. PSA findings also yield corrected estimates with stronger magnitude suggesting that the associations were underestimated in the crude analysis.

RÉSUMÉ

Contexte: Au cours des deux dernières décennies, la relation entre les indicateurs de santé bucco-dentaire et les cancers de la tête et du cou a suscité un intérêt croissant. Bien qu'il existe une forte plausibilité biologique à l'appui de ces associations, la littérature reste controversée, certaines études faisant état de fortes associations positives, tandis que d'autres ne font état d'aucune association ou d'associations négatives. Certains auteurs soutiennent que ces liens sont influencés par des médiateurs sous-jacents, des facteurs de risque non mesurés, des erreurs et des biais dans les données, ce qui explique ces divergences. L'utilisation de contrôles négatifs, un outil épidémiologique couramment utilisé pour détecter les biais dans les études d'observation, offre la possibilité de clarifier cette question. Bien que les contrôles négatifs permettent d'identifier et de distinguer les associations fallacieuses des associations réelles, ils n'ont pas été systématiquement utilisés dans les études épidémiologiques de la recherche sur la santé bucco-dentaire.

Objectif: Estimer dans quelle mesure l'association entre les indicateurs de santé bucco-dentaire et le risque de cancer du col de l'utérus est due à des facteurs de confusion non mesurés en utilisant les données de l'étude HeNCe life. En outre, nous prévoyons d'estimer l'ampleur du biais de mauvaise classification non différentielle dû à l'exposition en utilisant une analyse de sensibilité probabiliste (ASP).

Méthodes: Les données de cette enquête proviennent de l'étude HeNCe life study - site canadien. Cette étude cas-témoins en milieu hospitalier a recruté des cas incidents d'HNC (n=389) appariés par fréquence à des témoins (n=429) selon le sexe et l'âge dans un délai de cinq ans dans quatre grands hôpitaux de référence de Montréal, au Canada. Des entretiens en personne à l'aide de questionnaires basés sur la grille de vie ont permis de recueillir des informations sur un large éventail d'expositions au cours de la vie. Les échantillons de rinçage et de brossage buccaux ont été analysés pour déterminer la positivité et le génotype du VPH. Les principales variables d'exposition (indicateurs de santé bucco-dentaire) comprenaient le nombre déclaré de dents manquantes, l'utilisation de prothèses dentaires et l'utilisation de bains de bouche. Nous avons estimé les rapports de cotes (RC) et les intervalles de confiance à 95 % (IC) pour les associations entre les indicateurs de santé bucco-dentaire et les cancers du col de l'utérus à l'aide de modèles de régression logistique inconditionnelle. L'exposition de contrôle négative, les maladies sexuellement transmissibles (MST), a été sélectionnée en fonction de la présence ou de l'absence de l'une des maladies suivantes : syphilis, gonorrhée, chlamydia et

herpès. La littérature actuelle ne fournit pas de preuves substantielles du lien entre ces maladies et le cancer de la tête et du cou (CTC). Nous avons utilisé des modèles de régression logistique inconditionnelle et des OR et IC à 95 % pour estimer l'association entre les MST et le risque de cancer de la tête et du cou, en supposant une hypothèse nulle. Une analyse de sensibilité probabiliste a été réalisée pour estimer l'ampleur et la direction du biais de mauvaise classification. Les estimations corrigées ont été obtenues en utilisant des prieurs sélectionnés à partir d'études de validation antérieures de mesures de santé bucco-dentaire autodéclarées.

Résultats: L'utilisation de prothèses complètes et le fait d'avoir plus de 9 dents manquantes suggèrent une augmentation du risque d'HNC [OR= 1.33, 95%CI (0.93-1.90) & OR=1.31, 95%CI (0.93-1.83)], respectivement. Des résultats similaires ont été obtenus lors de la stratification par sous-site HNC. L'analyse de contrôle négatif a abouti à un résultat nul, indiquant qu'il n'y avait pas de biais détectable dû à des facteurs de confusion non mesurés. Les estimations corrigées du biais de l'association entre les indicateurs de santé bucco-dentaire et le risque de cancer du col de l'utérus pour les valeurs de sensibilité et de spécificité prédites ont montré des associations plus fortes qui s'éloignaient de la valeur nulle.

Conclusion: Nos résultats suggèrent que les associations entre les indicateurs de santé bucco-dentaire et le risque de cancer du col de l'utérus observées dans la littérature existante sont potentiellement vraies et ne sont pas influencées par des facteurs de confusion ou des biais non mesurés. Les résultats de l'EPS donnent également des estimations corrigées d'une plus grande ampleur, ce qui suggère que les associations ont été sous-estimées dans l'analyse brute.

CONTRIBUTION OF AUTHORS

Praveen Kumar Elango, MSc. Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, Quebec, Canada: Conceptualized study, performed the statistical analysis, interpreted findings, and wrote the manuscript.

Belinda Nicolau, Professor, Faculty of Dental Medicine and Oral Health Sciences, McGill

University, Montreal, Quebec, Canada: HeNCe Life study principal investigator, supervised the candidate, contributed to the conceptualization of the study, assisted with the interpretation of findings, and reviewed the manuscript.

Marie-Claude Rousseau, Professor, Epidemiology and Biostatistics Unit, Centre Armand Frappier Santé Biotechnologie, Institut national de la recherche scientifique, Laval, Canada:

Member of the candidate's thesis committee: reviewed the conceptual ideas and the manuscript.

Audrey Grant, Assistant Professor, Department of Anesthesia, Faculty of Medicine and Health Sciences, Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, Canada. Member of the candidate's thesis committee: reviewed the conceptual ideas and the manuscript.

Sreenath Madathil, Assistant Professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, Quebec, Canada: Co-supervised the candidate, contributed to the conceptualization of the study, oversaw the statistical analysis, assisted with the interpretation of findings, and reviewed the manuscript.

1 INTRODUCTION

Head and neck cancers (HNC) are the seventh most prevalent cancer globally. They have a complex multi-factorial etiopathogenesis and include a variety of tumours, affecting the upper aerodigestive tract including the mouth, throat, and larynx (1). Squamous cell head and neck carcinomas, originating from the epithelial lining of the oral cavity, pharynx, and larynx, are the most common histological type of HNC, accounting for about 90% of cases (1, 2). Recent trends indicate an increasing incidence of HNC in Canada, with 7,900 cases and 2,100 estimated deaths in 2023 (3). Global 5-year survival rates for HNC remain around 50-60% (4), while in Canada, the 5-year net survival rate is 64% (3). HNC survival depends on the stage at diagnosis, with early-stage detection significantly improving prognosis and resulting in lower functional impairment (4).

HNC are associated with significant functional morbidity and loss of quality of life due to the anatomical proximity of the cancerous lesions to vital structures (5). HNC patients experienced a higher psychological burden compared to other types of cancers, which explains the fact that these cancers seem to have a high suicide rate (6). The public health burden associated with these cancers is significant, especially in low-income and middle-income countries, which contributes to a higher percentage of mortality and morbidity rates (7). HNC are also one of the most expensive solid tumours to treat and have a high recurrence rate owing to their invasive nature (8).

Decades of research have identified key risk factors in the etiopathogenesis of HNC, leading to improvements in treatment and prevention (9). However, these risk factors alone cannot fully explain the variations in HNC risk. The primary HNC risk factors are tobacco and alcohol consumption and, for a subset of these cancers, human papillomavirus (HPV) infection (10). Additionally, the consumption of fruits and vegetables seems to have a protective effect against HNC (11). Moreover, low socioeconomic status, obesity, and poor oral health have been identified as potential risk factors (12, 13).

The relationship between oral health status (e.g., number of teeth, denture use, periodontal diseases) and behaviours (e.g., toothbrush frequency) and HNC has gathered significant attention in recent years (13). However, the results from studies investigating these associations have yielded inconsistent and contradicting findings (13, 14).

The reasons may be due to inconsistencies in the selection of oral health variables and different methods of measurement employed in various studies. While some researchers claim the associations are true and potentially causal, others believe these are spurious associations due to underlying unknown factors, mediators, and bias in the data.

Negative controls are validation tools commonly used to identify and differentiate non-causal associations from causal ones by detecting unmeasured confounding as well as other sources of bias (15). While their routine use in observational studies is increasing, the application of these techniques in oral epidemiology is limited.

There are still knowledge gaps in understanding how oral health affects HNC risk, especially regarding the nature of these associations. This thesis attempts to enhance our understanding of this link and aims to assess the association by using epidemiological validation tools. Using data from a hospital-based case-control study conducted in Montreal, we investigated the association between oral health and HNC risk, tested for unmeasured confounders, and performed a bias analysis to validate the findings. This work could pave the way for future studies in this domain and potentially have a significant public health impact.

2 LITERATURE REVIEW

2.1 Definition of HNC

HNC is a term that generally includes malignant tumours, mainly originating from the epithelium of the oral cavity, pharynx, larynx, paranasal sinuses, and nasal cavity, collectively referring to malignancies within these regions (10). The majority of cancers occurring in these areas are histopathologically categorised as squamous cell carcinomas. Based on the International Code for Disease Classification by the World Health Organization, the anatomical subsites categorized as HNC include ICD-10 codes C00-C14, C30-C32, and C76 (16). An outline of HNC epidemiology and a synopsis of its primary established risk factors and potentially implicated risk factors are described in this section.

2.2 Epidemiology of HNC

HNC are the seventh most common type of cancer, with a global incidence of around 660,000 cases and 325,000 deaths per year (17). According to the 2023 Canadian Cancer Society report, 7,900 incident HNC cases were identified across Canada, with an estimated death count of 2,100. There is a significant variation in the geographic distribution, with cases being more prevalent in Canada's eastern and central regions. Males are more prone to develop HNC, accounting for 5,800 cases in 2023, compared to 2,100 cases among females (3). HNC contribute to about 5.3% of the global cancer burden (18). Oral cavity cancers, including those of lips, contribute to roughly half of the total global incidence of HNC, while oropharyngeal (OPC) and laryngeal cancers collectively account for about 25% of the total HNC cases (2).

Global trends in HNC also show variation according to geographic distribution, with an increased incidence of oral cancers in Southeast Asia and the Asia-Pacific regions, especially due to the behaviour of chewing areca nut (betel quid), with or without tobacco (19). Moreover, recent global trends reveal geographic variation in HNC according to HPV positivity, with a higher prevalence of HPV-positive HNC in high-income countries and HPV-negative HNC being more common in low-income countries (20). Indeed, there has been a notable rise in HNC incidence rates over the past few decades, primarily driven by OPC, the most common site affected by HPV (21). Estimates suggest a 30% increase in the incidence of HNC by 2030 (17).

Despite new developments and advancements in the various treatment modalities, the 5-year survival rates for HNC remain low (22). Globally, the 5-year survival rates for oral and oropharyngeal cancers is approximately 50%, with the hypopharynx being the site with the poorest survival rates and prognosis owing to inaccessibility to the tumour (23). TNM staging (24) at the time of diagnosis impacts HNC 5-year survival rates, and the prognosis typically worsens with advanced stages of the disease. Importantly, over 60% of patients are diagnosed with locally advanced lesions (20). Significant differences in mortality-to-incidence ratios exist across different geographic regions, contributing to unequal economic and health-related burdens (25).

2.3 Risk factors of HNC

The etiological landscape of HNC is dynamic and, with varying risk factors contributing to the steadily increasing incidence. HNC is influenced by diverse genetic and environmental factors (26, 27). The main risk factors for HNC are smoking and alcohol and their joint effects (28, 29), accounting for about 75% of the aetiological burden of HNC (22). HPV is another established risk factor that in recent decades has been implicated in HNC aetiology, especially those occurring in the oropharynx subsite (22). While several studies have shown independent associations between HNC and the established risk factors, namely alcohol intake, tobacco, and HPV, others have explored a possible combined effect due to the interplay of three risk factors (30).

The aetiological landscape of HNC is dynamic, with varying risk factors contributing to its steadily increasing incidence over the decades. HNC is influenced by a combination of genetic and environmental factors (26, 27). The primary risk factors for HNC are smoking and alcohol, and their combined effects (28, 29), which account for approximately 75% of the aetiological burden of HNC (22). HPV has also emerged as an important risk factor in recent decades, particularly in the aetiology of OPC (22). While several studies have demonstrated independent associations between HNC and key risk factors—namely alcohol intake, tobacco use, and HPV—others have explored the potential combined effect of these three risk factors (30).

Besides the established risk factors, numerous potential risk factors have been involved in HNC pathogenesis, yielding positive estimates that imply a causal link (11, 19). Some of these risk factors, influenced by age and geographic region, may also contribute to variations in HNC incidence rates (17). Understanding the multi-dimensional nature of these risk factors helps researchers, healthcare professionals, and policymakers develop more effective strategies for

preventing, screening for, and treating HNC, ultimately improving health-related outcomes in affected populations.

2.3.1 Socio-demographic risk factors

It is well established that HNC are more predominant in males compared to females, mostly owing to higher levels of alcohol and tobacco consumption (31, 32). However, some studies indicate an increasing HNC trend in females, attributed to the adoption of male-like patterns of alcohol and tobacco consumption by women (33). In contrast, other studies have highlighted a strong male predilection in never-smokers and non-drinkers (31). HNC risk increases with age, with higher incidence observed in older age groups. HNC usually present in the later decades of life, with most patients diagnosed with late-stage disease in their fifth and sixth decades of life (34, 35).

Low socioeconomic status is generally considered a risk factor for HNC. The reasons for this are complex, but there is substantial evidence showing that smoking and alcohol habits are more prevalent in this social stratum. (36). Similarly, occupational exposures to carcinogenic agents like asbestos occur more frequently in certain jobs, making occupations like cleaners, welders, and mineworkers more prone to developing HNC (37).

2.3.2 Tobacco Consumption and HNC

Tobacco consumption is one of the strongest risk factors for HNC and plays a significant role in their development, despite changes in the patterns of HNC aetiology. Tobacco can be consumed in smokeless forms and smoking forms, the usage of the latter being more common. Tobacco consumption in the form of smoking habits such as cigarettes, bidis, pipes, and cigars has been a major contributor to the steadily increasing incidence of HNC (38). The dose-dependent relationship between tobacco exposure and HNC risk has also been studied in both HPV-negative and HPV-positive groups, showing a directly proportional effect in both cases (39). Smoking is strongly associated with all specific anatomic sites, including the oropharynx, larynx, and oral cavity, with strong associations that have consistently been replicated in several studies (40, 41).

Cigar and pipe smoking are independent risk factors of HNC, even in the absence of cigarette smoking (38). Smoking prevalence varies greatly according to geographic region, with higher rates among women in developed countries compared to those in low-income countries (42).

Tobacco smoking seems to have a greater impact on certain subsites of HNC, particularly oropharyngeal and laryngeal cancer, which have shown strong associations in past studies. In contrast, the oral cavity subsite has demonstrated comparatively weaker associations (43), suggesting other underlying risk factors may influence these sites. This variation is due to differences in exposure susceptibility to various risk factors across different anatomic subsites (40).

Studies have shown that smoking cessation post-diagnosis has resulted in better survival rates and prognosis after treatment (44). It is well established that quitting tobacco use reduces the risk of developing HNC, with past smokers having a significantly lower risk compared to current smokers (43).

2.3.3 Alcohol consumption and HNC

Besides smoking, alcohol is one of the most established risk factors for HNC (31, 45). Alcohol and smoking have a multiplicative effect, significantly increasing HNC risk in individuals who engage in both habits (28, 29). Alcohol intake is more strongly associated with oral cancers than with other HNC subsites such as the oropharynx and larynx (45). HNC risk due to alcohol intake depends on a variety of factors such as frequency, type of alcoholic beverage, quantity, and intensity. Increased risk is observed across all subsites with heavy drinking (>3 drinks per week) and consumption of spirits with higher alcohol concentration (46) (e.g., whisky or brandy containing around 40% alcohol). Alcohol's solubility increases the permeability of the oral mucosal membrane, making it more susceptible to carcinogens like nitrites and other chemicals in smoke. While the exact molecular pathways and mechanisms are not fully understood, acetaldehyde, a by-product of ethanol, can irreversibly damage DNA, inhibit DNA synthesis and repair, and disrupt DNA methylation (47). Other mechanisms include lipid peroxidation due to oxidative stress and inflammation-induced DNA damage (47).

2.3.4 Human papilloma virus (HPV) and HNC

HPV collectively comprises a family of circular, double-stranded DNA that has been established as an aetiological agent in around 5% of all cancers globally, mainly causing cervical, vaginal, penile, and oropharyngeal cancers (48). The HPV virus encodes three oncogenes (E5, E6, E7), two regulatory (E1 and E2), and two structural proteins

(L1 and L2). HPVs are phylogenetically classified into genera, species, and over 200 different genotypes. The three main HPV genera are alpha papillomavirus (α), beta papillomavirus (β),

and gamma papillomavirus (γ). The α-HPV contains several HPV types that can infect both the oral and genital mucosal epithelium (48). HPV-driven OPC has significantly contributed to the recent increase in the incidence of HNC (19, 22, 49, 50, 55, 56). Based on oncological potential, around 18 α-HPV types such as HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82 are categorised as high risk- HPV (HR-HPV) and the rest are categorised as low risk- HPV (LR-HPV) (51). Several α-HPV types, such as HPV16, classified as HR-HPV, have been confirmed to be major HNC risk factors (52, 53). Other HR-HPV types, such as HPV18, HPV33, and HPV52, have also been implicated in HNC development (54).

HPV-driven HNC represent a unique subtype with distinct molecular, clinical, and pathologic characteristics contrasting to HPV-negative HNC (56). As previously mentioned, HPV-related HNC are more common in the oropharynx, especially in the tonsils. (21, 56). They also seem to be poorly differentiated histologically and have basaloid morphology (20). These cancers are more prevalent in young adults from high-income countries (19). They usually have a better prognosis owing to different treatment strategies and targeted therapies compared to HPV-negative HNC (20). They are more prevalent in young adults who have initiated sexual activity at a young age, those who had a higher number of sexual partners and those who reported having oral sex more often (57). While some studies (58), suggest that patients with HPV-related HNC are less likely to have a strong tobacco and alcohol use history, others have shown these habits interact and increase HNC risk (30).

Furthermore, evidence suggests an increase in HPV-related subsites and a decrease in HPV-unrelated subsites, particularly in developed countries such as the USA, Canada, and the UK (21).

2.3.5 Chewing habits and HNC

A significant portion of the regional variations in HNC incidence, particularly subsite-specific incidence, can be explained by geographically limited habits such as betel nut chewing, which is common in South and Southeast Asian countries (59). Betel nut is an established environmental carcinogen with potent genotoxic effects, contributing to the etiopathogenesis of HNC by altering molecular mechanisms and causing DNA modifications (59-61). Betel nut chewing is strongly associated with pre-malignant lesions such as oral leukoplakia and oral submucous fibrosis, which can eventually transform into oral cancers (60). The most common anatomical subsite of HNC affected by betel nut chewing are various sites within the oral cavity, including the buccal mucosa, tongue, and gingiva. (61)

2.3.6 Diet and HNC

Although diet has been linked to HNC risk, the evidence is unclear due to the ambiguous definition of dietary constituents. The most consistent findings suggest a protective role for a diet rich in whole grains and higher non-starchy vegetable and fruit intake on HNC risk (62). Diet quality has also been assessed using diet quality indices by studies (11, 63) concluding that a poor-quality diet increased HNC incidence, especially in alcohol drinkers.

2.3.7 Obesity and HNC

Obesity is another potential emerging driver of HNC and a concerning public health issue. It is linked to the etiopathogenesis of HNC through molecular mechanism changes that lead to abnormal lipid metabolism (64). Surprisingly certain investigations, including both cohort and case-control studies, have also reported an inverse correlation between BMI and the occurrence of HNC (65-67). However, these findings remain inconclusive due to a lack of consistency and knowledge regarding molecular pathways and mechanisms that could potentially explain this link, as well as the possibility of reverse causation. In other words, people may lose weight because of cancer, thereby skewing the association (68)

2.3.8 Occupational Exposure and HNC

Certain occupational workers such as welders, painters, and cleaners are exposed to various carcinogenic agents due to inhalation of toxic fumes in their line of work implicated in the pathogenesis of HNC (37). A duration-response relationship with increasing risk in longer durations of exposure to harmful fumes or agents has also been noted in certain occupations (37). Studies also reported elevated risks for butchers, meat preparers, and bartenders, which could be explained by exposure to viral agents, nitrosamines, or polycyclic aromatic hydrocarbons (69). Certain occupations, such as factory workers and heavy machinery operators, make people exposed to smoke, which gets metabolized by the body, leading to DNA changes and epigenetic modifications (37, 69).

2.3.9 Oral health and HNC

Oral health has been increasingly linked to HNC aetiology, especially in the last two decades. Multiple studies involving various oral health indicators such as denture use, dental visits, gum bleeding, and mouthwash use have shown significant positive associations with HNC risk (14, 70, 71). Poor oral health status at the time of HNC diagnosis has also been found to further strengthen this association (72). Periodontal disease, a major putative risk factor for HNC, is the most frequently used measure in studies examining the oral health-HNC link (73). Recent epidemiologic evidence indicates higher HNC incidence in people with periodontal disease compared to people without periodontal disease (74).

This association has been explained by alterations in the oral microbiome and chronic inflammatory state in periodontitis, leading to dysregulation of the microbiome and immunological responses, promoting and growing cancerous cells (75).

As periodontal disease progresses, there is an increase in tooth mobility, which eventually leads to tooth loss. Therefore, missing teeth have been used as a proxy for periodontal disease, which has also been considered an oral health indicator in some studies. A smaller number of missing teeth, regular annual dental visits, no gum disease, and daily tooth brushing have been associated with HNC risk inversely indicating a protective effect (14, 71, 76).

Some studies have shown associations between complete denture use and HNC risk (76). Recurrent denture-related sores have also been associated with oral cancers. This could be explained by ill-fitting dentures causing chronic irritation and inflammation, leading to increased susceptibility to HNC (77).

Poor oral health has also been found to affect survival in HNC patients, exhibiting a direct proportionality to survival rates (78). A meta-analysis by Bai et al in 2023 investigating the association between oral hygiene and HNC risk concluded with positive associations for several oral health indicators (e.g., a greater number of missing teeth, gum bleeding, and periodontal disease) and oral cavity cancer risk. Conversely, mouthwash use, frequent dental visits, and tooth brushing twice daily had a protective effect (79). However, several authors have reported no associations between oral health indicators and HNC after adjusting for covariates such as smoking and alcohol intake (14). These contrasting findings may be due to significant variability in reporting measures for different oral health indicators, lack of consistency in selecting these indicators, and lack of appropriate control for confounders. Determining whether the associations observed in population studies are causal is crucial in

evaluating the role of oral health in cancer risk and setting oral health promotion and cancer prevention strategies (80).

2.4 Negative controls

Negative controls are defined as tools to detect confounding and bias in epidemiological studies, especially observational studies, which are prone to recall and selection biases. Such errors are mitigated due to standardisation and randomisation and do not threaten the integrity of causal inferences observed in experimental studies (15). The purpose of negative controls is to ascertain both potential and unconfirmed causes of incorrect causal inference, thereby validating the causality of those associations. Negative controls are routinely used in epidemiology but seldom seen in oral literature. These tools can also identify unknown confounders and resolve confounding issues in the associations.

There are two types of negative controls, exposure and outcome controls (15). These are categorised depending on their comparison with the confounding structures of the exposure or outcome of interest in the study.

2.4.1 Usage of Negative Control in Observational Studies

In epidemiological studies, risk factors and associations between exposure and outcomes are typically identified and estimated by adjusting for measured factors or variables through stratification, matching, multivariate modeling, and inverse probability weighting (81). Despite these stringent measures, bias due to unknown factors or underlying mediators can persist, leading to spurious associations. These spurious associations result in contested causality inferences and raise questions about the reliability of observational studies in establishing causal links. Using epidemiological tools such as negative controls can validate and strengthen the findings of observational studies if applied correctly and with adherence to proper selection criteria. Negative controls have been used to detect confounding, recall, and selection bias in various studies (82, 83).

2.4.2 Bias Analysis using Negative Controls

Negative control analyses are performed based on the assumption of null findings and can only detect the presence or absence of bias and its direction (15). It is mainly used to detect the presence of unmeasured confounder effects and selection bias. This epidemiological tool

cannot quantify or assess the magnitude of bias to help produce corrected estimates of the associations to infer causality.

2.5 Quantitative Bias Analysis

Quantitative bias analysis (QBA) focuses on identifying and estimating the direction and magnitude of bias and the uncertainty due to systematic errors (84). It identifies sources of errors and describes models to quantify them by assigning bias parameters to negate the human tendency towards overconfidence in research findings, results, syntheses, and critiques, as well as the inferences based on them (84).

QBA has been practiced for decades, and there have been calls for its widespread adaptation in epidemiological studies. Historically, the lack of availability of powerful computational tools and proper knowledge of QBA methods have been barriers to its use. However, these issues have been largely addressed in recent times. By assessing uncertainty in specific research domains and improving validity and accuracy through corrected results, QBA helps allocate resources to research domains appropriately. While QBA applications are still rare in epidemiological studies, they are definitely on the rise (85).

2.5.1 Types of Bias-

The three main issues in deriving causal inferences in epidemiology are: 1) bias 2) confounding 3) interaction. A bias is defined as "any systematic error in the design, conduct, or analysis of a study that results in a mistaken estimate of an exposure's effect on the risk of disease."(86)

There are many types of bias in epidemiological studies depending on their sources such as confounding bias, information bias, measurement errors, and selection bias.

2.5.1.1 Selection bias

Selection bias occurs when there is a difference in how cases and controls or exposed and unexposed individuals are selected for study participation. This might lead to spurious associations between exposure and disease due to selection bias, even if they are unrelated. For example, differences in response rates for participation among cases and controls can introduce bias. In general, people who respond to studies have different demographic, cultural, social and medical characteristics than those who do not (87). Selection bias can significantly affect and interfere with the internal validity of a study, leading to non-valid inferences about between associations of exposure and disease (86).

2.5.1.2 Information bias

Information bias can arise from inconsistencies and inaccuracies in collecting information or from inadequate methods of obtaining information. Some types of information bias include misclassification bias, surveillance bias, recall bias, and reporting bias (86). Misclassification bias occurs when there are inaccuracies in obtaining proper information such as exposure status, which can lead to an exposed participant being categorised as unexposed and vice versa. There are two types of misclassification bias: differential and non-differential. In differential misclassification, the rate of misclassification differs between cases and controls. For example, exposed individuals may be misclassified as unexposed more frequently than unexposed individuals are classified as exposed.

Recall bias is a type of information bias that occurs due to differences in recall accuracy, and mainly occurs in retrospective studies such as case-control studies. Participants may find it difficult to recall information about past exposures, with controls often having more difficulty than cases (88). Reporting bias occurs when subjects do not disclose known information about an exposure due to reluctance, beliefs, attitudes, and perceptions. This leads to underreporting and inaccuracies in the data (89).

