# Postoperative pulmonary complications in head and neck surgery: risk factors and prognostic risk prediction models

**Mohammad Al-Tamimi**

Faculty of Dental Medicine and Oral Health Sciences

McGill University

Montreal, Québec, Canada

September 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

## Dedication

This work is dedicated to my parents who I would not be able to achieve anything without their unconditional love and support. Also, I dedicate this work to my partner, significant other, and love of my life Marwa, who was there for me in the difficult times during the past few months.

## Acknowledgement

I would like to express my sincere gratitude to my brother from another mother and father, Mohammad Abu Samak. He has been for me the big brother that I never had and without his help and advice, I do not think I would be able to finish my thesis.

## Table of Contents:

## List of Tables:

## List of figures:

## List of Abbreviations:

**PPC: Postoperative pulmonary complications.**

**HNS: Head and Neck Surgery.**

**ASA: American society of Anesthesiologists.**

**ACS NSQIP: American College of surgeons National surgical quality improvement program.**

**PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.**

**C-Statistic: Concordance statistic**

**ROC: Receiver Operating Characteristic**

**AUC: Area Under the ROC Curve**

## Abstract:

**Background:** Postoperative pulmonary complications (PPCs) are common following head and neck surgeries (HNS). The incidence of PPCs in the HNS patients varies between 4.5% and 47%, and these complications are associated with increased morbidity, mortality, intensive care unit admission, and increased hospital length of stay. Furthermore, PPCs are associated with common risk factors in cardiac and general surgeries. However, evidence around the risk factors associated with PPCs in HNS patients is lacking. Also, the prediction models used to estimate the probability of having PPCs were found t perform poorly for HNS patients. The aim of this thesis project was to summarise and appraise the current evidence on PPC risk factors following HNS and to develop a validated prognostic risk prediction model for PPC among these patients. This thesis objectives are: 1) to review the literature on the PPC risk factors for HNS patients 2) to validate and update a risk prediction model on a sample of HNS patients.

**Methods:** First, we reviewed the published articles on PPC risk factors for HNS patients. We included the full-text of peer-reviewed publications that reported PPC risk factors for HNS patients. The risk of bias was appraised using Joana-Briggs Institute tool. This review identified a set of risk factors that can be easily recorded in clinical settings. In the second step of this thesis project, we validated the predictions of the Gupta model, which was developed for general surgery patients using a sample of the US hospitals on the dataset derived from the National Surgical Quality Improvement Program from the American College of Surgeons (NSQIP ACS). A cohort of 79,726 patients who had HNS procedures were identified in the period of 2018-2019. We replicated the Gupta model and tested its predictions on the 2019 dataset then updated the model by adding a set of risk factors identified from the review: age, sex, smoking history, body mass index, operation time, and reconstruction procedure. We evaluated this updated model's overall performance with scaled Brier score and R2 (Nagelkerke). The discrimination ability was evaluated using C-Statistics. The model's calibration was assessed by evaluating the calibration slope.

**Results:** The first step of this thesis identified thirty peer-reviewed studies that reported PPC risk factors for HNS patients. The majority of studies (70%) were retrospective cohorts. The

reported risk factors of PPCs in HNS patients were age, male sex, smoking history, respiratory comorbidity, smoking history, body mass index, operation time, site of the surgery, and reconstruction procedure. Majority of included studies had a high risk of bias. Although we identified articles reporting development or validation of PPC risk prediction model for the HNS population, none of these models were externally validated in other HNS cohorts. Gupta model multivariable logistic regression model had predictors: functional dependency status, emergency of the procedure, history of blood sepsis, surgical specialty doing the procedure, and American Society of Anesthesiologists (ASA) classification. The overall performance assessment of the original Gupta et al. model during the temporal validation using the 2019 dataset presented a scaled Brier score and R2 of -0.056 and -0.20, respectively, while C-Statistics (discrimination) was 0.72 (95%CI= 0.68-.075) and calibration slope was 0.94. The overall performance of the updated model after internal-external validation in 2019 display a scaled Brier score and R2 of 0.02 and 0.14, respectively, while C-statistic and calibration slope were 0.82 (95%CI=0.79-0.86) and 0.98, respectively.

**Conclusion:** Limited numbers of studies provide good quality evidence on PPC risk factors and discussed potential causality. However, the review identified few risk factors that might be more specific for the HNS patients. The updated model was found to preform better than the original model in HNS patients. Future studies are needed to assess the clinical applicability of updated model.

## RÉSUMÉ

**Contexte :** Les complications pulmonaires postopératoires (CPP) sont fréquentes après une chirurgie de la tête et du cou (CTC). L'incidence des CPP chez les patients CTC varie entre 4,5 % et 47 %, et ces complications sont associées à une augmentation de la morbidité, de la mortalité, de l'admission en unité de soins intensifs et de la durée du séjour à l'hôpital. Bien que les CPP soient associés à des facteurs de risque courants dans les chirurgies cardiaques et générales, les preuves manquent sur les facteurs de risque associés aux CPP chez les patients CTC. De plus, les modèles de prédiction utilisés pour estimer la probabilité d'avoir des CPP fonctionnent de manière sous-optimale pour les patients CTC. L'objectif de ce projet de thèse était de résumer et d'évaluer les preuves actuelles sur les facteurs de risque de CPP après une CTC et de développer un modèle validé de prédiction du risque pronostique de CPP chez ces patients. Cette thèse vise à : 1) faire une revue de la littérature sur les facteurs de risque de CPP pour les patients CTC et 2) valider et mettre à jour un modèle de prédiction de risque sur un échantillon de patients CTC

**Méthodes :** Tout d'abord, nous avons examiné les articles publiés sur les facteurs de risque de CPP pour les patients CTC. Nous avons inclus le texte complet des publications évaluées par des pairs qui rapportaient des facteurs de risque de CPP pour les patients CTC. Le risque de biais a été évalué à l'aide de l'outil Joana-Briggs Institute. Cette revue a identifié un ensemble de facteurs de risque qui peuvent être facilement enregistrés en milieu clinique. Dans la deuxième étape de ce projet de thèse, nous avons validé les prédictions du modèle Gupta, qui a été développé pour les patients en chirurgie générale en utilisant un échantillon d'hôpitaux américains sur l'ensemble de données dérivé de l'American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP). Une cohorte de 79 726 patients ayant subi des procédures CTC a été identifiée au cours de la période 2018-2019. Nous avons reproduit le modèle Gupta et testé ses prédictions sur l'ensemble de données de 2019, puis mis à jour le modèle en ajoutant un ensemble de facteurs de risque identifiés à partir de la revue de littérature : âge, sexe, antécédents de tabagisme, indice de masse corporelle, durée de l'opération et procédure de reconstruction. Nous avons évalué les performances globales de ce modèle mis à jour avec un score de Brier et un R2 (Nagelkerke) mis à l'échelle. La capacité de discrimination a

été évaluée à l'aide de C-Statistics. Le calibrage du modèle a été évalué en évaluant la pente de calibrage.

**Résultats** : La première étape de cette thèse a identifié 30 études évaluées par des pairs qui ont rapporté des facteurs de risque de CPP pour les patients CTC. La majorité des études (70 %) étaient des études de cohortes rétrospectives. Les facteurs de risque rapportés de CPP chez les patients CTC étaient l'âge, le sexe masculin, les antécédents de tabagisme, la comorbidité respiratoire, les antécédents de tabagisme, l'indice de masse corporelle, la durée de l'opération, le site de la chirurgie et la procédure de reconstruction. La majorité des études incluses présentaient un risque élevé de biais. Bien que nous ayons identifié des articles rapportant le développement ou la validation du modèle de prédiction du risque CPP pour la population CTC, aucun de ces modèles n'a été validé en externe dans d'autres cohortes CTC. Le modèle de régression logistique multivariable du modèle Gupta avait des prédicteurs : statut de dépendance fonctionnelle, urgence de la procédure, antécédents de septicémie sanguine, spécialité chirurgicale effectuant la procédure et classification de la Société Américaine des Anesthésiologistes (ASA). L'évaluation globale des performances du modèle original de Gupta lors de la validation temporelle à l'aide de l'ensemble de données de 2019 a présenté un score Brier échelonné et un R2 de -0,056 et -0,20, respectivement, tandis que la C-Statistique (discrimination) était de 0,72 (IC à 95 % = 0,68 -0,075) et la pente d'étalonnage était de 0,94. Les performances globales du modèle mis à jour après validation interne-externe en 2019 affichent un score Brier échelonné et un R2 de 0,02 et 0,14, respectivement, tandis que la C-Statistique et la pente d'étalonnage étaient de 0,82 (IC à 95 % = 0,79-0,86) et 0,98, respectivement.

**Conclusion :** Un nombre limité d'études fournissent des preuves de bonne qualité sur les facteurs de risque de CPP et discutent de la causalité potentielle. Cependant, la revue de littérature a identifié quelques facteurs de risque qui pourraient être plus spécifiques aux patients CTC. Le modèle mis à jour fonctionne mieux que le modèle original pour les patients CTC. Des études futures sont nécessaires pour évaluer l'applicabilité clinique du modèle mis à jour.

## Preface:

This project is a manuscript-based thesis written according to the updated standards established by McGill Graduate and Postdoctoral Studies for fulfilling the requirements of a master's degree in the Dental Sciences -Thesis program. The two manuscripts follow the primary goal of this thesis by providing insight into PPC risk prediction in HNS patients. The first manuscript reviews the literature on the PPC risk factors in HNS patients aiming to understand the current PPC risk factors in this surgical specialty and investigate the existing studies' strength, limitations, and risk of bias. The second manuscript investigates the generalizability of a risk prediction model identified in the literature in HNS patients while updating this model with identified predictors from the first manuscript. Manuscripts are logically coherent with this thesis work and share a unified theme. Based on the standards of McGill University, each manuscript comprised a separate set of appendices and reference lists. The first chapter of this thesis includes an introduction to the topic, which is then followed by reviewing the literature providing the current knowledge in the field. Supported by the second chapter, the rationale of this thesis project is then provided in the third chapter. The fourth chapter includes detailed objectives, while the fifth chapter provides the methods followed in conducting this thesis project.

*Manuscripts I and II* are the next two chapters that stand alone regarding tables, figures, appendices, and reference lists. The eighth chapter comprehensively discusses this thesis work, followed by the ninth chapter, where the overall conclusions are provided. The two manuscripts comprise multiple authors whose contributions to each manuscript are provided in the next section.

## Contribution of Authors:

**Manuscript I**

**Postoperative Pulmonary Complications Risk Factors in Head and Neck Surgery: A Scoping Review**

**Mohammad Al-Tamimi,** Master's Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Conceived objective of the investigation, screened and selected studies, extracted data, assessed the quality of studies, statistical analyses, performed statistical analysis, and wrote the manuscript.

**Sreenath Madathil,** Assistant professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, contributed to the conceptual development, selection process, quality assessment, statistical analyses, interpretation, and writing the manuscript.

**Ahmed Derbas,** Oral and maxillofacial surgery resident, Oral and Maxillofacial Surgery Department, Hamad Medical Corporation, Doha, Qatar: Participated in the screening, selection process and writing the manuscript.

**Hamed Ghanati,** Master's Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: participated in the conceptual development and contributed to writing the manuscript.

**Martin Morris,** Schulich Library of Physical Sciences, Life Sciences and Engineering, McGill University, Montréal, QC, Canada: Conducted the literature search process and contributed to writing the manuscript.

**Belinda Nicolau.**, Professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, contributed to the conceptual development, selection process, quality assessment, interpretation, and writing the manuscript (Corresponding author).

**Manuscript II**

**Temporal Validation and Extension of a Prediction Model for Postoperative Pulmonary Complications**

**Mohammad Al-Tamimi,** Master's Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Conceived objective of the investigation, prepared data, performed statistical analysis, interpreted the results, and wrote the manuscript.

**Belinda Nicolau**, Professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, contributed to the conceptual development, provided the databases, interpretation, and writing the manuscript.

**Nicholas Makhoul,** Associate Professor, Department of Dentistry and Oral and Maxillofacial Surgery, McGill: participated in the conceptual development and contributed to writing the manuscript.

**Sreenath Madathil,** Assistant professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, performed statistical analysis, contributed to the conceptual development, study design, interpretation of results, and writing the manuscript (Corresponding author).

## 1. Introduction:

Every year, around 300 million surgical procedures are performed worldwide[1]. Complication rates following general surgical procedures might reach 30% as reported in colorectal surgeries, for example[2,3]. These complications might change the course of recovery of a patient as it causes increased risk of reoperation, prolonged hospital length of stay (LOS), and increased mortality[4].

One category of postoperative complications, pulmonary complications (PPCs), can be broadly defined as conditions affecting the respiratory tract that can adversely influence the clinical course and recovery of patients after surgery[5]. The most important and morbid PPCs are atelectasis (a state of collapsed and non-aerated region of the lung parenchyma[6,7]), pneumonia, respiratory failure, and exacerbation of underlying chronic lung disease[7]. Around 30% of patients who undergo non-cardiac surgeries in the US, lasting at least 2 hours, under general anaesthesia and mechanical ventilation might develop PPCs[8]. In addition, PPCs can result in potentially significant morbidity and mortality when they include more severe atelectasis, bronchospasm, postoperative pneumonia, pneumothorax, acute respiratory distress syndrome, pulmonary embolism, or respiratory failure[5]. Around 3% of patients undergoing general surgery experience pulmonary respiratory failure and 25% of them die within 30 days[9]. Also, PPCs represent a significant economic burden because these complications may increase hospital LOS; a standard reference point for assessing the cost of service[5]. Khan et al[10] using data from a Canadian cohort observed that postoperative pneumonia in non-cardiac surgery patients increased hospital LOS by 89% and hospital costs by 55%.

Head and neck surgery (HNS) is a specialty within the field of ear, nose and throat medicine (otolaryngology) combined with oral and maxillofacial medicine that focuses on surgically treating head and neck disorders. The incidence of PPCs among HNS patients varies between 4.5% and 47%[11-16], and these complications are associated with increased morbidity, mortality, intensive care unit admission, and hospital LOS[14,17,18].

The American College of Physicians defines the risk factors into two groups: patient- and procedure-related factors[7]. Patient-related factors include age, comorbidities (chronic lung disease, congestive heart failure, obstructive sleep apnea), functional dependence, obesity, and

smoking[7]. Whereas procedure-related factors are site and duration of the surgery, anesthetic technique, and emergency of the surgery[7]. To mitigate PPCs incidence in high-risk patients, preoperative surgical risk assessment is utilized[19]. This evaluation objectively outlines risks prior to surgery, facilitating informed consent process and encourages exercising risk reduction measurements such as preoperative respiratory physiotherapy and smoking cessation[19]. Several preoperative surgical risk assessment algorithms have been developed to identify patients at high risk of PPCs. Among those algorithms is the Gupta Pulmonary Risk Index, which has used data from the American College of Surgeons –National Surgical Quality Improvement (ACS-NSQIP) database[9]. Although these risk assessment tools were developed using large surgical cohorts, only a few cases of HNS were included. Consequently, these tools, including the Gupta pulmonary risk index, have insufficient accuracy and poor discrimination ability to predict PPCs among HNS patients[20]. Moreover, utilizing these tools in the HNS patients may lead to risk misclassification and consequently, suboptimal care delivery[21]. To avoid this situation, tools specifically designed to evaluate the PPCs among patients undergoing HNS is imperative.

Risk prediction models, which has been increasingly used in medical decision making, can be used to estimate the individual probability of having PPCs, i.e., the predictions from these models can be used to identify patients at high-risk of developing PPCs[22]. However, prior to developing new models, existing models need to be updated and validated[22]. This process is significant as it would results in less redundancy in prediction models and continuous refinement for existing models[22].

More importantly, prior to validating the existing models, a comprehensive review of the literature is necessary to identify the risk factors for PPCs that will be included in the models. In our initial search, we could not identify such a comprehensive review. Also, we could not find any tool that specifically evaluates the risk of PPCs among patients undergoing HNS and was externally validated. Therefore, this thesis reviews the literature in PPCs risk factors and validate Gupta Pulmonary Index in patients undergoing HNS.

## 2. Literature review:

### 2.1 Definition of Postoperative pulmonary complications (PPCs):

Several PPCs definitions are available in the literature. Abbott and colleagues[23] recommended a definition of PPCs as "*Composite of respiratory diagnoses that share common pathophysiological mechanisms including pulmonary collapse and airway contamination: (i) atelectasis detected on computed tomography or chest radiograph, (ii) pneumonia using US Centers for Disease Control (CDC) criteria, (iii) Acute Respiratory Distress Syndrome using Berlin consensus definition, (iv) pulmonary aspiration (clear clinical history and radiological evidence)*." Also, they categorized the severity of PPCs as: (i) None: no planned use of supplemental oxygen or mechanical ventilation support as part of routine care, but not in response to a complication or deteriorating physiology; (ii) Mild: therapeutic supplemental oxygen support <0.6 FiO2; (iii) Moderate: therapeutic supplemental oxygen support => 0.6 FiO2; (iv) Severe: unplanned non-invasive mechanical ventilation, CPAP, or invasive mechanical ventilation requiring tracheal intubation. In addition, they authors suggest excluding outcomes sometimes reported as PPCs because they do not share the same biological mechanism as PPCs. Examples of these outcomes include: (a) pulmonary embolism, (b) pleural effusion, (c) cardiogenic pulmonary oedema, (c) pneumothorax, (e) bronchospasm.

Atelectasis is defined as the alveolar collapse due to the mechanical ventilation during the general anesthesia[24]. Abbott and colleagues[23] recommended the US CDC definition of pneumonia *"Two or more serial chest radiographs with at least one of the following (one radiograph is sufficient for patients with no underlying pulmonary or cardiac disease): (i) New or progressive and persistent infiltrates, (ii) consolidation, (iii) cavitation; and at least one of the following: (a) fever (>38$^{o}$C) with no other recognized cause, (b) leucopaenia (white cell count <4*10^9 litre^-1) or leucocytosis (white cell count >12*10^9 litre^-1), (c) for adults >70 year old,*

*altered mental status with no other recognized cause; and at least two of the following: (a) new onset of purulent sputum or change in character of sputum, or increased respiratory secretions, or increased suctioning requirements,*

*(b) new onset or worsening cough, or dyspnea, or tachypnoea, (c) rales or bronchial breath sounds, (d) worsening gas exchange (hypoxemia, increased oxygen requirement, increased ventilator demand)*."