Surveillance bias occurs when certain individuals in a population are observed more closely over a period of time such as national cohorts. For example, disease ascertainment in such a monitored population is better compared to the general population, introducing surveillance bias. This leads to over-reporting and overestimation of findings (90).

2.5.1.3 Confounding bias

Bias due to cofounders poses a major threat to the validity of findings from observational studies. A confounder is a factor (X) that is a known risk factor of outcome/disease (O) and is associated with exposure (E) but not the result of the exposure (E). This is considered a mixing of effects, where the effects of an exposure on an outcome are mixed with the effects of an additional factor, resulting in distorted associations (90). Figure 1 demonstrates a confounder and the association of interest.

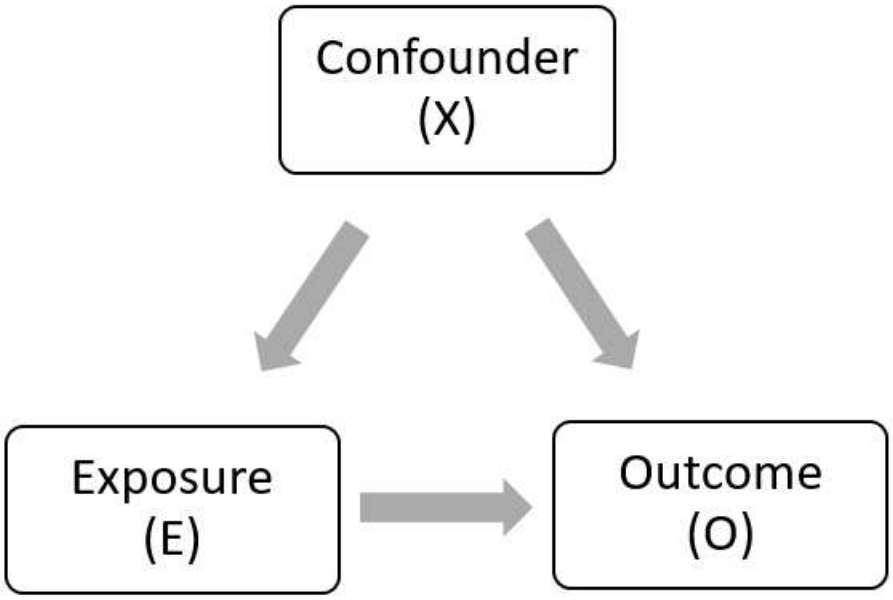


Fig-1 schematic representation of confounder (X) affecting the relationship between exposure (E) and outcome (O).

2.5.2 Probabilistic Sensitivity Analysis

Probabilistic bias analysis (PBS) involves assigning a distribution of values for the parameters instead of single deterministic values (92). In PSA, we assume that these values come from a known distribution, which could be based on values from previous literature, such as sensitivity and specificity values from validation studies. Specifying distributions for each probability can be a tedious task. Most PSA type of bias analysis is done by methods described by Fox et al. (84). Selecting and identifying distributions for parameters can be challenging as there might not always be sufficient findings from existing literature. The selection of distributions such as logit, or beta, depends mainly on the nature of the data, the type of model being used, and the research focus.

3 RATIONALE

More than 660,000 new HNC cases and 325,000 HNC-related deaths occur annually worldwide. HNC are the 7th most common cancer in the world (17), representing nearly 8% of male cancers, with 5-year survival rates below 60%. These low survival rates have not improved in the past 30 years despite advances in diagnosis and treatment. About three-quarters of HNC are associated with tobacco smoking, alcohol consumption, and, in a subset of HNC - oropharyngeal cancer, HPV infection.

Several studies have suggested an association between oral health status and behaviour and HNC. However, these associations remain highly contested despite extensive research. This is largely due to inconsistencies in the measurement of oral health indicators, knowledge gaps regarding the biological mechanisms underlying these associations, and the possibility that these associations may be due to the confounding effects of unknown factors and other biases.

From the public health point of view, it is, therefore, essential to determine if oral health and oral health behaviours are indeed risk factors for HNC so that preventive strategies such as oral health screening can be incorporated into routine cancer screening procedures to improve HNC detection at an early stage of the disease.

This study addresses this gap by testing and validating the association between oral health indicators and HNC risk using an epidemiological tool called negative controls and quantifying and correcting misclassification bias of the exposure potentially affecting this association, in a database from a hospital case-control study conducted in Montreal, Canada.

4 AIM AND OBJECTIVES

The overarching goal of this project is to investigate whether the associations between indicators of oral health and HNC risk are true or if they can be deemed spurious, potentially arising from underlying biases and mediation effects.

The specific objectives are:

1) To estimate the extent to which oral health indicators are associated with HNC using data from the HeNCe Life study, a hospital case-control study conducted in Montreal, Canada.

2) To test whether the observed associations are due to unmeasured confounders in the association between oral health indicators and HNC.

3) To estimate the extent to which the associations between oral health indicators and HNC are due to exposure misclassification bias.

5 METHODS

5.1 Study design

The data used for this study were obtained from the Canadian side of the Head and Neck Cancer Life Study (HeNCe Life study), an international hospital-based case-control study (93) investigating the aetiology of HNC using the life course epidemiology approach (94). The study was conducted in three countries: (i) 2005 to 2013 in Montréal, Canada; (ii) 2008 to 2012 in Kozhikode, India; and (iii) 2003 to 2005 in São Paulo, Brazil. All three countries used similar protocols tailored to the specific context of each country.

For this thesis work, only data from the Canadian site were utilized.

5.2 Eligibility criteria

1) Participants must be born in Canada and live within 50 km of the hospital area. This geographical restriction ensures that both the cases and controls represent the same secondary study base or catchment area. People born and living in different geographic regions tend to be exposed to different risk factors (environmental and genetic) and are less similar to the target population (95).

2) Participants must also be fluent in either French or English. The study collected data using questionnaire-based interviews, so proficiency in either one of the languages was essential to ensure the interviewer had proper rapport with the participants and improve the quality of the data collected.

3) Participants had to be at least 18 years old, HNC usually manifest in the later decades of life, most commonly the fifth and seventh decades of life (34, 35)

4) Participants should have no previous history of cancer. Prevalent cases might bias the risk estimates.

5) Participants with any previous history of mental, cognitive, or immunosuppressive disorders were excluded. This was done to ascertain the accuracy and quality of the retrospective data collected and to ensure effective and ease of communication in the interviews. People with immunosuppressive disorders such as HIV were excluded due to increased susceptibility to be diagnosed with HNC, thereby biasing the risk estimates (96).

5.3 Study population

A total sample of 918 participants, including 460 incident HNC cases and 458 controls, were recruited from the four major hospitals in Montreal as part of the HeNCe life study. We excluded 100 participants whose HPV status was not available either due to lack of HPV testing or ineligible samples. A total of 818 participants from the Canadian site were included for the analysis in this study.

5.4 Case and control definition and selection

Cases (n=389) were newly diagnosed squamous cell carcinomas at stages I to IV of the head and neck region involving the pharynx, larynx, and oral cavity. Only newly incident cases of HNC were included; prevalent cases might bias the risk estimates, and survivors may represent prevalent cases. Cases were confirmed by histological examination, the gold standard for diagnosing HNC. These included cancers of the oropharynx, hypopharynx, larynx, tongue, gums, floor of the mouth, and other sites in the oral cavity, corresponding to International Classification of Diseases (ICD-10) codes: [C00.3 – C06.9, C09, C10, C13, C14 and C32] (16). Cancers of the lip [C00], salivary glands [C7, C8] nasopharyngeal [C11], and oesophagus (C15) cancers were excluded due to different histology and aetiologies (97, 98).

Controls (n=429) were selected from several outpatient clinics (e.g., Nephrology, Gynaecology, Ophthalmology, Dentistry, Dermatology, and Family Medicine) at the same hospitals as the cases and were frequently matched to the cases by sex and age (5 years). Matching was employed for confounders that needed to be controlled for but were not of interest as independent risk factors (99).To mitigate potential biases, controls were carefully selected from a list of patients with non-chronic diseases unrelated to alcohol or tobacco use. To prevent the overrepresentation of any single diagnostic group, the contribution from each group was limited to 20% of the total controls.

5.5 Ethics approval and Informed consent

The HeNCe life study was approved by the institutional review boards (IRB) of McGill University, Institut National de la Recherche Scientifique (INRS), and the ethics boards of all participating hospitals where cases and controls were recruited. All participants were explained in detail about the information collected for medical research and the objectives of the study and signed consent forms were obtained. Data obtained were then anonymized and participants

were given nonrevealing unique identification numbers to conceal personal information and identity.

5.6 Study measurements

5.6.1 Interviews

Interviewers specifically trained for this study conducted face-to-face interviews using a structured questionnaire with the help of a life grid tool (100). This interactive interview (Appendix), designed to capture social and behavioural events across the individual's life course, is based on questions derived from previous studies such as the International Agency for Research on Cancer (101), British Civil Servants, Whitehall II Study (102), National Study of Health and Development British Birth Cohort 1946 (103), National Child Development Survey British Birth Cohort 1958 (104). The life grid used in this study (Appendix 1) was a modified adaptation of the life grid used in Blane's study (105).

The interview lasted for about 1.5 to 2 hours and collected detailed information on three time periods of participants' lives: childhood (birth to 16 years old), early adulthood (17-30 years old), and late adulthood (older than 30 years old). The questionnaire gathered data on participant demographics, detailed medical history, socioeconomic position indicators (e.g., education, occupation, living conditions, and other amenities), and behavioural characteristics of the participants and their parents e.g., smoking, alcohol intake; eating habits and dietary intake; overall health and physique; history of sexual practices; dental hygiene and oral health status including indicators such as dental caries, missing teeth, bleeding gums, mouthwash use and regular dental visits; familial cancer history; marital status; and social support.

The life grid technique (100) was used to collect retrospective data about the life course exposures, this is based on evidence from literature suggesting that when participants are provided with a temporal reference point it serves as a personal timeline and helps reduce recall bias (106). The interviewers provided the life grid to the participants at the start of the interview and enquired about important lifetime personal milestones such as birth or death of a family member, year of marriage, and year of divorce, among others. These events were marked down, and other personal and historical events were added to the life grid to serve as a timeline template to improve recall. For example, one might recall beginning to drink or smoke the year they lost a family member.

5.6.2 Biological Sample Collection

To perform genetic analyses and HPV genotyping, oral trans-epithelial cells were collected.

Epithelial samples are required for diagnosing HPV infections at a specific site, such as the oral cavity, and oropharynx, as they infect the basal layer of squamous epithelium most commonly. Although serum antibodies can be used as a reliable marker for lifetime HPV infections, antibodies to L1 oncoproteins can indicate cumulative exposure to all sites, and it is not site-specific (e.g., oral cavity). To detect signs of malignancy, serum antibodies to oncoproteins E6, and E7 can serve as markers (107). Epithelial samples can be obtained by collecting exfoliated epithelial cells from saliva, brush or mouthwash specimens, or tissue biopsies (108). Participants, especially controls, are more cooperative with less invasive methods such as brush or mouthwash samples.

Our study collected samples in three forms: mouth rinse specimens, brush-exfoliated cell specimens, and also tumour tissue samples obtained during the surgery. Brush and mouth rinse specimens were collected from all participants. Mouth rinse samples were collected by asking the patients to rinse their mouths for 15-30 seconds with an alcohol-based mouthwash solution and to spit the solution into a pre-labelled collection vial. For the brush-exfoliated samples, participants were given an OralCDx biopsy brush (109) and asked to apply 20-30 gentle strokes in different regions of the mouth. However, HNC cases were instructed to apply strokes in and around the tumor site, and brushes were pressed firmly (oral cavity lesions) until pinpoint bleeding occurred. After sample collection, brushes were swirled vigorously in a PreservCyt® buffer (110) bottle to transfer the epithelial cells to the solution and transported to the microbiology and immunology laboratory at the CHUM hospital in Montreal. Cell suspensions were centrifuged for 15 minutes at 22°C at 13,000 x g. The excess supernatant was disposed of the cell pellets were resuspended in 300 µL of 20 mmol/L Tris solution (pH 8.3), and the DNA was extracted utilizing the Epicentre MasterPureTM Kit (110). Extracted DNA was stored at -80° C until genotyping.

5.6.3 HPV DNA amplification and HPV genotyping

Currently, there are around 200 commercially available HPV assays in use for detecting HPV. Partial HPV typing focuses on higher sensitivity for detecting High-risk -HPV (e.g., HPV16, HPV18), whereas extended HPV typing requires higher analytical sensitivity to detect a large number of HPV genotypes in any sample. This is especially the case in HPV vaccine and HPV prevalence research.

HPV DNA detection and genotyping were done using a linear array. Beta globin testing (111) was done on 10 µl of extracted DNA for each sample in batches by performing PCR using PC04 and GH20 followed by agarose gel electrophoresis (112). This was done to ensure a sufficient number of epithelial cells for PCR analysis, DNA integrity, and the absence of inhibitors. Samples that were beta-globin negative were considered inadequate for genotyping. Following that, the beta-globin-positive DNA samples were amplified for HPV using PGMY09-PGMY11 primers (113).

Genotyping was done using reverse hybridisation and biotin-labelled probes for the following α-HPV genotypes: 6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73, 81, 82, 83, and 84 (114).

5.6.4 Quality control and data quality

Several measures were taken to ensure the highest data quality. Data collection using life-grid questionnaires required trained personnel. The training was given by the PI (Dr. B Nicolau) to all the interviewers through in-person visits, audio-visual aids, and a reference guide for interviewers, which was maintained at each recruitment site. In case of inconsistencies, phone calls were made to clarify the answers. Data collected were immediately transferred to the central database and scrutinized before analysis. Biological samples were stored appropriately at 4°C until analysed.

5.7 Statistical analysis

5.7.1 Selection of study variables

5.7.1.1 Outcome variable: HNC

The dependent variable or outcome variable is HNC, i.e., tumours of the upper aerodigestive tract, including the oral cavity, pharynx, and larynx. We used a histological diagnosis of stage I to IV squamous cell carcinomas of the oral cavity, oropharynx, hypopharynx, and larynx for confirmation to identify and confirm cases (115). Physicians use histological tests as a gold standard to diagnose oral malignancies and are therefore very reliable (116). It is a binomial variable indicating the presence or absence of HNC.

5.7.1.2 Main exposure variable: Oral Health

The exposure of interest variable in this study was oral health status, for which we had considered a variety of oral health indicators such as the number of missing teeth, complete denture use, and mouthwash use.

The number of missing teeth variable was created by combining data regarding missing teeth from each of the three time periods during the participant's life (up to 16 years of age, 17-30 years of age, and after 30 years of age) up until diagnosis of the disease. It is important to note that participants were asked how many teeth were missing during each specific period, meaning the number reflects teeth lost in that particular period rather than being a cumulative measure. The questionnaire had answers none, 1-5, 6-15, 16-20, 21-30, and more than 30 for each time period. The midpoint of each category was taken to calculate values for the continuous variable of each period; then these were summed up to create a continuous variable to represent the lifetime number of missing teeth. This variable was then divided into two categories: < 9 missing teeth and > 9 missing teeth, using the 50th percentile of the controls' distribution as a cut-off point. Number of missing teeth is often used as an oral health indicator variable in previous studies (117).

Complete denture use and mouthwash use variables were considered as categorical variables coded as Yes / No.

5.7.1.3 Other variables

<u>HPV status</u>

As previously indicated, typing was carried out for a variety of HPV types, and a single variable was made for each type. For example, variable HPV-18 indicated the presence or absence of HPV-18 among the participants. The HPV type variables were then combined to create a categorical variable for HPV status, which has two categories: No HPV or LR-HPV, indicating HPV absence or presence of low-risk HPV types [6, 11, 40, 42, 54, 55, 61, 62,50, 64, 67, 69, 70, 71, 72, 81, 83, and 84], and HR-HPV indicating the presence of high-risk HPV types which included [HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, and 82] (114,118,119).

Alcohol consumption

Alcohol intake is one of the established HNC risk factors. We considered the lifetime exposure of alcohol of an individual to calculate the total alcohol consumption of a participant. Participants were considered “ever drinkers” if they had the habit for at least one year.

Data on alcohol intake were gathered according to beverage type. Numerous beverage varieties were taken into consideration: These include (i) wine; (ii) beer; (iii) hard liquor (such as brandy, gin, arrack, whiskey, vodka, cachaça, and grappa); (iv) aperitif; and (v) any other kind of alcohol. Furthermore, for each type, five distinct consumption units were taken into account: (i) 50 ml for small glasses; (ii) 100 ml for medium glasses; (iii) 250 ml for large glasses; (iv) 330 ml for half bottles; and (v) 750 ml for full bottles. The proportion of ethanol in each type of beverage was calculated following the standards: Beer: 5%; hard liquor: 50%; various types of wine, toddy, and aperitifs: 10% (120).

For participants who reported drinking alcohol, the total amount of alcohol consumed per day was calculated in milliliters (ml) for each age, starting from the age at the start of the habit to the age at quitting. The formula used was: *Ethanol (ml) =Volume of beverage (ml) x Ethanol percentage*. For example, if a participant drinks 500 ml of beer, the ethanol content is: 500×0.05=25 ml of ethanol. Ages before the start of habit and after quitting were assigned zero values. A value of zero was assigned for all ages for those who were non-drinkers. Figure 2 shows the questionnaire used to collect the life course history of alcohol consumption.

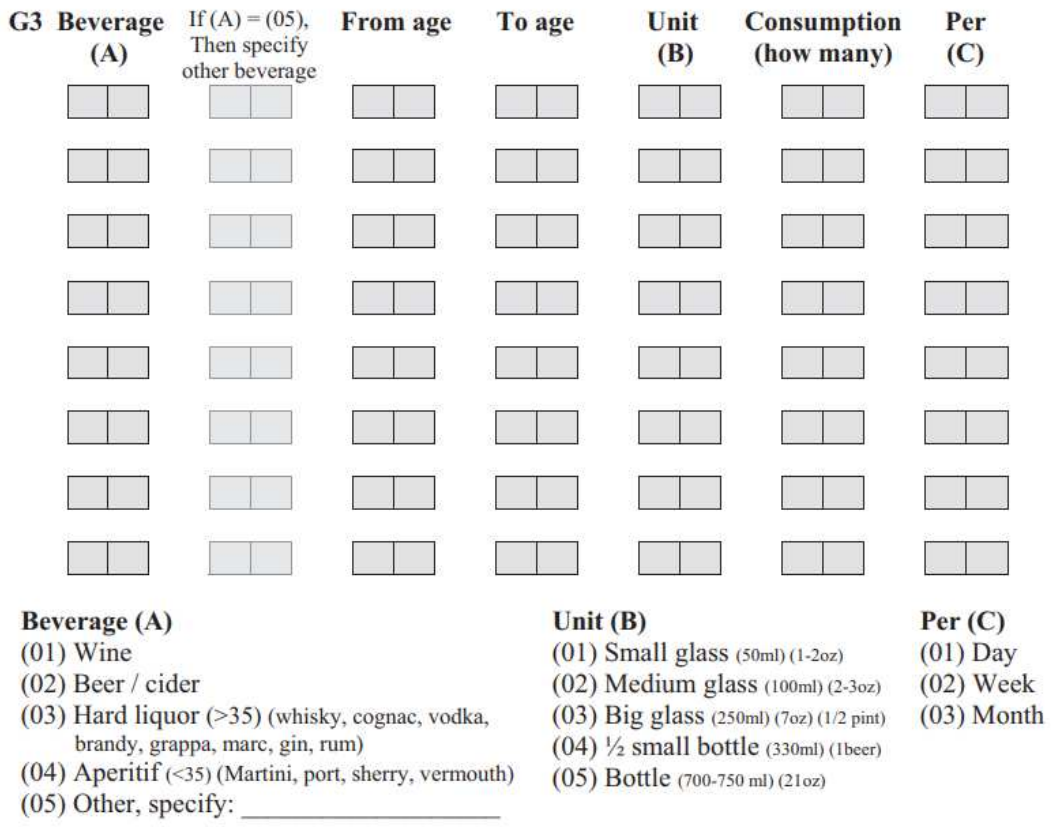


Fig - 2 Instrument used to collect life course history of alcohol intake

Tobacco consumption

The most important and established risk factor of HNC is tobacco consumption. Lifetime exposure was collected for tobacco as smoke pack years (121) in a similar manner as alcohol consumption. Information was collected on different forms of tobacco smoking: cigarette, cigar, and pipe smoking. Changes in exposure status of less than one year were not considered as participants were classified as 'Ever-smokers' only if they consumed any type of tobacco product for at least one year in their lifetime. Tobacco smoking history was collected in periods of the same type of product same frequency, additionally, commercial brand if applicable.

Interviewers were trained appropriately to use the life grid approach appropriately to collect data effectively such as to not provide leading questions and to cross-check changes in exposure. Figure 3 displays the instrument used to record tobacco history.

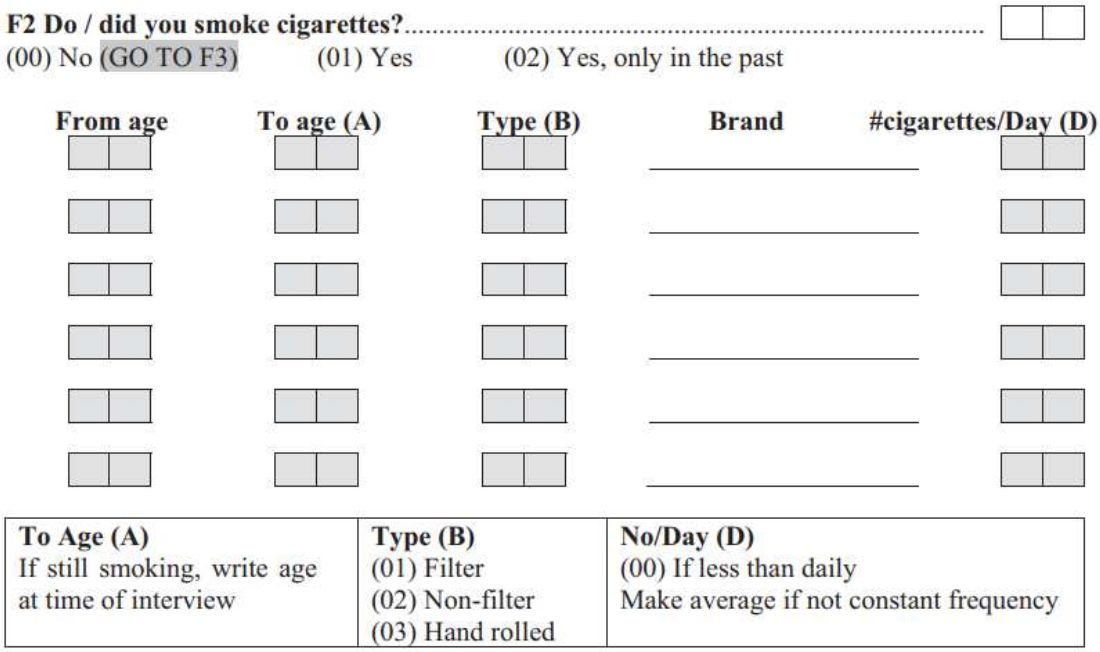


Figure-3 Instrument used to collect lifetime exposure to tobacco consumption (Cigarettes)

Standardisation of tobacco smoking

Commercial cigarettes are sold in packs of 20 in Canada. To standardise the consumption of different types of smoked tobacco products based on the estimated amount of tobacco they contain (122) we refer to it as a "standardised cigarette" and it is calculated as follows,

1 pack of standardized cigarettes = 20 commercial cigarettes (filtered/non-filtered) = 4 hand-rolled cigarettes = 13.3 bidis = 4 cigars = 5 pipes.

A total number of standardized cigarette packs represented the lifetime tobacco exposure of an individual and was calculated similarly at each age as done for alcohol intake, starting from the age of start of habit to the age at quitting for ever-smokers. Similarly, non-smokers were assigned a value of zero value for all ages until the date of the interview.

Avoiding dichotomizing the alcohol and smoking variables and avoiding assuming a linear relationship between these covariates and the outcome is crucial to prevent residual confounding and a loss of power, leading to less accurate results (123).

Education

We used the number of years of formal education as a proxy for SEP (124). The variable was used as a continuous variable in the statistical models.

Age and sex

Age was taken into account as a continuous variable and sex was considered as a categorical variable. Although these variables were taken into consideration in the study design by frequency matching, we further considered their residual effect by adjusting the model.

5.7.1.4 *Negative control variable*

To fulfil the second aim of my study, I used a negative control exposure analysis. As explained in the literature review section of this thesis, this analysis is used to detect any underlying bias due to unmeasured confounders. For a variable to be used as a negative control, it must meet specific criteria as illustrated in Figure 4. First, the negative control variable must not be associated with the outcome. Secondly, it must have a similar confounding structure to the exposure of interest (15). We selected sexually transmitted diseases (STD) as our negative control variable for the analysis. This binary variable was created based on the presence or absence of any of the following diseases: syphilis, gonorrhea, chlamydia, and herpes.

There has been no significant evidence in existing literature, linking any of these diseases to HNC. It is important to note that the status of HPV, an STD and known risk factor of HNC, was not included in this STD variable. According to the predetermined exclusion criteria, study participants with sexually transmitted diseases like AIDS, were excluded based on immunocompromised status. Information about STD was collected as part of the life course exposure variables in the medical conditions section of the HeNCe study questionnaire. As a result, the STD variable meets both requirements needed to be chosen as the best negative control exposure.

Negative control exposure analysis was conducted to test the negative control exposure (STD) for hypothesised null association with the outcome of interest. This was done by replacing the exposure (oral health) with the negative control variable (STD) in the crude and adjusted models and interpreting the associations obtained. An observed null finding would indicate the lack of detectable bias due to unmeasured confounders. If there is an association between the negative control exposure and HNC, then it indicates the presence of unmeasured confounding bias in the association being validated.

Figure 4 shows a graphical representation of the confounding structure of negative control. The figure shows how both measured (L) and unmeasured (U) confounders influence the relationship between the exposure (A) and the outcome (Y), in the case of this project oral

health and HNC. The negative control (N) can help identify whether the relationship between A and Y is distorted by these confounders. If an association exists between N and Y, it suggests that the observed relationship between A and Y may be confounded by the same factors, including those that are unmeasured (U).