As for the respiratory failure, Abbott and colleagues[23] recommended the Berlin definition of Respiratory Distress Syndrome along with the need for mechanical ventilation: *"Timing: within 1 week of a known clinical insult or new or worsening respiratory symptoms and… Chest imaging: bilateral opacities not fully explained by effusions, lobar/lung collapse or nodules AND… Origin of oedema: respiratory failure not fully explained by cardiac failure or fluid overload (requires objective assessment, e.g. echocardiography, to exclude hydrostatic oedema), and… Oxygenation: mild PaO2:FiO2 between 26.7 and 40.0 kPa (200-300 mm Hg) with PEEP or CPAP=>5 cm H2O; moderate PaO2:FiO2 between 13.3 and 26.6 kPa (100-200 mm Hg) with PEEP=>5 cmH2O; severe PaO2:FiO2=<13.3 kPa (100 mm Hg) with PEEP=>5 cm H2O".*

Mechanical ventilation: The need for tracheal re-intubation and mechanical ventilation after extubation, and within 30 days after surgery or mechanical ventilation for more than 24 hours after surgery. The inclusion of non-invasive ventilation may be considered on a study basis."

## 2.2 Pulmonary postoperative complications burden in general surgery:

The incidence of PPCs in patients undergoing general surgery ranges from 2.6% to 8%, and associated mortality rates varies between 8% and 25%[9,25,26]. This large variation in incidence and mortality rates may be attributed to different definitions of PPCs used among the studies. Despite these variations, PPCs poses major economic burden and impact individuals' quality of life. For instance, Khan and colleagues[10] in a cohort of Canadian patients undergoing non-cardiac surgeries found that pneumonia, as a postoperative complication, increased the length of stay at hospitals by 75% after adjustments for patient characteristics and comorbidities. They also reported an increased cost by 47%. While acute respiratory failure accounted for 35% increase for hospital LOS and 41% increase in costs.

In a US cohort Thompson et al[27] reported that postoperative pneumonia in abdominal surgical patients increased hospital LOS by 11 days and hospital charges by $31,000 ($2,000 US per year).

Finding from Smetana et al. review[7] showed that the mean increase in hospital LOS specifically attributable to postoperative pneumonia was approximately 8 days in general surgeries. Another study calculated the cost increase caused by PPCs using the entire U.S. population. This work conducted by Shander et al[5] used the methods proposed by Linde-Zwirble et al[28] and showed that PPCs added 717 USD to the average cost of elective general surgery with 92,200 additional ICU admissions, 584,300 additional ICU days, and 3.42 billion USD in additional costs[5]. Similarly, The 5-year and 10-year mortality rates among US veterans patients who had general surgical procedures and developed postoperative respiratory failure were >50% and 70% [29], respectively. They authors also found that the median survival of patients who developed PPCs was decreased by 87% compared with those who did not develop PPCs[29]. Moreover, Arozullah et al[30], in their study of US veterans patients who had major general surgeries, reported patients who develop PPCs have a 30-day postoperative mortality rate of 21% compared with 2% in those who do not with an overall case fatality rate of approximately 10%.

### 2.3 PPCs in head and neck surgery and their risk factors:

A large variation in the incidence rates of PPCs is observed among patients undergoing HNS compared to the general surgery. For example, McCulloch et al and Loeffeblein et al reported similar PPCs incidence, 15% [13] and 18.8%[17], respectively. However, Petrar and colleagues reported an incidence rate of 44.8%[14] which is similar to Rao et al (46%)[12] and Ong et al (47%)[15]. This variation can be due to differences in the populations used for each study. For instance, while McCulloch et al [13] and Rao et al[12] studied different US cohorts, Petrar et al[14], Loeffeblein et al[18] and Ong et al[15] studied Canadian, German and New Zeeland cohorts. Reported risk factors that predicts occurrence of PPCs in HNS patients vary in the literature; for example, while studies that reported some predisposing factors did not find other comorbidities and characteristics that significantly affect the development of PPC, other studies found the contrary. This includes obesity[18], grade 3 according to American Society of Anesthesiologists (ASA) classification[18], alcohol use[18,31], age[12,14,31], male gender[12], history of lung disease[12,15], history of heart disease[14,31]. However, there is no consensus is for true PPCs predictors in HNS and we did not find any review summarizing the main risk factors for PPCs.

Therefore, there is a need to systematically map the existing literature for PPC risk factors for HNS patients. Thus, the first objective of this thesis was conducting a scoping review to identify the main risk factor for PPCs (Manuscript I).

## 2.4 Risk prediction for PPC

### 2.4.1 Definition of Risk Prediction:

Prediction modelling is an empirical approach to estimating disease or disease outcome probabilities[22]. Risk prediction models use predictors (covariates) to estimate the absolute probability or risk that a specific outcome is currently present (diagnostic prediction model) or will occur within a specific time period (prognostic prediction model) in a person with a specific predictor profile[22,32]. A model is a (mathematical) function that connects the presence or incidence of a studied outcome to a set of predictors[33]. Individual characteristics (e.g., age, sex), medical history, and physical examination, as well as imaging, blood and urine tests, and genetic markers can all be used as predictors[33].

As a result, prediction models are frequently developed to assist healthcare professionals and patients in making decisions about management plan options and directions such as extensive testing, starting, or stopping treatment(s), and lifestyle changes. These discussions help to inform patients about their risks of currently having (diagnosis) or developing (prognosis) a specific disease or outcome[34]. That being said, prediction models are not intended to replace healthcare professionals' reasoning, override their clinical experience, or take over their jobs, but rather to supplement their reasoning, provide informed consent, and decision making by giving more objectively estimated probabilities[19,34-38]. All of this is consistent with the concept of "personalized medicine," which aims to provide more patient-specific effect estimates to aid in more individualized clinical decision-making[39]. In addition, prediction models can also aid to classify a group of individuals into high and low risk groups, allowing for group tailored interventions[39].

Prediction models are becoming more common as access to robust data sets expand in the Big Data era[22]. Scientific publications on modelling for prediction have increased dramatically in recent years, totaling 651,000 from 1993 to 2017[22]. From 7,400 in 1993 to 17,000 in 2003, 39,000 in 2013, and 53,000 in 2017[22], annual numbers more than doubled every decade.

### 2.4.2 Risk prediction models design:

In general, risk prediction models are divided into two main groups: (i) diagnostic risk prediction models (ii) prognostic risk prediction models. A diagnostic risk prediction model estimates an individual's likelihood of currently having a specific disease[22] while prognostic prediction models calculates an individual's likelihood of developing an outcome or medical condition in the future[22]. As an example of PPCs, consider a prognostic risk prediction model for predicting the likelihood of postoperative respiratory failure that uses patient characteristics (e.g., functional dependency, medical history, and comorbidities) and procedure factors (e.g., surgery site and urgency) as predictors[9].. Diagnostic models' predictors could be an individual's sociodemographic (e.g., sex, ethnicity, and comorbidities) and other characteristics related to the outcome of interest, and an example for such a model the study by Jehi et al[40] that developed a diagnostic model for COVID-19 positive test result based on predictors mentioned before .

In clinical practice, prediction models can provide patients and physicians with information about the likelihood of a diagnosis or a prognostic outcome[22]. Prognostic estimates may be helpful in planning an individual's palliative care in terminal disease, or they may provide hope for recovery if a good prognosis is expected following an acute event such as a stroke[22]. Furthermore, classifying a patient based on his or her risk may be useful for communication among physicians in referrals and consults, for example[22].

Predictions in diagnostic models can be useful for estimating the likelihood that a disease develops. When the probability is relatively high, intervention is advised; when the probability is low, no treatment is indicated, and more diagnostic testing may be considered[22]. The prediction model hypothesis holds that better judgments may be made with a model than without one[22].

To develop a multivariable risk prediction model, there are few steps that are recommended to follow[33] as described in the literature.

### 2.4.2.1 Source of data:

The sources for data can be summarized as mentioned in Steyerberg[22] book: 1) retrospective cohort, 2) prospective cohort, 3) registry database, 4) case-control designs.

A prospective cohort design would be ideal for gathering data for building a prediction model[22,33,]. Randomized clinical trials (RCTs) data can be used to develop prediction models as they collect data prospectively as cohort studies. Prognostic models derived from RCT data may be less generalizable due to factors such as rigorous eligibility requirements, necessitating testing in a non-randomized scenario[34]. Retrospective cohort studies use existing data that are often documented for other purposes (e.g., regular care hospital records). While these designs have long follow-up, they have the risk of inferior, less methodically acquired data[34]. Retrospective cohorts are limited by definition and completeness of predictors and usually outcome assessment may not be based on protocols[22]. Unfortunately, retrospective studies dominate the prognostic literature[33], because they are simple and have low costs[22].

Registry databases are also used for prediction modelling. While they have the same limitations as a retrospective cohort study, they have the advantage of broader coverage of populations[22]. Finally, traditional case-control studies, in which cases and controls are drawn from an hypothetical source population, are effective for identifying independent predictors of an outcome from a wider set, but not ideal for developing a prediction model[33]. Because the baseline risk or hazard cannot be extracted from the data alone, this design does not allow for the assessment of absolute risks[33]. Also, they are limited by selection of controls and completeness of predictors[22].

### 2.4.2.2 Outcomes:

The outcomes of prediction studies should ideally be those that are important to patients. These might include death, development of the disease, or remission of the disease[33]. To eliminate

potential bias, outcome assessment should ideally be blinded to or independent of any knowledge of the predictors being considered[33].

### 2.4.2.3 Candidate predictors:

Candidate predictors are variables chosen for the study based on their predictive performance. Subject demographics, clinical history, physical examination, illness features, test findings, and, as previously noted, past therapies can all be candidate predictors[33]. All factors suspected of being connected with the result of interest may theoretically be evaluated as potential predictors, although this relationship does not have to be causal [33]. Skin color in the Apgar score and tumor markers as indicators of cancer progression or recurrence are two examples of highly predictive but non-causal elements in prediction models[34].

Researchers typically collect more predictors than can be analysed, much alone incorporated, in the final model. The events per variable (EPV) 1 to 10 'rule of thumb' is frequently used to limit the probability of false positive findings (predictors)[33]. This rule, which is not founded on solid scientific logic, claims that at least ten persons who have (produced) the event of interest are required for successful predictive modelling[41].

Finally, predictors should be explicitly specified and measured in a consistent and reproducible manner in order to increase the applicability and prediction stability of the resulting model in new people[42].

### 2.4.2.4 Data quality:

There is no agreement on how to evaluate the quality of data used for prediction modelling, thus this process is left for each investigator's judjment[33]. Ideally, measurements of the candidate predictors and outcomes should be standardized[33].

However, when there is evidence of considerable measurement error or inter-observer variability in some predictors, they are deemed less suitable. Because these predictors will likely result in different predictive ability of the model when tested or applied in other or future cohorts[33].

### 2.4.2.5 Missing data:

In medical research, especially prediction research, missing values are frequent. The potential impact of missing values on study outcomes grows in proportion to the amount of missing data. Missing data are generally linked to other relevant information or factors, including the outcomes under inquiry, either directly or indirectly[33]. Hence, simply excluding participants with missing values from the analysis reduces the effective sample size and may result in inaccurate estimates of predictor outcome associations and predictive performance of the final model because individuals with completely observed data are no longer a random subsample of the original study sample[43,44]. Another factor to consider when dealing with missing data is whether a variable that is used to calculate the risk probability and is frequently missing in the dataset[33]. Hence, the variable will also be unavailable in populations to whom the prediction model will be applied later[33]. If this is the case, it is best to leave it out of the prediction model[33].

### 2.4.2.6 Modelling continuous predictors:

The practice of dichotomizing a continuous variable into categories should be avoided, mainly because it loses information compared with when the continuous form of the variable is used[45-47].

### 2.4.2.7 Developing the final model:

There is no agreement in the literature on the optimum approach for deciding on the best model, that is, how candidate predictors are selected for inclusion in the multivariable analyses, and then how predictors are chosen for inclusion in the final prediction model[33]. There are two major common methodologies identified in the literature, both with variants: complete model and predictor selection strategy.

In the full model approach, all candidate predictors are included in the final prediction model. Advocates of this approach claim that prevents overfitting and predictor selection bias[22,48]. On the other hand, it is sometimes difficult to select predictors because previous knowledge of the most promising candidate predictors (priori) is necessary, while researching too many candidate predictors must be avoided[22,48].

The other approach is the use of predictor selection in the multivariable analyses[33]. Here, candidate predictors that are not significant in the multivariable model are removed[33]. Backward elimination begins with all possible predictors in the multivariable model and conducts a series of tests to remove or maintain variables in the model depending on a pre-set nominal significance threshold for variable exclusion, such as when comparing two models using the log likelihood ratio test[33]. In the less desirable forward selection strategy, the model is built up in phases starting with the best candidate predictors. Forward selection, unlike backward elimination, does not allow for the simultaneous evaluation of the impacts of all candidate variables[49].

In this approach, so-called overfitted models may arise, specifically in small datasets[33]. Hence, unstable models might result because the selected predictors will vary depending on the specifics of the dataset used[33]. Therefore, regardless of which type of variable selection used, subsequent internal validation of the models using, bootstrapping techniques, in which this predictor selection process is repeated in every bootstrap sample (this will be discussed later), is recommended to gain insight into the likelihood of the model missing important variables, being overfitted or unstable[22,48].

The multivariable analysis calculates regression coefficients (log odds or Hazards Ratio) for each predictor in the final model, which are then adjusted for the other predictors[33]. The coefficients thereby measure each predictor's contribution to the outcome probability or risk estimation[33]. A regression coefficient, in other words, reflects the influence of a one-unit (or one-step in the case of categorical variables) increase in the level of the relevant predictor on the estimated outcome risk while all other predictors in the model are held constant[33]. An estimate of the baseline probability risk or the anticipated risk for an individual with all predictor variables set to zero is another essential statistic from a regression analysis in prediction modelling research[33]. The intercept of the logistic regression model indicates the baseline risk. As a result, predicted probabilities for developing the event within a specific time period for individuals can be calculated by combining the intercept or estimated baseline hazard, the observed values of the predictors, and the corresponding regression coefficients in mathematical functions specific to the statistical methods used to develop the model.

### 2.4.2.8 Assessing the predictive performance:

The two most significant parts of a prediction model are discrimination and calibration. Discrimination focuses on rating individuals from low to high risk, whilst the calibration focuses on the absolute chance of experiencing an event[50]. Calibration is best plotted by showing recorded result frequencies against mean expected outcome probabilities or hazards within subgroups of individuals arranged by increasing estimated likelihood[22,48].

Several statistics are available to summarize discrimination, but the c-index (equal to the area under the receiver operating characteristic curve for logistic models) appears to be the most employed. There have been produced generalized variants of the c-index for survival analysis that allow for censoring[51,52]. The c-index for a prognostic model is the likelihood that, given two individuals, one of whom will develop the event of interest and one who will stay event free, the prediction model will assign a higher probability of an event to the former.

### 2.4.2.9 Internal validation:

When prediction models are evaluated on new but comparable individuals, they can be anticipated to perform optimistically in the data sample from which they are created[33]. This is because the models were built to match the development sample optimally, but they get less accurate when evaluated on new but similar individuals, which is known as overfitting[33]. When the number of outcomes/events in the development sample drops and the number of candidate predictors in the development sample (compared to the number of events) grows, the potential for optimism in model performance increases[33]. Also, when no prediction selection procedures are applied, model performance is generally the best[22,48].

Internal validation means only the study sample data is used to estimate the potential for overfitting and optimism in model performance[33]. While internal validation is frequently done by randomly splitting the dataset into two subsets, that is, a development sample (such as three-fourths of the original dataset) and a validation sample, this approach is statistically inefficient because not all available data are used to develop the prediction model[33]. Thus, bootstrapping is

the preferred method for internal validation, especially when the development sample is relatively small and/or a high number of candidate predictors is studied[48].

Bootstrapping is a statistical method that aims to imitate the sampling process in medical research, as the study sample is supposed to be a representation of the source population[33]. Bootstrapping uses only the data at hand by sampling with replacement a study sample of the same size (to preserve the precision) from the original study sample in which the prediction model was developed[33]. Drawing samples with replacement imitates that random sampling component, making bootstrap samples similar, but not identical to the original study sample[33]. In every bootstrap sample (usually 100 or 500 samples), the data are analyzed as in the original study sample, repeating each step of the model development including applied predictor selection strategies[33]. This may create a different model developed from each bootstrap sample with corresponding c-statistic. Thereafter, prediction from each bootstrap model on the original study sample (imitating the source population), is used to calculate difference in c-index[33]. The average of all these 'c-index differences' indicates the optimism in the apparent c-index of the prediction model that was initially developed in the original study sample[48,53].

All data are, therefore, used for model construction with bootstrapping, and it offers insight into the extent to which the produced model (in the original development sample) is overfitted and overly optimistic[33]. Furthermore, by repeating the complete selection process in each bootstrap sample, bootstrapping approaches can account for the impacts of all predictor selection processes done in the studies[33]. As a result, bootstrap-adjusted performance (e.g., c-statistic) more accurately represents what can be expected when the model is tested or applied in fresh individuals from the same theoretical source population[33].

#### 2.4.2.10 Presentation of the model:

The final prediction model should always be given as the original regression model equation, that is, regression coefficients (including the intercept for a logistic model). As a result, future researchers and users will be able to apply the model to new individuals to predict the risk of the outcome.

A model can also be presented as predicted probabilities for developing the outcome within a specific time period for individuals by combining the intercept or estimated baseline hazard, the observed values of the predictors, and the corresponding regression coefficients in mathematical functions specific to the statistical methods used to develop the model[33].

### 2.4.2.11 External Validation:

It is not enough to demonstrate a produced model's decent or good performance on the development sample alone, simply because most models there display optimistic results, even after corrections from internal validation techniques[54]. It is critical to validate that any established model predicts well in 'similar but different' persons outside the development set, and hence is generalizable to them[54]. The greater the difference between these other cases and the development study, the stronger the model's generalizability test[54]. Internal validation uses only development data and hence does not give the degree of variability that would be experienced in real-world implementations of the model[54].

Model external validation involves taking the original model or simplified score, with its predictors and assigned weights (regression coefficients), as estimated from the development study; measuring the predictor and outcome values in the new individuals; applying the original model to these data; and quantifying the model's predictive performance[55-58]. As discussed before, discrimination, calibration and classification are important features of predictive performance of prediction models to be quantified in external validation studies.

### 2.4.2.12 Temporal validation:

Temporal validation, which is simply applying the model's equation as done in external validation, is occasionally performed using an existing dataset that has Individuals from the same institution but from a subsequent time period[54]. They still use the same inclusion and exclusion criteria as well as the same definitions and assessment methodologies for predictors and outcomes[54]. When it involves a prospective design explicitly planned for the validation purpose that begins after the model has been built, temporal validation may allow for more variance, if not simply due to changes in healthcare through time[54].

### 2.4.2.13 Updating a prediction model:

When a model with a lower predictive performance is discovered, 'validation investigators' prefer to simply reject it and design or fit a new one, often by fully redoing the predictor selection process[54]. This results in the loss of earlier scientific knowledge gathered in the preceding (developing) study, which goes against the idea that conclusions and guidelines to improve evidence-based treatment should be based on as much information as feasible[54]. Furthermore, clinicians are forced to choose which model to apply in their patients, despite the fact that several have been designed for the same objective[54]. Instead of constructing new models for each new patient sample, updating current prediction models and adjusting or recalibration them to the local conditions or setting of the validation sample at hand is a far better option[22,54]. As a consequence, the updated models add information from new individuals with that gathered in the original model[59-61]. As a result, the updated models are tailored to the features of new individuals and are likely to be more transportable to other persons. Methods to update models were suggested[22,60], one of them focuses on calibration improvement by adapting the original prediction model's baseline risk or hazard (if known) to the people in the validation sample[55,62]. This needs only one parameter (intercept) of the original model to be changed. Additional updating methods range from adjusting all predictor weights at once to adjusting a specific predictor weight to adding an entirely new predictor to the existing mode[22,62].

The use of the aforementioned procedures results in updated models that are tailored to the circumstances of the validation sample. However, just as with a newly produced model, it was proposed that updated models to be validated for transportability and impact before being used in everyday practice [59]. This thesis work adopts the recommended steps described in section 2.4.2 to achieve it objective.

## 2.4.3 Current PPCs models and what is used in HNS:

Several preoperative surgical risk assessment models were developed to identify high PPCs risk patients among those who undergo general surgery, namely the Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT)[25] and the Gupta Pulmonary Index[9].

The latter utilizes data from the American College of Surgeons –National Surgical Quality Improvement (ACS-NSQIP) database[9]. Although these risk assessment tools were developed using large surgical cohorts in both Europe (59 Spanish hospitals)[25] and the United States (211 hospitals)[9], only a few HNS cases were included. In fact, the percentage of HNS evaluated in both ARISCAT and Gupta et al.'s study cohorts were 6% and 0.3%, respectively[9,25]. Consequently, ARISCAT score displayed poor ability to discriminate those with and without a PPCs in 794 HNS patients in southeastern of the United Satates[21]. Similarly, the Gupta Pulmonary Index had an insufficient accuracy and poor discrimination ability in predicting complications of 128 HNS patients in Portugal [20]. On another note, the only model that was developed specifically for HNS patient from our literature search was the one by Smith and colleagues[31]. The model was developed using their hospital's data in the US for 794 patients who underwent head and neck tumor resections. Their model had a good discrimination ability (AUC 0.75), but was not validated externally on any other cohort of HNS patients. Hence, we cannot assess this model's generalizability without external validation reports. Unfortunately, Smith et al[31] did not report their model comprehensively (i.e., intercept and coefficients), which limits the external validation for future studies. Another limitation was the predictors used in their model, as some those predictors were not available in a database registry such as the ACS NSQIP (e.g., preoperative metabolic equivalents)[31].

Despite the limitations of both models (ARISCAT and Gupta Pulmonary risk Index), they are widely employed in preoperative surgical risk assessment prior to major HNS. However, by utilizing these models in the HNS patients, experts may risk misclassification of patients' preoperative assessment and thus, to a suboptimal care delivery[21].

## 3 Rationale:

Around 300 million patients undergo general surgical procedures in the world, 30% of these patients experience postoperative complications[1-3]. One of the main categories of these complications is PPCs. The incidence rates for PPCs in general surgery vary from 2.6% to 8%[9,25,26], with mortality rates that can reach 25%[9]. In HNS, the incidence rates of PPCs vary from 4.5% to 47%[11,14,15,17]. PPCs are associated with increased mortality, morbidity, hospital LOS, and costs[10,14,18]. Reported risk factors that predicts occurrence of PPCs in HNS patients vary in the literature and no consensus is found for definitive predictors. Therefore, there is a need to systematically map the existing literature for PPC risk factors for HNS patients.

To mitigate PPCs, risk prediction can be used to identify patients at high risk of developing the outcome, facilitate informed consent process and encourage exercising risk reduction measurements, such as preoperative respiratory physiotherapy and smoking cessation[19]. Further, risk prediction can also guide strategies that prompt considerations of nonsurgical alternatives or advise postponing surgical procedures[25]. To best of our knowledge, there are no PPCs risk prediction models that are externally validated and have good performance, discrimination ability, and calibration for HNS patients.

Steyerberg[22], in the book "Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating," suggests examining the existing risk prediction models' performance and check their clinical applicability prior to developing new risk prediction models. This practice can prevent redundancy and facilitate the process of refining the existing models. In this study, we adopted a diagram published by Maarten van Smeden over his Twitter social media in 2018[63] highlights the best practices of developing a risk prediction. The diagram also provides insights to a set of questions that are rought to be answered prior to conducting a modeling study. These questions could be framed in PPCs risk prediction modeling in HNS as follows:

1. Is there a need for a risk prediction model for PPCs in HNS patients?
2. What is the target population of interest?

3. Are sufficient data on HNS and are the predictors of interest available?
4. Can any existing model for PPCs or its categories be validated or updated for HNS patients?
5. Is there a large dataset for HNS patients?
6. Are the model predictors identified by solid research on the risk factors?

The first question's answer is "Yes" based on what is discussed in the previous paragraphs. Answering the second question, the model is needed for all HNS patients at high risk of developing PPCs[21]. Considering the third question, different studies have explored PPCs and their risk factors in HNS patients[11,14,15,17], providing a good literature background. Among these studies, the Gupta Pulmonary risk Index specifically investigated the risk factors of postoperative respiratory failure using the huge ACS NSQIP datasets for all types of surgeries, including HNS. They develop their model on the 2007 patients and validate it on the 2008 patients[9]. Therefore, the updated NSQIP dataset (2018-2019) can be used to develop or validate the model. To answer the fourth question, it is essential to understand the status of current PPCs risk prediction modeling and identify the models that are potentially generalizable to HNS patients. Therefore, prior to planning to develop new models for HNS patients, there is a need to review the literature in PPCs risk factors for HNS patients and explore the potentially significant risk factors and their clinical feasibility to record in the HNS patients. Finally, the Gupta pulmonary model is outdated and is warranted to be updated to ensure he quality of performance of this model in HNS patients. This thesis project was designed to address the above gaps in knowledge.

## 4. Objectives:

The overarching goal of this thesis project is to develop a validated prognostic risk prediction model for PPCs in HNS patients. The specific objectives to achieve this goal are:

1. To conduct a scoping review to systematically map the literature on PPCs risk factors in HNS patients and assess the risk of bias of the literature on this topic. Such a review will help in identifying the reported risk factors for PPC in HNS patients.
2. To temporally validate and update an identified PPC risk prediction model using data from NSQIP ACS.

*Manuscript I* titled "Postoperative Pulmonary Complications Risk Factors in Head and Neck Surgery: A Scoping Review" addresses the first aim of this thesis by systematically reviewing the published papers on the PPCs risk factors in HNS. The specific objectives of this manuscript are to: 1) identify types of study designs and data sources used to examine PPCs risk factors in HNS, and 2) identify risk factors associated with PPCs in HNS patients. The results of the scoping review (*Manuscript I*) can provide valuable information on the reported risk factors for PPCs in HNS patients that can be used as predictors in a risk prediction model. Importantly, it identifies candidate predictors that can be used to update a current PPC risk prediction model to be optimized for HNS patients, which is an essential information to fulfill the second objective of my thesis. The model has been developed and validated by Gupta et al. (2011) on a sample of general surgery patients using ACS NSQIP dataset[9].

*Manuscript II* titled "Temporal Validation and Extension of a Prediction Model for Postoperative Pulmonary Complications" aims to temporally validate Gupta original model using data derived from ACS NSQIP database over the years following the development of the model. Then update the model with predictors recorded in the dataset and were reported to predict PPCs in HNS patients.

## 5. Methods:

### 5.1 Objective 1: Scoping review methodology:

The first objective of this thesis work was to conduct a scoping review to identify the main PPCs risk factors. As previously discussed in the literature review, while there is a plenty of studies evaluating PPCs risk factor, the results of these studies are not congruent, and demonstrate a wide range of variation. More importantly, there are no studies synthesizing this evidence while evaluating the strengths and limitations of each of the included studies. Below is an overview and a detailed description of the scoping review's:

Scoping reviews are conducted to identify, map, collate and summaries the existing literature that covers the field of question[64]. They aim to assist researchers to recognize fundamental ideas, theories, evidence sources and gaps in knowledge in that field[65]. In contrast to systemic reviews, scoping reviews incorporate the "Big Picture" in the underlying literature rather than answering a narrowly defined specific question[66]. More details of the steps followed to conduct the review are provided in Manuscript I. Here I will follow some methodological foundations for conducting scoping reviews.

According to Joanna Briggs Institute (JBI) manual[66], scoping reviews can be conducted to address several objectives

1. A preparatory exercise prior to conducting a systematic review. Scoping reviews can create a map of the range of available evidence might be conducted, hence, allowing researchers to pose more precise questions in systematic reviews. Thus, as mentioned before in the Literature Review Chapter, a comprehensive review that reports all risk factors for PPC in HNS patients was not found.
2. Investigating broad topics in order to uncover evidence gaps, explain essential concepts, and report evidence that addresses and guides practice in a given area. Scoping reviews can be used to organize evidence by time (when it was published), location (country), source (peer-reviewed or grey literature), and/or origin (healthcare or academic discipline).

3. Scoping reviews, as important tools for evidence reconnaissance, may give a wide perspective of a topic. For example, a scoping study aimed at developing a "concept map" may strive to investigate how, by whom, and for what purpose a certain phrase is used in a given field. A review of this type would attempt to map how the phrase is used in the literature, what it refers to, and what it entails. As such, PPCs risk factors in HNS is a broad topic that needs to be systematically mapped to know where we stand in the literature with regards to a given topic.

This is generally done according to JBI manual in the following framework steps[66]:

1. Identifying the research question: clarifying and linking the study's objectives and research questions: The scoping review question leads and directs the scoping review's inclusion criteria. The clarity in the review question can facilitate a successful literature search more successful, and provides clear structure for the preparation of the scoping review report. The question, similar to the title, should include the PCC (Population, Concept, Context) components. A scoping review will typically include one main question. Thus, our research question in *Manuscript I* was "What is known in the literature about the risk factors for PPCs following HNS?"
2. Identifying relevant studies: balancing feasibility with breadth and comprehensiveness of the scoping review process

A scoping review's search strategy should attempt to be broad to discover both published and unpublished (grey literature) primary research and reviews. Our research objective was to identify risk factors in the selected studies using methodological and statistical rigor. Hence, we considered only peer-reviewed literature, arising primarily from scientific journal publications. A medical librarian trained in knowledge synthesis techniques conducted a systematic scoping search of the literature to identify candidate articles[67].

3. Study selection: using an iterative team approach to selecting studies and extracting data

This can be achieved using an inclusion criteria that explicitly identifies the basis on which sources was evaluated for inclusion in the scoping review. These criteria should help readers to

properly comprehend what the reviewers recommend, as well as guide the reviewers themselves to base judgments on the sources to be included in the scoping review. Thus, in *Manuscript I,* we included the papers that were only related to HNS, did include PCCs as the outcome, and did mention risk factors of PPCs. Furthermore, only original peer-reviewed studies were included. Due to limited resources for translation, only articles published in English were included, while there was no restriction related to time of publication.

The data extraction technique is referred to as charting the outcomes in scoping reviews. This procedure provides a coherent and detailed overview of the results that corresponds to the scoping review's purpose and question. At this stage, we charted a table to record each included study critical information, such as author, reference, and results or conclusions related to the review question(s). This was refined further throughout the review stage, and the charting table was updated accordingly. We opted to chart the following critical information: (a) Author(s); (b) Year of publication; (c) Aims/purpose; (d) Study population and sample size; (e) Methodology; (f) Intervention type and details of it; (g) Outcomes and details of these (e.g., how measures); (h) Key findings that relate to the scoping review question.

4. Charting the data: incorporating a numerical summary and qualitative thematic analysis

Usually, the reviewers state how many studies were retrieved and chosen. Hence, a search decision flowchart was supported with a narrative explanation of the search decision process. The flow chart clearly illustrated the review decision process, including the search results, citation removal, study selection, and final summary presentation. The summary aimed to convey the included studies goals or intentions, ideas used, and outcomes that relate to the review question logically and systematically. The findings were grouped into the following major conceptual categories (tabulating data): "surgical intervention type," "study population", "aims," "study design used," "important findings" (evidence established), and "gaps in the research." A detailed explanation was provided for each category which is referred to as thematic analysis.

5. Collating, summarizing and reporting the results: identifying the implications of the study findings for policy, practice or research

The findings were presented considering the existing literature, practice, and policy. Scoping reviews are subject to the limits of any review; as some relevant sources of information might be ignored, and the review was contingent on the availability of information on the review question. A scoping review usually, as our review in *Manuscript I*, does not give an evaluation of the quality or degree of evidence, hence recommendations for practice cannot be assessed. However, based the review's findings of knowledge gaps allowed us to provide clear and specific suggestions for further research.

## 5.2 Objective 2: Validation and extension of the model:

The second objective of this thesis was to temporally validate and update a PPCs risk prediction to be optimized for HNS patients. As mentioned in the literature review, there are certain steps and methods to validate and update an existing model. Hence, a brief review of the methodology used in *Manuscript II* is presented here. However, detailed methodology steps are reported in *Manuscript II*.

### 5.2.1 Data acquisition: ACS NSQIP database:

The ACS-NSQIP database is designed to support surgeons and hospitals in understand ingthe quality of care they provide for patients in similar situations in North American hospitals and other participant hospitals worldwide[68]. It was launched by the ACS in 2001 with eighteen participant hospitals as a pilot program. The program was found to function well in terms of reducing morbidity and mortality in participating hospitals. In 2004, ACS NSQIP started enrolling other private sector hospitals as participants. Now, the database has more than 700 participant hospitals from forty-nine out of the fifty states in the US and more than a hundred hospitals from eleven other countries worldwide. ACS NSQIP collects reliable and validated data on patient demographics, laboratories, comorbidities, and thirty-day postoperative outcomes for patients undergoing a broad range of operations across all surgical subspecialties. These data are used to provide hospitals with risk-adjusted 30-day outcomes comparisons, and we have previously leveraged these data to develop a risk prediction tool.

### 5.2.2 The identification of HNS patients:

Through our collaborations with the Montreal General Hospital, we obtained data from hospitals participating in ACS NSQIP. Patients who underwent at least one procedure in the head and neck region spanning all surgical subspecialties from January 1, 2018 to December 31, 2019 were identified. Head and neck surgery is a specialty within the field of ear, nose and throat medicine (otolaryngology) combined with oral and maxillofacial medicine that focuses on surgical treatement of head and neck disorders. Head and neck surgeons treat cancerous and non-cancerous tumors including oral cancer, pharynx cancer, nasopharyngeal cancer hypopharyngeal cancer, salivary gland cancer, paranasal and sinus cavity cancer, thyroid cancer, nasopharyngeal, laryngeal or larynx cancer, and sarcomas of the head or neck.

These procedures were identified using one of the ACS NSQIP database records coding named Current Procedural Terminology (CPT®), which refers to a set of medical codes used by physicians, allied health professionals, nonphysician practitioners, hospitals, outpatient facilities, and laboratories to describe the procedures and services they perform.

CPT® codes ranges included: 21010-21499 (musculoskeletal head operations), 21501-21899 (musculoskeletal neck operations), 31300-31599 (larynx operations), 30000-30999 (nose operations), 40490-42999 (Mouth, tongue, and pharynx operations), 60000-60659 (Thyroid, parathyroid, carotid bodies operations), and 69000-69979 (ear operations).

### 5.2.3 Variables of interest in the database:

The dataset contains an array of information (275 variables) including sociodemographic (e.g., age, sex), lifestyle (e.g., smoking history) clinical and laboratorial factors (e.g., medical history, surgical procedure code, preoperative lab tests, postoperative course, length of stay, complications occurred in 30-day period postoperatively).

The model by Gupta et al. used 5 predictors: type of surgery, emergency case, dependent functional status, sepsis, and ASA class[9]. We used these predictors in our base multivariable logistic regression model to build a priori based on the reported predictive values. Subsequently, an extension of this model including predictors identified in the scoping review was performed using multivariable logistic regression model.

### 5.2.4 PPCs definition according to Gupta et al.:

The primary outcome was the occurrence of PPC within the first 30 days after the operation. According to Gupta's[9] definition, PPCs occurred if a patient had, within the 30 days after the primary procedure, the first occurrence of any of the following: (1) required placement of an endotracheal tube or other similar breathing tube or ventilator support, which was not intended or planned and (2) have a total cumulative duration of ventilator-assisted respirations greater than 48 hours during the postoperative hospitalization and any subsequent hospitalizations

### 5.2.5 Data analysis:

### 5.2.5.1 Temporal validation of Gupta et al.'s model on HNS patients:

The Gupta et al. model was developed and validated using NSQIP datasets for 2007 and 2008 patients. Thus, we used the model's predictions to validate it on HNS patients from 2019 NSQIP dataset as temporal validation[54]. We assessed the Gupta et al. model performance using scaled Brier score and $R^2$ (Negelkerke), calibration by calibration slope and discrimination ability by c-statistic.

### 5.2.5.2 Updating of Gupta et al. model:

One method to update an existing model is to extend the model with selected predictors using a process of intercept recalibration [22]. We used dataset from year 2018 for development of the updated model. In this step, we fitted Gupta et al.'s predictors' coefficients as reported in the original model without their intercept. We then extended the model with our selected predictors: age, sex, operation time, BMI, history of smoking, history of COPD, reconstruction procedure, and site of the procedure while recalibrating the model's intercept. We assessed the updated model performance using scaled Brier score and R2 (Negelkerke) and discrimination ability by c-statistic. Also, calibration slope was plotted to assess the calibration of the updated model.

External validation was done using the2019 data, which was saved for later use during the development step.

To elaborate and explain these metrics used in *Manuscript II*, here is a brief description for each one:

A) The Brier score: first proposed by Glenn W. Brier in 1950[69] as the square of the differences between the predicted values and the actual values of the outcome. The Brier score shows how accurate the model predictions are. It can be explained as 'the lower the Brier score, the better the overall predictive performance of a model'. The best risk prediction model performance that has no error in the predictions receives a Brier score of 0. If a model is non-informative, which means it assigns the same probability to all the predictions, the model receives a Brier score of 0.25. The Brier score is usually scaled and reported as a percentage. The scaled Brier score ranges from 0% for the non-informative model to 100% for the perfect model[22,70].

B) Nagelkerke $R^2$: another useful metric in overall performance measurement. Measuring the $R^2$ (i.e. Explained variation) is a common method of performance assessment for continuous outcomes[70]. Nagelkerke $R^2$ is similar to Pearson's $R^2$, but estimated for a generalized linear models such as logistic regression[22,70]. For binary outcomes Y, we scored a model predictions as follows p: $Y*\log(p) + (Y-1)*(\log(1 - p))$[22,70].

C) C-index: it is a measure of discriminative ability of a model which is defined as the capability of a model in discriminating between participants with and without the outcome. Discrimination capability of a model can be evaluated by measuring the Concordance statistic (C-Statistic) and discrimination slope (index). The C-statistic of the risk prediction model is defined as the probability that a randomly selected participant with the outcome of interest will have a higher predicted probability compared to a randomly selected participant without the outcome[22]. Moreover, C-statistic indicates how good the model classifies the participants into two groups. With a binary outcome, the value of C-statistic equals to the area under the receiver operating characteristic (ROC) curve (AUC). ROC curve can be derived from plotting the sensitivity of the model (true positive rate) over 1–specificity (false positive rate)[71].