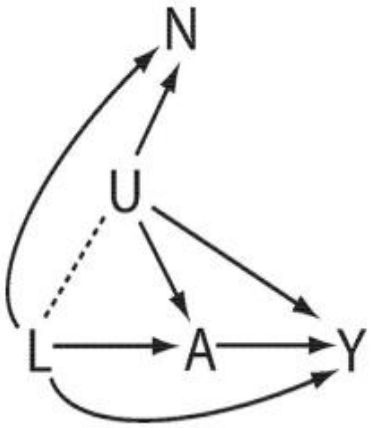


Figure-4 Schematic representation of the confounding structure of negative control (N) and exposure (A) through measured confounders (L) and unmeasured confounders (U) when the association to be tested is between exposure (A) and outcome (Y) (15)

5.7.2 Directed Acyclic Graphs and Confounder Selection

We used logistic regression models to estimate the associations between oral health indicators and HNC risk. All models were adjusted for various confounders, which were identified using (Directed acyclic graphs) DAGs (125). DAGs are graphical tools used in epidemiology to represent assumptions about the relationships between variables and to help identify and understand causal relationships (126). They are directed: Each arrow in the graph has a direction, indicating the direction of the relationship (cause to effect), acyclic: The graph does not contain any cycles; you cannot start at one variable and follow the arrows to return to the same variable (each variable cannot be its predecessor or successor), graph: a visual representation of variables, including outcome, exposure, confounders (nodes) and the arrows (edges) indicate causal relationships between them (127).

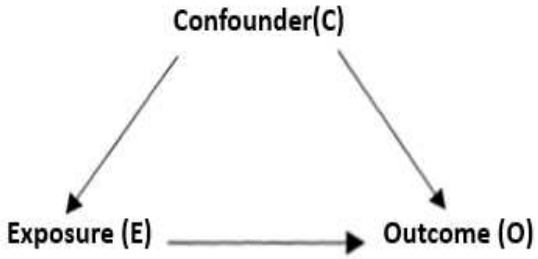


Figure-5 Schematic DAG Exposure of interest (E) affects outcome (O), and (C) is a confounder affecting both E and O. Path E← C→ O is a backdoor path.

DAGs are non-parametric methods that have been increasingly used in oral epidemiology. They do not require distribution parameters; they only require authors' expertise and proficiency in the assumed relationships between different variables that would occur in nature (128). Figure 6 illustrates the DAG used in this study to identify the potential confounders. It provides an overview of the relationships between the various confounders affecting the oral health-HNC association.

We used three criteria to identify confounders: 1) a confounder must be associated with the exposure; 2) It must be a cause of the outcome either through direct effect or indirectly (being a parent of a cause); 3) a confounder cannot be affected by the exposure of the outcome (should not be a mediator or a collider).

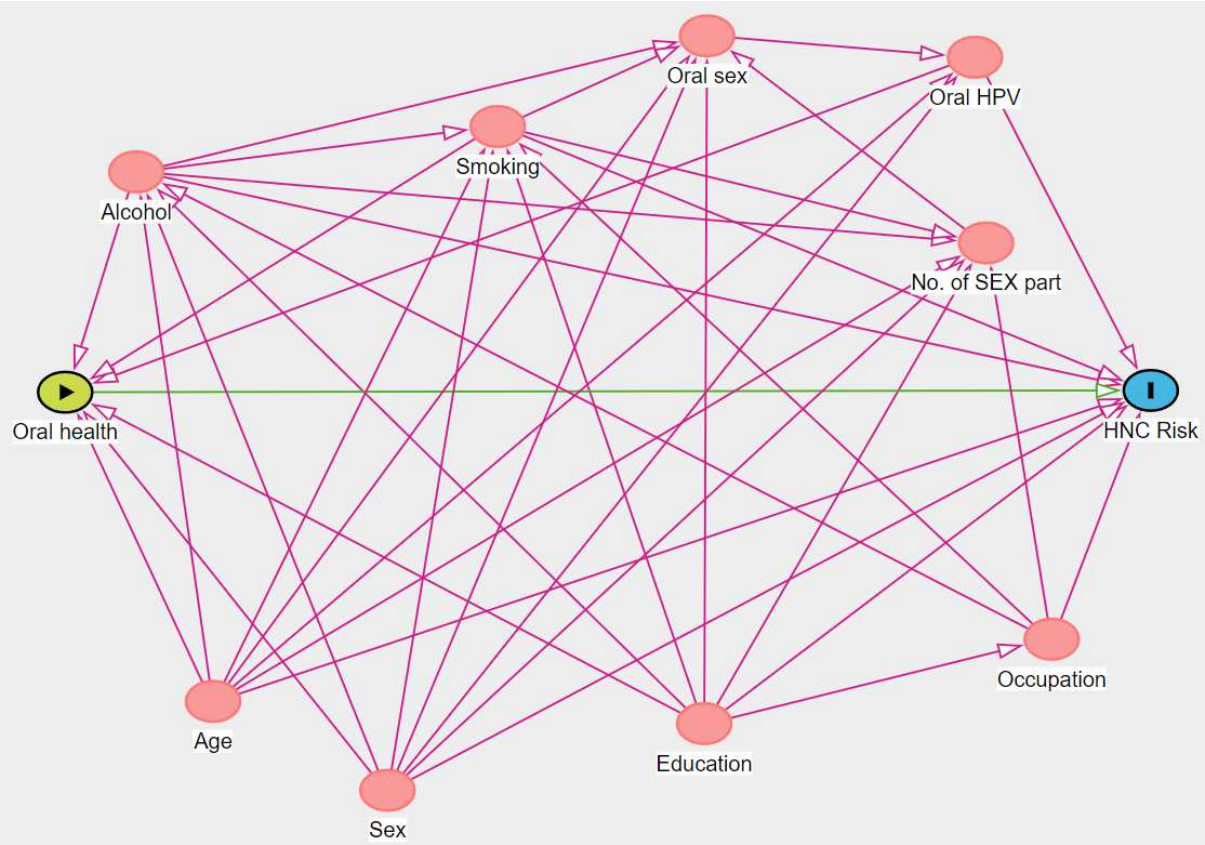


Figure-6 Directed Acyclic Graph describing the relationship between the study variables.

Once identified, we chose the minimum sufficient set of variables to close the backdoor/indirect path between exposure and outcome which might lead to bias in the association of interest. The final set of confounders selected for adjustment were age (years), sex, total education years, HPV status, lifetime smoking (pack-years), and lifetime alcohol consumption (litre years).

5.7.3 Data analysis

All data analysis was performed using Rstudio programming language (129), data wrangling, recoding, to run logistic models, to obtain descriptive statistics, DAGs, and tables, and to perform the PSA. The priors for the beta distribution of bias parameters assumed in (PSA) probabilistic sensitivity analysis were obtained from the SHELF expert elicitation package by Oakley et al (130).

5.7.3.1 Descriptive statistics

Descriptive statistics was performed to describe and compare the population characteristics of the cases and controls. The variables described were age, sex, education years, alcohol intake,

tobacco consumption, HPV status, sexually transmitted diseases (negative control), and oral health indicators such as complete denture use, missing teeth, and mouthwash use. All the variables were categorical except age and education years which were continuous. Lifetime exposure variables, lifetime ethanol exposure, and smoke pack years were also treated as continuous variables. For continuous variables, t-tests were performed to determine whether there were differences in the means between cases and controls. Mean and standard deviation measures were used for the continuous variables described, for the categorical variables we use cross-tabulations to compare the distribution among cases and controls. Table 1 represents the descriptive statistics of the selected variables of the study participants.

5.7.3.2 Logistic regression models

Logistic regression is a statistical method used to model a dichotomous or binary outcome variable (Yes/No) and independent variables (exposure) (131). Confounding variables can be added as independent variables to the models to adjust for their effect (132). The main idea is to model the odds of the outcome, (usually a disease) occurring as a function of the independent variables. It is used for predicting the probability of occurrence of an outcome or disease as a logistic function (131). The following equation represents the probability of an outcome as a function of the exposure and confounder variables.

$$Pr\ (\text{Outcome} \mid \text{E}, C) = \frac{e^{(\beta 0 + \beta 1 X + \beta' C)}}{1 + e^{(\beta 0 + \beta 1 X + \beta' C)}}$$

Where E = exposure of interest (oral health), and C is a vector representing the minimum sufficient set of confounders that need to be adjusted for to close backdoor paths.

In this study, we used unconditional logistic regression models rather than conditional regression models as the case-control study design included frequency matching rather than one-by-one matching. Data matched on a few demographic variables are loose-matching data, and in such cases, the unconditional logistic regression model is a proper reliable method to perform (133). The HeNCe life study data were frequency matched on two demographic variables (age and sex), so unconditional logistic regression models can be used reliably. We performed unconditional logistic regression models to estimate the associations between various oral health indicators and HNC risk. Logistic regression models use odds ratio (OR) and 95% confidence intervals (CI) as a measure of association (131).

5.7.3.3 Negative Control Analysis

We also ran unconditional logistic regression models for the negative control exposure analysis. This epidemiological concept involved assuming a null hypothesis between negative control exposure variable (STD) and HNC risk, to test for unmeasured confounding effects. The exposure in the unconditional logistic models was replaced by the negative control variable (STD) Multivariate analysis was also done to yield adjusted estimates by adding the confounders identified by the DAGs to the model. This provides more accurate adjusted estimates of the tested associations. These analyses are also described in the manuscript (Chapter 6) with detailed results and discussion.

5.7.3.4 Probabilistic Sensitivity Analysis

PSA was done to test for the robustness of the model and to include uncertainty in the model. For the final objective, I used PSA to assess the misclassification bias which might have affected our exposure (oral health) due to the self-reported nature of the questionnaire and the retrospective nature of data collection. The misclassification bias was assumed to be non-differential. QBA methods and techniques generally involve assuming single deterministic values for the bias parameters, whereas bias parameters for PSA have values from known distributions, which is much more common in reality (92). PSA assumes all parameter values have uncertainty that are simultaneous and each of them are considered as probabilities. Next, we do sampling simulations such as Monte Carlo simulations from these distributions, selecting bias parameters from these distributions each time, and performing a merged analysis. Finally, we summarise these simulations and take the mean and 95% CI of simulation intervals to provide the corrected estimates for the associations between oral health and HNC risk. We followed the guidelines and steps for PSA as recommended by Fox et al (84).

(a) Identifying potential sources of bias in the data

(b) Identifying bias parameters required for adjustment

(c) Assigning distribution for each bias parameter based on existing knowledge and assumptions

(d) Use sampling techniques to sample bias parameters from distribution to incorporate uncertainty in the parameters.

(e) Generating estimates through modelling sampled bias parameters from distribution and repeat steps (a) – (d) for N no. of repetitions as appropriate.

(f) Summarising the bias-adjusted estimates with the mean or median of summary statistics with confidence intervals to yield corrected estimates of associations.

5.7.4 Missing values

Although the number of missing values in the HeNCe Life study database is low, the oral health indicators and the STD variables had a few missing values. Smoking was missing for one case, alcohol consumption for one case and one control, and the education (years) variable was missing for one control. Because individuals with missing values can differ from those who answered every question, subjects with missing values were not excluded to prevent bias and decrease the sample size. Missing values that could be imputed were imputed using multivariate imputation by chained equations (MICE) (134,135). This technique involves imputing each variable with missing values conditional on other variables in the dataset.

6 RESULTS: (MANUSCRIPT)

Preface

Oral health-HNC associations have been highly contested, gaining interest, especially in recent decades, yet conclusive evidence remains elusive. It has been argued that confounding factors, mediators, and measurement errors may account for these inconsistencies. Importantly, to the best of our knowledge, no studies have investigated this association from an unmeasured confounder perspective. This chapter includes the manuscript from the study which investigates oral health as a potential risk factor of HNC and in determining the credibility of this association. This manuscript addresses this significant knowledge gap paving the way for future studies in this domain. It consists of the negative control exposure analysis and the probabilistic sensitivity analysis with results and interpretations for our findings in detail. Our goal was to find if the oral health-HNC associations are true or distorted by unmeasured confounders and biases in the data. The sample is not representative of the general population of Canada or Quebec, but results from our study have the potential to inform future etiologic studies that aim to investigate the association between oral health and HNC.

Oral Health in Head and Neck Cancers Associations: Addressing Confounding through Negative Control and Quantitative Bias Analyses

Elango P[1], Nicolau B[1], Farsi N[2], Grant AV[3], Rousseau MC[4], Madathil S[1]

[1] Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, QC, Canada

[2] Department of Preventive Dental Sciences, King Abdul Aziz University, Jeddah, Saudi Arabia

[3] Department of Anaesthesia, Faculty of Medicine and Health Sciences, Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, Canada

[4] Epidemiology and Biostatistics Unit, Centre Armand-Frappier Santé Biotechnologie, Institut national de la recherche scientifique, Laval, Canada

Type of manuscript: Original research article

Corresponding author:
Dr. Belinda Nicolau
Faculty of Dental Medicine and Oral Health Sciences, McGill University,
McGill University
2001 McGill College Ave, Suite 527
Montreal, Quebec H3A 1G1
Email: belinda.f.nicolau@mcgill.ca
Tel: +1-514-3961719

Keywords: Head and Neck Cancers, Oropharyngeal Cancers, Oral Cancers, Epidemiology, Negative controls, Bias Analysis

Abbreviations:

HNC - Head and Neck Cancers

OPC - Oropharyngeal Cancers

HNSCC - Head and Neck Squamous Cell Carcinomas

HeNCe – Head and Neck Cancer Life study

HPV - Human Papilloma Virus

STD - Sexually Transmitted Diseases

HIV - Human Immunodeficiency Virus

AIDS - Acquired Immunodeficiency Syndrome

OR - Odds Ratio

aOR – Adjusted Odds Ratio

CI - Confidence Interval

PSA - Probabilistic Sensitivity Analysis

Abstract:

Background: The connection between oral health and head and neck cancers (HNC) has sparked significant interest in recent years. While some studies have reported strong links, others argue that these associations are influenced by mediators, unmeasured risk factors, and other biases. In this study, we use a negative control exposure technique to test whether the association between oral health indicators and HNC risk is due to unmeasured confounders. Additionally, we employ probabilistic sensitivity analysis (PSA) to estimate the extent of non-differential misclassification bias due to exposure.

Methods: The HeNCe study, a hospital-based case-control study, recruited incident cases of HNC (n=389) frequency matched to controls (n=429) by sex and age (within five years) from four referral hospitals in Montreal, Canada. In-person interviews using questionnaires and the life grid technique collected information on an array of life course exposures. Oral rinse and oral brush specimens were collected and analyzed for HPV positivity and genotyping. We estimated the odds ratios (OR) and 95% confidence intervals (CI) for the associations between oral health and HNC using unconditional logistic regression models controlling for confounders identified using directly acyclic graphs. Using negative control exposure analysis, we assessed bias due to unmeasured confounders in the associations. Moreover, we conducted a PSA using predetermined bias parameters from previous studies to estimate the magnitude and direction of bias due to exposure misclassification.

Results**:** Participants using complete dentures and having more than nine missing teeth had an increased HNC risk [OR= 1.33, 95%CI (0.93-1.90) & OR=1.31, 95%CI (0.93-1.83)], respectively. Similar results were obtained when stratified by HNC subsite. Negative control

analysis yielded a null finding indicating no detectable presence of significant bias due to unmeasured confounders. Bias-corrected estimates of the association between oral health indicators, and HNC risk further moved away from the null.

Conclusion: Using negative control exposure analysis, our findings suggested a positive association between oral health indicatorsand HNC risk, as observed in previous studies. PSA findings yielded corrected estimates with increased magnitude, suggesting an underestimation of the crudes and showing the associations persist even after corrections.

1. Introduction:

Head and neck cancers (HNC) are malignant neoplasms primarily of epithelial origin, affecting the oral cavity, pharynx, nasal cavity, paranasal sinuses, and larynx (1). They are the seventh most common cancer worldwide, with about 660,000 new cases and 325,000 deaths annually (2). HNC is a complex, multifactorial disease with established risk factors, such as tobacco and alcohol consumption and human papillomavirus (HPV) infection. Potential HNC risk factors also include oral health status, dietary habits, obesity, and occupational exposure (3-5).

While several studies showed positive associations between oral health indicators (e.g., periodontal disease, oral hygiene, frequency of dental visits, tooth loss, denture use, and mouthwash use) and HNC risk (6,7), some argue that these associations are spurious and might be due to unmeasured factors, mediators, and underlying biases. This concern arises from the fact that most evidence on the oral-HNC link comes from observational studies (e.g., case-control and cohort), which are subject to the effect of unmeasured confounders and selection bias (8). Despite this, the extent to which these associations can be attributed to unmeasured confounding and systemic biases have not yet been thoroughly investigated.

To address this gap, we employ negative controls and probabilistic sensitivity analysis (PSA). Negative controls help identify biases, including those resulting from unmeasured confounding by distinguishing spurious associations from true ones (9). PSA estimates the extent of systemic bias and corrects for these biases accordingly by considering a probability distribution for each parameter rather than a single deterministic value (10).

In this study, we estimate the association between oral health indicators and HNC risk using data from a hospital-based case-control study. By employing the negative control validation tool and PSA, we evaluate the effect of unmeasured confounders, account for uncertainty in the model, and ascertain the extent of misclassification bias.

2. Methods:

The data come from the HeNCe Life study, a hospital-based case-control study conducted in four main referral hospitals in Montreal, Canada, Between September 2005 and November 2013. This study has been previously described (11). Briefly, cases (n=389) were consecutive incident histologically confirmed head and neck squamous cell carcinoma (HNSCC), comprising cancers of oral cavity, oropharynx, and larynx identified using the International Classification of Disease version 10 (ICD-10) (12). Controls (n=429) frequency matched by

age (5 years bracket) and sex to cases were selected from several outpatient clinics unrelated to major HNC risk factors (e.g., tobacco, alcohol) in the same hospitals as the cases and during the same period.

Face-to-face interviews using a questionnaire with a life grid (13) collected information on an array of life course exposures including sociodemographic characteristics (e.g., age, education, socioeconomic position) and behavioural factors (e.g., tobacco smoking, alcohol consumption). Oral rinse and oral brush protocols collected oral trans-epithelial cells from various sites in the oral cavity. Samples were stored in PreservCyt® buffer solution (Hologic, Bedford, MA) at 4°C until testing. HPV DNA detection and genotyping were done using Linear Array (Roche Molecular Diagnostics, Pleasanton, CA) (14).

Main exposure: Oral Health Indicators

We collected self-reported information on several oral health indicators, including complete denture use, mouthwash use, and the number of missing teeth (15). The first two variables were measured as binary (yes/no). Missing teeth were assessed over three time periods (childhood, early adulthood, and late adulthood) using an ordinal scale with five categories: none, 1-5, 6-15, 16-20, 21-30, and >30. To calculate the lifetime number of missing teeth, we used the midpoint of each category and summed the values across the three periods. The resulting continuous variable was then categorized into ≤9 or >9 missing teeth, using the 50th percentile among the control group as the cut-off point.

Negative control exposure

The criteria for selecting a negative control exposure variable are that it must not be associated with the outcome but must have a similar confounding structure to the exposure of interest, i.e., it should be exposed to potential confounders in a similar manner Fig-1(9). We used sexually transmitted diseases (STD) as the negative control exposure in our study. Our comprehensive questionnaire on STD included the presence and absence of syphilis, gonorrhoea, chlamydia, and herpes. None of these diseases has previously been linked to HNC risk, making them suitable for this purpose. It is essential to highlight that the STD variable did not include HPV, which is an established HNC risk factor. Also, participants with an immunocompromised status, such as those who were HIV AIDS-positive, were not invited to participate in the HeNCe study based on predefined exclusion criteria.

Probabilistic Sensitivity analysis:

Self-reported data are prone to misclassification and recall biases. PSA was performed to determine the extent of misclassification bias, which was assumed to be non-differential. The objective is to provide a reasonable estimate of the potential effects of misclassification bias on the observed results. Assuming a deterministic single value for bias parameters is unrealistic due to the inherent uncertainty in measurement and reporting (16). Instead, PSA assumes a probability distribution of these bias parameters; in our case, a beta distribution (17) was used. The priors or the probability distribution parameters (sensitivity and specificity) of each oral health variable were determined by expert opinion and previous validation studies (18-24). The mouthwash use variable was excluded from this analysis since it cannot be clinically diagnosed or determined, and there is no gold standard for its measurement. For the other variables, predetermined priors were obtained by assuming a best guess and considering an appropriate estimate from the existing validation studies of self-reported measures, as described in Supplemental Table 1(18-23). The assumptions need not be perfect but should be reasonable and well-described to be informative. Using these predetermined priors, the extent of bias in the associations was estimated and corrected estimates were obtained.

Statistical analysis:

Unconditional logistic regression with odds ratios (OR) and 95% confidence intervals (CI) estimated the associations between oral health indicators and HNC risk (Table 2). All models assessing HNC risk overall and stratified by subsite (oropharyngeal, laryngeal and oral cancer) were adjusted for the minimal sufficient set of potential confounders, which were identified using Directed Acyclic Graphs (DAGs) DAGs (Supplementary material). DAGs were constructed using the dagitty package in R to facilitate the identification of confounders (25).

The next step involved conducting negative control exposure analysis to test the association between STD and HNC. These analyses were done by substituting oral health with STD variable in both crude and adjusted models unconditional logistic regression models, for the same set of confounders. An observed null finding would indicate the absence of detectable bias due to unmeasured confounders. Conversely, an association between STD and HNC risk would suggest the presence of bias in the original association, which must then be estimated and quantified.

For the PSA, we obtained priors from several validation studies on self-reported oral health questionnaires according to existing literature (18-24). Using expert opinion and mean estimates from the validation studies (Supplemental Table 1), we determined the optimal values for each variable's sensitivity and specificity. The range of these estimates was established by taking conservative confidence intervals, adding and subtracting five points from the point estimate. These values were entered into the model as sensitivity and specificity parameters to assess misclassification bias, assuming it was non-differential. We conducted 20,000 simulations. All Statistical analysis, including the PSA was done using R Core Team 2022 (26).

3. Results:

The mean (standard deviation) age of cases and controls was 61.7 yrs [10.4] and 61.1 yrs [10.9] years respectively, with most of the cases being males [74%]. Oropharyngeal cancer was the most common [n=188], followed by laryngeal [n=128] and oral cavity cancers [n=73]. Controls had a higher number of education years than cases. More cases were smokers, consumed alcohol and had high-risk HPV exposures compared to controls. Approximately 60% of cases and 45% of controls reported having more than 9 missing teeth, respectively. Complete denture use was also more prevalent in cases (46%) than controls (31%) (Table 1).

We reported the associations between oral health indicators and HNC risk overall and by anatomical site in Table 2. Compared to those who did not wear dentures and had less than nine teeth missing, the aORs for complete denture use [aOR 1.33(0.93-1.90)] and for more than nine teeth [aOR 1.31(0.93-1.83)] suggested a 33% and 31% in the odds of HNC, respectively, although the CIs were wide.

The aORs for both oropharyngeal cancer [complete denture: aOR=1.43(0.89-2.28); missing teeth: aOR 1.34(0.86-2.08)] and laryngeal cancers [complete denture: aOR=1.48(0.92-2.38); missing teeth: aOR=1.50(0.93-2.42)] were also in the positive direction. There appeared to be no significant association between mouthwash use and HNC risk.

Table 3 displays the results of the negative control exposure. Compared to those without STD, the aOR for STD presence was close to the null value: overall HNC (aOR=0.99 (0.97-1.03)), oropharyngeal cancer (aOR= 1.00(0.96-1.01)) and for laryngeal cancer (aOR= 0.87(0.46-1.00)), indicating no significant detectable bias due to unmeasured confounders.

Table 4 represents the bias-corrected estimates of the oral health indicators, accounting for misclassification bias and random errors. The OR for complete denture use rises substantially from 1.83 to 2.97 (95% CI: 2.48-4.30) after accounting for misclassification bias with the selected priors (Se- 90%, Sp- 80%), indicating previous underestimation. This estimate is further adjusted to 3.01 (95% CI: 2.12-4.72) with the addition of random error correction, indicating a strong effect. Similarly, for missing teeth variable, the crude OR of 1.82 (95% CI: 1.38-2.40) reflects a significant positive association. Bias correction with predetermined priors (Se- 80%, Sp- 90%) increases the OR to 2.47 (95% CI: 2.25-2.85), indicating a stronger association than initially estimated. When random error is also considered, the OR adjusts to 2.48 (95% CI: 1.85-3.37), maintaining the strength of the association.

4. Discussion:

Studies have shown independent associations between various oral health indicators and HNC risk (27). Based on our findings and previously reported results from HeNCe Life study (15), complete denture use and missing more than 9 teeth may increase the risk of HNC overall, oropharyngeal and laryngeal cancer risk, aligning with existing literature (27-29). Contrary to some studies, our findings did not indicate any associations between mouthwash use, and HNC risk (30-31). Possible explanations for the observed associations could be due to cumulating evidence pointing towards a link between inflammation and cancer through alterations in the inflammatory pathway (32). Other study investigating the link between oral health and HNC have shown inconsistent results, with some reporting no associations for several oral health measures (27-33). Findings from the negative control exposure analysis yield null findings supporting the positive association between complete denture use and high numbers of missing teeth and the risk of overall HNC, oropharyngeal and laryngeal cancers.

Results from the PSA indicate how these associations change when accounting for systemic errors such as misclassification bias and random errors. Overall, the PSA highlights the impact of such errors on the estimated associations, often revealing stronger associations than initially indicated by crude models. This underscores the importance of adjusting for these biases to obtain more accurate estimates in epidemiological research rather than just acknowledging them as a limitation.

One of the strengths of this study is the availability of detailed lifetime exposure information regarding established risk factors of HNC which allows us to model for precise estimates. Null estimates obtained from the negative control analysis indicates no detectable unmeasured

confounder bias. However, residual confounding might still persist and the associations might be subject to some inconsistencies. The retrospective nature of the study might lead to recall bias which might be at play affecting the true estimates. Since all the oral health indicators and variables used in the models are self-reported and therefore prone to information and measurement bias (34). Other sources of bias might still affect these associations leading to distorted estimates.

Data were obtained from a hospital-based case-control study. Hospital controls are prone to selection bias, so to reduce this risk, control subjects were chosen from the same healthcare facilities with medical conditions not linked to smoking or alcohol consumption for etiological diversity (3, 5). Control recruitment was also capped at 20% from each department to reduce overrepresentation. The lack of significant associations between oral cancer and oral health variables in our analysis could be due to the lack of a larger sub-sample size.

5. Conclusion:

PSA bias analysis and epidemiological validation concepts such as negative controls and its use in epidemiological studies could prove to be great assets in strengthening study findings and justifying the hypotheses as presented and demonstrated in our study. Thus, the practice of using negative controls routinely in oral literature to attain epidemiologically validated results should be implemented and promoted. The utility of different types of bias analysis has been on the rise recently in epidemiology since publications by lash et al (35). Although it is advisable to evaluate results cautiously as there might be some residual bias, measurement error, misclassifications, and other unknown factors that could lead to underlying problems and distorted estimates. With regard to oral health-HNC associations, the negative control exposure analysis yielded a null finding which strengthens and provides direction towards a positive association, in alignment with previous studies. PSA yielded corrected estimates that revealed stronger associations, suggesting that the true relationships were more pronounced than originally reported crudes. However, it is necessary to conduct future long-term longitudinal studies that replicate these trends in order to definitively validate these relationships and demonstrate causation.

6. References:

1. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. The Lancet. 2008 May 17;371(9625):1695-709.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021 May;71(3):209-49.

3. Brugere J, Guenel P, Leclerc A, Rodriguez J. Differential effects of tobacco and alcohol in cancer of the larynx, pharynx, and mouth. Cancer. 1986 Jan 15;57(2):391-5.

4. Graham S, Mettlin C, Marshall J, Priore R, Rzepka T, Shedd D. Dietary factors in the epidemiology of cancer of the larynx. American Journal of Epidemiology. 1981 Jun 1;113(6):675-80.

5. Vučičević Boras V, Fučić A, Baranović S, Blivajs I, Milenović M, Bišof V, Rakušić Z, Ceppi M, Bruzzone M. Environmental and behavioural head and neck cancer risk factors. Central European journal of public health. 2019 Jun 26;27(2):106-9.

6. Zheng T, Boyle P, Hu H, Duan J, Jiang P, Ma D, Shui L, Niu S, Scully C, MacMahon B. Dentition, oral hygiene, and risk of oral cancer: a case-control study in Beijing, People's Republic of China. Cancer Causes & Control. 1990 Nov; 1:235-41.

7. Velly AM, Franco EL, Schlecht N, Pintos J, Kowalski LP, Oliveira BV, Curado MP. Relationship between dental factors and risk of upper aerodigestive tract cancer. Oral oncology. 1998 Jul 1;34(4):284-91.

8. Mezei G, Kheifets L. Selection bias and its implications for case–control studies: a case study of magnetic field exposure and childhood leukaemia. International journal of epidemiology. 2006 Apr 1;35(2):397-406.

9. Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology (Cambridge, Mass.). 2010 May;21(3):383.

10. Oakley JE, O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2004 Aug; 66(3):751-69.

11. Laprise C, Madathil SA, Schlecht NF, Castonguay G, Soulières D, Nguyen-Tan PF, Allison P, Coutlée F, Hier M, Rousseau MC, Franco EL. Increased risk of oropharyngeal cancers mediated by oral human papillomavirus infection: Results from a Canadian study. Head & neck. 2019 Mar; 41(3):678-85.

12. American Joint Committee on Cancer. AJCC cancer staging manual. Edge SB, editor. New York: Springer; 2010.

13. Parry O, Thompson C, Fowkes G. Life course data collection: qualitative interviewing using the life grid. Sociological research online. 1999 Jul;4(2):102-12.

14. Flores-Miramontes MG, Torres-Reyes LA, Alvarado-Ruíz L, Romero-Martínez SA, Ramírez-Rodríguez V, Balderas-Peña LM, Vallejo-Ruíz V, Piña-Sánchez P, Cortés-Gutiérrez EI, Jave-Suárez LF, Aguilar-Lemarroy A. Human papillomavirus genotyping by Linear Array and Next-Generation Sequencing in cervical samples from Western Mexico. Virology Journal. 2015 Dec; 12:1-1.

15. Farsi N. Epidemiology of Human Papilloma virus related head and neck cancers. McGill University (Canada); 2015.

16. Greenland S. Multiple-bias modelling for analysis of observational data. J R Stat Soc Ser A Stat

Soc. 2005;168(2):267–306.

17. Owen CB. Parameter estimation for the beta distribution. Brigham Young University; 2008.

18. Balappanavar AY, Sardana V, Nagesh L, Ankola AV, Kakodkar P, Hebbal M. Questionnaire vs clinical surveys: the right choice?–A cross-sectional comparative study. Indian J Dent Res.2011;22(3):494

19. Pitiphat W, Garcia RI, Douglass CW, Joshipura KJ. Validation of self-reported oral health measures. Journal of public health dentistry. 2002 Jun;62(2):122-8.

20. Ramos RQ, Bastos JL, Peres MA. Diagnostic validity of self-reported oral health outcomes in population surveys: literature review. Revista brasileira de epidemiologia. 2013; 16:716-28.

21. Sekundo C, Stock C, Jürges H, Listl S. Patients' self-reported measures of oral health—A validation study on basis of oral health questions used in a large multi-country survey for populations aged 50+. Gerodontology. 2019 Jun; 36(2):171-9.

22. Arenas-Márquez MJ, do Nascimento Tôrres LH, da Silva DD, Hilgert JB, Hugo FN, Neri AL, de Souza MD. Validity of self-report of oral conditions in older people. Brazilian Journal of Oral Sciences. 2019 Nov 18; 18:e191670-.

23. Pinelli C, de Castro Monteiro Loffredo L. Reproducibility and validity of self-perceived oral health conditions. Clinical Oral Investigations. 2007 Dec; 11:431-7.

24. Komagamine Y, Kanazawa M, Kaiba Y, Sato Y, Minakuchi S. Reliability and validity of a questionnaire for self-assessment of complete dentures. BMC oral health. 2014 Dec; 14:1-7.

25. Textor J, Van der Zander B, Gilthorpe MS, Liśkiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. International journal of epidemiology. 2016 Dec 1; 45(6):1887-94.

26. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

27. Laprise C, Shahul HP, Madathil SA, Thekkepurakkal AS, Castonguay G, Varghese I, Shiraz S, Allison P, Schlecht NF, Rousseau MC, Franco EL. Periodontal diseases and risk of oral cancer in Southern India: Results from the HeNCe Life study. International journal of cancer. 2016 Oct 1;139(7):1512-9.

28. Ahrens W, Pohlabeln H, Foraita R, Nelis M, Lagiou P, Lagiou A, Bouchardy C, Slamova A, Schejbalova M, Merletti F, Richiardi L. Oral health, dental care and mouthwash associated with upper aerodigestive tract cancer risk in Europe: the ARCAGE study. Oral oncology. 2014 Jun 1;50(6):616- 25.

29. Hashim D, Sartori S, Brennan P, Curado MP, Wünsch-Filho V, Divaris K, Olshan AF, Zevallos JP, Winn DM, Franceschi S, Castellsagué X. The role of oral hygiene in head and neck cancer: results from International Head and Neck Cancer Epidemiology (INHANCE) consortium. Annals of Oncology. 2016 Aug 1; 27(8):1619-25.

30. Boffetta P, Hayes RB, Sartori S, Lee YC, Muscat J, Olshan A, Winn DM, Castellsagué X, Zhang ZF, Morgenstern H, Chen C. Mouthwash use and cancer of the head and neck: a pooled analysis from the International Head and Neck Cancer Epidemiology Consortium. European journal of cancer prevention. 2016 Jul 1; 25(4):344-8.

31. Chang JS, Lo HI, Wong TY, Huang CC, Lee WT, Tsai ST, Chen KC, Yen CJ, Wu YH, Hsueh WT, Yang MW. Investigating the association between oral hygiene and head and neck cancer. Oral oncology. 2013 Oct 1;49(10):1010-7.

32. Greten FR, Grivennikov SI. Inflammation and cancer: triggers, mechanisms, and consequences. Immunity. 2019 Jul 16; 51(1):27-41.

33. Garrote LF, Herrero R, Reyes RM, Vaccarella S, Anta JL, Ferbeye L, Munoz N, Franceschi S. Risk factors for cancer of the oral cavity and oro-pharynx in Cuba. British journal of cancer. 2001 Jul;85(1):46-54.

34. Durán D, Al-Soneidar WA, Madathil SA, Kaufman JS, Nicolau B. Quantitative Bias Analysis of misclassification in case-control studies: an example with Human Papillomavirus and Oropharyngeal Cancer.

35. Fox MP, MacLehose RF, Lash TL. Applying quantitative bias analysis to epidemiologic data. Cham (Switzerland): Springer; 2021.

Table 1 Selected characteristics of the study population (n=818)

Variables	**Cases** n = 389	Controls n= 429
Age Mean (SD)	61.7 yrs [10.4]	61.1 yrs [10.9]
Sex		
Male	288 (74.0)	297 (69.2)
Female	101 (26)	132 (30.8)
Tumor site		
Oropharynx	188	-
Larynx	128	-
Oral cavity	73	-
Education yrs, Mean (SD)	12.1 [3.88]	13.9 [4.37]
Lifetime smoking, Mean (SD)	40.5 [45.6]	24.2 [38.5]
Lifetime alcohol consumption, Mean (SD)	659.8 L[1390.3]	366.8 L[828]
HPV		
HR-HPV individuals	150 (38.5)	50 (11.65)
LR-HPV individuals	239 (61.5)	379 (88.35)
Missing teeth		
9 or less	157 (40.3)	237 (55.2)
More than 9	232 (59.7)	192 (44.8)
Denture use		
Yes	178 (45.7)	135(31.4)
No	209 (53.7)	294 (68.5)
Mouthwash use		
Yes	204 (52.4)	218 (50.8)
No	182 (46.7)	209 (48.7)
Negative control exposure-STD		
Presence of STD	64 (16.4)	73 (17)
Absence of STD	322(82.7)	354(82.5)

Table 2 Associations between indicators of oral health and HNC overall and stratified by subsite

Oral Health Indicator	Controls n=429	HNC n=389	OR (95% CI)	Oropharyngeal n=188	OR (95%CI)	Larynx n=128	OR (95%CI)	Oral n=73	OR (95%CI)
Complete Denture									
No	294	209	1	110	1	59	1	40	1
Yes	135	178	1.33(0.93-1.90)	78	1.43(0.89-2.28)	68	1.48(0.92-2.38)	32	1.01(0.56-1.83)
Mouthwash use									
No	209	186	1	86	1	66	1	30	1
Yes	218	204	0.99(0.96-1.01)	101	1.00(0.98-1.03)	61	0.84(0.56-1.26)	42	0.97(0.88-1.71)
Missing teeth									
9 or less	237	157	1	83	1	42	1	32	1
More than 9	192	232	1.31(0.93-1.83)	105	1.34(0.86-2.08)	86	1.50(0.93-2.42)	41	0.99(0.55-2.0)

All models adjusted for age (continuous), sex, education (continuous), lifetime smoking (pack-years), lifetime alcohol consumption (ethanol litre), HPV status categorised as high-risk HPV and low-risk HPV groups

Table 3 Associations between STD and HNC risk factors overall and stratifying by anatomical site: Negative control exposure analysis

	Cases	Controls	Univariate analysis	Multivariate analysis
			OR (95% CI)	OR (95% CI)
HNC overall	389	429	1.00(0.98-1.02)	0.99(0.96-1.01)
Oropharyngeal cancers	188	429	1.00(0.97-1.02)	1.00(0.97-1.03)
Laryngeal cancers	128	429	1.00(0.97-1.02)	0.87(0.46-1.03)
Oral cancers	73	429	1.01(0.97-1.03)	0.80(0.34-1.04)

All Multivariate models for the negative control association models were adjusted for age (continuous), sex, education (continuous), lifetime smoking (pack-years), lifetime alcohol consumption (ethanol litre), HPV status categorized as high and low-risk groups

Table 4 Systematic and random error corrected estimates from Probabilistic sensitivity analysis (PSA)

Oral health indicator	Sample size (n)	Unadjusted OR (95% CI)	OR (95% CI) Sys err corrected	OR (95% CI) Sys & random err corrected
Complete denture				
Se- 90 Sp- 80	818	1.83(1.38-2.44)	2.97 (2.48-4.30)	3.01(2.12-4.72)
Missing teeth				
Se- 80 Sp- 90	818	1.82(1.38-2.40)	2.47(2.25-2.85)	2.48(1.85-3.37)

All models were univariate crude models which yielded corrected crude estimates of association for the selected sensitivity and specificity priors. Se - Sensitivity Sp- Specificity

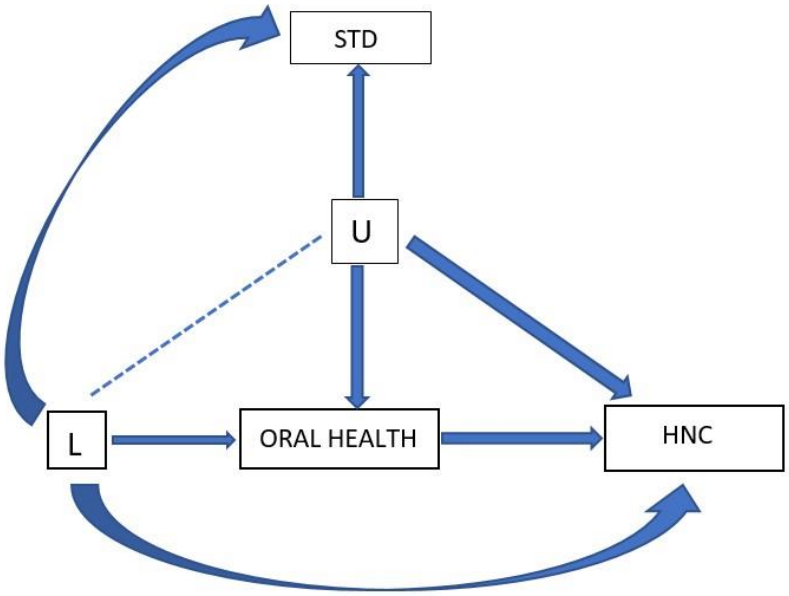


FIG 1- A Diagrammatic representation of the causal diagram illustrating the relationship between Exposure (oral health), Negative control exposure (STD), Outcome (HNC), L- Measured confounders (smoking, alcohol, HPV) and U (Unmeasured confounders).[9]

7 DISCUSSION

This chapter includes a summary of the main results of this study and highlights the novel findings. It also discusses the research gaps in the literature that this project aims to address as well as its potential contributions. This dissertation focuses on oral health as a possible risk factor for HNC and investigates the credibility of this link.

7.1 Summary of Findings

The overall goal of this thesis is to investigate the association between oral health indicators and HNC risk. We used unconditional logistic regression models with several oral health indicators such as missing teeth, denture use, and mouthwash use as exposures, to achieve this goal. Findings yielded positive associations between denture use-HNC risk and missing teeth-HNC risk. This is in alignment with previous literature (14, 70, 71, 76). Missing teeth are considered a proxy for periodontal health, as tooth loss often results from bone loss and compromised periodontium as periodontal disease progresses (73). While several studies have reported an inverse association between mouthwash use and HNC risk, suggesting a protective factor (79), mouthwash use did not yield significant estimates in our dataset. The associations between denture use-HNC risk and missing teeth-HNC risk seemed to persist even when stratified for subsite except for the oral cavity subsite.

To further investigate the credibility of these associations, the second part of the analysis involved an epidemiological tool-based negative control exposure analysis to detect potential bias due to unmeasured confounding in these associations. This analysis, based on the null hypothesis described in the methodology chapter (15), compared individuals with STD and those without STD. The aOR for STD presence was close to the null value for overall HNC risk, indicating no association between the negative control variable and HNC. Similar results were obtained for the subsite-specific negative control analyses, supporting our hypothesis that the STD variable was not related to HNC risk. These findings suggest that oral health-HNC associations are not affected by detectable unmeasured confounding, owing to similar confounding structure of these associations as described in methodology chapter. However, other sources of bias due to other sources might still distort these associations.

For example, the retrospective nature of data collection and the self-reported data of the study might have led to information bias, resulting in misclassification (measurement error) of our exposure (oral health indicators). Although we assumed that the misclassification bias was non-

differential due to similar data collection amongst cases and controls, we estimate the extent of this misclassification bias using PSA analysis to evaluate the direction and magnitude of this bias. To perform the PSA, we assumed priors from previous validation studies to detect the accuracy of self-reported oral health measures (136). The priors were assumed to have a beta distribution and sampling iterations were done. Using these predetermined priors, the magnitude of the bias in the associations was estimated, and corrected estimates were obtained. PSA utilises a more realistic range of values for priors rather than single deterministic values such as in the case of other QBA, making the results more accurate (92). The bias-corrected estimates revealed an underestimation of the crude values and the estimates for both complete denture use and missing teeth moved away from the null. This further strengthened oral health-HNC associations, which is in alignment with previous findings (70, 71, 76, 137)

Findings from the PSA analysis, further suggest that the effect of oral health on HNC may be underestimated in the literature if systematic biases are not properly accounted for in study design or data analysis. Potential biases, such as selection bias, measurement error, and unmeasured confounding, could theoretically occur and might distort and mask true associations. It is also essential to improve the accuracy and reliability of measurement methods and instruments to prevent measurement error which might affect future studies.

The significance of our study findings highlights the need to utilise simple epidemiological tools, whenever feasible to validate and strengthen study findings, especially in the case of observational studies. This would prove to be an effective measure to strengthen the hypothesis while identifying potential causative links. Validated results improve the credibility and accuracy of study findings.

7.2 Strengths and limitations

This research study has a number of strengths. It includes data on life course exposures for various risk factors, giving more accurate and detailed information (138). Additionally, this enabled the creation of a more detailed DAG for the selection of confounders, ensuring proper adjustment in the analysis. Data collection was verified and validated by several methods to ensure reliability and quality. Re-interviews with a small portion of participants indicated moderate to high-reliability scores. Only newly diagnosed HNC cases were recruited in the HeNCe life study to ensure that the diagnosis did not influence their lifestyle behaviours and habits. Oral health data included variables such as denture use, number of missing teeth, and mouthwash use, allowing us to model multiple exposures as proxies for oral health status when assessing HNC risk.

The retrospective nature of data collection could introduce recall bias. However, the use of the life-grid tool (100) was used to improve recall accuracy and reduce bias, although some residual bias might persist. Oral health data were self-reported and did not include oral health examinations by trained dental personnel. Existing literature indicates that self-reported oral health measures vary in sensitivity (28% to 75%) depending on the type of oral health measure. Specificity also varied from 62% to 95% with most being highly specific (>90 percent) but poorly sensitive (136). This could potentially lead to misclassification and measurement error, which we addressed in the third part of the analysis. These methodological limitations call for improved validation studies on oral health questionnaires and the incorporation of oral health examination during medical diagnoses.

Stratified analysis involving the subsites lacked sufficient sample size and statistical power, especially oral cancers (n=75), which could explain the fact that some of the existing associations might have gone undetected. However, various studies in the literature have reported positive associations between oral health status and subsite-specific HNC risk (75, 77) Berkson's bias, a form of selection bias occurring when the exposure of interest is linked to the likelihood of presenting at the hospital, is a potential concern in hospital-based case-control studies (139). The use of hospital-based controls can affect the external validity compared to population-based controls. (140). However, the HeNCe life study mitigated this risk by recruiting controls from different hospital departments with diseases unrelated to HNC risk factors, reducing the possibility of Berkson's bias. Recruitment was also capped at 20% from each department to limit overrepresentation.

7.3 Future direction

Long-term longitudinal studies are required to confirm and replicate our results, which is essential to increase the generalisability and improve credibility of the findings. Future research should also investigate the molecular mechanisms involved in the oral health-HNC link, helping to clarify the biological pathways involved. This will help address current gaps and contested findings present in this domain. Incorporating oral health examination into cancer screening could provide valuable data on the relationship between oral health and HNC risk, while genetic and molecular studies would help to establish the biological plausibility of oral health as a conceivable risk factor.

Epidemiological studies, especially observational studies that aim to investigate exposures or risk factors must adhere to strict methodological guidelines and whenever feasible must be epidemiologically validated. As demonstrated in our study, simple epidemiological tools can be used for validation without significant resource requirements. However, lack of knowledge about these techniques, time constraints and lack of availability of data from existing literature serve as barriers that prevent researchers from performing validation studies. Whenever possible rather than just acknowledging potential bias, quantitative bias analysis can be performed to obtain bias-corrected estimates to further strengthen study findings.

7.4 Knowledge translation plan

I plan to disseminate the study results and findings to the faculty of Dental Medicine and Oral Health Sciences at McGill University and the Institut National de la Recherche Scientifique, Laval, through the positions of my supervisors Dr Nicolau and Dr Madathil, both at McGill and my committee member Dr Rousseau. They hold teaching positions at their respective institutions and teach cancer epidemiology. This information dissemination effort will create awareness about the utility of various epidemiological tools in epidemiological studies and potentially promote researchers to take up more validation studies and QBA. These changes could have a significant impact in the credibility and reliability of causal inferences derived from observational study designs.

Furthermore, I have and will continue disseminating information in various research conferences and workshops. The first one was a poster presentation at the Celebration of Research and Training in Oncology (CORTO) 2023 in Montreal, and the most notable one was International Association of Dental Research (IADR) 2024 in New Orleans. This research work was also one of the four projects selected for an oral presentation at the prestigious Epiforum symposium at IADR 24. I also presented this work at the Canadian Oral Health Summit (COHS 24) in Halifax.

Finally, I will submit the manuscript in Chapter -6 which includes my study findings and results to a peer-reviewed scientific journal to further improve this research work's distribution and translation of knowledge to a wider audience.

8 CONCLUSION

The overall conclusion of this project is that oral health may be a true risk factor for HNC, according to our findings from a Canadian sample, as supported by bias-corrected estimates. However, future long-term longitudinal studies are required to further explain and understand this link. My research work has demonstrated that there are some biases, that distort these associations, and they must be acknowledged in future studies in this domain. PSA bias analysis and epidemiological validation concepts such as negative controls, and their use in observational studies, could prove to be great assets in strengthening study findings and justifying the hypotheses as presented and demonstrated in our study. Therefore, it is important to adopt and encourage the practice of regularly utilizing such epidemiological validation tools in oral literature to achieve validated results, whenever feasible.

9 REFERENCES

1) Mody MD, Rocco JW, Yom SS, Haddad RI, Saba NF. Head and neck cancer. The Lancet**.** 2021 Dec 18; 398(10318):2289-99.

2) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021 May;71(3):209-49.

3) Canadian Cancer Statistics Advisory Committee in collaboration with the Canadian Cancer Society, Statistics Canada and the Public Health Agency of Canada. Canadian Cancer Statistics 2023. Toronto, ON: Canadian Cancer Society; 2023.

4) Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis. The oncologist. 2010;15(9):994-1001

5) De Graeff A, de Leeuw JR, Ros WJ, Hordijk GJ, Blijham GH, Winnubst JA. Long-term quality of life of patients with head and neck cancer. The laryngoscope. 2000 Jan;110(1):98-106.

6) Osazuwa-Peters N, Simpson MC, Zhao L, Boakye EA, Olomukoro SI, Deshields T, Loux TM, Varvares MA, Schootman M. Suicide risk among cancer survivors: Head and neck versus other cancers. Cancer. 2018 Oct 15; 124(20):4072-9.

7) Zhang S, Wang B, Ma F, Tong F, Yan B, Liu T, Xie H, Song L, Yu S, Wei L. Characteristics of B lymphocyte infiltration in HPV+ head and neck squamous cell carcinoma. Cancer science. 2021 Apr; 112(4):1402-16.

8) Bhat GR, Hyole RG, Li J. Head and neck cancer: Current challenges and future perspectives. InAdvances in cancer research 2021 Jan 1 (Vol. 152, pp. 67-102). Academic Press.

9) Hashim D, Genden E, Posner M, Hashibe M, Boffetta P. Head and neck cancer prevention: from primary prevention to impact of clinicians on reducing burden. Annals of Oncology. 2019 May 1;30(5):744-56.

10) Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. The Lancet. 2008 May 17; 371(9625):1695-709.

11) Clinton SK, Giovannucci EL, Hursting SD. The world cancer research fund/American institute for cancer research third expert report on diet, nutrition, physical activity, and cancer: impact and future directions. The Journal of nutrition. 2020 Apr 1;150(4):663-71.

12) Wang K, Yu XH, Tang YJ, Tang YL, Liang XH. Obesity: an emerging driver of head and neck cancer. Life sciences. 2019 Sep 15;233:116687.
13) Chang JS, Lo HI, Wong TY, Huang CC, Lee WT, Tsai ST, Chen KC, Yen CJ, Wu YH, Hsueh WT, Yang MW. Investigating the association between oral hygiene and head and neck cancer. Oral oncology. 2013 Oct 1;49(10):1010-7.
14) Divaris K, Olshan AF, Smith J, Bell ME, Weissler MC, Funkhouser WK, Bradshaw PT. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. Cancer Causes & Control. 2010 Apr; 21:567-75.
15) Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010 May 1; 21(3):383-8.
16) Karjalainen A, World Health Organization. International statistical classification of diseases and related health problems (ICD-10) in occupational health. World Health Organization; 1999.
17) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021 May;71(3):209-49.
18) Sharaf Z, Behzadifar M, Behzadifar M, Fitzmaurice C, Abate D. Global, regional, and National Cancer Incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017. Global Burden of Cancer. 2021 Jan 21.
19) Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. British Dental Journal. 2022 Nov 11;233(9):780-6.
20) Lechner M, Liu J, Masterson L, Fenton TR. HPV-associated oropharyngeal cancer: Epidemiology, molecular biology and clinical management. Nature reviews Clinical oncology. 2022 May;19(5):306-27.
21) Vigneswaran N, Williams MD. Epidemiologic trends in head and neck cancer and aids in diagnosis. Oral and Maxillofacial Surgery Clinics. 2014 May 1;26(2):123-41.
22) Thomas SJ, Penfold CM, Waylen A, Ness AR. The changing aetiology of head and neck squamous cell cancer: A tale of three cancers? Clinical Otolaryngology. 2018 Aug;43(4):999-1003
23) Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral oncology. 2009 Apr 1;45(4-5):309-16.