D) Calibration slope: as discussed in the literature review chapter, calibration of a risk prediction model is the agreement between the predicted and observed values of the outcome variable[50]. A well-calibrated model correctly predicts a decile of probability 100P % of the time with P confidence. For instance, a perfectly calibrated model that predicts 10% risk of developing PPCs for a number of HNS participants, 10% of these participants should be truly having PPCs. Calibration of a risk prediction model can also be visually evaluated by plotting the predictions and observed values in a two-dimensional graph. If the outcome is binary, such as predicting the risk of developing PPCs after HNS, the calibration plot's Y-axis will be the observed proportion of participants with the outcome of interest at a specific predicted risk level. The x-axis of such calibration plot contains the groups of predicted probabilities. A diagonal line in the calibration plot represents the ideal calibration. That is for a perfectly calibrated model, the predicted probabilities(x-axis) are equal to observed proportions (y-axis). For example, if the smoothed calibration line is above the ideal (diagonal) line, the model is underestimating the probabilities of the outcome at that point. While if the smoothed calibration line is below the ideal line the model is overestimating the probabilities of the outcome at that point.

## 6. Results:

### Preface for Manuscript I:

According to the literature review chapter of this thesis, PPCs are associated with increased mortality, morbidity, hospital LOS, and costs. Reported risk factors that predicts occurrence of PPCs in HNS patients vary in the literature and no consensus was found for definitive predictors. Different risk factors for PPCs have been reported for HNS. Nonetheless, little is known about the studies reporting risk PPC's factors in terms of risk of bias and PPC definition. *Manuscript I* will fill this knowledge gap by reviewing the papers discussing these risk factors in light of a clear PPCs definition. Based on McGill University's guidelines for a manuscript-based thesis, the next chapter stands alone as *Manuscript I* with an independent reference list and appendices.

## 6. Manuscript I:

**Risk factors for postoperative pulmonary complications after head and neck surgery: A scoping review**

Mohammad Al-Tamimi[1], Sreenath Madathil[1], Ahmed Derbas[2], Hamed Ghanati[1], Martin Morris[3], Belinda Nicolau[*1]

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

[2] Oral and Maxillofacial Surgery Department, Hamad Medical Corporation, Doha, Qatar

[3] Schulich McGill Library of Physical Sciences, Life Sciences and Engineering, McGill University, Montreal, Quebec, Canada

[*]Corresponding author: Dr. Belinda Nicolau, PhD; 2001 Avenue McGill College, Montreal, Quebec, H3A 1G1, Canada; Tel: +1 514 835 8681, Email: belinda.nicolau@mcgill.ca

**Funding statement:** Authors received no funds to conduct this project.

**Data availability:** No new data were generated in this project.

**Conflict of interest:** Authors declare that they have no conflict of interest.

**Running title:** H&N Postoperative Pulmonary Complications

**Keywords:** Postoperative Complications; Risk Factors; Head and Neck Neoplasms; Surgical Procedures, Operative; Review Literature as Topic

## Abstract:

**Background**: Postoperative pulmonary complications (PPCs) are common following head and neck surgeries (HNS). The incidence of PPCs in the HNS patients varies between 4.5% and 47%, and these complications are associated with increased morbidity, mortality, intensive care unit admission, and hospital length of stay. While PPCs are associated with common risk factors in cardiac and general surgeries, evidence lacks on risk factors associated with PPCs in HNS patients. Therefore, we undertook a scoping review of research on PPCs risk factors in HNS patients.

**Objectives:** This review aims to systematically map the literature on the PPC risk factors for HNS patients and identify the gaps in knowledge.

**Methods**: We used the Arksey and O'Malley's five-stage scoping review framework to carry out our review. This included: (1) identifying the research question; (2) identifying the relevant literature; (3) selecting the studies; (4) charting the data and (5) collating, summarizing and reporting the results. With the help of librarian we developed a list of keyword combinations related to head and neck surgery, pulmonary complications, and risk factors to inform the search strategy. The following databases were searched: Embase, Medline(Ovid), CINAHL and Scopus. We limited the article in English with no restriction by date. The risk of bias was appraised using Joana-Briggs Institute tool.

**Results**: We identified thirty peer-reviewed studies that reported PPC risk factors for HNS patients. The majority of studies (70%) were retrospective cohorts. The reported risk factors of PPCs in HNS patients were age, male sex, smoking history, respiratory comorbidity, smoking history, body mass index, operation time, site of the surgery, and reconstruction procedure. Majority of included studies had a high risk of bias mainly in the analytical domain.

**Conclusion**: Limited numbers of studies provide good quality results on PPCs risk factors and discussed casualty. However, the review identified few risk factors that might be more specific for the HNS patients.

## Introduction:

Every year, around 300 million surgical procedures are performed worldwide[1]. Several patients undergoing these procedures experience postoperative complications with rates reaching as much as 30% in some diseases such as colorectal cancer[2,3]. Apart from the burden on patients, these complications may change the course of recovery by increasing the risks of reoperation, prolonged hospital length of stay (LOS), and mortality[4]. One of these complications is postoperative pulmonary complication (PPC) that comprises substantial and morbid conditions affecting the respiratory tract that can adversely influence the clinical course and recovery of patients after surgery[5]. Among different PPCs, atelectasis (the collapsed and non-aerated region of the lung parenchyma[6]), pneumonia, respiratory failure, and exacerbation of underlying chronic lung disease are the most substantial and morbid conditions[7]. Thirty percent of patients undergoing surgeries lasting at least 2 hours with general anesthesia and mechanical ventilation may experience PPCs[8]. Importantly, PPCs can result in potentially significant morbidity and mortality when they include more severe atelectasis, bronchospasm, postoperative pneumonia, pneumothorax, acute respiratory distress syndrome, pulmonary embolism, or respiratory failure[5]. Three percent of patients experience pulmonary respiratory failure after surgery with 25% of them dying within the 30 days after surgery[9].

PPCs also impose a significant economic burden as it increases hospital LOS, a standard reference point for assessing the cost of medical services[5]. For example, Khan et al.[10] observed an 89 % increase of hospital LOS and a 55% rise in hospital costs for non-cardiac patients experiencing postoperative pneumonia.

Head and neck surgery (HNS) is a specialty within the field of ear, nose and throat medicine (otolaryngology) combined with oral and maxillofacial medicine that focuses on surgically treating head and neck disorders. The incidence of PPCs in HNS patients varies between 4.5% and 47% [11-16], and these complications are associated with increased morbidity, mortality, intensive care unit admission, and hospital LOS[14,17,18].

Years of research have identified different risk factors for PPCs occurrence. While some of these factors are patient-related (e.g., age, comorbidities (chronic lung disease, congestive heart

failure, obstructive sleep apnea), functional dependence, obesity, and smoking), others are procedure-related (e.g., site of surgery, duration of surgery, anesthetic technique, and emergency of the surgery were the procedure-related[7]).

Despite different reports of PPCs risk factors in HNS patients[14,16], there is no agreement between the studies about relative contribution of these factors (i.e., effect sizes). Also, a considerable difference in the PPCs' measurement methods is evident among the studies, which might explain the significant variation in the reported PPCs incidence rates. Furthermore, we could not find a comprehensive summary in the literature that helps to better understand the different risk factors and their roles. Importantly, a comprehensive review of known risk factors is essential in designing interventions and further investigate the mechanism of PPCs development.

Therefore, we aim to systematically (1) identify and map the literature about risk factors of PPCs in patients undergoing HNS; and (2) identify the gaps in knowledge for PPCs in patients undergoing HNS.

## **Methods**:

This scoping review started by establishing a research team that encompassed researchers, librarians and graduate students with expertise in epidemiology, head and neck cancer and research synthesis. Our methodology was based on Arksey and O'Malley's framework for scoping reviews[19] that includes the 5 phases: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarizing, and reporting the results. Following these phases our team discussed and refined the research question through several iterations before finalizing the precise and most relevant research question to maximize the expected contribution of the scoping review. Further, with the continued support from librarian in the team specialized in scoping reviews, we developed a detailed protocol for scoping review. Further details of each of these phases in this process are provided below.

### ***Research question:***

The current review systematically mapped the literature and identified gaps in knowledge on PPCs risk factors following HNS based on the research question of: "What is known from the literature about the risk factors for PPCs following HNS?"

### *Identifying relevant studies:*

Risk factors identification is a research objective that requires methodological and statistical rigor. Hence, we decided to consider only peer-reviewed literature, arising primarily from scientific journal publications. A medical librarian trained in knowledge synthesis techniques (MM) conducted a systematic scoping search of the literature to identify candidate articles[20]. A strategy (Table 1) for Ovid Medline was constructed using Medical Subject Headings (MeSH), combining the concepts of head and neck surgery, pulmonary complications, and risk factors. This was then translated to Embase (Ovid), CINAHL and Scopus. Initial searches were carried out on April 20, 2021. The language of articles was limited to English, while no restrictions were placed on publication year.

All citations were imported into EndNote 20[21] and the duplication was resolved. Citations were then imported into the web-based software Rayyan web[22] for screening of title and abstract and data characterization of full articles.

### *Study Selection:*

Figure 1 presents the literature selection process using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for Scoping Reviews[23] diagram. A two-stage screening process was followed to evaluate the relevance of articles identified in the search. In the first stage, two independent reviewers (MT, AD) screened the studies with regard to their title and abstract. The reviewers excluded the papers that were not related to HNS, did not include PPCs as the outcome, or did not mention risk factors of PPCs. Furthermore, only original peer-reviewed studies were included. Because of limited resources for translation, only articles published in English were included. In the second stage, the discrepancies between the two reviewers' screening were resolved during the discussion with two experts. The overall agreement between the reviewers was assessed by calculating Cohen's kappa score.

Following the screening process, articles that met the inclusion criteria were included in the analysis.

### *Data charting and synthesis:*

Two reviewers (MT and AD) extracted the data regarding the included studies' title, authors, publication year, study design, population, sample size, exposure (surgical intervention), outcome, variables included in the models, and results. Next, MT collated the collected data by combining both quantitative (reported odds ratios and numbers of studies) and qualitative (classification under conceptual categories) approaches, which is suitable for amalgamating heterogeneous studies and data[24]. Two blinded reviewers (MT and HG) independently performed quality critical appraisal to assess internal validity and risk of bias in the included studies using the Joanna Briggs Institute (JBI) assessment tools [25].

## Results

### *Study selection:*

Following the screening process, 30 articles were retained after applying the inclusion and exclusion criteria. The overall agreement between the reviewers was very good as indicated by a Cohen's kappa score of 0.72[26]. Table1 provides a summary of the final screened articles. The majority of the studies were cohort (n=21, 71%) followed by case-control (n=4, 13%), randomized clinical trials (n=2, 6.4%), cross-sectional (n=1, 3.2%) , one case report(n=1, 3.2%), and quasi-experimental (n=1, 3.2%) design. Among the cohort studies, 20 were retrospective cohorts and one was a prospective cohort. The number of articles published per year was not evenly distributed. Two general categories of risk factors emerged: Patient-related factors and procedure-related factors. The definition of PPCs was another independent theme that also emerged from the review.

Overall, the cohort studies showed moderately low risk of bias. However, approximately 60% of them did not report dropouts or strategies to deal with missing values. Around 30% failed to conduct proper statistical analysis (Figure 2). A similar pattern was observed for case-control studies; while generally exhibiting a low risk of bias, they did not identify confounding variables

nor adjust for these variables. Moreover, two studies did not match cases and controls properly according to the patient's characteristics (age, sex, and comorbidities) (Figure 3). Overall, the two RCTs were of low risk of bias but both studies fail to report on incomplete follow up and how they dealt with it (Figure 4). Lastly, the included cross-sectional study and the case report exhibited low risk of bias based on the JBI checklist (Figures 5 and 6). The quasi-experimental study did not have a control group and the participants were not similar (i.e., matched according to the patient's characteristics); thus, we failed to follow up on how they dealt with this heterogeneity (Figure 7).

## Definition of postoperative pulmonary complications (PPCs):

The definition of PPCs varied among the studies. While the majority of the studies (n= 20) included pneumonia[11-15,17,27,28,30,32,38,40,42-46,49-51], several studies also included respiratory failure(n=5) , atelectasis (n=9) [14,15,17,27,40,41,44,46,47,49,50], pulmonary embolism (n=10)[14,17,26,29,33,36-38,44,49], acute respiratory distress syndrome (n=6) [13,14,28,34,40,49], pulmonary edema (n=6) [14,17,27,28,42,46], pneumothorax (n=3)[17,28,40], reintubation (n=2)[38,43], and the need for assisted ventilation (n=5)[12,13,17,33,38].

### *Patient-related factors*

### *Obesity:*

Obesity - often measured by higher Body Mass Index (BMI)- has been associated with incidence of PPCs in HNS patients[17,28,29,31,37,39,50]. Patients with a BMI of equal and more than 30 kg/m$^2$ have an increased risk for PPCs after major HNS with microvascular reconstruction (OR= 3.24, 95% CI 1.80-5.82)[17]. Similar results were reported by others [28,37]. Interestingly, Fung[29] et al observed obese children (BMI above 27.7 kg/m$^2$) who undertook adenotonsillectomy for sleep-disordered breathing had odds ratio of 8.54 (95% CI 3.44-21.19) to develop PPC compared to non-obese children. In the Kanzaki[31] case report, the patient who developed PPC had a BMI of 31 kg/m$^2$ after parotid gland resection with radical neck dissection.

### *Smoking:*

Smoking status and smoking history have also been associated with PPCs. Patients who quit smoking had a lower risk of developing PPCs (OR= 0.58, 95%CI: 0.31-1.1) than current smokers[28]. However, Liu[33] found that current smokers were 4.37-fold more likely to have PPCs than patients who had never smoked or had quit smoking. Similar findings showing an association between PCCs and smoking status and smoking history have been reported in the literature[13,12,51].

***Male sex:***

Males are more likely to develop PPCs than female patients. Fung[29] found the male sex to be significantly based on the proportion of male patients having complications compared with female patients. Also, Li[32] and Xu[11] reported male sex as an independent risk factor for developing PPCs with an OR=1.94 (95% CI: 1.12-3.37) and OR= 16.73 (95% CI:1.67-167.81), respectively.

***Preoperative oxygen saturation:***

Low preoperative oxygen saturation is another risk factor for developing PPCs. Patients with lower baseline PaO2/FIO2 (partial pressure of oxygen in the arterial blood to a fraction of inspired oxygen) had an OR=0.99 (95% CI: 0.99-0.99) to develop PPCs[28]. Similar findings were reported by Smith[44]; patients having lower preoperative O2 saturation had an OR=0.93 (95% CI: 0.84, 1.03) to develop PPCs. While Rao[12] reported a significant relationship between PaO2 (partial pressure of oxygen in the arterial blood) and developing PPCs.

***Age:***

Several studies have reported an association between age and PPCs[12,14, 31,33, 35, 39, 42, 44, 49]. For example, patients who developed PPCs were significantly older, with a median age of 66 years and a relative risk of 1.66 (95% CI: 1.03–2.68), compared to median age of 59 years in patients who did not develop PPC[14]. Similarly, Joo[49] patients in the age groups 60-69 years (OR= 3.8 95% CI: 1.2-11.7) and 70 to 79 years (OR= 7.1, 95% CI:1.3-37.6) were more at risk of developing PPCs compared to patients younger than 60 years of age. Also, Spires[39] and Rao[12] found increasing age as a risk factor for developing PPCs. Liu [33] found that older patients were more likely to experience PPCs in their cohort, as the likelihood of the outcome increased by more than 1.15-fold per one-year increase in age.

### *History of Chronic Lung Disease*

Liu[33] found patients with history of Chronic Obstructive Pulmonary Disease have OR= 2.35 (95% CI: 1.30−4.77) to develop PPCs, while Buitelaar[42] reported OR= 5.46 (95% CI: 2.56–11.61) for patients with COPD. However, Semenov[45] reported an OR= 1.47 (95% CI: 1.22–1.77) for patients having Chronic Obstructive Pulmonary Disease to develop PPC. Also, Rao[12] reported dyspnea (shortness of breath) as a risk factor for developing PPCs while Joo[49] only mentioned chronic lung disease as risk factor. Ong[15] defined chronic lung disease as pulmonary decreased ratio of forced expiratory volume in 1 second to vital capacity and reported an OR= 0.01 (95% CI: 0.0-0.90) and OR=0.31 (95% CI: 0.36-0.53) to developing pulmonary infection and atelectasis, respectively

### *American Society of Anesthesiologists* (*ASA) grade:*

ASA classification aims to predict perioperative mortality, but it has also been proven to predict both postoperative pulmonary and cardiac complications[53]. The 5 ASA classes are: (1) an ordinarily healthy patient (class I), (2) a patient with mild systemic disease (class II), (3) a patient with systemic disease that is not incapacitating (class III), (4) a patient with an incapacitating systemic disease that is a constant threat to life (class IV), and (5) a moribund patient who is not expected to survive for 24 hours with or without operation (class V)**[54]** patients with ASA grade 3 or 4 have a higher risk of developing PPCs, compared to patients with ASA 1. [17] [42] By contrast, Menezes[38] found no significant association between ASA score and the occurrence of PPCs. However, having a higher ASA score was positively associated with 1-year mortality.

### *History of Heart Disease:*

History of myocardial infarction is another risk factor for PPCs. Patients with history of myocardial infarction and congestive heart failure have an OR= 3.82 (95% CI: 1.37–10.70) and OR=1.62 (95% CI: 0.64, 3.76) to develop PPCs, respectively[42,44]. Also, Spires[39] reported history of a myocardium-infection, arrhythmia, congestive failure, or mural thrombi as risk factors for PPCs development. Petrar[14] reported a preoperative diagnosis of hypertension as the only medical comorbidity significantly associated with PPCs.

### *Alcohol Abuse:*

Loeffelbein[17] found patients with history of Alcohol abuse to have OR= 1.71 (95% CI: 1.06-2.75), while Smith[44] reported patients with alcohol abuse OR= 1.72 (95% CI: 1.11- 2.69) to develop PPCs.

### *History of DVT and PE:*

Only Spires[39] identified deep venous thrombosis and pulmonary embolism as risk factors for PPC.

### *Albumin level:*

Xu[11] et al reported a lower preoperative serum albumin level (less than 35 mg/dL) associated with PPCs development.

### *Weight loss:*

Semenov[45] reported an OR= 2.85 (95% CI:2.34–3.48) for patients with a history of weight loss who develop PPCs.

### *Procedure-related factors:*

Several clinical and surgical procedures have been associated with PPCs. Below we describe most often reported factors. A full list of extracted data is available at https://github.com/mohdtamimi/Scoping_Review.