24) Lydiatt WM, Patel SG, O'Sullivan B, Brandwein MS, Ridge JA, Migliacci JC, Loomis AM, Shah JP. Head and neck cancers—major changes in the American Joint Committee on cancer eighth edition cancer staging manual. CA: a cancer journal for clinicians. 2017 Mar;67(2):122-37.
25) Patterson RH, Fischman VG, Wasserman I, Siu J, Shrime MG, Fagan JJ, Koch W, Alkire BC. Global burden of head and neck cancer: economic consequences, health, and the role of surgery. Otolaryngology–Head and Neck Surgery. 2020 Mar;162(3):296-303.
26) Cadoni G, Giraldi L, Petrelli L, Pandolfini M, Giuliani M, Paludetti G, Pastorino R, Leoncini E, Arzani D, Almadori G, Boccia S. Prognostic factors in head and neck cancer: a 10-year retrospective analysis in a single-institution in Italy. Acta Otorhinolaryngologica Italica. 2017 Dec;37(6):458.
27) Cadoni G, Boccia S, Petrelli L, Di Giannantonio P, Arzani D, Giorgio A, De Feo E, Pandolfini M, Galli P, Paludetti G, Ricciardi G. A review of genetic epidemiology of head and neck cancer related to polymorphisms in metabolic genes, cell cycle control and alcohol metabolism. ACTA otorhinolaryngologica italica. 2012 Feb;32(1):1.
28) Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg RS, Preston-Martin S, Bernstein L, Schoenberg JB, Stemhagen A, Fraumeni Jr JF. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer research. 1988 Jun 1;48(11):3282-7.
29) Tuyns AJ, Esteve J, Raymond L, Berrino F, Benhamou E, Blanchet F, Boffetta P, Crosignani P, Moral AD, Lehmann W, Merletti F. Cancer of the larynx/hypopharynx, tobacco and alcohol: IARC international case-control study in Turin and Varese (Italy), Zaragoza and Navarra (Spain), Geneva (Switzerland) and Calvados (France). International journal of cancer. 1988 Apr 15;41(4):483-91.
30) Farsi NJ, Rousseau MC, Schlecht N, Castonguay G, Allison P, Nguyen-Tan PF, Soulières D, Coutlée F, Hier M, Madathil S, Franco EL. Aetiological heterogeneity of head and neck squamous cell carcinomas: the role of human papillomavirus infections, smoking and alcohol. Carcinogenesis. 2017 Dec 7;38(12):1188-95.
31) Park JO, Nam IC, Kim CS, Park SJ, Lee DH, Kim HB, Han KD, Joo YH. Sex differences in the prevalence of head and neck cancers: a 10-year follow-up study of 10 million healthy people. Cancers. 2022 May 20;14(10):2521.
32) Xie SH, Mattsson F, Lagergren J. Incidence trends in oesophageal cancer by histological type: an updated analysis in Sweden. Cancer Epidemiology. 2017 Apr 1;47:114-7.

33) Marcu LG, Yeoh E. A review of risk factors and genetic alterations in head and neck carcinogenesis and implications for current and future approaches to treatment. Journal of cancer research and clinical oncology. 2009 Oct;135:1303-14.
34) Dhull AK, Atri R, Dhankhar R, Chauhan AK, Kaushal V. Major risk factors in head and neck cancer: a retrospective analysis of 12-year experiences. World journal of oncology. 2018 Jun;9(3):80.
35) Halmos GB, Bras L, Siesling S, van der Laan BF, Langendijk JA, van Dijk BA. Age-specific incidence and treatment patterns of head and neck cancer in the Netherlands—A cohort study. Clinical otolaryngology. 2018 Feb;43(1):317-24.
36) Conway DI, Brenner DR, McMahon AD, Macpherson LM, Agudo A, Ahrens W, Bosetti C, Brenner H, Castellsague X, Chen C, Curado MP. Estimating and explaining the effect of education and income on head and neck cancer risk: INHANCE consortium pooled analysis of 31 case-control studies from 27 countries. International journal of cancer. 2015 Mar 1;136(5):1125-39.
37) Khetan P, Boffetta P, Luce D, Stucker I, Curado MP, Menezes A, Wunsch-Filho V, Ahrens W, Lagiou P, Serraino D, Richiardi L. Occupations and the risk of head and neck cancer: a pooled analysis of the International Head and Neck Cancer Epidemiology (INHANCE) consortium. Journal of occupational and environmental medicine. 2019 May;61(5):397.
38) Wyss A, Hashibe M, Chuang SC, Lee YC, Zhang ZF, Yu GP, Winn DM, Wei Q, Talamini R, Szeszenia-Dabrowska N, Sturgis EM. Cigarette, cigar, and pipe smoking and the risk of head and neck cancers: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. American journal of epidemiology. 2013 Sep 1;178(5):679-90.
39) Madathil S, Rousseau MC, Joseph L, Coutlée F, Schlecht NF, Franco E, Nicolau B. Latency of tobacco smoking for head and neck cancer among HPV-positive and HPV-negative individuals. International Journal of Cancer. 2020 Jul 1;147(1):56-64.
40) Maasland DH, van den Brandt PA, Kremer B, Goldbohm RA, Schouten LJ. Alcohol consumption, cigarette smoking and the risk of subtypes of head-neck cancer: results from the Netherlands Cohort Study. BMC cancer. 2014 Dec;14(1):1-4.
41) Chauhan R, Trivedi V, Rani R, Singh U. A study of head and neck cancer patients with reference to tobacco use, gender, and subsite distribution. South Asian Journal of Cancer. 2022 Jan;11(01):046-51.

42) Freedman ND, Abnet CC, Leitzmann MF, Hollenbeck AR, Schatzkin A. Prospective investigation of the cigarette smoking–head and neck cancer association by sex. Cancer: Interdisciplinary International Journal of the American Cancer Society. 2007 Oct 1;110(7):1593-601.

43) Maasland DH, van den Brandt PA, Kremer B, Goldbohm RA, Schouten LJ. Alcohol consumption, cigarette smoking and the risk of subtypes of head-neck cancer: results from the Netherlands Cohort Study. BMC cancer. 2014 Dec;14:1-4.

44) Dhull AK, Atri R, Dhankhar R, Chauhan AK, Kaushal V. Major risk factors in head and neck cancer: a retrospective analysis of 12-year experiences. World journal of oncology. 2018 Jun;9(3):80.

45) Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, Islami F, Lemmens VE, Rehm J, Soerjomataram I. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. The Lancet Oncology. 2021 Aug 1;22(8):1071-80.

46) Di Credico G, Polesel J, Dal Maso L, Pauli F, Torelli N, Luce D, Radoï L, Matsuo K, Serraino D, Brennan P, Holcatova I. Alcohol drinking and head and neck cancer risk: the joint effect of intensity and duration. British journal of cancer. 2020 Oct 27;123(9):1456-63.

47) Rumgay H, Murphy N, Ferrari P, Soerjomataram I. Alcohol and cancer: epidemiology and biological mechanisms. Nutrients. 2021 Sep 11;13(9):3173.

48) De Villiers E-M, Fauquet C, Broker TR, Bernard H-U, Zur Hausen H. Classification of papillomaviruses. Virology. 2004;324(1):17-27.

49) Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. Medical Sciences. 2023 Jun 13;11(2):42

50) Bosetti C, Carioli G, Santucci C, Bertuccio P, Gallus S, Garavello W, Negri E, La Vecchia C. Global trends in oral and pharyngeal cancer incidence and mortality. International journal of cancer. 2020 Aug 15;147(4):1040-9.

51) Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. Human papillomavirus molecular biology and disease association. Reviews in medical virology. 2015;25:2-23.

52) Johnson DE, Burtness B, Leemans CR, Lui VW, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nature reviews Disease primers. 2020 Nov 26;6(1):92.

53) Dayyani F, Etzel CJ, Liu M, Ho CH, Lippman SM, Tsao AS. Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC). Head & neck oncology. 2010 Dec;2(1):1-1.
54) Michaud DS, Langevin SM, Eliot M, Nelson HH, Pawlita M, McClean MD, Kelsey KT. High-risk HPV types and head and neck cancer. International journal of cancer. 2014 Oct 1;135(7):1653-61.
55) Rettig EM, D'Souza G. Epidemiology of head and neck cancer. Surgical oncology clinics. 2015 Jul 1;24(3):379-96.
56) Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE, Shah KV. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. Journal of the National Cancer Institute. 2000 May 3;92(9):709-20.
57) D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, Westra WH, Gillison ML. Case–control study of human papillomavirus and oropharyngeal cancer. New England Journal of Medicine. 2007 May 10;356(19):1944-56.
58) Deschler DG, Richmon JD, Khariwala SS, Ferris RL, Wang MB. The "new" head and neck cancer patient—young, nonsmoker, nondrinker, and HPV positive: evaluation. Otolaryngology--Head and Neck Surgery. 2014 Sep;151(3):375-80.
59) Chen YJ, Chang JT, Liao CT, Wang HM, Yen TC, Chiu CC, Lu YC, Li HF, Cheng AJ. Head and neck cancer in the betel quid chewing area: recent advances in molecular carcinogenesis. Cancer science. 2008 Aug;99(8):1507-14.
60) Chang MC, Chiang CP, Lin CL, Lee JJ, Hahn LJ, Jeng JH. Cell-mediated immunity and head and neck cancer: with special emphasis on betel quid chewing habit. Oral oncology. 2005 Sep 1;41(8):757-75.
61) Lee YC, Li S, Chen Y, Li Q, Chen CJ, Hsu WL, Lou PJ, Zhu C, Pan J, Shen H, Ma H. Tobacco smoking, alcohol drinking, betel quid chewing, and the risk of head and neck cancer in an East Asian population. Head & neck. 2019 Jan;41(1):92-102.
62) Capozzi LC, Lau H, Reimer RA, McNeely M, Giese-Davis J, Culos-Reed SN. Exercise and nutrition for head and neck cancer patients: a patient oriented, clinic-supported randomized controlled trial. BMC cancer. 2012 Dec;12:1-9.
63) Saraiya V, Bradshaw P, Meyer K, Gammon M, Slade G, Brennan P, Abedi-Ardekani B, Olshan A. The association between diet quality and cancer incidence of the head and neck. Cancer Causes & Control. 2020 Feb;31:193-202.

64) Tan X, Nelson HH, Langevin SM, McClean M, Marsit CJ, Waterboer T, Pawlita M, Kelsey KT, Michaud DS. Obesity and head and neck cancer risk and survival by human papillomavirus serology. Cancer Causes & Control. 2015 Jan;26:111-9.
65) Maasland DH, van den Brandt PA, Kremer B, Goldbohm RA, Schouten LJ. Alcohol consumption, cigarette smoking and the risk of subtypes of head-neck cancer: results from the Netherlands Cohort Study. BMC cancer. 2014 Dec;14(1):1-4.
66) Tarleton HP, Park SL, Zhu WM, Lee YC, Hashibe M, Morgenstern H, Tashkin DP, Mao JT, Cozen W, Mack TM, Zhang ZF. Body mass index change in adulthood and lung and upper aerodigestive tract cancers. International journal of cancer. 2012 Sep 15;131(6):1407-16.
67) Kreimer AR, Randi G, Herrero R, Castellsagué X, Vecchia CL, Franceschi S. Diet and body mass, and oral and oropharyngeal squamous cell carcinomas: analysis from the IARC multinational case–control study. International journal of cancer. 2006 May 1;118(9):2293-7.
68) Langius JA, Bakker S, Rietveld DH, Kruizenga HM, Langendijk JA, Weijs PJ, Leemans CR. Critical weight loss is a major prognostic indicator for disease-specific survival in patients with head and neck cancer receiving radiotherapy. British journal of cancer. 2013 Sep;109(5):1093-9.
69) Paget-Bailly S, Guida F, Carton M, Menvielle G, Radoï L, Cyr D, Schmaus A, Cénée S, Papadopoulos A, Févotte J, Pilorget C. Occupation and Head and Neck Cancer Risk in Men. Journal of occupational and environmental medicine. 2013 Sep 1;55(9):1065-73.
70) Chang CC, Lee WT, Hsiao JR, Ou CY, Huang CC, Tsai ST, Chen KC, Huang JS, Wong TY, Lai YH, Wu YH. Oral hygiene and the overall survival of head and neck cancer patients. Cancer medicine. 2019 Apr;8(4):1854-64.
71) Kawakita D, Lee YC, Li Q, Chen Y, Chen CJ, Hsu WL, Lou PJ, Zhu C, Pan J, Shen H, Ma H. Impact of oral hygiene on head and neck cancer risk in a Chinese population. Head & neck. 2017 Dec;39(12):2549-57.
72) Critchlow SB, Morgan C, Leung T. The oral health status of pre-treatment head and neck cancer patients. British dental journal. 2014 Jan 11;216(1):E1-E1.
73) Gopinath D, Kunnath Menon R, K. Veettil S, George Botelho M, Johnson NW. Periodontal diseases as putative risk factors for head and neck cancer: systematic review and meta-analysis. Cancers. 2020 Jul 14;12(7):1893.

74) Nwizu N, Wactawski-Wende J, Genco RJ. Periodontal disease and cancer: Epidemiologic studies and possible mechanisms. Periodontology 2000. 2020 Jun;83(1):213-33.
75) Radaic A, Ganther S, Kamarajan P, Grandis J, Yom SS, Kapila YL. Paradigm shift in the pathogenesis and treatment of oral cancer and other cancers focused on the oralome and antimicrobial-based therapeutics. Periodontology 2000. 2021 Oct;87(1):76-93.
76) Hashim D, Sartori S, Brennan P, Curado MP, Wünsch-Filho V, Divaris K, Olshan AF, Zevallos JP, Winn DM, Franceschi S, Castellsagué X. The role of oral hygiene in head and neck cancer: results from International Head and Neck Cancer Epidemiology (INHANCE) consortium. Annals of Oncology. 2016 Aug 1;27(8):1619-25.
77) Rotundo LD, Toporcov TN, Biazevic GH, Carvalho MB, Kowalski LP, Antunes JL. Are recurrent denture-related sores associated with the risk of oral cancer? A case control study. Revista brasileira de epidemiologia. 2013;16:705-15.
78) Farquhar DR, Divaris K, Mazul AL, Weissler MC, Zevallos JP, Olshan AF. Poor oral health affects survival in head and neck cancer. Oral oncology. 2017 Oct 1;73:111-7.
79) Bai X, Cui C, Yin J, Li H, Gong Q, Wei B, Lu Y. The association between oral hygiene and head and neck cancer: A meta-analysis. Acta Odontologica Scandinavica. 2023 Jul 4;81(5):374-95.
80) Chung M, York BR, Michaud DS. Oral health and cancer. Current oral health reports. 2019 Jun;6:130-7.
81) Greenland S. Introduction to regression models. Modern epidemiology. 2008:381-417.
82) Fox MP, MacLehose RF, Lash TL. Applying quantitative bias analysis to epidemiologic data. Cham (Switzerland): Springer; 2021.
83) Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. International journal of epidemiology. 2021 Oct 1;50(5):1708-30.
84) Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. International journal of epidemiology. 2014 Dec 1;43(6):1969-85.
85) Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. International journal of epidemiology. 2021 Oct 1;50(5):1708-30.
86) Gordis L. Epidemiology e-book. Elsevier Health Sciences; 2013 Nov 14.

87) Rönmark E, Lundqvist A, Lundbäck B, Nyström L. Non-responders to a postal questionnaire on respiratory symptoms and diseases. European journal of epidemiology. 1999 Mar;15:293-9.
88) Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases?. Nephrology. 2020 Jun;25(6):435-41.
89) McGauran N, Wieseler B, Kreis J, Schüler YB, Kölsch H, Kaiser T. Reporting bias in medical research-a narrative review. Trials. 2010 Dec;11:1-5.
90) Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. Jama. 2011 Jun 15;305(23):2462-3.
91) Fadnes LT, Taube A, Tylleskär T. How to identify information bias due to self-reporting in epidemiological research. The Internet Journal of Epidemiology. 2009 Jan;7(2):28-38.
92) Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Medical decision making. 1985 Jun;5(2):157-77.
93) Laprise C, Madathil SA, Schlecht NF, Castonguay G, Soulières D, Nguyen-Tan PF, Allison P, Coutlée F, Hier M, Rousseau MC, Franco EL. Human papillomavirus genotypes and risk of head and neck cancers: Results from the HeNCe Life case-control study. Oral oncology. 2017 Jun 1;69:56-61.
94) Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. Journal of epidemiology and community health. 2003 Oct;57(10):778.
95) Anantharaman D, Abedi-Ardekani B, Beachler DC, Gheit T, Olshan AF, Wisniewski K, Wunsch-Filho V, Toporcov TN, Tajara EH, Levi JE, Moyses RA. Geographic heterogeneity in the prevalence of human papillomavirus in head and neck cancer. International journal of cancer. 2017 May 1;140(9):1968-75.
96) Beachler DC, Abraham AG, Silverberg MJ, Jing Y, Fakhry C, Gill MJ, Dubrow R, Kitahata MM, Klein MB, Burchell AN, Korthuis PT. Incidence and risk factors of HPV-related and HPV-unrelated Head and Neck Squamous Cell Carcinoma in HIV-infected individuals. Oral oncology. 2014 Dec 1;50(12):1169-76.
97) Burket LW, Lynch MA, Brightman VJ, Greenberg MS. Burket's oral medicine: diagnosis and treatment. 2003.
98) Zhang Y. Epidemiology of esophageal cancer. World journal of gastroenterology: WJG. 2013 Sep 9;19(34):5598.

99) Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. II. Types of controls. Am J Epidemiology. 1992 May 1;135(9):1029-41.

100) Parry O, Thompson C, Fowkes G. Life course data collection: qualitative interviewing using the life grid. Sociological research online. 1999 Jul;4(2):102-12.

101) Herrero R, Castellsagué X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T, Sridhar H, Rose B, Pintos J, Fernández L. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. Journal of the National Cancer Institute. 2003 Dec 3;95(23):1772-83.

102) Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, White I, Brunner E, Feeney A. Health inequalities among British civil servants: the Whitehall II study. In Stress and the Brain 2013 Oct 23 (pp. 61-67). Routledge.

103) Wadsworth ME. The imprint of time: childhood, history, and adult life. (No Title). 1991.

104) Power C, Manor O, Fox J. Health and Class: The Early Years: London: Chapman Hall; 1991

105) Blane DB. Collecting retrospective data: development of a reliable method and a pilot study of its use. Social science & medicine. 1996 Mar 1;42(5):751-7.

106) Berney LR, Blane DB. Collecting retrospective data: accuracy of recall after 50 years judged against historical records. Social science & medicine. 1997 Nov 1;45(10):1519-25.

107) Spector ME, Sacco AG, Bellile E, Taylor JM, Jones T, Sun K, Brown WC, Birkeland AC, Bradford CR, Wolf GT, Prince ME. E6 and E7 antibody levels are potential biomarkers of recurrence in patients with advanced-stage human papillomavirus–positive oropharyngeal squamous cell carcinoma. Clinical Cancer Research. 2017 Jun 1;23(11):2723-9.

108) Adelstein DJ, Ridge JA, Gillison ML, Chaturvedi AK, D'Souza G, Gravitt PE, Westra W, Psyrri A, Martin Kast W, Koutsky LA, Giuliano A. Head and neck squamous cell cancer and the human papillomavirus: summary of a National Cancer Institute State of the Science Meeting, November 9–10, 2008, Washington, DC. Head & Neck: Journal for the Sciences and Specialties of the Head and Neck. 2009 Nov;31(11):1393-422.

109) Kosicki DM, Riva C, Pajarola GF, Burkhardt A, Grätz KW. OralCDx brush biopsy--a tool for early diagnosis of oral squamous cell carcinoma. Schweizer Monatsschrift fur

Zahnmedizin= Revue Mensuelle Suisse D'odonto-stomatologie= Rivista Mensile Svizzera di Odontologia e Stomatologia. 2007 Jan 1;117(3):222-7.

110) MasterPureTM Complete DNA and RNA Purification Kit | Lucigen [Internet]. [cited 2020 Dec 16].Available from: https://www.lucigen.com/MasterPure-Complete-DNA-and-RNAPurification

111) Tristão W, Ribeiro RM, Oliveira CA, Betiol JC, Bettini JD. Epidemiological study of HPV in oral mucosa through PCR. Brazilian journal of otorhinolaryngology. 2012;78:66-70.

112) Coutlée F, Gravitt P, Richardson H, Hankins C, Franco E, Lapointe N, Voyer H, Canadian Women's HIV Study Group‡. Nonisotopic detection and typing of human papillomavirus DNA in genital samples by the line blot assay. Journal of clinical microbiology. 1999 Jun 1;37(6):1852-7.

113) Coutlée F, Gravitt P, Kornegay J, Hankins C, Richardson H, Lapointe N, Voyer H, Franco E. Use of PGMY primers in L1 consensus PCR improves detection of human papillomavirus DNA in genital samples. Journal of clinical microbiology. 2002 Mar;40(3):902-7.

114) Bernard HU. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. Journal of clinical virology. 2005 Mar 1;32:1-6.

115) O'Sullivan B, Shah J. New TNM staging criteria for head and neck tumors. InSeminars in surgical oncology 2003 (Vol. 21, No. 1, pp. 30-42). Hoboken: Wiley Subscription Services, Inc., A Wiley Company.

116) Mehrotra R, Gupta A, Singh M, Ibrahim R. Application of cytology and molecular biology in diagnosing premalignant or malignant oral lesions. Molecular cancer. 2006 Dec;5:1-9.

117) Haworth S, Shungin D, Kwak SY, Kim HY, West NX, Thomas SJ, Franks PW, Timpson NJ, Shin MJ, Johansson I. Tooth loss is a complex measure of oral disease: Determinants and methodological considerations. Community dentistry and oral epidemiology. 2018 Dec;46(6):555-62.

118) De Villiers EM, Fauquet C, Broker TR, Bernard HU, Zur Hausen H. Classification of papillomaviruses. Virology. 2004 Jun 20;324(1):17-27.

119) Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, et al Epidemiologic classification of human papillomavirus types associated with cervical. cancer. N Engl J Med. 2003 Feb 6;348(6):518-27.

120) Gajalakshmi V, Hung RJ, Mathew A, Varghese C, Brennan P, Boffetta P. Tobacco smoking and chewing, alcohol drinking and lung cancer risk among men in southern India. International journal of cancer. 2003 Nov 10;107(3):441-7.

121) Bernaards CM, Twisk JW, Snel J, Van Mechelen W, Kemper HC. Is calculating pack-years retrospectively a valid method to estimate life-time tobacco smoking? A comparison between prospectively calculated pack-years and retrospectively calculated pack-years. Addiction. 2001 Nov;96(11):1653-61.

122) Schlecht NF, Franco EL, Pintos J, Kowalski LP. Effect of smoking cessation and tobacco type on the risk of cancers of the upper aero-digestive tract in Brazil. Epidemiology. 1999 Jul 1:412-8.

123) Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Statistics in medicine. 2006 Jan 15;25(1):127-41.

124) Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S. Socioeconomic status in health research: one size does not fit all. Jama. 2005 Dec 14;294(22):2879-88.

125) Tennant PW, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, Tomova GD. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. International journal of epidemiology. 2021 Apr 1;50(2):620-32.

126) Hernan M, Robins J. Causal inference: What if. boca raton: Chapman & hill/crc. 2020.

127) Merchant AT, Pitiphat W. Directed acyclic graphs (DAGs): an aid to assess confounding in dental research. Community dentistry and oral epidemiology. 2002;30(6):399-404.

128) Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. International journal of epidemiology. 2021 Oct 1;50(5):1708-30.

129) R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

130) Oakley JE, O'Hagan A. SHELF: the Sheffield elicitation framework (version 2.0). Sheffield, UK: School of Mathematics and Statistics, University of Sheffield. 2010 Mar 4;560.

131) Kleinbaum DG, Klein M, Kleinbaum DG, Klein M. Introduction to logistic regression. Logistic regression: a self-learning text. 2010:1-39.

132) Bergerud W. Introduction to Regression Models: with worked forestry examples. Biom. Info. Hand. 7. Res Br, BC Min For, Victoria, BC Work Pap. 1996;26(1996.1996).

133) Kuo CL, Duan Y, Grady J. Unconditional or conditional logistic regression model for age-matched case–control data?. Frontiers in public health. 2018 Mar 2;6:57.

134) Acock AC. Working with missing values. Journal of Marriage and family. 2005 Nov;67(4):1012-28.

135) White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine. 2011 Feb 20;30(4):377-99

136) Farmer J, Ramraj C, Azarpazhooh A, Dempster L, Ravaghi V, Quiñonez C. Comparing self-reported and clinically diagnosed unmet dental treatment needs using a nationally representative survey. Journal of Public Health Dentistry. 2017 Sep;77(4):295-301.

137) Lizano M, Berumen J, García-Carrancá A. HPV-related carcinogenesis: basic concepts, viral types and variants. Archives of medical research. 2009 Aug 1;40(6):428-34.

138) Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. Journal of epidemiology and community health. 2003 Oct;57(10):778.

139) Berkson J. Limitations of the application of fourfold table analysis to hospital data. International journal of epidemiology. 2014 Apr 1;43(2):511-5.

140) Neupane B, Walter SD, Krueger P, Loeb M. Community controls were preferred to hospital controls in a case–control study where the cases are derived from the hospital. Journal of clinical epidemiology. 2010 Aug 1;63(8):926-31.