### *Anesthesia Technique and Dose:*

TIVA technique seems to have a protective effect compared to inhalation[27]. Also, Zhou[40] found that sevoflurane for maintenance anesthesia, in comparison with propofol, significantly reduced the incidence of PPCs in patients undergoing HNS. Finally, two studies reported increased anesthesia duration as a risk factor for PPC[12,34].

### *Operation time:*

Several studies have reported that prolonged operation time is a risk factor for developing PPCs after HNS[11,29,46,51,43].

### *Tracheostomy:*

Patients who had undergone tracheostomy and duration of tracheotomy, had OR= 0.72 (95% CI: 0.55−0.98) and OR=1.69 (95% CI: 1.36- 2.12).to develop PPCs, respectively[32,33], while Dillon[43] found OR= 3.00 (95% CI: 0.70–12.00).

### *Blood loss and transfusion:*

Only Logan[34] found massive blood transfusion as a risk factor for developing PPCs.

### *Immobilization after surgery:*

Moreano[35] observed that patients who used the Kendel pneumatic compression device after the surgery had an OR= 0.24 (95% CI: 0.07- 0.75) which means this device protects from pulmonary embolism. Hence, the risk of immobilization is evident as Spires[39] found in his study. Also, Yeung[46] found patients with delayed mobilization four or more days after surgery had OR= 4.2 (95% CI: 1.05–17.0) to develop PPCs. While Kanzaki[31] mentioned in his case report extended immobilization is a risk factor for PPCs.

### *Size of the flap with reconstruction :*

The incidence of PPCs was higher among patients with rectus abdominis flap reconstruction larger than 120 $cm^2$ [47]. Similarly, increased incidence of PPCs was observed among patients with latissimus dorsi myocutaneous flaps larger than 120 $cm^2$ [41]. Semenov[45] reported patients with pedicled or free flap reconstruction OR= 1.43 (95% CI: 1.15–1.78) to develop PPCs.

### *Neck dissection :*

Gallo[30] found that patients who had neck dissection procedures had an OR= 2.52 (95% CI:1.13-5.63) to develop PPCs.

### *Tongue base resection:*

Smith[48] reported that resection of up to half of the base of the tongue resection of more than half of the base of the tongue with resulted 22% and 75% incidence of PPCs, respectively.

### *Cricohyoidopexy:*

Patients who underwent a Cricohyoidopexy, a reconstruction method to repair and restore the function of the larynx after surgical ablation of cancer in which the hyoid bone is fixed to the cricoid cartilage, have a higher risk of PPCs (OR= 4.4, 95%CI: 1.1-18.1)[49].

### *Discussion:*

PPCs risk factors for HNS patients identified in this work fall in line with what was found in the review by Smetana[7], which was adopted by the American College of Physicians guidelines for general surgery postoperative pulmonary risk factors. Our review found several patient related risk factors: advanced age, obesity, smoking status and history, male gender, preoperative oxygen saturation levels, history of lung disease, ASA classification, history of heart disease, alcohol use, albumin levels, and weight loss. The clinical and surgical related risk factors were anesthesia, operation time, blood loss, tracheostomy, size of flap and reconstruction, and surgical site.

Smetana et al's review reported advanced age an independent predictor for PPCs after adjusting for comorbid conditions[7]. This is an important finding as previous studies hypothesized that the increased risk with age due to age was due to accumulating comorbid conditions[7]. Smetana et al reported varying obesity definitions from body mass index (BMI) of 25 kg/m$^2$ or greater to "morbid obesity"[7]. They found that PPC rates for patients with BMI of 43 kg/m$^2$ or less were 10%. In comparison, rates for patients above 43 kg/m$^2$ were 12% with no significant difference[7], which might be due to the patients in this category (i.e., morbid obesity) being already susceptible to high risk of PPC for other morbidities. This also aligns with other studies reporting that patients with higher morbidity have higher risk for developing PPCs[45].

Finding from the Smetana et al' review reveal a similar magnitude of the odds ratios for smoking history,[7] alcohol use, and patients with chronic obstructive pulmonary disease and low serum albumin level (below the range 30-39 g/L)[7]. This might be because these patients already have pulmonary disease which can be propagated after surgery and general anesthesia. The review by

Canet and Gallart mentioned low preoperative oxygen saturation levels as measured by pulse oximetry to be a good candidate for predicting PPCs in general population as it is regularly monitored and can predict hypoxaemic events[52]. They also identified weight loss as a PPCs risk factor[52].

Smetana review reported patients with ASA grade 3 or 4 were more likely to develop PPC. As mentioned before, patients with higher morbidity scores might be more susceptible to developing PPCs, as they might have more than one risk factor for PPC development (like COPD, heart disease, or advanced age).

While patients who had general anesthesia had more chance to develop PPC as reported in Smetana et al review[7]. This might be due to longer assisted ventilation periods that accompany the general anesthesia.

Smetana review had a definition of prolonged surgery as 2.5-4 hours and patients having prolonged surgery were at more risk of developing PPCs[7]. Longer operation time means longer anesthesia duration and possibly more blood loss, which increases the risks of PPCs from other factors.

Moreover, blood transfusion preoperatively was considered a risk factor for developing PPCs according to Smetana et al[7], while Canet et al[52] mentioned 4 units or more as the threshold for excessive blood loss that might lead to develop PPC. This might be due to the use of prophylactic heparin, so patients will not develop pulmonary embolism, which might cause the excessive bleeding[34].

On another level, head and neck surgeries had higher risk of developing PPC as reported in Smetana et al review[7]. As this might be the proximity of HNS sites to the airway and some HNS procedures include the larynx and pharynx.

There is a considerable discrepancy between the included studies that assessed PPCs risk factors, the suggested reasons also noted by Canet[52] might be: (1) different definitions of the PPCs outcome; as some might use single complication [such as pneumonia, atelectasis, ARDS, pulmonary embolism, or respiratory failure] or composite complications, (2) risk factors have not

been defined in the same way, (3) study designs have different inclusion and exclusion criteria and statistical analysis approaches which might affect the results that were reported.

One of the limitations of the of the included studies was the statistical analysis as most studies did not adjust for confounding factors. Another limitation was that most study designs were retrospective cohorts, which means causality inferences cannot be reported with these studies. Also, only few studies adjusted for confounders in the analysis, which increased the risk of bias. On another note, there are various validated risk of bias assessment tools. JBI risk of bias assessment tool (used in our study) comprises various checklists for all study designs, which gives it an advantage for such a broad topic and research question. Also, JBI is comparable and validated, but it is really matter of preference and convenience[55]. Despite the fact that we included peer-reviewed original studies, the risk of bias was moderately high among these studies in general, which raises the alarm on the quality of peer-review process in this field of research.

Scoping reviews are conducted to identify, map, collate and summarize the existing literature that covers a field of question[19], and aim to assist researchers to recognize fundamental ideas, theories, evidence sources and gaps in knowledge in that field[24]. In contrast to systemic reviews, scoping reviews take a "Big Picture" approach to examining literature rather than centring on a specific research question[25]. Thus, we took the approach of scoping review to answer our 'broad' research question.

To the best of our knowledge, this is the first review to map the existing literature for risk factors of PPCs in head and neck surgery patients.

## Conclusion:

Risk factors for PPCs in HNS patients can be mapped as follows: a) patient-related risk factors: obesity, advanced age, history of chronic lung disease, history of heart disease, smoking, male gender, alcohol use, preoperative oxygen saturation, ASA classification, serum albumin level, weight loss, history of DVT or PE b)procedure-related risk factors: Operation time, blood loss and transfusion, anesthesia technique and dose, tongue base resection, tracheostomy,

immobilization after surgery, size of the flap with reconstruction, neck dissection, and cricohyoidopexy. This review helps identifying the current status of the literature for PPC risk factors in HNS and can inform future researchers, who want to conduct a systematic review and meta-analysis, about the feasibility of this.

Authors contributions:

**Mohammad Al-Tamimi,** Master's Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Conceived objective of the investigation, screened and selected studies, extracted data, assessed the quality of studies, statistical analyses, performed statistical analysis, and wrote the manuscript.

**Sreenath Madathil,** Assistant professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, contributed to the conceptual development, selection process, quality assessment, interpretation, and writing the manuscript.

**Ahmed Derbas,** Oral and maxillofacial surgery resident, Oral and Maxillofacial Surgery Department, Hamad Medical Corporation, Doha, Qatar: Participated in the screening and selection process and extracted data.

**Hamed Ghanati,** Master's Candidate, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: participated in the conceptual development, quality assessment and contributed to writing the manuscript.

**Martin Morris,** Schulich Library of Physical Sciences, Life Sciences and Engineering, McGill University, Montréal, QC, Canada: Conducted the search process, contributed to writing the manuscript.

**Belinda Nicolau.**, Professor, Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montréal, QC, Canada: Supervised all steps, contributed to the conceptual development, selection process, quality assessment, interpretation, and writing the manuscript (Corresponding author).

## Tables and figures:

**Table 1. Search strategy for MEDLINE(Ovid) database.**

| ***Steps*** |
| --- |
| 1. "Head and Neck Neoplasms"/ or exp Facial Neoplasms/ or exp Mouth Neoplasms/ or exp Otorhinolaryngologic Neoplasms/ or exp Tracheal Neoplasms/ |
| 2. ((cancer* or tumo?r* or neoplas* or metaplas* or carcinoma* or metastasi* or squamous cell carinoma? or SCC or HNSCC or malignan*) adj5 (head or neck or uadt or "upper aero-digestive" or "upper aerodigestive" or face or facial or oral* or intra-oral* or intraoral* or mouth or buccal or gingiv* or gum* or lip? or labial* or palat* or lingual* or mandib* or maxill* or jaw? or tongue* or glossal* or otor?inolaryngolog* or throat or ear? or auricle* or auricular or larynx* or laryngeal* or nose* or nasal* or paranasal* or sinus or hypopharynx or hypopharyngeal* or nasopharynx or nasopharyngeal* or oropharynx or oropharyngeal or tonsil* or trachea* or cheek* or pharynx or pharyngeal or retromolar or alveolar or tonsil* or sinonasal or sinus* or vestib* or piriform or post-cricoid or glottic or subglottic or superglottic or transglottic or "unknown primary" or trigone or maxillofacial*)).tw,kf. |
| 3. (exp Head/ or exp Neck/) and exp Neoplasms/ |
| 4. (head and neck surgery).mp. |
| 5. 1 or 2 or 3 or 4 |
| 6. laryngeal diseases/ or laryngeal edema/ or exp laryngeal nerve injuries/ or exp laryngitis/ or laryngocele/ or laryngomalacia/ or laryngopharyngeal reflux/ or supraglottitis/ or tuberculosis, laryngeal/ or exp vocal cord dysfunction/ or vocal cord paralysis/ or lung diseases/ or exp lung diseases, fungal/ or exp lung diseases, interstitial/ or exp lung diseases, obstructive/ or exp lung diseases, parasitic/ or exp pulmonary atelectasis/ or pulmonary edema/ or exp pulmonary embolism/ or exp pulmonary eosinophilia/ or exp respiratory distress syndrome/ or exp tuberculosis, pulmonary/ or exp respiration disorders/ or exp respiratory hypersensitivity/ or exp respiratory tract fistula/ or exp respiratory tract infections/ |
| 7. ((respiratory or pleural or pulmonary or vocal c?ord or laryinx or laryngeal) adj3 (disease? or effusion or infection? or syndrome or atalectas?s or oedema? or edema? or embolism? or failure? or insufficien* or distress or disorder? or hypersensitiv* or injury or injuries or dysfunction or eosinophil* or empyema?)).tw,kf. |

8. (pneumonia or pneumothora* or hydrothora* or hydropneumothora* or hemopneumothora* or hemothora* or bronchospasm* or bronchitis or pleurisy).tw,kf.

9. or/6-8

10. exp Postoperative Complications/

11. ((postoperative or post-operative or surgical) adj3 (complication? or infection? or sequela?)).tw,kf.

12. 10 or 11

13. 5 and 9 and 12

14. exp oral surgical procedures/ or exp otorhinolaryngologic surgical procedures/ or exp thoracic surgical procedures/

15. (exp Head/ or exp Neck/) and (su.fs. or exp Radiotherapy/ or radiotherap*.tw,kf. or radiation.tw,kf.)

16. ((head or neck or uadt or "upper aero-digestive" or "upper aerodigestive" or face or facial or oral* or intra-oral* or intraoral* or mouth or buccal or gingiv* or gum* or lip? or labial* or palat* or lingual* or mandib* or maxill* or jaw? or tongue* or glossal* or otor?inolaryngolog* or throat or ear? or auricle* or auricular or larynx* or laryngeal* or nose* or nasal* or paranasal* or sinus or hypopharynx or hypopharyngeal* or nasopharynx or nasopharyngeal* or oropharynx or oropharyngeal or tonsil* or trachea* or cheek* or pharynx or pharyngeal or retromolar or alveolar or tonsil* or sinonasal or sinus* or vestib* or piriform or post-cricoid or glottic or subglottic or superglottic or transglottic or "unknown primary" or trigone or maxillofacial*) adj3 (surgery or surgeries or surgical* or operation? or procedure? or radiotherap* or radiation*)).mp.

17. exp Surgical Flaps/ or ((free or regional or local) adj1 (flap? or tissue?)).tw,kf. or neck dissect*.tw,kf.

18. or/14-17

19. 5 and (9 or 18) and 12

20. exp risk/

21. risk?.tw,kf.

22. (tobacco or smoke or smoking or cigarette?).mp.

23. alcohols/ or exp ethanol/ or exp Alcoholic Beverages/ or exp Alcoholism/ or alcohol*.tw,kf.

24. exp Age Factors/ or age.tw,kf.

25. (oxygen adj3 (home or therap*)).mp.

26. ((link? or linked or associat*) adj1 (with or to)).tw,kf.

27. or/20-26

28. 19 and 27

29. 20 or 21 or 22 or 23 or 24 or 25

30. 19 and 29

31. 19 and 24

32. 20 or 21

33. 19 and 32

**Table 2**. Overview of included studies in the scoping review discussing PPCs risk factors in HNS

| Citation (Author, Year) | Design | Recruitment period | Sample size (N) | Surgical intervention | PPC definition | PPC Risk factors |
|---|---|---|---|---|---|---|
| Chang 2016[27] | Retrospective Cohort | 2012-2013 | 156 | Head and neck cancer surgery with free flap surgery | Pulmonary edema, pneumonia or atelectasis | TIVA anesthesia technique as protective factor |
| Damian 2016[28] | Retrospective cohort | 2005-2011 | 110 | Head and neck cancer surgery with free flap reconstruction | Pulmonary edema, pneumonia, pneumothorax, pulmonary embolism, and ARDS. | BMI, quit smoking status, and baseline PaO2/FIO2 |
| Fung 2010[29] | Case-control study | 2002-2007 | 49 | Adenoidectomy performed using the suction monopolar diathermy technique | Oxygen desaturation, coughing episode, bronchospasm, Airway obstruction, Respiratory depression, Admission to ICU (an unplanned admission, and a prolonged duration of stay more than 24 hours) | Male, tonsillectomy, and BMI. |
| Gallo, 2009[30] | Retrospective cohort | 1980-2000 | 416 | Partial laryngectomy | Diagnosis of pneumonia | BMI > 30,age > 60 and neck dissection procedure were significant. |
| Kanzaki, 2004[31] | Case report | 2003 | 2 | Head and neck surgery | Pulmonary embolism | Obesity, immobilization, age, heart disease, and length of surgery |
| Li, 2016[32] | Retrospective cohort | 2012-2013 | 482 | Oral oncology resection, free flap transplantation, | Diagnosis of pneumonia | Male gender, long duration of tracheotomy, and smoking |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | and tracheostomy after surgery | | |
| Liu, 2017[33] | Case-control study | 2011-2014 | 465 | Major oncological SHNC | Postoperative mechanical ventilation for 48 h or longer with the ventilator in place at the time of or 24 h before the event. | Age, smoking, immunosuppression, COPD, mean SAPS II on admission, serum albumin level (g/dl), and tracheostomy |
| Loeffelbein, 2016[17] | Retrospective cohort | 2007- 2013 | 648 | Major surgery in the head and neck area with microvascular reconstruction | Pneumonia, atelectasis, pleural effusions, pulmonary embolism, pulmonary oedema, pneumothorax and respiratory failure. Prolonged mechanical ventilation, defined as more than 15 h postoperatively, or the need for re-intubation. | Obesity, ASA grade 3 and alcohol |
| Logan, 1998[34] | Retrospective cohort | 1985-1995 | 418 | Tumor ablation, neck dissection, and reconstruction using microvascular free tissue transfer | Acute Respiratory Distress Syndrome | Massive blood transfusion, pneumonia, general anesthesia duration, deep circumflex iliac artery flaps. |
| Moreano,1998[35] | Retrospective cohort | 1987-1994 | 12,805 | Otolaryngology-head and neck surgery | Deep Vein Thrombosis, Pulmonary Embolism | Age and immobilization |
| Rahman, 2009[36] | Retrospective cohort | 1989-2003 | 262 | Thyroidectomy | Respiratory distress characterized by dyspnea, stridor, tension hematoma, laryngeal edema, and/or tracheomalacia. | Thyroid swelling and giant goitre |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | Nonobstructive respiratory complications included cough, excessive sputum, and chest signs of consolidation with or without fever. | |
| Rao, 1992[12] | Prospective cohort | N/A | 73 | Head and neck surgery | Ventilator dependency more than 12 h, pneumonia. | Age, PaO2, Roizen classification of dyspnea, smoking history, anesthesia duration, forced expiratory volume in 1 second, and peak flow. |
| Shaw, 2021[37] | Retrospective cohort | 2013-2014 | 60 | Tracheostomy for airway protection and free flap reconstruction for head and neck cancer surgery | A positive diagnosis of PPC was confirmed by the presence of four or more variables of Melbourne Group Scale (MGS) | High BMI, oxygen therapy and additional physiotherapy sessions |
| Menezes, 2021[38] | Retrospective cohort | 2016-2017 | 128 | Head and neck surgery | Pneumonia, pulmonary embolism, deep venous thrombosis, unplanned intubation, ventilator support > 48 h. | Higher ASA score (ASA 3 & 4) and higher preoperative ARISCAT score |
| Spires, 1989[39] | Retrospective cohort | 1950-1987 | 30 | Head and neck cancer surgery | Pulmonary embolism | History of venous thromboembolism, obesity, immobility, increasing age, heart disease |
| Xu, 2017[11] | Retrospective cohort | 2014-2016 | 331 | Oral cancer surgery with or without reconstruction | Diagnosis of pneumonia | Male, preoperative serum albumin level, operation time and postoperative hospital stay |