10 APPENDIX

10.1 Supplemental Material

Supplemental Table 1 Validation Studies of self-reported oral health measures

Oral Health variable	Study	Sensitivity (Se) value	Specificity (Sp) value
Complete denture	Arenas-Márquez et al 2019	99.3(97.9-99.8)	84.4(79.7-89.3)
	Ramos et al 2014	100%	100%
	Pitiphat et al 2002	100%	93%
Missing teeth	Balappanavar et al 2011	83.8%	83.3%
	Sekundo et al 2019	93.4(84.1-98.2)	92(85.8-96.1)
	Ramos et al 2014	88-91%	97%
	Pinelli et al 2007	96.7%	94.12%

Supplemental Material Fig 1 Directed Acyclic Graph Acyclic Graph

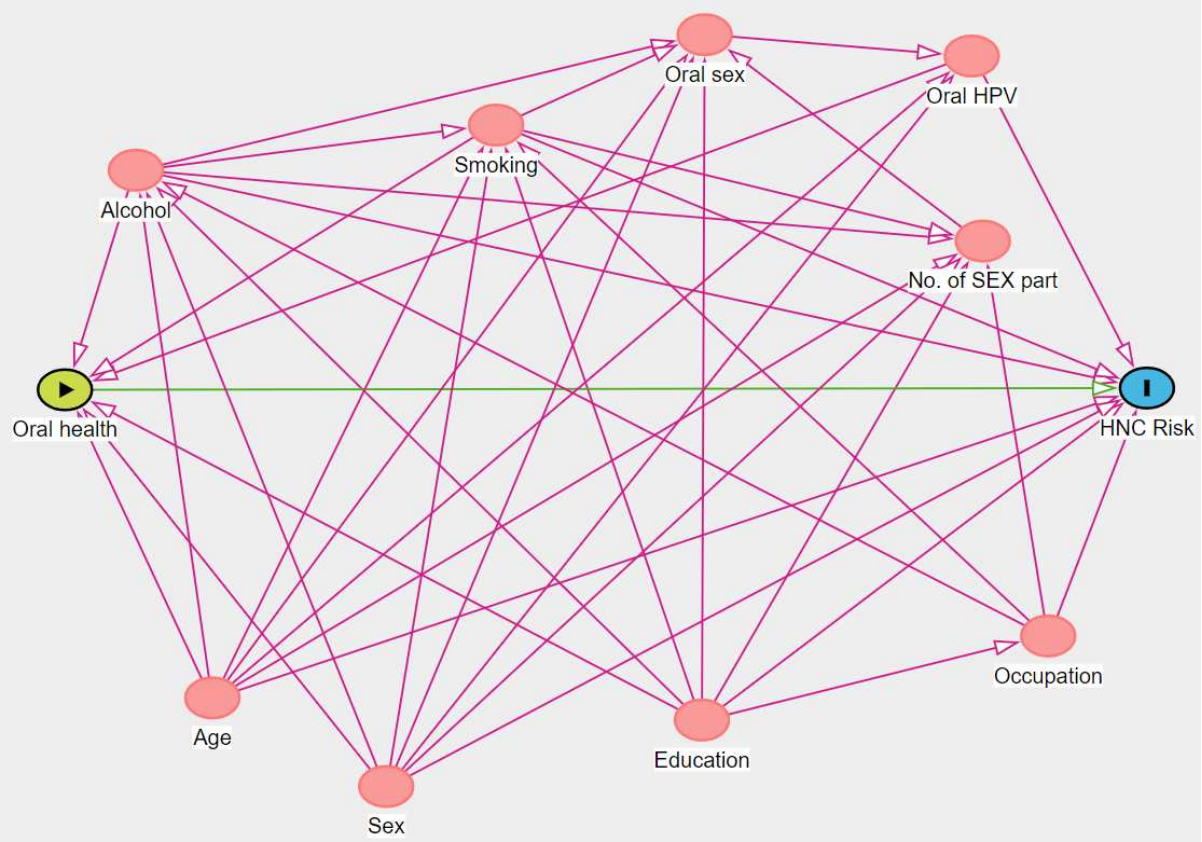
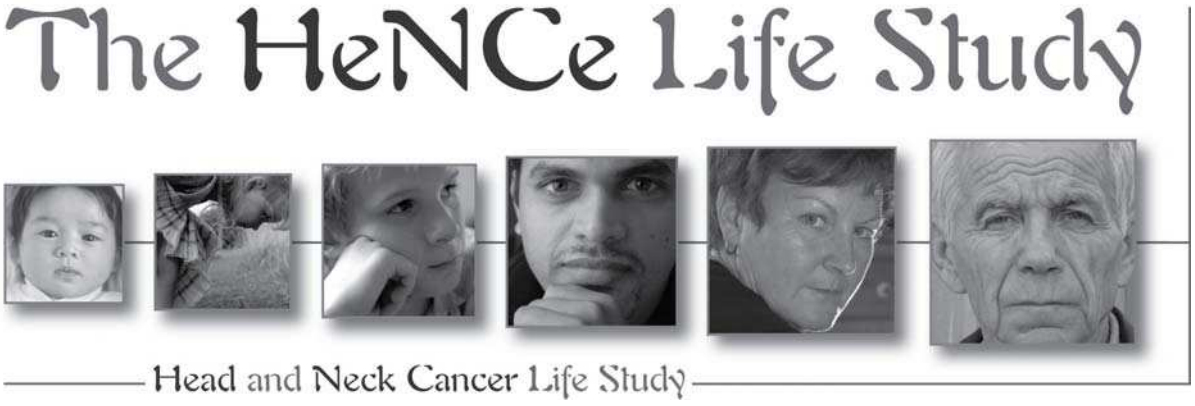


Fig 1 DAG used to identify sufficient set of potential confounders to adjust for in models

10.2 HeNCe Life study questionnaire

CONFIDENTIAL

MULTI CENTER STUDY OF ORAL CANCER: A LIFE COURSE APPROACH



UNIT OF EPIDEMIOLOGY & BIOSTATISTICS
INRS-INSTITUT ARMAND-FRAPPIER – LAVAL – CANADA

FACULTY OF DENTISTRY & DEPARTMENT OF EPIDEMIOLOGY
MCGILL UNIVERSITY – MONTREAL – CANADA

DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY
CENTRE DE RECHERCHE DU CHUM – MONTREAL - CANADA

DEPARTMENT OF EPIDEMIOLOGY AND POPULATION HEALTH
ALBERT EINSTEIN COLLEGE OF MEDICINE - NEW YORK – USA

HOSPITAL DO CÂNCER-DEPARTAMENTO DE CIRURGIA DE CABEÇA E PESCOÇO
SÃO PAULO - BRASIL

GOVERNMENT DENTAL COLLEGE –MEDICAL COLLEGE CAMPUS
KOZHIKODE – SOUTH INDIA

DEPARTMENT OF CLINICAL VIROLOGY – CHRISTIAN MEDICAL COLLEGE
VELLORE – SOUTH INDIA

2010

TABLE OF CONTENTS

A. MEDICAL INFORMATION

Interviewer Reminder: Prior to interview, obtain information below from research file or medical records.

Identification Number..... 0 2 - [ ][ ][ ] - [ ]
Country **Participant**

Country: (01) Brazil (04) United Kingdom
(02) Canada (05) India
(03) South Africa

Medical file N°..... [ ][ ][ ][ ][ ][ ][ ][ ][ ][ ]

A1 Status.....
(01) Case (02) Control

A2 Subject's Initials (Surname, Name).....

A3 Hospital / recruitment site.....
(01) Jewish General Hospital MUHC (02) Hôpital Notre-Dame CHUM
(03) Montreal General Hospital (04) Royal Victoria Hospital

A4 Control Department *(Code 88 for cases)*.....
(01) Neurology (04) Rheumatology (07) Urology
(02) Ear, Nose, Throat (05) Orthopaedics (08) Other, specify:
(03) Endocrinology (06) Gastroenterology _____

For controls only:

A5 Main Diagnosis for being seen at this department (LC)..... [ ][ ][ ] - [ ]
Condition description: _____ (I.C.D.10)

For cases only:

A6 Cancer site.....
(01) Pharynx (C146,148,149) (02) Larynx (C161) (03) Oral cavity (C141,143,144,145)

A7 Global TNM stage T_____ N_____ M____ ➔ Global Staging **(LC)** _____

A8 Date of Diagnosis..... [ ][ ] - [ ][ ] - [ ][ ][ ][ ]
(99-99-9999) Don't know Day Month Year

A9 Time since Diagnosis (months).....

Initial treatment modality(ies)

A10 Surgery..... | 8 | 8 |
(01) No (02) Yes

A11 Date of surgery..... | 8 | 8 | - | 8 | 8 | - | 8 | 8 |
Day Month Year

A12 Radiotherapy..... | 8 | 8 |
(01) No (02) Yes

A13 Date of radiotherapy..... | 8 | 8 | - | 8 | 8 | - | 8 | 8 |
Day Month Year

A14 Chemotherapy..... | 8 | 8 |
(01) No (02) Yes

A15 Date of chemotherapy..... | 8 | 8 | - | 8 | 8 | - | 8 | 8 |
Day Month Year

For all subjects:

A16 Initials of the person who collected the medical data (Surname, Name)..... | | |

A17 Date medical data collected..... | | | - | | | - | | | | |
(99-99-9999) Don't know Day Month Year

B. GENERAL INFORMATION

B1 Date of Interview..... [ ][ ] - [ ][ ] - [ ][ ][ ][ ]
Day Month Year

B2 Time of beginning of Interview..... [ ][ ] - [ ][ ]
Hour **Minute**

B3 Interview..... [ ][ ]
(01) Original (02) Duplicate *(6-12 weeks later)* (03) Duplicate *(+12 weeks later)*

B4 Sex..... [ ][ ]
(01) Female (02) Male

Interviewer Reminder: **Present life grid here**. See instructions in guidebook.

B5 What is your date of birth?..... [ ][ ] - [ ][ ] - [ ][ ][ ][ ]
(99-99-9999) Don't know Day Month Year

B6 How old are you?..... [ ][ ][ ]

B7 Do you consider yourself living in a rural (farm) or an urban (city) area?... [ ][ ]
(01) Urban (02) Rural (GO TO B9)

B8 What city do you live in? (LC)..... [ ][ ]
Name of City: _____ Postal Code: _____ - _____

Interviewer Reminder: Confirm name of city from list of codes. Rural area is in the farm

B9 How many years have you been living there? (Last consecutive years)..... [ ][ ]
(00) Less than one year

B10 In which city / place did you live in just before?(LC)..... [ ][ ]
Name of city: _____ Postal Code: _____ - _____
(00) Rural area

B11 Were you born in a rural (farm) or an urban (city) area?..... [ ][ ]
(01) Urban (02) Rural (GO TO B13)

B12 In what city were you born in? (LC)..... [ ][ ]
Name of city: _____ Postal Code: _____ - _____
(00) Other country

B13 How many years did you live there?..... [ ][ ]
(00) Less than one year

B14 In this list, which group best describe you? [ ][ ]

(01) White (Caucasian)
(02) Black
(03) Asian Indian
(04) Asian Pakistani
(05) Asian Bangladeshi
(06) Chinese
(07) Mixed ethnic group
(08) Aboriginal
(09) Other, specify: _____ [ ][ ]

B15 To which of these religions do you identify with? [ ][ ]

(00) None (GO TO B18)
(01) Muslim
(02) Christian
(03) Jewish
(04) Buddhist
(05) Hindu
(06) Other, specify: _____ [ ][ ]

B16 Do you practice this religion?..... [ ][ ]

(00) No (GO TO B18) (01) Yes

B17 How old were you when you started practicing this religion?..... [ ][ ]

(00) My whole life

B18 *For the interviewer:* **(Language used in this questionnaire)?**..... [ ][ ]

(01) English
(02) French

C. EDUCATION

This section is about your education. Firstly,

C1 Did you ever attend school?..... ☐☐
(00) No (GO TO SECTION D)
(01) No, but I can read and write (GO TO SECTION D)
(02) Yes

Let's start by looking at when you started school, when you stopped and interruptions in between. We will use this grid to help us out. I will ask you more specific questions about your education afterwards.

Interviewer Reminder: Collect general information using the **life grid**, referring to it later when asking questions C2 through C9.
- Situate years of **formal** education i.e. that were successfully completed at school.
- Do **NOT** consider regular interruptions (ex.: summer time) or kindergarten. But DO consider interruptions for medical reasons, evacuations, etc...

C2 How many years of formal education do you have? (Subtract years failed)..... ☐☐

C3 What was the highest degree or qualification that you obtained?..... ☐☐
(00) None (GO TO C5)
(01) Elementary / primary school
(02) High school
(03) Technical qualification
(04) CEGEP (non-technical)
(05) University
(06) Post-graduate

C4 How old were you when you obtained this degree?..... ☐☐
(99) Don't know

C5 Have you ever failed a school year?..... ☐☐
(00) No
(01) Yes, once
(02) Yes, twice
(03) Yes, 3 or more times

C6 Have you ever interrupted your full time education?..... ☐☐
(00) No (GO TO SECTION D)
(01) Yes

C7 How many years of formal education did you have when you FIRST interrupted your full time education?..... ☐☐

C8 How old were you when you FIRST interrupted you full time education?..... ☐☐

D. OCCUPATIONS & EMPLOYMENT

In this section I would like to ask you a few questions about jobs you may have had.

Interviewer Reminder: A job is a **continuous period of time of ONE YEAR OR MORE working and paid by the same employer** even though the participant may have had different positions during that period. If the participant was self-employed, a job is considered to be a period of time doing the same type of self-employed work.

D1 Have you ever had a job in your life?..... [ ][ ]
(00) No (GO TO SECTION E) (01) Yes

D2 Which of the options below best describes your work situation in the past 7 days?..... [ ][ ]
(01) Full time work (30+ hours / week)
(02) Part time work (< 30 hours / week)
(03) Unemployed
(04) Fully retired from work
(05) Permanently sick or disabled
(06) On sick leave
(07) Other, specify: _____ [ ][ ]

Let's look at the different jobs you've had, the different positions you may have held and the times you may have been unemployed. Again, we will use this grid to help us out and refer to it for the specific questions I will have afterwards.

Interviewer Reminder: Collect general information using the **life grid**, referring to it later when asking questions D2 through D112.

- Going back to an old employer, even if for more than one year, is considered to be another, separate job and should be counted as such.
- Seasonal work (6 months full time) done for more than 2 years in a row counts as 1 job.
- Include army service IF paid or compensated for.
- Include "informal work", i.e. direct selling, itinerant seller, undeclared work. Count different contracts, odd jobs, etc… as one job IF done continuously over at least one year. Subject should consider all different work related activities in this period as a whole whilst describing this job through the related questions.
- Mark periods of unemployment on life grid (refer to description in box below).

Do **NOT** include:

- Summer or holiday time jobs while at school or full-time education.
- Part-time jobs done at the same time as full-time education.
- Part time jobs done at the same time as a full-time job.

D3 Since you started working how many jobs have you had?..... [ ][ ]
(01) (02) (03) (04) (05) (06) (07) (08) (09 or more)

Interviewer Reminder: Unemployment means being out of a work for **at least 3 MONTHS**. You do not have to be registered as unemployed BUT you must be enabled to work and actively or passively looking for work.

Do **NOT** include:
- Holidays or vacations while attending full-time education
- Interruptions due to seasonal work
- Maternity leaves, Sabbatical leaves
- Deliberate choice to exclude oneself from the workforce, i.e. living on inheritance, housewife...

D4 Since you started working how many times have you been unemployed?.... ☐☐
(00) None (GO TO D6)
(01) (02) (03) (04) (05) (06) (07) (08) (09 or more)

D5 Please describe the longest periods of your life in which you were unemployed.

From age?	**To age?**	**# Months**
☐☐	☐☐	☐☐
☐☐	☐☐	☐☐
☐☐	☐☐	☐☐
☐☐	☐☐	☐☐
☐☐	☐☐	☐☐

FIRST JOB

Interviewer Reminder: Confirm which job is 1st job with life grid.

I would like to ask you a few questions about your **first job**. So,

D6 You were doing that job...

From age? ___ **To age?** ___ **i.e. # Years** ___

D7 Did you occupy different positions at that job?..... ___
(00) No (Fill in FIRST column only) (01) Yes

D8 Please describe your job / different positions (LC)..... **FIRST** ___ **LAST** ___

FIRST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

LAST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

D9 What did the company you worked for specialise in? (LC)..... ___
_____

Interviewer Reminder: Confirm job / position code with list of codes for Q D8 and D9.

D10 Were you an employee or self-employed?..... ___ ___
(01) Employee (02) Self-employed (GO TO D12)

D11 Were you an employee? (GO TO D13)..... ___ ___
(01) Not supervising others
(02) Foreman, supervisor, team leader
(03) Manager: Firm of <25 employees
(04) Manager: Firm of 25+ employees
(05) Professional

D12 Were you self employed?.....

(01) Without incorporated business
(02) With incorporated business but without employees other than family members
(03) With <25 employees
(04) With 25+ employees
(05) Professional

D13 Did you work?.....

(01) Full time (30 hours +) (02) Part time (<30 hours)

D14 How many hours a week?.....

D15 How much were you paid PER YEAR at that time?.....

FIRST: $

LAST: $

Describe: _____

- Calculate average amount in Canadian dollars
- Average: hourly rate x 35 hours x 50 weeks OR Min + Max / # yrs, prorated
- Self-employed: average earnings per year as per income tax declarations if submitted

Now I would like to ask you a few questions about work environmental hazards. Consider your job in general, regardless of the different positions you may have occupied.

D16 Did your work often involve exposure to chemical hazards such as dust, oils, solvents or thinners, smoke, gas, etc…?.....

(00) No (GO TO D24) (01) Yes (99) Don't know (GO TO D24)

Did it involve exposure to...?

D17 Dust (Silica dust, saw dust, sanding dust, epoxy-resins, welding...).....

(00) No (01) Yes

D18 Oils (Mineral oil, lubricants...).....

(00) No (01) Yes

D19 Solvents or thinners (acetone, paint thinners, chlorinated solvent (trichloroethylene), solvent of cellulose...).....

(00) No (01) Yes

D20 Smoke (Gas from motors, coal, wood, rubber...).....

(00) No (01) Yes

D21 Gas (Oxygen, ammonia...).....

(00) No (01) Yes

D22 Did your work involve working with substances such as: asphalt, alcohol, gasoline, glue, mercury, kerosene, etc?.....

(00) No (01) Yes

D23 Did your work <u>often</u> involve exposure to other chemicals?.....

(00) No (01) Yes, specify (ex.: cigarette smoke): _____

_____

D24 Did your work <u>often</u> involve exposure to physical hazards such as humidity, high temperatures, pressure (physiological), electro-magnetic radiations, etc…?.....

(00) No (GO TO **<u>Interviewer Reminder</u>** preceding D30) (01) Yes

(99) Don't know (GO TO **<u>Interviewer Reminder</u>** preceding D30)

Did it involve exposure to...

D25 Humidity?.....

(00) No (01) Yes

D26 High temperatures?.....

(00) No (01) Yes

D27 Pressure (physiological; ex.: loud noise, underwater work, gravity changes)?.....

(00) No (01) Yes

D28 Electromagnetic radiations (x-rays, microwaves, radioactive substances)?

(00) No (01) Yes

D29 Did your work <u>often</u> involve exposure to other physical hazards?.....

(00) No (01) Yes, specify: _____

_____

<u>Interviewer Reminder</u>: If D16 **<u>OR</u>** D24 are (01) Yes, then ask D30. If not, go to D31.

D30 Did you use any kind of protection for chemical / physical hazards?.....

(00) No (02) Yes, sometimes

(01) Yes, most of the time (03) Yes, rarely

D31 Was your first job the same one as your longest job?.....

(00) No (01) Yes, the same one as my longest job (GO TO D58)

(02) Yes, the same one my whole life (GO TO SECTION E)

LONGEST JOB

Now I would like to ask you some questions about your **longest job**. I will be using the same set of questions I used in the previous section. So,

Interviewer Reminder: Confirm which job is longest job with life grid.

D32 You were doing that job...

From age? ___ **To age?** ___ **i.e. # Years** ___

D33 Did you occupy different positions at that job?..... ___
(00) No (Fill in FIRST column only) (01) Yes

D34 Please describe your job / different positions (LC)..... **FIRST** ___ **LAST** ___

FIRST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

LAST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

D35 What did the company you worked for specialise in? (LC)..... ___
_____

Interviewer Reminder: Confirm job / position code with list of codes for Q D34 and D35.

D36 Were you an employee or self-employed?..... ___ ___
(01) Employee (02) Self-employed (GO TO D39)

D37 Were you an employee...?..... ___ ___
(01) Not supervising others
(02) Foreman, supervisor, team leader
(03) Manager: Firm of <25 employees
(04) Manager: Firm of 25+ employees
(05) Professional

D38 Were you self employed?..... ___

(01) Without incorporated business
(02) With incorporated business but without employees other than family members
(03) With <25 employees
(04) With 25+ employees
(05) Professional

D39 Did you work...?..... ___

(01) Full time (30 hours + / week) (02) Part time (<30 hours / week)

D40 How many hours a week?..... ___

D41 How much were you paid PER YEAR at that time?.....

FIRST: $ ___

LAST: $ ___

Describe: _____

- Calculate average amount in Canadian dollars
- Average: hourly rate x 35 hours x 50 weeks OR Min + Max / # yrs, prorated
- Self-employed: average earnings per year as per income tax declarations if submitted

Now I would like to ask you a few questions about work environmental hazards. Consider your job <u>in general</u>, regardless of the different positions you may have occupied.

D42 Did your work <u>often</u> involve exposure to chemical hazards such as dust, oils, solvents or thinners, smoke, gas, etc…?..... ___

(00) No (GO TO D50) (01) Yes (99) Don't know (GO TO D50)

Did it involve exposure to...?

D43 Dust (Silica dust, saw dust, sanding dust, epoxy-resins, welding...)..... ___

(00) No (01) Yes

D44 Oils (Mineral oil, lubricants...)..... ___

(00) No (01) Yes

D45 Solvents or thinners (acetone, paint thinners, chlorinated solvent (trichloroethylene), solvent of cellulose...)..... ___

(00) No (01) Yes

D46 Smoke (Gas from motors, coal, wood, rubber...)..... ___

(00) No (01) Yes

D47 Gas (Oxygen, ammonia...)..... ___

(00) No (01) Yes

D48 Did your work involve working with substances such as: asphalt, alcohol, gasoline, glue, mercury, kerosene, etc?.....

(00) No (01) Yes

D49 Did your work <u>often</u> involve exposure to other chemicals?.....

(00) No (01) Yes, specify (ex.: cigarette smoke): _____

_____

D50 Did your work <u>often</u> involve exposure to physical hazards such as humidity, high temperatures, pressure (physiological), electro-magnetic radiations, etc…?.....

(00) No (GO TO **<u>Interviewer Reminder</u>** preceding D56) (01) Yes

(99) Don't know (GO TO **<u>Interviewer Reminder</u>** preceding D56)

Did it involve exposure to...

D51 Humidity?.....

(00) No (01) Yes

D52 High temperatures?.....

(00) No (01) Yes

D53 Pressure (physiological; ex.: loud noise, underwater work, gravity changes)?.....

(00) No (01) Yes

D54 Electromagnetic radiations (x-rays, microwaves, radioactive substances)?

(00) No (01) Yes

D55 Did your work <u>often</u> involve exposure to other physical hazards?.....

(00) No (01) Yes, specify: _____

_____

<u>Interviewer Reminder</u>: If D42 **<u>OR</u>** D50 are (01) Yes, then ask D56. If not, go to D57.

D56 Did you use any kind of protection for chemical / physical hazards?.....

(00) No (02) Yes, sometimes

(01) Yes, most of the time (03) Yes, rarely

D57 Was your longest job the same one as your latest or current job?.....

(00) No

(01) Yes, the same one as my latest / current job (GO TO SECTION E)

LAST / LATEST JOB

Finally about your last / latest job...

Interviewer Reminder: Confirm which job is last/latest job with life grid.

D58 You were doing that job...

From age? [ ][ ] **To age?** [ ][ ] **i.e. # Years** [ ][ ]

D59 Did you occupy different positions at that job?..... [ ][ ]
(00) No (Fill in FIRST column only) (01) Yes

FIRST **LAST**

D60 Please describe your job / different positions (LC)..... [ ][ ][ ] [ ][ ][ ]

FIRST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

LAST POSITION

Job Title: _____
Work environment: _____
Most frequent tasks: _____
_____
_____

D61 What did the company you worked for specialise in? (LC)..... [ ][ ][ ]
_____

Interviewer Reminder: Confirm job / position code with list of codes for Q D60 and D61.

D62 Were you an employee or self-employed?..... [ ][ ] [ ][ ]
(01) Employee (02) Self-employed (GO TO D66)

D63 Were you an employee...?..... [ ][ ] [ ][ ]
(01) Not supervising others
(02) Foreman, supervisor, team leader
(03) Manager: Firm of <25 employees
(04) Manager: Firm of 25+ employees
(05) Professional

D64 Were you self employed...?.....

(01) Without incorporated business
(02) With incorporated business but without employees other than family members
(03) With <25 employees
(04) With 25+ employees
(05) Professional

D65 Did you work...?.....

(01) Full time (30 hours + / week)
(02) Part time (<30 hours / week)

D66 How many hours a week?.....

D67 How much were you paid PER YEAR at that time?.....

FIRST: $

LAST: $

Describe: _____

- Calculate average amount in Canadian dollars
- Average: hourly rate x 35 hours x 50 weeks OR Min + Max / # yrs, prorated
- Self-employed: average earnings per year as per income tax declarations if submitted

Now I would like to ask you a few questions about work environmental hazards. Consider your job in general, regardless of the different positions you may have occupied.

D68 Did your work often involve exposure to chemical hazards such as dust, oils, solvents or thinners, smoke, gas, etc…?.....

(00) No (GO TO D76)
(01) Yes
(99) Don't know (GO TO D76)

Did it involve exposure to...?

D69 Dust (Silica dust, saw dust, sanding dust, epoxy-resins, welding...).....

(00) No
(01) Yes

D70 Oils (Mineral oil, lubricants...).....

(00) No
(01) Yes

D71 Solvents or thinners (acetone, paint thinners, chlorinated solvent (trichloroethylene), solvent of cellulose...).....

(00) No
(01) Yes

D72 Smoke (Gas from motors, coal, wood, rubber...).....

(00) No
(01) Yes

D73 Gas (Oxygen, ammonia...).....

(00) No
(01) Yes

D74 Did your work involve working with substances such as: asphalt, alcohol, gasoline, glue, mercury, kerosene, etc?..... [ ][ ]
(00) No (01) Yes

D75 Did your work <u>often</u> involve exposure to other chemicals?..... [ ][ ]
(00) No (01) Yes, specify (ex.: cigarette smoke): _____ [ ][ ]
_____

D76 Did your work <u>often</u> involve exposure to physical hazards such as humidity, high temperatures, pressure (physiological), electro-magnetic radiations, etc…?..... [ ][ ]
(00) No (GO TO **<u>Interviewer Reminder</u>** preceding D85) (01) Yes
(99) Don't know (GO TO **<u>Interviewer Reminder</u>** preceding D85)

Did it involve exposure to...

D77 Humidity?..... [ ][ ]
(00) No (01) Yes

D78 High temperatures?..... [ ][ ]
(00) No (01) Yes

D79 Pressure (physiological; ex.: loud noise, underwater work, gravity changes)?..... [ ][ ]
(00) No (01) Yes

D80 Electromagnetic radiations (x-rays, microwaves, radioactive substances)?..... [ ][ ]
(00) No (01) Yes

D81 Did your work <u>often</u> involve exposure to other physical hazards?..... [ ][ ]
(00) No (01) Yes, specify: _____ [ ][ ]
_____

<u>Interviewer Reminder</u>: If D68 **<u>OR</u>** D76 are (01) Yes, then ask D82. If not, GO TO SECTION E.

D82 Did you use any kind of protection for chemical / physical hazards?..... [ ][ ]
(00) No (02) Yes, sometimes
(01) Yes, most of the time (03) Yes, rarely

E. HOUSING CONDITIONS & RESIDENTIAL ENVIRONMENT

In this section I would like to ask you a few questions about your housing conditions and residential environment at different times in your life. We will use the grid first to look at the different addresses you lived at, noting the times you moved from one place to another.

Interviewer Reminder: Collect general information using the **life grid**, referring to it later when asking questions E1 through E181.

- **An address is a place where the participant lived for at least 1 YEAR.**
- Moving back to an old address *within the same time period* is considered to be a separate place of residence and should be counted as such as long as it is for at least one more year.
- Moving back to an old address in *another time period* is always considered a separate place of residence as long as it is for a longer period of time than previously.
- If an address overlaps two time periods, consider it the main residence in a period only if the participant lived there for the longest time.
- If "boarding school" (E9), answers should pertain to the residence when child was back home.
- If person changed living place many times within the same year or over many years (ex.: gypsies, travelers, musicians touring, homeless) do not count any addresses. Rather, record the number of years spent with this housing pattern in E2, E4 and E6. If this pattern is present for the longest time in one period of life, note age span for that period and answer (06) to E9.

E1 Up until you were 16 years old (incl.) at how many *different* addresses did you live?
(01) (GO TO E3)... (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |

E2 Up until you were 16 years old (incl.) how many times (total) did you spend changing living places more than once in the same year?
(00) (01) (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |

E3 Between the ages of 17 and 30 (incl.) at how many *different* addresses did you live?
(01) (GO TO E5)… (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |

E4 Between the ages of 17 and 30 (incl.) how many times (total) did you spend changing living places more than once in the same year?
(00) (01) (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |

E5 From the age of 30 (excl.) until today at how many *different* addresses did you live?
(01) (GO TO E7)… (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |
If the respondent is less than 30 years old, mark (88) and GO TO E7

E6 From the age of 30 (excl.) until today how many times (total) did you spend changing living places more than once in the same year?
(00) (01) (02) (03) (04) (05) (06) (07) (08) (09 or more)..... | |

CHILDHOOD RESIDENCE

I would like to ask you a few questions about the residence / home in which you lived **for the longest time during your childhood**. By childhood I mean up to age 16 (incl.).

Interviewer Reminder: Identify and confirm longest residence in childhood using the life grid.

E7 You lived at that place...?

From age?	**To age?**	**i.e. # Years**

E8 Do you remember what the POSTAL CODE is for this residence? _____ - _____

For all the following questions, refer to the situation that was present "MOST OF THE TIME" while living in that residence.

Interviewer Reminder: Immediate family means: husband / wife & children and extended family means mother, father & own family.

E9 What type of setting were you living in at that place?.....

(01) With immediate family
(02) With extended family
(03) Foster home (GO TO E43)
(99) Don't know
(04) Boarding school, monastery (GO TO E43)
(05) Institution (ex.: psychiatric hospital, rehabilitation centre) (GO TO E43)
(06) Pattern of many different living places (GO TO E43)
(07) Other, specify: _____

E10 Who was the owner of the place?.....

(01) My family or a member of my family
(02) State or municipality
(03) Private owners / company (renting)
(04) Other, specify: _____
(99) Don't know

E11 How many people lived in the household? (At once, for the longest period of time)..
(Include borders, live-in maids, roommates...) (99) Don't know

Interviewer Reminder:
QE11: Include people who were permanent residents and those who were living in the house for the longest period of time.
QE12: Rooms include: kitchen, living room, dining room, bedroom, furnished basement. Do **NOT** include: toilet, bathroom, laundry room, hallway, garage, patio.

E12 How many rooms did your place have? (If renovated, count # rooms during longest period living there)..... | | |

(99) Don't know

E13 Were some or all of these rooms damp / humid / wet? (For example: wallpaper peels of wall, mould grows on internal walls, clothes stem when aired after storage)..... | | |

(00) No (01) Yes, all (02) Yes, some (99) Don't know

Now, I will read a list of facilities you may have had in the place where you lived. We would like to know **which of these facilities were present inside your childhood residence.**

E14 Did your home have a bathroom (indoor toilet, bath and/or shower)?..... | | |

(00) No (GO TO E16) (01) Yes (99) Don't know (GO TO E16)

E15 How many?..... | | |

E16 Did your home have a sewage system?..... | | |

(00) No (02) Yes, a septic tank

(01) Yes, a central public system (99) Don't know

E17 Did your home have running cold water?..... | | |

(00) No (02) Yes, an independent one (rural) i.e. outside the house

(01) Yes, a central public system (urban) i.e. inside the house (99) Don't know

E18 Did your home have electricity?..... | | |

(00) No (02) Yes, by a generator / battery only

(01) Yes, by a central system (99) Don't know

E19 Did your home have running hot water?..... | | |

(00) No (01) Yes (99) Don't know

E20 Did your house have a wood (or coal) stove?..... | | |

(00) No (GO TO E26) (01) Yes (99) Don't know (GO TO E26)

E21 Was the stove located inside the house?..... | | |

(00) No (GO TO E26) (01) Yes (99) Don't know (GO TO E26)

E22 Was the stove located in an area with any ventilation / windows?..... | | |

(00) No (01) Yes (99) Don't know

E23 Did the stove have a chimney?..... | | |

(00) No (01) Yes (99) Don't know

E24 How often did you use the stove to <u>cook</u>?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week (05) Only during the winter

E25 How often did you use the stove to <u>heat</u> your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

E26 Did you use any other kind of method to <u>heat</u> your home?.....

(00) No (GO TO E30) (01) Yes

E27 What kind of material did you use?.....

(01) Electricity (03) Gas (05) Wood (06) Other, specify: _____
(02) Petrol (04) Coal (99) Don't know

E28 In what kind of appliance was this material used?.....

(01) Furnace with chimney (05) Fireplace with chimney
(02) Furnace without chimney (06) Baseboards
(03) Open fire (07) Radiators
(04) Fireplace without chimney (08) Other, specify: _____
(99) Don't know

E29 How often did you use this method to heat your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

I will now read a **list of household goods** you may have had in your childhood residence or not. You may find that some of these appliances were not applicable to the epoch you were a child. Chose the answer that best represents your situation, regardless.

E30 Did your place have a refrigerator?.....

(00) No, it had no appliance to cool food (02) Yes
(01) No, it had an ice box (99) Don't know

E31 Did your place have a radio?.....

(00) No (01) Yes (99) Don't know

E32 Did your place have a TV?.....

(00) No (02) Yes, color
(01) Yes, black and white (99) Don't know

E33 Did your place have a machine to wash clothes (inside own dwelling)**?**.....

(00) No, it had no appliance to wash clothes (02) Yes
(01) Yes, it had a clothes ringer (99) Don't know

E34 Did your place have a system to play recorded music?.....

(00) No, it had nothing to play recorded music
(01) Yes, it had a gramophone
(02) Yes, a record player
(03) Yes, a cassette player
(04) Yes, a CD player
(99) Don't know

E35 Did your place have a vacuum cleaner?.....

(00) No, it had no appliance to vacuum
(01) No, it had a non-electric device to vacuum
(02) Yes
(99) Don't know

E36 Did your place have a VCR?.....

(00) No, it had no appliance to watch recorded images
(01) No, it had a less sophisticated image viewing machine
(02) Yes (VCR or DVD)
(99) Don't know

E37 Did your place have a computer?.....

(00) No, that did not exist at the time (01) No (02) Yes (99) Don't know

Also, I would like to ask you...

E38 Did your household have a car?.....

(00) No (GO TO E40) (01) Yes (99) Don't know (GO TO E40)

E39 How many?.....

Finally, I would like to ask you a few questions about the **residential area** where you lived during your childhood. Could you tell me **how common** each of these situations were in your neighbourhood... (Use Answer Sheet)

(00) Not common (01) Common (02) Very common (99) Don't know

E40 Noise from neighbouring apartments, streets, trains, airplanes, industry, etc….

E41 Smoke, dust or smell from industry, traffic, sewage or from other sources.....

E42 Cigarette, cigar and/or pipe smoke from residents in this household.....

LONGEST RESIDENCE IN EARLY ADULT LIFE (17-30 yrs)

Now I would like to ask you a few questions about the residence / home in which you lived **for the longest time during your early adult life, that is between the ages of 17 (incl.) and 30 (incl.)**. I will use the same set of question I used in the previous sections.

Interviewer Reminder: Identify / confirm longest residence in early adulthood using life grid.

E43 Is this residence the same one as your childhood residence?..... ☐☐
(00) No (01) Yes (GO TO E80)

E44 You lived at that place...?

From age?	**To age?**	**i.e. # Years**
☐☐	☐☐	☐☐

E45 Do you remember what the POSTAL CODE is for this residence? _____ - _____

For all the following questions, refer to the situation that was present "MOST OF THE TIME" while living in that residence.

E46 What type of setting were you living in at that place?..... ☐☐
(01) With immediate family / alone
(02) With extended family
(03) Foster home (GO TO E80)
(99) Don't know
(04) Boarding school, monastery (GO TO E80)
(05) Institution (ex.: psychiatric hospital, rehabilitation centre) (GO TO E80)
(06) Pattern of many different living places (GO TO E80)
(07) Other, specify: _____ ☐☐

E47 Who was the owner of the place?..... ☐☐
(00) Myself (even if bought after renting)
(01) My family or a member of my family
(02) State or municipality
(03) Private owners / company (renting)
(04) Other, specify: _____ ☐☐
(99) Don't know

E48 How many people lived in the household? (At once, for the longest period of time).. ☐☐
(Include borders, live-in maids, roommates...) (99) Don't know

Interviewer Reminder:
QE48: Include people who were permanent residents and those who were living in the house for the longest period of time.
QE49: Rooms include: kitchen, living room, dining room, bedroom, furnished basement. Do **NOT** include: toilet, bathroom, laundry room, hallway, garage, patio.

E49 How many rooms did your place have? (If renovated, count # rooms during longest period living there)..... | | |

(99) Don't know

E50 Were some or all of these rooms damp / humid / wet? (For example: wallpaper peels of wall, mould grows on internal walls, clothes stem when aired after storage)..... | | |

(00) No (01) Yes, all (02) Yes, some (99) Don't know

Now, I will read a list of facilities you may have had in the place where you lived. We would like to know **which of these facilities were present inside your early adulthood residence.**

E51 Did your home have a bathroom (indoor toilet, bath and/or shower)?..... | | |

(00) No (GO TO E53) (01) Yes (99) Don't know (GO TO E53)

E52 How many?..... | | |

E53 Did your home have a sewage system?..... | | |

(00) No (02) Yes, a septic tank

(01) Yes, a central public system (99) Don't know

E54 Did your home have running cold water?..... | | |

(00) No (02) Yes, an independent one (rural) i.e. outside the house

(01) Yes, a central public system (urban) i.e. inside the house (99) Don't know

E55 Did your home have electricity?..... | | |

(00) No (02) Yes, by a generator / battery only

(01) Yes, by a central system (99) Don't know

E56 Did your home have running hot water?..... | | |

(00) No (01) Yes (99) Don't know

Could you please tell me…

E57 Did your house have a wood (or coal) stove?..... | | |

(00) No (GO TO E63) (01) Yes (99) Don't know (GO TO E63)

E58 Was the stove located inside the house?..... | | |

(00) No (GO TO E63) (01) Yes (99) Don't know (GO TO E63)

E59 Was the stove located in an area with any ventilation / windows?..... | | |

(00) No (01) Yes (99) Don't know

E60 Did the stove have a chimney?..... | | |

(00) No (01) Yes (99) Don't know

E61 How often did you use the stove to <u>cook</u>?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week (05) Only during the winter

E62 How often did you use the stove to <u>heat</u> your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

E63 Did you use any other kind of method to <u>heat</u> your home?.....

(00) No (GO TO E67) (01) Yes

E64 What kind of material did you use?.....

(01) Electricity (03) Gas (05) Wood (06) Other, specify: _____
(02) Petrol (04) Coal (99) Don't know

E65 In what kind of appliance was this material used?.....

(01) Furnace with chimney (05) Fireplace with chimney
(02) Furnace without chimney (06) Baseboards
(03) Open fire (07) Radiators
(04) Fireplace without chimney (08) Other, specify: _____
(99) Don't know

E66 How often did you use this method to heat your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

I will now read a **list of household goods** you may have had in your early adulthood residence or not. You may find that some of these appliances were not applicable to the epoch you were 17 to 30 years old. Chose the answer that best represents your situation, regardless.

E67 Did your place have a refrigerator?.....

(00) No, it had no appliance to cool food (02) Yes
(01) No, it had an ice box (99) Don't know

E68 Did your place have a radio?.....

(00) No (01) Yes (99) Don't know

E69 Did your place have a TV?.....

(00) No (02) Yes, color
(01) Yes, black and white (99) Don't know

E70 Did your place have a machine to wash clothes (inside own dwelling)**?**.....

(00) No, it had no appliance to wash clothes (02) Yes
(01) Yes, it had a clothes ringer (99) Don't know

E71 Did your place have a system to play recorded music?.....

(00) No, it had nothing to play recorded music
(01) Yes, it had a gramophone
(02) Yes, a record player
(03) Yes, a cassette player
(04) Yes, a CD player
(99) Don't know

E72 Did your place have a vacuum cleaner?.....

(00) No, it had no appliance to vacuum
(01) No, it had a non-electric device to vacuum
(02) Yes
(99) Don't know

E73 Did your place have a VCR?.....

(00) No, it had no appliance to watch recorded images
(01) No, it had a less sophisticated image viewing machine
(02) Yes (VCR or DVD)
(99) Don't know

E74 Did your place have a computer?.....

(00) No, that did not exist at the time (01) No (02) Yes (99) Don't know

Also, I would like to ask you...

E75 Did your household have a car?.....

(00) No (GO TO E77) (01) Yes (99) Don't know (GO TO E77)

E76 How many?.....

Finally, here are a few questions about the **residential area** where you lived during your early adulthood. **How common** was it in your neighbourhood to have... (Use Answer Sheet)

(00) Not common (01) Common (02) Very common (99) Don't know

E77 Noise from neighbouring apartments, streets, trains, airplanes, industry, etc….

E78 Smoke, dust or smell from industry, traffic, sewage or from other sources.....

E79 Cigarette, cigar and/or pipe smoke from residents in this household.....

LONGEST RESIDENCE IN LATER ADULTHOOD (30 yrs +)

Now let's talk about your **longest residence in later adulthood**, that is after age 30 (excl.).

Interviewer Reminder: Identify / confirm longest residence in later adulthood using life grid.

E80 Is this residence the same one as the residence you lived in for the longest time between the ages of 17 and 30 or your childhood residence?..... ☐☐

(00) No (01) Yes, same as longest residence between ages of 17-30 (GO TO SECTION F)
(02) Yes, same as childhood residence (GO TO SECTION F)
(03) Yes, same one in the three periods of my life (GO TO SECTION F)
(88) None of the above, ex.: subject is less than 30 yrs old (GO TO SECTION F)

E81 You lived at that place…?

From age? ☐☐ **To age?** ☐☐ **i.e. # Years** ☐☐

E82 Do you remember what the POSTAL CODE is for this residence? _____ - _____

For all the following questions, refer to the situation that was present "MOST OF THE TIME" while living in that residence.

E83What type of setting were you living in at that place?..... ☐☐

(01) With immediate family / alone
(02) With extended family
(03) Foster home (GO TO SECTION F)
(04) Boarding school, monastery (GO TO SECTION F)
(05) Institution (ex.: psychiatric hospital, rehabilitation centre) (GO TO SECTION F)
(06) Pattern of many different living places (GO TO SECTION F)
(07) Other, specify: _____ ☐☐
(99) Don't know

E84 Who was the owner of the place?..... ☐☐

(00) Myself (even if bought after renting)
(01) My family or a member of my family
(02) State or municipality
(03) Private owners / company (renting)
(04) Other, specify: _____ ☐☐
(99) Don't know

E85 How many people lived in the household? (At once, for the longest period of time).. ☐☐
(Include borders, live-in maids, roommates…) (99) Don't know

Interviewer Reminder:
QE85: Include people who were permanent residents and those who were living in the house for the longest period of time.
QE86: Rooms include: kitchen, living room, dining room, bedroom, furnished basement.
Do **NOT** include: toilet, bathroom, laundry room, hallway, garage, patio.
If renovated, count # rooms during longest period living there.

E86 How many rooms did your place have? (If renovated, count # rooms during longest period living there)..... | | |

(99) Don't know

E87 Were some or all of these rooms damp / humid / wet? (For example: wallpaper peels of wall, mould grows on internal walls, clothes stem when aired after storage)..... | | |

(00) No (01) Yes, all (02) Yes, some (99) Don't know

Now, I will read a list of facilities you may have had in the place where you lived. We would like to know **which of these facilities were present inside your later adulthood residence.**

E88 Did your home have a bathroom (indoor toilet, bath and/or shower)?..... | | |

(00) No (GO TO E90) (01) Yes (99) Don't know (GO TO E90)

E89 How many?..... | | |

E90 Did your home have a sewage system?..... | | |

(00) No (02) Yes, a septic tank

(01) Yes, a central public system (99) Don't know

E91 Did your home have running cold water?..... | | |

(00) No (02) Yes, an independent one (rural) i.e. outside the house

(01) Yes, a central public system (urban) i.e. inside the house (99) Don't know

E92 Did your home have electricity?..... | | |

(00) No (02) Yes, by a generator / battery only

(01) Yes, by a central system (99) Don't know

E93 Did your home have running hot water?..... | | |

(00) No (01) Yes (99) Don't know

Could you please tell me…

E94 Did your house have a wood (or coal) stove?..... | | |

(00) No (GO TO E100) (01) Yes (99) Don't know (GO TO E100)

E95 Was the stove located inside the house?..... | | |

(00) No (GO TO E100) (01) Yes (99) Don't know (GO TO E100)

E96 Was the stove located in an area with any ventilation / windows?..... | | |

(00) No (01) Yes (99) Don't know

E97 Did the stove have a chimney?..... | | |

(00) No (01) Yes (99) Don't know

E98 How often did you use the stove to <u>cook</u>?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week (05) Only during the winter

E99 How often did you use the stove to <u>heat</u> your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

E100 Did you use any other kind of method to <u>heat</u> your home?.....

(00) No (GO TO E104) (01) Yes

E101 What kind of material did you use?.....

(01) Electricity (03) Gas (05) Wood (06) Other, specify: _____
(02) Petrol (04) Coal (99) Don't know

E102 In what kind of appliance was this material used?.....

(01) Furnace with chimney (05) Fireplace with chimney
(02) Furnace without chimney (06) Baseboards
(03) Open fire (07) Radiators
(04) Fireplace without chimney (08) Other, specify: _____
(99) Don't know

E103 How often did you use this method to heat your home?.....

(00) Never (02) 5-6 times a week (04) 1-2 times a week
(01) Everyday (03) 3-4 times a week

I will now read a **list of household goods** you may have had in your later adulthood residence or not. You may find that some of these appliances were not applicable to the epoch you were in later adulthood. Chose the answer that best represents your situation, regardless.

E104 Did your place have a refrigerator?.....

(00) No, it had no appliance to cool food (02) Yes
(01) No, it had an ice box (99) Don't know

E105 Did your place have a radio?.....

(00) No (01) Yes (99) Don't know

E106 Did your place have a TV?.....

(00) No (02) Yes, color
(01) Yes, black and white (99) Don't know

E107 Did your place have a machine to wash clothes (inside own dwelling)**?**.....

(00) No, it had no appliance to wash clothes (02) Yes
(01) Yes, it had a clothes ringer (99) Don't know

E108 Did your place have a system to play recorded music?.....

(00) No, it had nothing to play recorded music
(01) Yes, it had a gramophone
(02) Yes, a record player
(03) Yes, a cassette player
(04) Yes, a CD player
(99) Don't know

E109 Did your place have a vacuum cleaner?.....

(00) No, it had no appliance to vacuum
(01) No, it had a non-electric device to vacuum
(02) Yes
(99) Don't know

E110 Did your place have a VCR?.....

(00) No, it had no appliance to watch recorded images
(01) No, it had a less sophisticated image viewing machine
(02) Yes (VCR or DVD)
(99) Don't know

E111 Did your place have a computer?.....

(00) No, that did not exist at the time (01) No (02) Yes (99) Don't know

Also, I would like to ask you...

E112 Did your household have a car?.....

(00) No (GO TO E114) (01) Yes (99) Don't know (GO TO E114)

E113 How many?.....

Finally, here are some questions about the **residential area** where you lived during your later adulthood. **How common** was it in your neighbourhood to have... (Use Answer Sheet)

(00) Not common (01) Common (02) Very common (99) Don't know

E114 Noise from neighbouring apartments, streets, trains, airplanes, industry, etc...

E115 Smoke, dust or smell from industry, traffic, sewage or from other sources.....

E116 Cigarette, cigar and/or pipe smoke from residents in this household.....

Country: 0 2 - ID N°: _ _ _ - _

F. SMOKING AND CHEWING HABITS

Now I would like to ask you some questions about your smoking and/or chewing habits.

F1 Have you ever smoked in your life? (or chewed, any product, any amount)..... _ _
(00) Never (GO TO F6) (01) Yes (I still do) (02) Yes, but only in the past

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.

Interviewer Reminder: Use **life grid** if necessary to help answer Q F2 to F8.
- Avoid overlapping years for the same product, type of cigarette or amount smoked, i.e. record 30-40, 41-45 rather than 30-40, 40-45.
- Only note changes occurring for **one year or more**.
- Exclude quitting during pregnancy(ies) if for less than one year.

F2 Do / did you smoke cigarettes?..... _ _
(00) No (GO TO F3) (01) Yes (02) Yes, only in the past

From age	To age (A)	Type (B)	Brand	#cigarettes/Day (D)
_ _	_ _	_ _	_____	_ _
_ _	_ _	_ _	_____	_ _
_ _	_ _	_ _	_____	_ _
_ _	_ _	_ _	_____	_ _
_ _	_ _	_ _	_____	_ _
_ _	_ _	_ _	_____	_ _

To Age (A)	Type (B)	No/Day (D)
If still smoking, write age at time of interview	(01) Filter (02) Non-filter (03) Hand rolled	(00) If less than daily Make average if not constant frequency

Section F – Smoking and Chewing Habits

0	2	-				-	
Country			**ID N°**				

F3 Do / did you smoke cigar?..... ☐☐

(00) No (GO TO F4) (01) Yes (02) Yes, only in the past

From age	To age (A)	Brand	#cigars/Day (D)
☐☐	☐☐	_____	☐☐
☐☐	☐☐	_____	☐☐
☐☐	☐☐	_____	☐☐

To Age (A)	No/Day (D)
If still smoking, write age at time of interview	(00) If less than daily Make average if not constant frequency

F4 Do / did you smoke pipe?..... ☐☐

(00) No (GO TO F5) (01) Yes (02) Yes, only in the past

From age	To age (A)	Brand	Unit (C)	#/Day(D)
☐☐	☐☐	_____	☐☐	☐☐
☐☐	☐☐	_____	☐☐	☐☐
☐☐	☐☐	_____	☐☐	☐☐

To Age (A)	Unit (C)	No/Day (D)
If still smoking, write age at time of interview	(01) Grams (02) Pipes	(00) If less than daily Make average if not constant frequency

F5 Do / did you <u>smoke or inhale</u> drugs (marijuana, grass, dope, joints...) at least once a week for at least 6 months in your lifetime?..... ☐☐

(00) No (GO TO F6) (01) Yes (02) Yes, only in the past

From age	To age (A)	Type (B)	Unit (C)	#/Day(D)
☐☐	☐☐	☐☐	☐☐	☐☐
☐☐	☐☐	☐☐	☐☐	☐☐
☐☐	☐☐	☐☐	☐☐	☐☐

To Age (A)	Type (B)	Unit (C)	No/Day (D)
If still smoking, write age at time of interview If less than one year, write same age From and To	(01) Marijuana (02) Grass (03) Crack (04) Hashish	(01) Grams (02) Joints	(00) If less than daily Make average if not constant frequency

F6 Do / did you <u>use any other drugs</u> (cocaine, heroin, LSD...) at least once a week for at least 6 months in your lifetime?..... ☐☐

(00) No (GO TO SECTION G) (01) Yes (02) Yes, only in the past

From age	To age (A)	Type (B)	Unit (C)	#/Day(D)
☐☐	☐☐	☐☐	☐☐	☐☐
☐☐	☐☐	☐☐	☐☐	☐☐
☐☐	☐☐	☐☐	☐☐	☐☐

To Age (A)	**Type (B)**	**Unit (C)**	**No/Day (D)**
If still using, write age at time of interview If less than one year, write same age From and To	(01) Cocaine (02) Acid / LSD (03) Speed (04) Heroin	(01) Grams (02) Joints (03) Injections (04) Pills	(00) If less than daily Make average if not constant frequency

G. DRINKING HABITS

Now I would like to ask you some questions about your drinking habits.

G1 Have you ever drunk alcoholic beverages <u>at least once a month</u>?..... _ _

(00) No (GO TO SECTION H) (01) Yes, I do (02) Yes, only in the past

We can use the grid to help us 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.

<u>Interviewer Reminder</u>: Use **life grid** if necessary to help answer Q G3.

- Avoid overlapping years for the same beverage i.e. record 30-40, 41-45 rather that 30-40, 40-45. Ask about each beverage separately.
- Note only changes occurring for **<u>one year or more</u>.**
- Exclude quitting during pregnancy(ies) if for less than one year.

G2 When do / did you usually drink alcoholic beverages?..... _ _

(01) With meals (03) Both

(02) Between meals (04) Only at social events

G3 Beverage (A)	If (A) = (05), Then specify other beverage	From age	To age	Unit (B)	Consumption (how many)	Per (C)
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _
_ _	_ _	_ _	_ _	_ _	_ _	_ _

Beverage (A)

(01) Wine

(02) Beer / cider

(03) Hard liquor (>35) (whisky, cognac, vodka, brandy, grappa, marc, gin, rum)

(04) Aperitif (<35) (Martini, port, sherry, vermouth)

(05) Other, specify: _____

Unit (B)

(01) Small glass (50ml) (1-2oz)

(02) Medium glass (100ml) (2-3oz)

(03) Big glass (250ml) (7oz) (1/2 pint)

(04) ½ small bottle (330ml) (1beer)

(05) Bottle (700-750 ml) (21oz)

Per (C)

(01) Day

(02) Week

(03) Month

H. DIETARY HABITS

Now, I have some questions about your dietary habits from your childhood (0-16 yrs).

H1 Please name 5 foods (any type) which you ate the most often during your childhood, starting with the most frequent.

1 _____
2 _____
3 _____
4 _____
5 _____

H2 If applicable, please name 5 foods (any type) which you did <u>not</u> eat during your childhood for any reason (religious beliefs, dislike, allergies, etc…).

1 _____
2 _____
3 _____
4 _____
5 _____

I would like to ask you a few questions about a list of foods that you ate during your childhood. Could you please tell me how often you ate the following foods during the ages of 0-16 yrs... (Use <u>Answer Sheet</u>)

(00) Sometimes	(01) Often	(02) Very Often	(99) I don't know

H3 Meat (all kinds)..... ☐☐

H4 Fish..... ☐☐

H5 Dairy products (milk, yogurt, cheese)..... ☐☐

H6 Vegetables..... ☐☐

H7 Fruits..... ☐☐

H8 Candies & Desserts..... ☐☐

H9 Chips & Fried Snacks..... ☐☐

H10 Did you eat spicy foods during your childhood?..... ☐☐

(00) No
(01) Yes, a little bit (mild)
(02) Yes, moderately spicy
(03) Yes, very spicy

Section H – Dietary Habits

0	2	-				-	
Country			**ID N°**				

Now, I have some questions about your dietary habits. As these habits may have changed somewhat according to your health status, **please tell me about your usual habits before diagnosis of the disease / being seen at this clinic**. How frequently did you consume the following foods and beverages?

Interviewer Reminder: Adapt portions to ones in table below.
- If less than once a week, **code (98)**.
- If not consumed at all, **code (00)**.
- **If don't not know code (99)**.