| | | | | | | |
|---|---|---|---|---|---|---|
| Zhou, 2020[40] | RCT | 2018-2019 | 220 | Head and neck cancer with free flap surgery | Pulmonary infection, pleural effusion, Atelectasis, pneumothorax, bronchospasm, pulmonary edema, pulmonary embolism, respiratory failure, acute respiratory distress syndrome | Sevoflurane anesthetic agent reduced PPC incidence |
| Wax, 1996[41] | Case-control study | 1991-1994 | 36 | Head and neck oncologic and reconstructive procedures | atelectasis | Size of the flap |
| Buitelaar, 2006[42] | Retrospective cohort | 1993-1998 | 469 | Primary surgery for head and neck tumors | Pneumonia, bronchitis, pleural effusion, respiratory depression | COPD , history of myocardial infarction, ASA grade, and age |
| Dillon, 2011[43] | Retrospective cohort | 2005-2008 | 92 | Major oral cancer surgery | Respiratory complications including respiratory distress requiring and not requiring reintubation, and pneumonia. | Tracheostomy and operation time |
| Smith, 2021[44] | Retrospective cohort | 2005-2017 | 794 | Head and neck surgery | Postoperative pneumonia and postoperative respiratory failure. | Age, alcohol, history of congestive heart failure, preoperative packed cell volume, preoperative oxygen saturation, and preoperative metabolic equivalents |
| Petrar, 2012[14] | Retrospective cohort | 2005- 2008 | 110 | Major head and neck cancer surgery | Pneumonia, pulmonary edema, bronchospasm, Acute Respiratory Distress Syndrome, | Age, preoperative diagnosis of hypertension. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | pulmonary embolism, atelectasis, and respiratory failure | |
| Semenov, 2012[45] | Retrospective cross-sectional | 2003-2008 | 93663 | Ablative procedures for a malignant oral cavity, laryngeal, hypopharyngeal, or oropharyngeal neoplasm | Codes for pneumonia were obtained from ICD-9 codes for infectious pneumonia, aspiration pneumonia, and VAP. | Dysphagia, chronic pulmonary disease (COPD) and weight loss. |
| Yeung, 2013[46] | Retrospective cohort | 2005-2009 | 62 | Oral cancer resection, followed by immediate free flap reconstruction | Pneumonia, Pulmonary edema, Atelectasis, Pulmonary embolism. | Operative time and delayed mobilization (immobilization) |
| Wax, 2002[47] | Case-control study | 1999-2000 | 106 | Abdominis rectus flap reconstruction | Atelectasis | Flap sizes (120 cm2 and more) |
| Ong, 2004[15] | RCT | N/A | 73 | Major head and neck surgery with tracheostomy | Pulmonary infection or atelectasis | Obstructive lung disease |
| Smith, 2008[48] | Retrospective cohort | 1996 – 2006 | 100 | Resection and free flap reconstruction of the oral cavity or oropharyngeal tumors | Tracheal aspiration | History of radiation therapy and tongue base resection procedure |
| Joo, 2009[49] | Retrospective cohort | 1993-2008 | 111 | Supracoracoide Partial Laryngectomy (SCPL) | Atelectasis, pneumonia, respiratory failure, and exacerbation of underlying chronic lung disease | Age, cricohyoidopexy, chronic lung disease, and smoking status, FEV1/FVC, |

| | | | | | | |
|---|---|---|---|---|---|---|
| McCulloch, 1997[13] | Retrospective cohort | 1985-1991 | 144 | Excision and repair of oral, pharyngeal, or laryngeal lesions; requiring more than 2 hours of operative time; and usually included neck dissection and/or flap reconstruction. | Postoperative pneumonia, adult respiratory distress syndrome (ARDS), and prolonged ventilation | Smoking history and antibiotic choice, and weight loss |
| Manzoor, 2007[50] | Quasi-experimental study | 2005-2006 | 70 | Surgery for cancer of larynx, nose, hypopharynx, nose and paranasal sinuses, ear mastoid, oral cavity, skin of the head and neck, salivary gland | Bronchopneumonia, Pulmonary embolism. | Smoking, assisted ventilation. and prolonged surgery. |

**Figure 1:** PRISMA flowchart displaying the selection process.

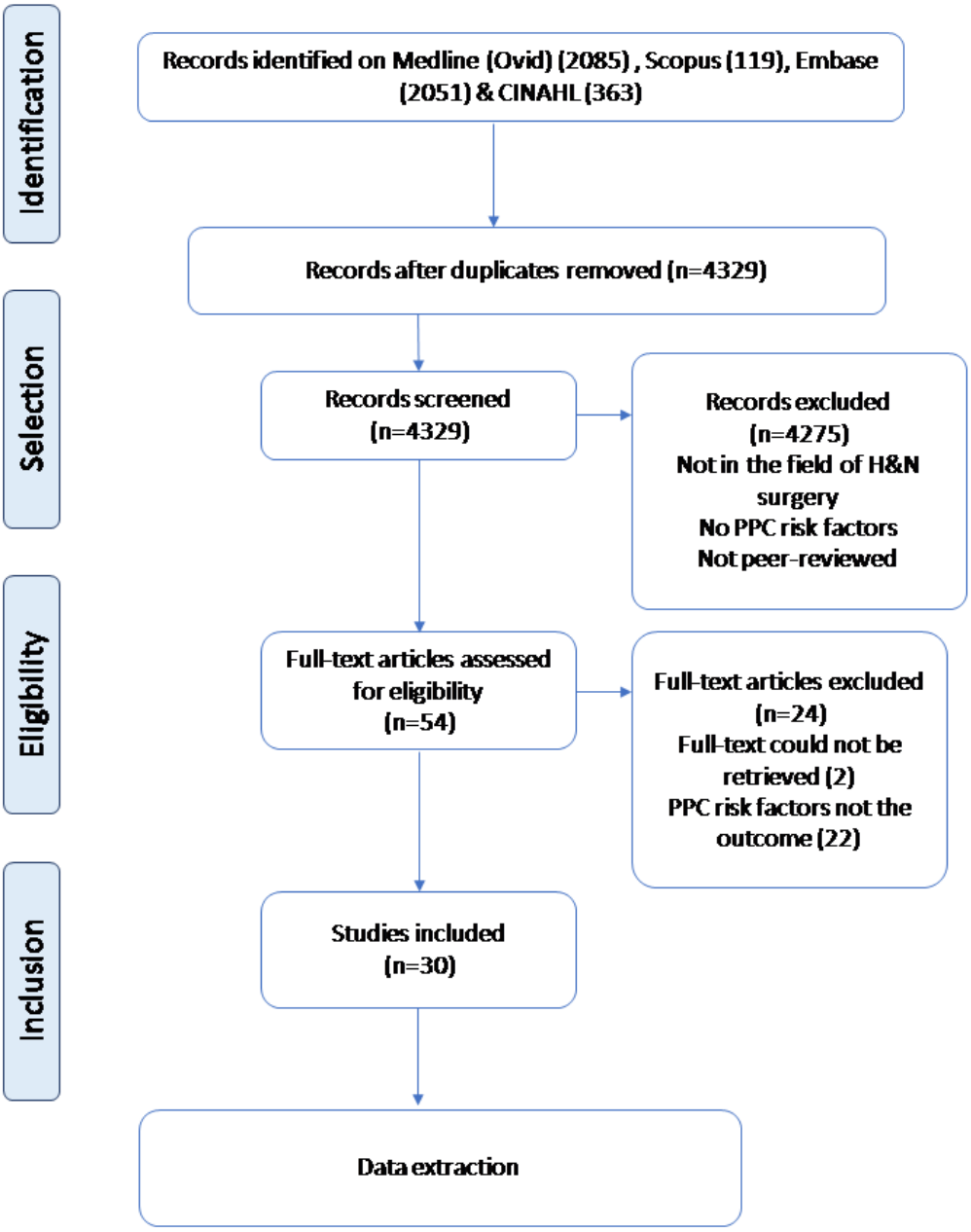
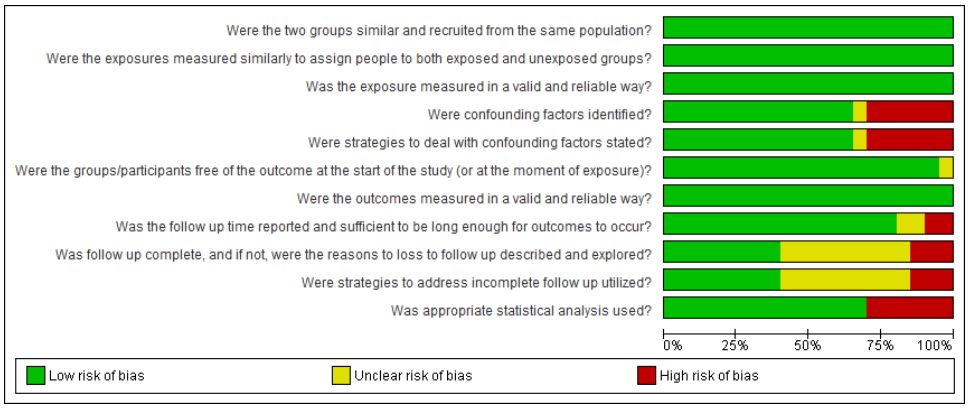


**Figure 2:** Risk of bias graph and summary: review authors' judgements about each risk of bias item presented as percentages across all cohort studies.



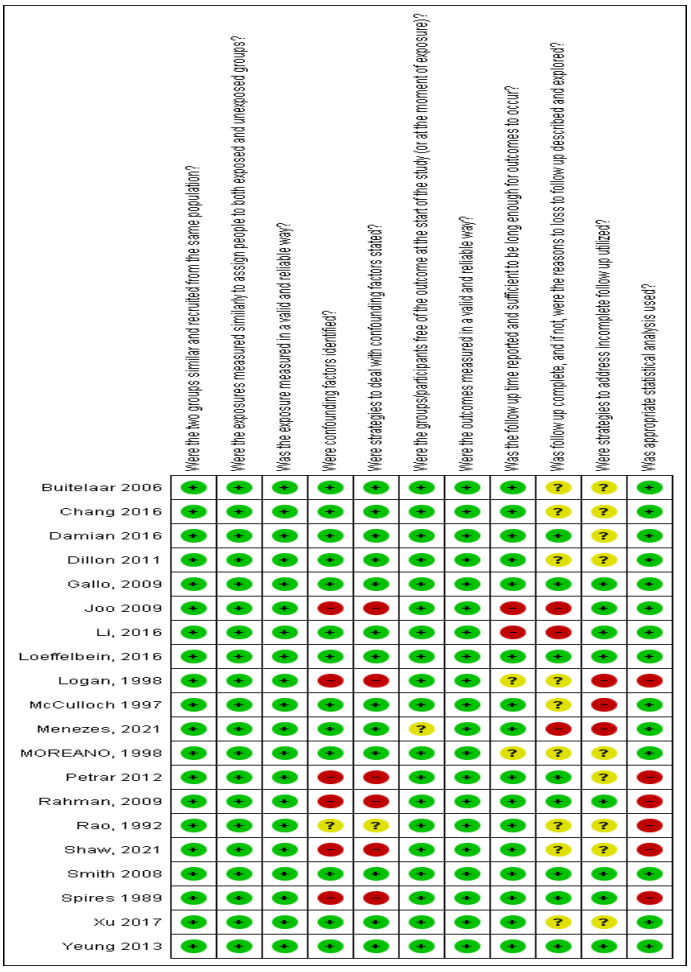



**Figure 3: Summary of** Risk of bias assessment: review authors' judgements about each risk of bias item presented as percentages across all case-control studies.

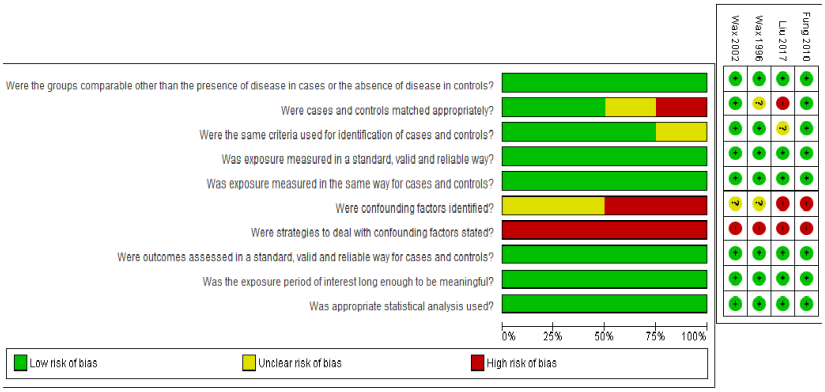



**Figure 4:** Risk of bias assessment: review authors' judgements about each risk of bias item presented as percentages across all RCT studies.

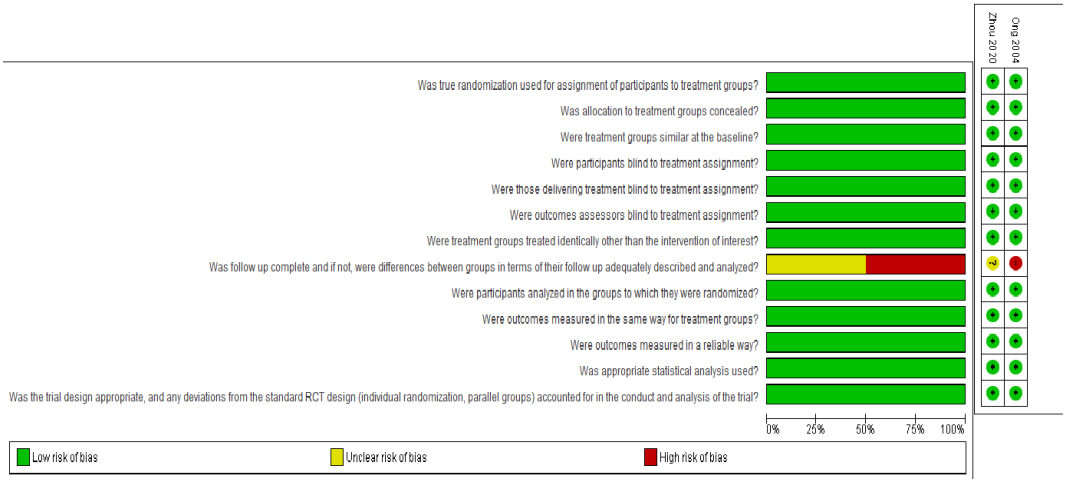


**Figure 5:** Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across the cross-sectional included study.

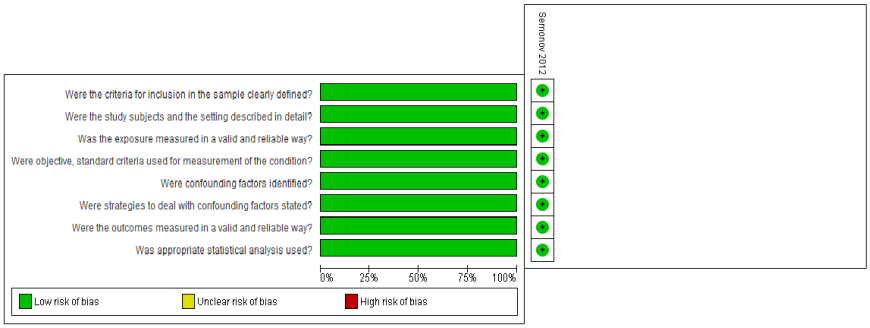



**Figure 6:** Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across the case report.

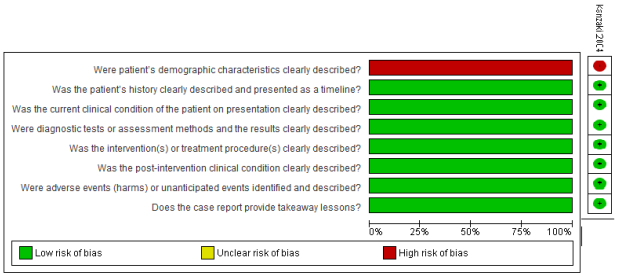
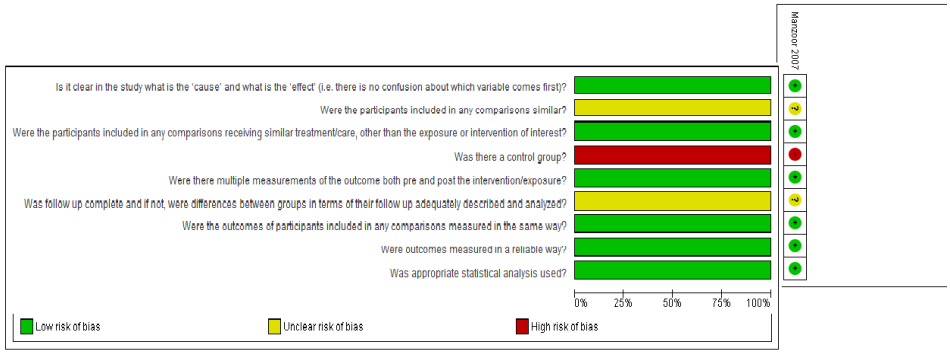



**Figure 7:** Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across the quasi-experimental included study.



**References:**


1. Meara JG, Leather AJ, Hagander L, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; **386**(9993): 569-624.

2. Longo WE, Virgo KS, Johnson FE, et al. Risk factors for morbidity and mortality after colectomy for colon cancer. *Dis Colon Rectum* 2000; **43**(1): 83-91.

3. Mayo NE, Feldman L, Scott S, et al. Impact of preoperative change in physical function on postoperative recovery: argument supporting prehabilitation for colorectal surgery. *Surgery* 2011; **150**(3): 505-14.

4. Tevis SE, Kennedy GD. Postoperative complications and implications on patient-centered outcomes. *J Surg Res* 2013; **181**(1): 106-13.

5. Shander A, Fleisher LA, Barie PS, Bigatello LM, Sladen RN, Watson CB. Clinical and economic burden of postoperative pulmonary complications: patient safety summit on definition, risk-reducing interventions, and preventive strategies. *Crit Care Med* 2011; **39**(9): 2163-72.

6. Peroni DG, Boner AL. Atelectasis: mechanisms, diagnosis and management. *Paediatr Respir Rev* 2000; **1**(3): 274-8.

7. Smetana GW, Lawrence VA, Cornell JE, American College of P. Preoperative pulmonary risk stratification for noncardiothoracic surgery: systematic review for the American College of Physicians. *Ann Intern Med* 2006; **144**(8): 581-95.

8. Fernandez-Bustamante A, Frendl G, Sprung J, et al. Postoperative Pulmonary Complications, Early Mortality, and Hospital Stay Following Noncardiothoracic Surgery: A Multicenter Study by the Perioperative Research Network Investigators. *JAMA Surg* 2017; **152**(2): 157-66.

9. Gupta H, Gupta PK, Fang X, et al. Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest* 2011; **140**(5): 1207-15.

10. Khan NA, Quan H, Bugar JM, Lemaire JB, Brant R, Ghali WA. Association of postoperative complications with hospital costs and length of stay in a tertiary care center. *J Gen Intern Med* 2006; **21**(2): 177-80.

11. Xu J, Hu J, Yu P, et al. Perioperative risk factors for postoperative pneumonia after major oral cancer surgery: A retrospective analysis of 331 cases. *PLoS ONE* 2017; **12**(11): e0188167.

12. Rao MK, Reilley TE, Schuller DE, Young DC. Analysis of risk factors for postoperative pulmonary complications in head and neck surgery. *Laryngoscope* 1992; **102**(1): 45-7.

13. McCulloch TM, Jensen NF, Girod DA, Tsue TT, Weymuller EA, Jr. Risk factors for pulmonary complications in the postoperative head and neck surgery patient. *Head Neck* 1997; **19**(5): 372-7.

14. Petrar S, Bartlett C, Hart RD, MacDougall P. Pulmonary complications after major head and neck surgery: A retrospective cohort study. *Laryngoscope* 2012; **122**(5): 1057-61.

15. Ong SK, Morton RP, Kolbe J, Whitlock RM, McIvor NP. Pulmonary complications following major head and neck surgery with tracheostomy: a prospective, randomized, controlled trial of prophylactic antibiotics. *Arch Otolaryngol Head Neck Surg* 2004; **130**(9): 1084-7.

16. de Melo GM, Ribeiro KC, Kowalski LP, Deheinzelin D. Risk factors for postoperative complications in oral cancer and their prognostic implications. *Arch Otolaryngol Head Neck Surg* 2001; **127**(7): 828-33.

17. Loeffelbein DJ, Julinek A, Wolff KD, Kochs E, Haller B, Haseneder R. Perioperative risk factors for postoperative pulmonary complications after major oral and maxillofacial surgery with microvascular reconstruction: A retrospective analysis of 648 cases. *J Craniomaxillofac Surg* 2016; **44**(8): 952-7.

18. Jones NF, Jarrahy R, Song JI, Kaufman MR, Markowitz B. Postoperative medical complications--not microsurgical complications--negatively influence the morbidity, mortality, and true costs after microsurgical reconstruction for head and neck cancer. *Plast Reconstr Surg* 2007; **119**(7): 2053-60.

19. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International journal of social research methodology* 2005; **8**(1): 19-32.

20. Morris M, Boruff JT, Gore GC. Scoping reviews: establishing the role of the librarian. *Journal of the Medical Library Association: JMLA* 2016; **104**(4): 346.

21. Team TE. EndNote. EndNote 20 ed. Philadelphia, PA: Clarivate; 2013.

22. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016; **5**(1): 210.

23. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine* 2018; **169**(7): 467-73.

24. Peters M, Godfrey C, McInerney P, Soares C, Khalil H, Parker D. The Joanna Briggs Institute reviewers' manual 2015: methodology for JBI scoping reviews. 2015.

25. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. *JBI Evidence Implementation* 2015; **13**(3): 179-87.

26. COHEN J. A COEFFICIENT OF AGREEMENT FOR NOMINAL SCALES. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT* 1960; **20**.

27. Chang YT, Wu CC, Tang TY, Lu CT, Lai CS, Shen CH. Differences between Total Intravenous Anesthesia and Inhalation Anesthesia in Free Flap Surgery of Head and Neck Cancer. *PLoS ONE* 2016; **11**(2): e0147713.

28. Damian D, Esquenazi J, Duvvuri U, Johnson JT, Sakai T. Incidence, outcome, and risk factors for postoperative pulmonary complications in head and neck cancer surgery patients with free flap reconstructions. *J Clin Anesth* 2016; **28**: 12-8.

29. Fung E, Cave D, Witmans M, Gan K, El-Hakim H. Postoperative respiratory complications and recovery in obese children following adenotonsillectomy for sleep-disordered breathing: A case-control study. *Otolaryngology - Head and Neck Surgery* 2010; **142**(6): 898-905.

30. Gallo O, Deganello A, Gitti G, et al. Prognostic role of pneumonia in supracricoid and supraglottic laryngectomies. *Oral Oncol* 2009; **45**(1): 30-8.

31. Kanzaki S, Kunihiro T, Imanishi T, Yamashita D, Ogawa K. Two cases of pulmonary embolism after head and neck surgery. *Auris Nasus Larynx* 2004; **31**(3): 313-7.

32. Li L, Yuan W, Zhang S, Wang K, Ruan H. Analysis of Risk Factors for Pneumonia in 482 Patients Undergoing Oral Cancer Surgery With Tracheotomy. *J Oral Maxillofac Surg* 2016; **74**(2): 415-9.

33. Liu Y, Di Y, Fu S. Risk factors for ventilator-associated pneumonia among patients undergoing major oncological surgery for head and neck cancer. *Front* 2017; **11**(2): 239-46.

34. Logan SW, Rogers SN, Richardson D, Vaughan ED, Brown JS. Adult respiratory distress syndrome after microvascular free tissue reconstruction in head and neck malignancy. *Br J Oral Maxillofac Surg* 1998; **36**(5): 371-4.

35. Moreano EH, Hutchison JL, McCulloch TM, Graham SM, Funk GF, Hoffman HT. Incidence of deep venous thrombosis and pulmonary embolism in otolaryngology-head and neck surgery. *Otolaryngol Head Neck Surg* 1998; **118**(6): 777-84.

36. Rahman GA. Possible risk factors for respiratory complications after thyroidectomy: an observational study. *ENT: Ear, Nose & Throat Journal* 2009; **88**(4): 890-2.

37. Shaw LM, Iseli TA, Wiesenfeld D, Ramakrishnan A, Granger CL. Postoperative pulmonary complications following major head and neck cancer surgery. *Int J Oral Maxillofac Surg* 2021; **50**(3): 302-8.

38. Sousa Menezes A, Fernandes A, Rocha Rodrigues J, et al. Optimizing classical risk scores to predict complications in head and neck surgery: a new approach. *Eur Arch Otorhinolaryngol* 2021; **278**(1): 191-202.

39. Spires JR, Byers RM, Sanchez ED. Pulmonary thromboembolism after head and neck surgery. *South Med J* 1989; **82**(9): 1111-5.

40. Zhou D, Zhu X, Wang L, Yang X, Liu Y, Zhang X. Which Anesthesia Regimen Is Best to Reduce Pulmonary Complications After Head and Neck Surgery? *Laryngoscope* 2020.

41. Wax MK, Hurst J. Pulmonary atelectasis after reconstruction with a latissimus dorsi myocutaneous flap. *Laryngoscope* 1996; **106**(3 Pt 1): 268-72.

42. Buitelaar DR, Balm AJ, Antonini N, van Tinteren H, Huitink JM. Cardiovascular and respiratory complications after major head and neck surgery. *Head Neck* 2006; **28**(7): 595-602.

43. Dillon JK, Liu SY, Patel CM, Schmidt BL. Identifying risk factors for postoperative cardiovascular and respiratory complications after major oral cancer surgery. *Head Neck* 2011; **33**(1): 112-6.

44. Smith DK, Freundlich RE, Shinn JR, Wood CB, Rohde SL, McEvoy MD. An improved predictive model for postoperative pulmonary complications after free flap reconstructions in the head and neck. *Head Neck* 2021; **30**: 30.

45. Semenov YR, Starmer HM, Gourin CG. The effect of pneumonia on short-term outcomes and cost of care after head and neck cancer surgery. *Laryngoscope* 2012; **122**(9): 1994-2004.

46. Yeung JK, Harrop R, McCreary O, et al. Delayed mobilization after microsurgical reconstruction: an independent risk factor for pneumonia. *Laryngoscope* 2013; **123**(12): 2996-3000.

47. Wax MK, Rosenthal EL, Takaguchi R, Cohen JI, Andersen PE, Futran N. Pulmonary atelectasis after reconstruction with a rectus abdominis free tissue transfer. *Arch Otolaryngol Head Neck Surg* 2002; **128**(3): 249-52.

48. Smith JE, Suh JD, Erman A, Nabili V, Chhetri DK, Blackwell KE. Risk factors predicting aspiration after free flap reconstruction of oral cavity and oropharyngeal defects. *Arch Otolaryngol Head Neck Surg* 2008; **134**(11): 1205-8.

49. Joo YH, Sun DI, Cho JH, Cho KJ, Kim MS. Factors that predict postoperative pulmonary complications after supracricoid partial laryngectomy. *Arch Otolaryngol Head Neck Surg* 2009; **135**(11): 1154-7.

50. Wood CB, Shinn JR, Rees AB, et al. Existing Predictive Models for Postoperative Pulmonary Complications Perform Poorly in a Head and Neck Surgery Population. *J Med Syst* 2019; **43**(10): 312.

51. Manzoor T, Ahmed Z, Sheikh NA, Khan MM. Pulmonary complications associated with head and neck cancer surgery. *J Coll Physicians Surg Pak* 2007; **17**(9): 558-61.

52. Canet J, Gallart L. Predicting postoperative pulmonary complications in the general population. *Curr Opin Anaesthesiol* 2013; **26**(2): 107-15.

53. Gerson MC, Hurst JM, Hertzberg VS, Baughman R, Rouan GW, Ellis K. Prediction of cardiac and pulmonary complications related to elective abdominal and noncardiac thoracic surgery in geriatric patients. *Am J Med* 1990; **88**(2): 101-7.

54. Owens WD, Felts JA, Spitznagel EL, Jr. ASA physical status classifications: a study of consistency of ratings. *Anesthesiology* 1978; **49**(4): 239-43.

55. Ma, LL., Wang, YY., Yang, ZH. *et al.* Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better?.*Military Med Res* **7,** 7 (2020). https://doi.org/10.1186/s40779-020-00238-8

## Preface for Manuscript II:

Based on what is discussed in chapters 2 and 3 of this thesis, the first step in providing a population of interest with a risk prediction model is to select relevant predictors to be part of the prediction model. Next step, investigating the applicability, clinical usefulness, and generalizability of the existing models. The *Manuscript I* identified a set of predictors and risk factors relevant to PPC development after HNS. Nevertheless, it identified studies reporting the use of Gupta et al. model that was developed and validated on the US population using ACS NSQIP dataset. This model has poor performance and discrimination in HNS patients; however, it was fully reported (e.g., values of intercept and coefficients) and can be systematically validated and updated using future years datasets from ACS NSQIP. While the only developed model for PPC risk prediction in HNS patients was not externally validated and the used predictors are not commonly available in data registries. Gupta et al. model is for predicting the 30-day risk of developing PPC. *Manuscript II* uses data from 2018 and 2019 ACS NSQIP, to assess this model's applicability in HNS patients and updating Gupta et al. model to be more optimized for HNS patients. The next chapter provides the full text of this manuscript, and, similar to the previous chapter, comprises independent appendices and a reference list, according to the requirements specified by the McGill Graduate and Postdoctoral Studies.

## 7. Manuscript II:

**Temporal Validation and Extension of a Prediction Model for Postoperative Pulmonary Complications**

Mohammad Al-Tamimi[1], Belinda Nicolau[1], Nicholas Makhoul[2], Sreenath Madathil[1]*

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

[2]Department of Dentistry and Oral and Maxillofacial Surgery, McGill University Health Centre, Montreal, Quebec, Canada

*Corresponding author: Dr. Sreenath Madathil

Faculty of Dental Medicine and Oral Health Sciences, McGill University, 2001 Avenue McGill College, Montréal, QC, H3A 1G1, Canada

**Keywords:** Head and Neck Surgery; postoperative pulmonary complications, predictive modeling

**Conflict of interest** None of the authors had conflict of interest regarding this project.

**Contributions** MT and SM conceived objective of the investigation, prepared data, and performed statistical analysis. BN and SM supervised all steps of the project. NM participated in the conceptual development and provided access to the dataset. All authors contributed to the conceptual development and writing the manuscript.

## Abstract:

**Background**: Postoperative pulmonary complications (PPCs) are common following head and neck surgeries (HNS). The incidence of PPCs in the HNS patients varies between 4.5% and 47%. These complications are often associated with increased morbidity, mortality, intensive care unit admission, and hospital length of stay. Prediction models used to estimate the probability of having PPCs in HNS are performing poorly. Therefore, we used one of the existing models to validate on HNS patients then updated the model with specific risk factors for these patients.

**Objectives:** This study aims to validate and update a PPCs risk prediction model on a sample of HNS patients using the National Surgical Quality Improvement Program from the American College of Surgeons (NSQIP ACS) database.

**Methods**: we validated the predictions of the Gupta pulmonary model developed in a sample of the US hospitals for general surgery patients using the NSQIP ACS dataset. This model, developed using multivariable logistic regression analysis, included the following functional dependency status, emergency of the procedure, history of blood sepsis, surgical specialty doing the procedure, and ASA classification. To validate our models, we used a cohort of 79726 patients who had HNS procedures in the period of 2018-2019. We replicated the Gupta model and tested its predictions on the derived dataset year-by-year then updated the model with a set of risk factors identified from a scoping review: age, sex, smoking history, body mass index, operation time, and reconstruction procedure. We evaluated the updated model's overall prediction performance by measuring Brier score scaled and $R^2$ (Negelkerke). The discrimination ability was tested using C-Statistics. The model's calibration was assessed by evaluating the calibration slope.

**Results**: the Gupta pulmonary risk prediction model developed for general surgery patients is reproducible and potentially applicable in the HNS patients. During the temporal validation (year 2019), overall performance assessment of the Gupta model was presented by Brier score scaled and $R^2$ of -0.0562 and -0.1963, respectively. The model's discrimination by C-Statistics was 0.715 (0.677-.0753) and calibration slope was 0.9389. The updated model after internal-external validation in year 2019 participants had a performance by Brier score scaled and $R^2$ of 0.0234 and 0.1435, respectively. Further, the C-statistic was 0.822 (0.785-0.858) for the updated model and

calibration slop was 0.9790. In summary, the updated model showed a high level of discrimination and high overall performance and calibration.

**Conclusion**: The updated model preforms better than the original model in HNS patients. Future studies are needed to understand the updated model's applicability in HNS clinical settings.

## Introduction:

Worldwide, approximately 300 million surgical procedures are performed annually, and 30% of patients undergoing these procedures suffer from postoperative complications[1-3]. One of the categories of these complications, postoperative pulmonary complications (PPCs), is broadly defined as conditions affecting the respiratory tract that adversely influence the prognosis and patients' recovery after surgery[4]. PPCs could be further subcategorized into (1) atelectasis (i.e., a state of collapsed and non-aerated region of the lung parenchyma[5]), (2) pneumonia, (3) pulmonary respiratory failure, and (4) exacerbation of underlying chronic lung disease[6]. These subcategories cause a significant risk of morbidity and mortality postoperatively. For instance, 3% of patients develop pulmonary respiratory failure postoperatively, and 25% of them die within 30 days[7]. Further, postoperative pneumonia may increase the length of hospital stay and associated costs by 89% and 55%, respectively[8].

One-third of patients who undergo general surgeries in the US lasting at least 2 hours with general anesthesia and mechanical ventilation might develop PPCs[17]. In head and neck surgeries (HNS), the incidence of PPCs worldwide varies between 4.5% and 47%[9-14] and these complications are associated with increased morbidity, mortality, intensive care unit admission, and length of hospital stay[11,15,16]. Preoperative surgical risk assessment is utilized to objectively outline risks before surgery, facilitating an informed consent process and encouraging exercising risk reduction measurements, such as preoperative respiratory physiotherapy and smoking cessation[18]. This risk prediction also guides strategies that prompt considerations of nonsurgical alternatives or advisability of postponing surgical procedures[19]. Several preoperative surgical risk assessment algorithms have been developed to identify patients at high risk for PPC, namely the Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT)[19] and the Gupta pulmonary risk index. The latter utilizes data from the American College of Surgeons –National Surgical Quality Improvement (ACS-NSQIP) database[7]. Although these risk assessment tools were developed using large surgical cohorts in both Europe year 2010 (59 Spanish hospitals) and the United States year 2011 (211 hospitals), only a few HNS patients were included. In fact, the

percentage of HNS evaluated in both ARISCAT and Gupta study cohorts were 6% and 0.3%, respectively[7,19]

Consequently, the ARISCAT score displayed poor ability to discriminate between those with and without a PPC in 794 HNS patients with (AUC) of 0.596 (95% CI: [0.542-0.649])[20]. Similarly, the Gupta Pulmonary risk Index has insufficient accuracy and poor discrimination ability in predicting complications in the previous cohort of HNS patients with AUC of 0.649 (95%CI: [0.589-0.701][20]. Despite these limitations, both tools (ARISCAT and Gupta Pulmonary risk Index) are widely employed in preoperative surgical risk assessment prior to major HNS[20]. Also, these tools were developed more than 10 years ago (Gupta in 2011 and ARISACT in 2010[7,19]) without any updates or customization for HNS patients. This customization might start from applying the recent literature reporting specific PPCs predictors for HNS patients, our recent scoping review for example. Moreover, utilizing these algorithms in the HNS patients lead to risk misclassification and, thus, to a suboptimal care delivery[20]. A recent model was developed for HNS patients by Smith and colleagues[22] which shows good discrimination (AUC of 0.75 [CI 0.69–0.80]) when compared to Gupta and ARISCAT, the AUCs for ARISCAT: 0.60, 95%CI: [0.54-0.65] and the Gupta index: 0.65, 95%CI: [0.59-0.71], models in that cohort of the US patients who had HNS. While their model's calibration was moderately good, it may underestimate the probability of a PPCs occurring at high probabilities >0.30[22]. Moreover, this model was not externally validated, limiting its generalizability[22].

Overall, there is a need to update the current models or develop a new model to predict PPCs after HNS[23]. We have previously conducted a systematic review to identify risk factors for PPCs in HNS patients. Using this updated evidence of PPC risk factors for HNS, our objective was to validate and then extend Gupta pulmonary index[7], specifically for patients undergoing HNS.

## Methods

### Source of Data:

We used data from the ACS-NSQIP from the years 2006 through 2019. This database is designed to help surgeons and hospitals better understand the quality of their care to patients within

similar situations in North American hospitals and other participant hospitals worldwide[25]. It was launched by the ACS in 2001 with 18 participant hospitals as a pilot program that functioned very well in reducing morbidity and mortality in participant hospitals. In 2004, ACS NSQIP started enrolling other private sector hospitals in their database. Currently, more than 700 participant hospitals across 49 states in the US and more than 100 hospitals from 11 different countries worldwide. ACS NSQIP collects reliable and validated data on patient demographics, laboratories, comorbidities, and 30-day postoperative outcomes for patients undergoing a broad range of operations across all surgical subspecialties. These datasets are used to provide hospitals with risk-adjusted 30-day outcome comparisons[25]. Gupta et. al used the 2007.'s and 2008's datasets from this database to develop and validate their PPC prediction model, respectively[7].

We created yearly cohorts of patients who underwent operations in the head and neck region from January 1, 2006, to December 31, 2019, spanning all surgical subspecialties using the following strategy.