	Unit	Food item	Frequency (Per week)
H11	1 glass (200ml).....	Milk.....	
H12	1 pot (125g).....	Yoghurt.....	
H13	1 teaspoon.....	Butter.....	
H14	1 serving (50g) (2 slices).....	Bread.....	
H15	1 serving (4 full tablespoons).....	Pasta or rice.....	
H16	1 serving (100g) (1 side dish).....	Maize (Corn based dishes, polenta).....	
H17	1 serving (80g) (medium piece)..	Red meat (beef).....	
H18	1 serving (100g) (medium piece)	Pork.....	
H19	1 serving (160g) (medium piece)	Chicken.....	
H20	1 serving (80g) (medium piece)..	Lamb.....	
H21	1 serving (150g) (medium piece)	Fish.....	
H22	1 serving.....	Ham (2 slices), salami (4 slices), sausages (1)	
H23	1.....	Egg.....	
H24	1 serving (50g).....	Cheese.....	
H25	1 medium.....	Potatoes.....	
H26	1 serving (50g) (1 side dish).....	Raw green vegetables and salads.....	
H27	1 serving (50g) (1 side dish).....	Cruciferae (broccoli, cabbage, Brussels sprouts)	

H28	1 medium.....	Carrots.....	
H29	1 medium.....	Fresh tomatoes (in season).....	
H30	1 serving (4 full tablespoons).....	Pulses (chickpeas, beans, lentils, etc.).....	
H31	1 serving (50g) (1 side dish).....	*As a summary, how often would you say you eat any kind of vegetable (except potatoes)?*.....	
H32	1 glass (200ml).....	**Fresh** fruit juices.....	
H33	1 medium.....	Apples or pears.....	
H34	1 medium.....	Citrus fruit (oranges, grapefruit, lemons) (in season).....	
H35	1 medium.....	Bananas.....	
H36	1 medium.....	*As a summary, how often would you say you eat any kind of fresh fruit (including fruit salads)?*.....	
H37	1 slice or cup.....	Cake and desserts.....	
H38	1 portion.....	Chips and fried snacks.....	

H39 Which type of fat did you predominantly use to season vegetables?.....

(00) I don't use any fat	(04) Raisin oil	(08) Other vegetable oil	(12) I don't use animal fat
(01) Olive oil	(05) Corn oil	(09) Margarine	
(02) Dandelion oil	(06) Sunflower oil	(10) Butter	(13) Other fat
(03) Coconut oil	(07) Soy bean oil	(11) Pork fat	(99) Don't know

H40 Which type of fat did you predominantly use for cooking?.....

(00) I don't use any fat	(04) Raisin oil	(08) Other vegetable oil	(12) I don't use animal fat
(01) Olive oil	(05) Corn oil	(09) Margarine	
(02) Dandelion oil	(06) Sunflower oil	(10) Butter	(13) Other fat
(03) Coconut oil	(07) Soy bean oil	(11) Pork fat	(99) Don't know

H41 On average, how often did you eat barbecued food in the summer?..... ___

(00) I never eat BBQ (03) Less than once a week (06) More than 5 times a week
(01) Less than once a month (04) Once or twice a week (99) Don't know
(02) Once a month (05) 3 to 5 times a week

H42 On average, how often did you eat barbecued food in the winter?..... ___

(00) I never eat BBQ (03) Less than once a week (06) More than 5 times a week
(01) Less than once a month (04) Once or twice a week (99) Don't know
(02) Once a month (05) 3 to 5 times a week

H43 Did you drink coffee?..... ___

(00) No (GO TO H44) (01) Yes (02) Yes, only in the past

From age	**To age**	**# Cups**	**Per (C)** (01) Day, (02) Week, (03) Month

H44 How many cups of tea do you drink per day?..... ___

(00) I don't drink tea (98) Less than one a day

H45 How many cans of regular soda do you drink per day?..... ___

(00) I don't drink regular soda (98) Less than one a day

H46 How many cans of diet soda do you drink per day?..... ___

(00) I don't drink diet soda (98) Less than one a day

Section H – Dietary Habits

0	2	-				-	
Country			**ID N°**				

Interviewer Reminder: Note weight and height in measure used by participant. Later, use conversions to record weight in **kgs** and height in **cms**. See Interviewer's guide for conversions.

H47 If you remember, can you tell me what your weight was two years ago?..... ☐☐☐
(_____ lbs) , i.e. _____**kgs** (999) Don't know

H48 Can you tell me what your weight was at age 30?..... ☐☐☐
(_____ lbs) , i.e. _____**kgs** (999) Don't know

H49 Can you tell me what your weight was at age 20?..... ☐☐☐
(_____ lbs) , i.e. _____**kgs** (999) Don't know

H50 What is your height?..... ☐☐☐
(_____ feet _____ inches) , i.e. _____**cm** (999) Don't know

H51 Physical Activity and Hobbies
We would like to know which activities and hobbies you have during your adulthood. Please indicate if you have participated in the following activities regularly i.e. for at least 6 months.

							Frequency				
Activities	**Y**	**N**	**Don't know**	**Age at start**	**Age at end**	**# months**	**# times**	**per day**	**per week**	**per month**	**Total years**
Walking (for exercise)											
Jogging or running											
Aerobics											
Golf											
Racket sports (tennis, squash, etc…)											
Bowling or curling											
Swimming											
Skiing or skating											
Biking											
Dancing											
Gardening											
Outdoor physical work (mowing the lawn, shovelling, raking…)											
Household work											
Construction work (sawing, sanding, etc…)											
Car maintenance / work											
Other physical activities											
1)											
2)											
3)											
4)											

I. ORAL HEALTH

I am going to ask you some questions about your oral health **before your diagnosis / being seen at this clinic** and at a different time in your lifetime.

I1 Did you wear complete dentures?..... [ ][ ]
(00) No (GO TO I4)
(01) Yes, bottom only (GO TO I3)
(02) Yes, top only
(03) Yes, top AND bottom (GO TO I10)

I2 At what age did you start wearing complete top dentures? (Years)..... [ ][ ][ ]

I3 At what age did you start wearing complete bottom dentures? (Years)..... [ ][ ][ ]
Code (888) if QI1 = (02)

Interviewer Reminder: If both top AND bottom complete dentures, i.e. (03) to Q I1, skip to I5.

I4 Did you wear partial dentures?..... [ ][ ]
(00) No
(01) Yes, top only
(02) Yes, bottom only
(03) Yes , top AND bottom

Interviewer Reminder: Refer to life grid to separate each life period.

I5 How often did you clean your teeth?..... [ ][ ]
(00) Never
(01) Less than once a week
(02) 1-2 time a week
(03) Every other day (3-4 times a week)
(04) Once a day
(05) Twice or more a day

I6 Did you use dental floss?..... [ ][ ]
(00) No
(01) Yes, daily
(02) Yes, once a week
(03) Rarely

I7 Did you use toothpicks / sticks?..... [ ][ ]
(00) No
(01) Yes, daily
(02) Yes, once a week
(03) Rarely

I8 Did you use any kind of substance to clean your teeth?..... [ ][ ]
(00) No
(01) Toothpaste
(02) Other, specify: _____ [ ][ ]

I9 Did your gums bleed when you cleaned your teeth?..... [ ][ ]
(00) No
(01) Sometimes
(02) Always or almost always

I10 Did you use mouthwash?..... ___
(00) No (GO TO I13) (01) Yes

I11 How often did you use mouthwash?..... ___
(01) Less than once a week (03) Every other day (3-4 times a week) (04) Once a day
(02) 1-2 times a week (05) Twice or more a day

I12 What was the brand name of the mouthwash? (LC)..... ___
Brand name: _____

Now, let's look at your oral health habits and oral health at different periods of your life.

I13 In the <u>last 20 years</u>, how often did you see a dentist?..... ___
(00) Never (03) Every 2 –5 years
(01) Every 6 months (04) Once every 5 years
(02) Every year (05) Only when I had pain

I14 Have you ever had a tooth extracted?..... ___
(00) No (GO TO I16) (01) Yes

I15 How many teeth extractions had you had?
Up until you were 16 of age..... ___
Between 17-30 years of age..... ___
After 30 years of age but before the diagnosis of the disease..... ___

(00) None	(02) 6–15	(04) 21-30	(99) Don't know
(01) 1-5	(03) 16-20	(05) More than 30	

I16 Have you ever had a filling?..... ___
(00) No (GO TO SECTION J) (01) Yes

I17 How many fillings had you had?
Up until you were 16 of age..... ___
Between 17-30 years of age..... ___
More than 30 years of age..... ___

J. FAMILY HISTORY OF CANCER

Interviewer Reminder:
- Family includes these **biological** relatives: father, mother, brother, sister, son, daughter, aunt, uncle, grandmother, grandfather.
- Input one person per line in chart below.

J1 Has any member of your family ever had cancer?..... _ _

(00) No (GO TO SECTION K) (01) Yes (99) Don't know

J2

Relationship (A)	Status (B)	Current / last Age (C)	Type of cancer	Type of tumour (LC)	Age at Diagnosis (D)
_ _	_ _	_ _ _	_____	_ _ _ - _	_ _ _
_ _	_ _	_ _ _	_____	_ _ _ _	_ _ _
_ _	_ _	_ _ _	_____	_ _ _ _	_ _ _
_ _	_ _	_ _ _	_____	_ _ _ _	_ _ _
_ _	_ _	_ _ _	_____	_ _ _ _	_ _ _
_ _	_ _	_ _ _	_____	_ _ _ _	_ _ _

Relationship (A)	Status (B)	Current / last Age (C)	Age at diagnosis (D)
(01) Mother (02) Father (03) Sister (04) Brother (05) Daughter (06) Son (07) Grand-mother (08) Grand-father (09) Aunt / uncle	(00) Deceased (01) Alive	(999) Don't know If alive, give present age If deceased, give age at death	(999) Don't know

Country: 0 2 - ID N°: ___ - ___

K. FAMILY ENVIRONMENT IN CHILDHOOD

I would like to ask you a few questions about your parents (mother and father), or the women or men who cared for you during your childhood, that is from your birth until you were 16 (incl.). If you were cared for by only one person, please respond only to the questions related to that person. We may refer to the grid to help us out at times.

This first set of questions is related to their level of education and their occupation.

K1 At your birth, how old was your natural father?..... ___
(99) Don't know

K2 At your birth, how many years of education did your father / the man who cared for you most of your childhood have?..... ___
(99) Don't know

K3 What was his longest occupation during your childhood? (LC)..... ___
Describe: _____
(999) Don't know

K4 At your birth, how old was your natural mother?..... ___
(99) Don't know

K5 At your birth, how many years of education did your mother / the woman who cared for you most of the time during your childhood have?..... ___
(99) Don't know

K6 What was her longest occupation during your childhood? (LC)..... ___
Describe: _____
(999) Don't know

Interviewer Reminder: Confirm occupation codes in K3 and K6 with list of codes.

Now I have a few questions on family environment during your childhood.

K7 In total, how many brothers and sisters do you have? (natural only)..... ___

K8 What was your birth order in your family (at time you were 16 years old)?..... ___
(00) Only child (02) Second child (04) Fourth child or more
(01) First child (03) Third child

K9 Did your family have continuous financial difficulties during your childhood? ___
(00) No (01) Yes (99) Don't know

K10 Did your parents argue or fight during your childhood?.....
(00) Never (02) Often
(01) Sometimes (99) Don't know

K11 How often did your father use to drink alcohol during your childhood?.....
(00) Never (02) Once a week / weekends (04) Everyday
(01) Occasionally (03) 3-4 times a week (99) Don't know

K12 How often did your mother use to drink alcohol during your childhood?.....
(00) Never (02) Once a week / weekends (04) Everyday
(01) Occasionally (03) 3-4 times a week (99) Don't know

K13 Did your father smoke? (any product).....
(00) No (01) Yes (99) Don't know

K14 Did your mother smoke? (any product).....
(00) No (01) Yes (99) Don't know

The following six questions relate to your natural parents.

K15 Were you ever separated from your <u>biological mother</u> <u>for a year or more</u> during your childhood?.....
(00) No (GO TO K18) (01) Yes (99) Don't know

K16 How old were you?
From age? **To age?** *(max = 16)* **i.e. # Years**

K17 Why did the separation happen?.....
(00) Parents separated / divorced (03) Adoption
(01) Mother died (04) Other, specify: _____
(02) Mother ill

K18 Were you ever separated from your <u>biological father</u> <u>for a year or more</u> during your childhood.....
(00) No (GO TO K21) (01) Yes (99) Don't know

K19 How old were you?
From age? **To age?** *(max = 16)* **i.e. # Years**

K20 Why did the separation happen?.....
(00) Parents separated / divorced (03) Adoption
(01) Father died (04) Other, specify: _____
(02) Father ill

Now I would like to ask you a few questions about your mother / father figure during your childhood.

K21 Who was the woman who cared for you most of your life during your childhood? ☐☐

(00) None (GO TO K29)
(01) Biological mother
(02) Step mother
(03) Adoptive mother
(04) Grand-mother
(05) Other, specify: _____ ☐☐

Here are some questions about how you remember your **MOTHER** (or the woman who cared for you) during the years you were growing up, that is, until you were age 16 – incl. (Use Answer Sheet)

(01) A great deal (02) Quite a lot (03) Little (04) Not at all

K22 How much did she understand your problems and worries?..... ☐☐

K23 How much could you confide in her about things that were bothering you?.... ☐☐

K24 How much love and affection did she give you?..... ☐☐

K25 How much time and attention did she give you when you needed it?..... ☐☐

K26 How strict was she with the rules for you?..... ☐☐

K27 How harsh was she when she punished you?..... ☐☐

K28 How much did she expect you to do your best in everything you did?..... ☐☐

Now I would like to ask you how you remember your **FATHER** (or the man who cared for you) during the years you were growing up that is, until you were 16 years old. (Use Answer Sheet)

K29 Who was the man who cared for you most of your life during your childhood?..... ☐☐

(00) None (GO TO K37)
(01) Biological father
(02) Step father
(03) Adoptive father
(04) Grand-father
(05) Other, specify: _____ ☐☐

(01) A great deal (02) Quite a lot (03) Little (04) Not at all

K30 How much did he understand your problems and worries?..... ☐☐

K31 How much could you confide in him about things that were bothering you?... ☐☐

K32 How much love and affection did he give you?..... ☐☐

K33 How much time and attention did he give you when you needed it?..... ☐☐

K34 How strict was he with the rules for you?..... ☐☐

K35 How harsh was he when he punished you?..... ☐☐

K36 How much did he expect you to do your best in everything you did?..... ☐☐

I have only a few more questions about your childhood. You do not have to answer if you do not feel comfortable doing so. **Did any of the following things happen during your childhood...? (0-16 yrs)**

K37 Were you physically abused?..... ☐☐
(00) No (01) Yes

K38 Were you sexually abused?..... ☐☐
(00) No (01) Yes

K39 Were your parents divorced?..... ☐☐
(00) No (01) Yes

Finally,

K40 Can you remember any life event(s) in your childhood that have either positively or negatively impacted upon you?..... ☐☐
(00) No (GO TO SECTION L) (01) Yes

K41 Can you tell me what? (Describe) (**LC**).....
1 _____ ☐☐
2 _____ ☐☐
3 _____ ☐☐
4 _____ ☐☐
5 _____ ☐☐

K42 Could you please tell me how much impact this (se) event (s) had on your life?
(Use Answer Sheet).....

-4 -3 -2 -1 0 1 2 3 4
Very negative no impact Very positive

Event 1 ……..score: _____…….. ☐☐
Event 2 ……..score: _____…….. ☐☐
Event 3 ……..score: _____…….. ☐☐
Event 4 ……..score: _____…….. ☐☐
Event 5 ……..score: _____…….. ☐☐

K43 For each of the following diseases, can you tell me if you ever had it and, if so, how often?

Presence (A)	Frequency (B)
(00) No	(01) Once
(01) Yes	(02) Sometimes
(99) Don't know	(03) Often

	Presence (A)	Frequency (B)
Measles	[ ][ ]	[ ][ ]
Mumps	[ ][ ]	[ ][ ]
Chicken pox	[ ][ ]	[ ][ ]
Whooping cough	[ ][ ]	[ ][ ]
Scarlet fever	[ ][ ]	[ ][ ]
Rheumatic fever	[ ][ ]	[ ][ ]
Infectious hepatitis	[ ][ ]	[ ][ ]
Tuberculosis	[ ][ ]	[ ][ ]
Asthma attack	[ ][ ]	[ ][ ]
Disease of the ear(s)	[ ][ ]	[ ][ ]
Disease of the nose	[ ][ ]	[ ][ ]
Disease of the throat	[ ][ ]	[ ][ ]
Other diseases: Specify (ex.: chronic heartburn, bulimia):		
_____	[ ][ ]	[ ][ ]
_____	[ ][ ]	[ ][ ]
_____	[ ][ ]	[ ][ ]
_____	[ ][ ]	[ ][ ]
_____	[ ][ ]	[ ][ ]

L. MARRIAGE, INTIMACY & LIFE AS A COUPLE

Now, I would like to ask you some questions about marriage and living as a couple.

L1 What is your marital status?..... _ _

(01) Single
(02) Living with a husband / wife (married)
(03) Living with a partner in common law
(04) Divorced
(05) Widowed
(06) Separated

Interviewer Reminder: Use **life grid** if necessary to help answer Q L2 to L26.

L2 How many times have you been married or lived in common law?..... _ _

(01) Once (Fill in first column only)
(02) More than once

At the time you FIRST / LAST got married or FIRST / LAST lived in common law...

	FIRST	LAST
L3 How old were you?.....	_ _	_ _
L4 How many years did your partner go to school for? (until today)	_ _	_ _
L5 What was your partner's longest occupation? (until today) **(LC)**	_ _ _	_ _ _

FIRST partner: _____

LAST partner: _____

L6 How did the relationship end?.....	_ _	_ _
L7 How old were you when the relationship ended?.....	_ _	_ _

(00) Still ongoing! (GO TO L8)
(01) Divorce
(02) Separation
(03) Partner deceased

L8 In your whole life, how many (biological) children have you had?..... _ _

(00) None (GO TO L10) *(Do NOT include miscarriage or stillborn)*

L9 With how many <u>different</u> partners?..... _ _

(00) All with the same one

I will ask you some questions regarding your sexuality. The reason I am asking these questions is because medical science has found some links between viruses that are sexually transmitted and some types of cancers. <u>You have no obligation to answer these questions if you do not feel comfortable doing so.</u>

L10 Have you ever had sexual intercourse?..... _ _

(00) No (GO TO L14)
(01) Yes
(99) Prefer not to say / Don't know

L11 How old were you when you had sexual intercourse for the first time?.....
(99) Prefer not to say / Don't know

L12 How many sexual partners have you had in total in your life? (regular and casual)
Up to 16 yrs old.....
Between 17-30 yrs old.....
After 30 yrs old.....

Answer's options L12 and L13		
(00) None	(03) 06-10	(06) 51-100
(01) One	(04) 11-20	(07) More than 100
(02) 2-5	(05) 21-50	(99) Prefer not to say / Don't know

L13 How many of these were prostitute? (99) Prefer not to say / Don't know
Up to 16 yrs old.....
Between 17-30 yrs old.....
More than 30 yrs old.....

L14 Have you ever had oral sex? (your mouth and a woman / man genitals).....
(00) No (GO TO L17) (99) Prefer not to say / Don't know (GO TO L17)
(01) Yes

L15 How old were you when you had oral sex for the first time?.....
(99) Prefer not to say / Don't know

Answer's options Q16	
(00) Occasionally	(02) Most of the time
(01) Frequently	(99) Prefer not to say / Don't know

L16 How often?
Up to 16 yrs old.....
Between 17-30 yrs old.....
After 30 years old.....

L17 Have you ever had non-consenting sex?.....
(00) No (GO TO L19) (99) Prefer not to say / Don't know (GO TO L19)
(01) Yes

L18 How old were you or from what age to what age? *(mark same age if one episode or if during less than one year)* (99) Prefer not to say / Don't know

From age? **To age?** **i.e. # Years**

L19 Have you ever had skin warts?.....
(00) No (GO TO (GO TO L22) (99) Prefer not to say / Don't know (GO TO L22)
(01) Yes

L20 If yes, where? (01) Yes (00) No (99) Prefer not to say / Don't know

Hands.....
Feet.....
Head and Neck.....
Other, specify: _____.....

L21 At which age, were you? (99) Prefer not to say / Don't know

Hands.....
Feet.....
Head and Neck.....
Other, specify: _____.....

L22 Since you started your sexual life have you ever had Candida Albicans?.....

(00) No (GO TO L24) (99) Prefer not to say / Don't know (GO TO L24)
(01) Yes

L23 If yes, where? (01) Yes (00) No (99) Prefer not to say / Don't know

Genital.....
Mouth.....
Other, specify: _____.....

L24 Have you ever had a sexually transmitted disease?.....

(00) No (GO TO SECTION M) (99) Prefer not to say / Don't know (GO TO SECTION M)
(01) Yes

L25 If yes, which ones? (01) Yes (00) No (99) Prefer not to say / Don't know

Gonorrhea.....
Syphilis.....
Herpes.....
Chlamydia.....
AIDS.....

L26 At which age, were you? (99) Prefer not to say / Don't know

Gonorrhea.....
Syphilis.....
Herpes.....
Chlamydia.....
AIDS.....

(*Note other types of sexually transmitted diseases in the* ***Participant's comments*** *on page 53.*)

M. SOCIAL SUPPORT

Finally I would like to ask you some questions about your friends, relatives and the people you live with.

M1 Is there <u>someone in particular</u> in your life that you think would listen to you and give you emotional support if you needed it?..... ☐☐

(01) Yes (00) No (GO TO M9)

	1st PERSON	2nd PERSON
M2 What is your relationship with this person?.....	☐☐	☐☐
(01) Spouse / partner (living together) (05) Neighbour		
(02) Boyfriend / girlfriend (06) Colleague		
(03) Parent (07) Son / daughter		
(04) Brother / sister (08) Other family member (cousin, etc)		
(09) Friend		
(10) Other, specify: _____	☐☐	☐☐
M3 Does he/she live near enough to come around if something came up? (01) Yes (00) No	☐☐	☐☐
M4 On average how often have you <u>seen</u> him / her in the last year?.....	☐☐	☐☐
(01) Not in the last year (04) 1 or 2 times a week		
(02) Less than once a month (05) 3+ times a week		
(03) Less than once a week		
M5 Would you prefer to see him / her more / less often or is this about right for you?..... (01) Less often (02) About right (03) More often	☐☐	☐☐
M6 How long have you known him / her for? (Years).....	☐☐	☐☐
M7 Would you say that you could talk frankly and share your feelings with him / her?.....	☐☐	☐☐
(00) No (02) Yes, about most things		
(01) Yes, about some things (03) Yes, about anything		

M8 Apart from this person / these two people, is there anyone else in particular that you think would listen to you and be supportive if you needed it?..... ☐☐

(00) No (01) Yes

M9 In your life in general, do you think you have enough opportunities to talk openly and share your feelings about things?..... ☐☐

(00) No (01) Yes

M10 In general, do you prefer to keep your feelings to yourself?..... ☐☐

(00) No (01) Yes

M11 Can you remember any life event(s) in your adulthood that have either positively or negatively impacted upon you?..... | |

(00) No (GO TO M14) (01) Yes

M12 Can you tell me what? (Describe) **(LC)**.....

1_____ | |

2_____ | |

3_____ | |

4_____ | |

5_____ | |

M13 Could you please tell me how much impact did this (se) event (s) have in your life?
(Use Answer Sheet).....

-4	-3	-2	-1	0	1	2	3	4
Very negative				no impact				Very positive

Event 1 ……..score: _____…….. | |

Event 2 ……..score: _____…….. | |

Event 3 ……..score: _____…….. | |

Event 4 ……..score: _____…….. | |

Event 5 ……..score: _____…….. | |

M14 Do you have any brothers or sisters of a similar age (±5 yrs) that would be interested in participating in this interview?..... **8 8**

(00) No (01) Yes

M15 10% of participants of this study will be re-interviewed. Do you agree to be re-contacted for you to participate a second time?..... | |

(00) No (01) Yes

M16 Incomplete questionnaire?..... | |

(00) No (01) Yes

If YES, reason:_____

M17 Time of end of interview..... | | - | |

Hour **Minute**

M18 Interviewer's initials?..... | |

M19 Initials of data enterer into FileMaker?..... | |

Participant's comments:

_____

_____

_____

N. BIOLOGICAL SAMPLING

N1 Was a mouthwash sample taken?..... ___
(00) No (01) Yes (02) Yes, but taken with water

N2 Was a sample for HPV analysis taken?
(this sample is taken from the lesion site for cases, from healthy buccal cells for controls).... ___
(00) No (01) Yes

N3 Was a sample for genetic analysis taken?
(this sample is taken from healthy buccal cells from both the cases and controls)..... ___
(00) No (01) Yes

N4 Please document below if there was any comments from the biological sampling (e.g., occurrence of untoward / adverse events such as patient discomfort, bleeding).

_____

_____

N5 Were all 3 above samples delivered to Dr Coutlée's laboratory?..... ___
(00) No (01) Yes

N6 Date of Sample Delivery..... ___ - ___ - ___
Day Month Year

N7 Please document below if there was any comments from the state of the sample (e.g., leaking of vials, etc…).

_____

_____

N8 HPV ANALYSIS

		Mouthwash		HPV		GEN	
	HPV type	**Present**	**Not-present**	**Present**	**Not-present**	**Present**	**Not-present**
N8a	6						
N8b	11						
N8c	16						
N8d	18						
N8e	26						
N8f	31						
N8g	33						
N8h	35						
N8i	39						
N8j	40						
N8k	42						
N8l	45						
N8m	51						
N8n	52/33/35/58						
N8o	52tm						

		Mouthwash		HPV		GEN	
	HPV type	**Present**	**Not-present**	**Present**	**Not-present**	**Present**	**Not-present**
N8p	53						
N8q	54						
N8r	55						
N8s	56						
N8t	58						
N8u	59						
N8v	61						
N8w	62						
N8x	64						
N8y	66						
N8z	67						
N8aa	68						
N8bb	69						
N8cc	70						
N8dd	71						
N8ee	72						
N8ff	73						
N8gg	81						
N8hh	82						
N8ii	83						
N8jj	84						
N8kk	IS39						
N8ll	CP6108						

N9 Mouthwash comments: _____

N10 HPV Sample comments: _____

N11 GEN Sample comments: _____

How many different types of HPV were found in…

N12 Mouthwash?..... ___

N13 HPV sample?..... ___

N14 GEN sample?..... ___

N15 GENETIC ANALYSIS

		Mouthwash	HPV	GEN
N15a	Concentration (ng/ul)			
N15b	PCR (+/-)			
N15c	Notes			