**Participants**:

Head and neck surgery is a specialty in ear, nose and throat medicine (otolaryngology) combined with oral and maxillofacial medicine that focuses on surgically treating head and neck disorders. Head and neck surgeons treat cancerous and non-cancerous tumors such as oral cancer, pharynx cancer, nasopharyngeal cancer hypopharyngeal cancer, salivary gland cancer, paranasal and sinus cavity cancer, thyroid cancer, Nasopharyngeal, laryngeal or larynx cancer, and sarcomas of the head or neck. These procedures can be identified using the Current Procedural Terminology (CPT®) in the ACS NSQIP database. The CPT codes are five-digit numerical codes created by the American Medical Association (AMA) in 1966, to standardize reporting of medical, surgical, and diagnostic services and procedures performed in inpatient and outpatient settings. Each CPT® code represents a specific medical procedure or service, eliminating the subjective interpretations and non-standardized data entry[26].

To identify participants who had at least one HNS, during the period from January 2006 to December 2019, we used the following CPT codes: 21010-21499 (musculoskeletal head operations), 21501-21899 (musculoskeletal neck operations), 31300-31599 (larynx operations),

30000-30999 (nose operations), 40490-42999 (Mouth, tongue, and pharynx operations), 60000-60659 (Thyroid, parathyroid, carotid bodies operations), and 69000-69979 (ear operations).

**Outcome:**

Our primary outcome was an occurrence of PPC within the first 30 days after the operation. With the interest of replicability and extension of Gupta et al.'s model, we followed their definition of PPCm[7], which was: (1) occurrence of a pulmonary event that required placement of an endotracheal tube or other similar breathing tube or ventilator support, which was not intended or planned within 30 days after primary surgical procedure; OR (2) having a total cumulative duration of ventilator-assisted respirations greater than 48 hours during the postoperative hospitalization and any subsequent hospitalizations within 30 days postoperatively.

**Predictors:**

The dataset contains an array of information (275 variables), including sociodemographic (e.g., age, sex, weight, height), lifestyle (e.g., smoking history), clinical, and other (e.g., medical history, surgical procedure code, preoperative lab tests, postoperative course, length of stay, complications occurred in 30-day period postoperatively).

The model by Gupta et al. used 5 predictors: site of surgery (depending on which surgical department did the procedure), emergency case, dependent functional status, blood sepsis, and ASA class[7]. We extended the model with our selected predictors from our previous scoping review which were: age, operation time, and BMI included as continuous numerical variables and sex, history of smoking, history of COPD, reconstruction procedure, and site of the procedure included as binary variables.

We coded the variables based on the Gupta et al.'s study[7]. Accordingly, ASA classification, surgical specialty, functional dependance status, blood sepsis, emergency procedures were coded as categorical variables with multiple classes, while blood sepsis, emergency of the procedure were considered as binary variables with (Yes/No) outcomes.

**Sample Size:**

We calculated the minimum sample size following the method published by Riley et al.[7227] considering the ideal Observed/Expected (O/E) ratio of 1 and the outcome event proportion of 0.04 in the ACS NSQIP dataset. We aimed for a 95% confidence interval of 0.2 for O/E to ensure a good calibration in the external validation. Hence, 4000 participants (and about 160 events) are required to target a 95% confidence interval width of 0.2 for O/E is 1 in the external validation population. Given this requirement, we had sufficient events in the 2019 NSQIP dataset to serve as a validation dataset.

**Missing data:**

Participants with missing values for the predictors was excluded from the analysis. A total of 2.1% of the participants had missing values for the predictors included in the analysis. This is less than 5% of the total participants and thus we assume that the missing values are completely at random (MCAR)[7328]. Furthermore, the sample size after removing the missing values was still higher than the required sample size for a validation study.

**Statistical analysis methods:**

The Gupta model was developed and validated using NSQIP datasets for patients in years 2007 and 2008. Hence, we used the model's equation to validate it on 2019 dataset for HNS patients. We replicated the Gupta et al model[7] as below multivariable logistic regression equation:

$Logit(Probability\ of\ developing$ PPC) =

−1.7397 + (0.7678*Partially dependent) + (1.4046* Totally dependent) -

(3.5265* ASA I) - (2.0008* ASA II) - (0.6201* ASA III)+

(0.2441* ASA IV)- (0.7840* No sepsis) + (0.2752* Preoperative sepsis)+

(0.9035* Preoperative septic shock) - (0.5739* Emergency) +

(0.7336* Neurosurgery) + (0.2744* Cardiac) + (0.1060* ENT) -

(0.5271* thyroid) - (1.2431* Gynecology) - (0.2389* Vascular) +

(0.3093* Urology)+ (0.6715* Thoracic) - (0.8577* Orthopedics) -

(0.3206* Plastic) **Equation (1)**

We assessed the Gupta et al.[7] model performance using data from years 2018 and 2019 by calculating Brier score scaled and $R^2$ (Negelkerke) for overall performance and discrimination ability by C-statistic (area under the curve AUC). Calibration was measured by calibration slope.

## Model extension:

We extended the Gupta et al's[7] model by including additional predictors, identified via our scoping review, and considering the linear combination (output from equation 1) as an offset to a multivariable logistic regression model. Along with model extension we also performed calibration in-large by recalibrating the intercept of Gupta et al.[7] model.

The development and internal validation of the updated model was done using the dataset for year 2018, while 2019 dataset was kept aside for internal-external validation. These 2 years were chosen as they were comparable and similar in all predictors and contained more than the needed sample size for validation.

Updated model performance and discrimination were assessed using measurements of Brier score scaled and $R^2$ (Negelkerke) for overall performance, calibration slope for calibration, and discrimination ability by C-statistic (area under the curve AUC).

External validity was assessed by using dataset of 2019 year that was kept aside. The updated model was validated on these patients and assessed the performance using Brier score scaled and $R^2$ (Negelkerke). Calibration slope for calibration measurement and discrimination ability by C-statistic (AUC).

All statistical analyses were done using R Studio® (version 4.0.5) software.

**Results:**

We identified 79,726 patients who had at least one HNS procedures during the period 2018-2019. The incidence of PPC occurred in 343 cases from the total participants (0.4%). A brief description of the participants characteristics is provided in Table 1. The mean age for participants was 50.1 years. The majority of participants were females (67%) and the mean BMI and the mean operation time were 30.2 kg/m2 and 110 minutes, respectively. Only 14.5% of the participants reported history of smoking and 2.3% had history of COPD. Around 5.5% and 1% of the participants had neck dissection procedures and reconstruction procedures, respectively. Around 53.1% of the patients had ASA class II and 34.1% had ASA class III, while only 2.3% had ASA class IV. The majority of patients (99.3%) were completely independent in terms of their functional dependance status, while only 0.6% and 0.1% were partially dependent and totally dependent, respectively.

Figure 1 displays the calibration slope (0.94) for the external validation of the for Gupta et al.[7] model using patient data from year 2019 (39,598 participants). The model had a moderate discrimination ability as AUC was 0.72 (95% CI: 0.68-0.75) (Figure 2), while performance was poor that indicated a non-informative model as scaled Brier score was -0.06 and $R^2$(Negelkerke) was -0.20. These performance measures indicated the Gupta et al. model did not follow the trend of the HNS patients data, so it fits worse than a horizontal line (negative $R^2$(Negelkerke)).

Table 2 shows the results of validation on year 2019 dataset for original Gupta et al .model, recalibration of Gupta et al. model intercept, and the updated model. This table shows a comparison between the measures of the 3 steps of validation. The updated model showed better discrimination ability, better calibration and better performance on year 2019 patients compared to the original Gupta et al. model. While only recalibrating the intercept of Gupta model had slight improvement over the original Gupta et al. model, as AUC was 0.71 (95% CI: 0.68-0.75), scaled Brier score was 0.01, $R^2$(Negelkerke) was 0.05, and calibration slope was 0.81. Which means only recalibrating the intercept makes the model's performance better and better fitted to the HNS patients data.

The coefficients of the updated model are shown in table 3. The updated model was developed on 40,128 participants from 2018 and the models estimates for the added predictors is presented in table 3 with the re-estimated intercept.

The Performance of the updated model was good as the scaled Brier score for model development using year 2018 patients was 0.03 and $R^2$(Negelkerke) was 0.16, while discrimination ability was good as AUC was 0.82 (CI: 0.78-0.85). Calibration slope was 1.00.

Figure 3 shows the updated model had a C-statistic (AUC) after internal-external validation of the model on 2019 patients (39598 participants) of 0.82 (95% CI: 0.79-0.86). Scaled Brier score was 0.02 and $R^2$(Negelkerke) was 0.14 while the calibration slop was 0.98 (Figure 4). When compared to the external validation of the original Gupta on this same dataset (year 2019 participants), the updated model is showing better overall performance, discrimination ability and calibration.

**Discussion:**

In this study, we temporally validated an existing model on HNS patients derived using the same database the original model was built with. Also, we extended the model with predictors available in the database that were reported to be significant for developing PPCs in HNS patients: advanced age[14,22], male sex[14], smoking history[31], history of COPD[9], higher BMI[15], longer operation time[14], procedure site (tongue[32]), and reconstruction procedures[33].

Our results from validating Gupta's model on HNS patients yielded AUC 0.72 (95% CI: 0.68-.075) showing lower discrimination ability of the model in HNS patients than the original scores in the development study on general surgery patients. Smith and colleagues had an AUC 0.65 (95%CI: 0.59, 0.71) when they validated Gupta's model on their cohort of HNS patients[22]. Also, Wood and colleagues had similar results when validated the Gupta model on their cohort AUC 0.65 (95%CI: 0.59-0.70)[20]. The moderate discrimination ability of our model might be due to the fact we used the same database that Gupta's model was developed, while other studies converted their cohort's data to be coded as NSQIP predictors.

Interestingly, we were able to come up with a model that has a moderately good performance, discriminatory abilities, and calibration. This is the first model for PPC prediction in HNS patients (AUC: 0.82) with such metrics near the original Gupta index metrics (AUC: 0.90)[7].

Aside from identifying high-risk patients, we believe this updated model can be useful in the informed consent process. The process of patient-centered informed consent necessitates the presentation of adequate risk and benefit information[34]. To meet this goal, accurate individualized assessment of PPC, which contributes significantly to morbidity and mortality. Giving individualized assessment has not always been an easy task because each patient is unique[7], with their own set of risk factors. Thus, by estimating the risk of PPC, this updated model can simplify the informed consent process, and we foresee it being used preoperatively by head and neck surgeons.

Complex head and neck surgical cases can present the surgical team with a unique set of challenges[15,35]. As Smith and colleagues[22] showed that risk factors for PPCs studied in other settings may differ significantly from those important in patients undergoing free-flap reconstructions. As a result, decision support tools developed for general surgical populations perform poorly when applied to HNS population and are unlikely to be useful these patients[20]. When applied to this patient population specifically, the updated model presented in this study outperforms other models. In practice, patients who are at high risk of PPCs require aggressive preoperative management, which allows the patient to be better optimized for surgery. Thus, exercising risk reduction measurements, such as preoperative respiratory physiotherapy and smoking cessation[18].

This study has many strengths. First, we took advantage of a recent scoping review of the literature to select our predictors making them more informed by the recent evidence. Another advantage was to update instead of developing a new model using an available and widely used model, reducing the redundancy in the literature. For instance, this updated model can be used later in the future modelling research to be further updated if new evidence for predictors comes in light. Also, the methodology used here is explained in details, which might help future

researchers to adapt it more as an extra hands-on example for updating methods that were suggested by Steyerberg[22].

Despite its many strengths, this study has a few shortcomings. We were able to do analyze patients with complete data for the selected predictors assuming missing values were MCAR as discussed in the methods. Although this is justifiable as they were less than 5% of the cohort, this is not the ideal practice in predicting modelling and multiple imputations should be considered[22,28]. The variables examined were restricted to those recorded by NSQIP. Even though the data set was fairly comprehensive, with more than 50 preoperative variables, some comorbidities were not included, such as obstructive sleep apnea[7]. Similarly, although pulmonary function test results may be relevant to many comorbidities and surgeries, they are not available in NSQIP. The findings of this study may not be applicable to hospitals that are not part of the NSQIP. However, given its diversity (more than 700 participating hopitals[25]), this is unlikely. Finally, while data collection in NSQIP is prospective, these datasets were analyzed retrospectively for the development and validation of the updated model.

**Conclusion**

This updated model is a unique predictive model for PPCs in HNS patients. Further studies are necessary to externally validate its performance and generalizability on other cohorts of HNS. However, our results suggest that this model is a moderate improvement over the original risk index developed in the general surgical population[9].

**Table 1:** Overview of the participants from years 2018 and 2019 who had HNS

| Predictors | 2018 (N=40128) | 2019 (N=39598) | Total (N=79726) |
|---|---|---|---|
| **Age** | | | |
| Mean [SD] | 49.9 [17.4] | 50.3 [17.4] | 50.1 [17.4] |
| Median [Q25, Q75] | 51.0 [36.0,64.0] | 52.0 [36.0,64.0] | 52.0 [36.0,64.0] |
| Min, Max | 18.0, 90.0 | 18.0, 90.0 | 18.0, 90.0 |
| **SEX** | | | |
| female | 26882 (67.0%) | 26751 (67.6%) | 53633 (67.3%) |
| Male | 13246 (33.0%) | 12847 (32.4%) | 26093 (32.7%) |
| **Operation time** | | | |
| Mean [SD] | 110 [94.9] | 109 [92.3] | 110 [93.6] |
| Median [Q25,Q75] | 89.0 [54.0,138] | 89.0 [55.0,137] | 89.0 [54.0,137] |
| Min, Max | 0, 1400 | 0, 1430 | 0, 1430 |
| **Smoking** | | | |
| No | 34100 (85.0%) | 34049 (86.0%) | 68149 (85.5%) |
| Yes | 6028 (15.0%) | 5549 (14.0%) | 11577 (14.5%) |
| **Neck dissection** | | | |
| No | 37938 (94.5%) | 37439 (94.5%) | 75377 (94.5%) |
| Yes | 2190 (5.5%) | 2159 (5.5%) | 4349 (5.5%) |
| **Reconstruction** | | | |
| No | 39770 (99.1%) | 39178 (98.9%) | 78948 (99.0%) |

| Predictors | 2018 (N=40128) | 2019 (N=39598) | Total (N=79726) |
|---|---|---|---|
| Yes | 358 (0.9%) | 420 (1.1%) | 778 (1.0%) |
| **Body mass index** | | | |
| Mean [SD] | 30.2 [7.41] | 30.2 [7.41] | 30.2 [7.41] |
| Median [Q25,Q75] | 28.9 [25.0,34.1] | 28.9 [25.0,34.0] | 28.9 [25.0,34.0] |
| Min, Max | 11.5, 87.0 | 11.5, 92.9 | 11.5, 92.9 |
| **COPD** | | | |
| No | 39161 (97.6%) | 38700 (97.7%) | 77861 (97.7%) |
| Yes | 967 (2.4%) | 898 (2.3%) | 1865 (2.3%) |
| **Functional status** | | | |
| Independent | 39803 (99.2%) | 39351 (99.4%) | 79154 (99.3%) |
| Partially Dependent | 273 (0.7%) | 217 (0.5%) | 490 (0.6%) |
| Totally Dependent | 52 (0.1%) | 30 (0.1%) | 82 (0.1%) |
| **ASA classification** | | | |
| 1-No Disturb | 4350 (10.8%) | 4020 (10.2%) | 8370 (10.5%) |
| 2-Mild Disturb | 21175 (52.8%) | 21155 (53.4%) | 42330 (53.1%) |
| 3-Severe Disturb | 13643 (34.0%) | 13531 (34.2%) | 27174 (34.1%) |
| 4-Life Threat | 953 (2.4%) | 888 (2.2%) | 1841 (2.3%) |
| 5-Moribund | 7 (0.0%) | 4 (0.0%) | 11 (0.0%) |
| **Preoperative sepsis** | | | |
| None | 39766 (99.1%) | 39352 (99.4%) | 79118 (99.2%) |
| Sepsis | 179 (0.4%) | 91 (0.2%) | 270 (0.3%) |

| Predictors | 2018 (N=40128) | 2019 (N=39598) | Total (N=79726) |
|---|---|---|---|
| Septic Shock | 9 (0.0%) | 6 (0.0%) | 15 (0.0%) |
| SIRS | 174 (0.4%) | 149 (0.4%) | 323 (0.4%) |
| **Emergency** | | | |
| No | 39799 (99.2%) | 39374 (99.4%) | 79173 (99.3%) |
| Yes | 329 (0.8%) | 224 (0.6%) | 553 (0.7%) |
| **Surgical Specialty** | | | |
| Cardiac Surgery | 8 (0.0%) | 6 (0.0%) | 14 (0.0%) |
| General Surgery | 16233 (40.5%) | 16841 (42.5%) | 33074 (41.5%) |
| Gynecology | 7 (0.0%) | 16 (0.0%) | 23 (0.0%) |
| Neurosurgery | 17 (0.0%) | 22 (0.1%) | 39 (0.0%) |
| Orthopedics | 35 (0.1%) | 31 (0.1%) | 66 (0.1%) |
| Otolaryngology (ENT) | 22873 (57.0%) | 21840 (55.2%) | 44713 (56.1%) |
| Plastics | 590 (1.5%) | 511 (1.3%) | 1101 (1.4%) |
| Thoracic | 85 (0.2%) | 72 (0.2%) | 157 (0.2%) |
| Urology | 196 (0.5%) | 195 (0.5%) | 391 (0.5%) |
| Vascular | 84 (0.2%) | 64 (0.2%) | 148 (0.2%) |
| **PPC** | | | |
| No | 39945 (99.5%) | 39438 (99.6%) | 79383 (99.6%) |
| Yes | 183 (0.5%) | 160 (0.4%) | 343 (0.4%) |

**Table 2**: Validation results on year 2019 patients during each step

| Index | Gupta et al. model | Re-calibration of intercept alone | Model extension & recalibration of intercept |
|---|---|---|---|
| **AUC (95% CI)** | 0.72 (0.68-.075) | 0.71 (0.68-.075) | 0.82 (0.79-0.86) |
| **$R^2$** | -0.20 | 0.05 | 0.14 |
| **Brier scaled** | -0.06 | 0.01 | 0.02 |
| **Calibration slope** | 0.94 | 0.81 | 0.98 |

**Table 3**: Updated model intercepts, coefficients with SE (standard error) and odds ratio

| Extended and recalibrated of Gupta et al. model | | | |
|---|---|---|---|
| | Coefficients | SE | Odds Ratio (95% CI) |
| Intercept | -20.32 | 383.55 | |
| Age | 0.024 | 0.005 | 1.02 (1.02-1.03) |
| Sex: Male | 0.07 | 0.16 | 1.07 (0.78-1.47) |
| Operation time | 0.003 | 0.0004 | 1.00 (1.00-1.00) |
| Smoking: Yes | 0.46 | 0.18 | 1.58 (1.10-2.27) |
| Surgical site: Head bones | 14.99 | 383.555 | |
| Larynx | 14.99 | 383.55 | |
| Mouth, tongue, Pharynx | 15.13 | 383.55 | |
| Neck tissue | 16.65 | 383.55 | |
| Thyroid | 15.11 | 383.55 | |
| Neck dissection | -0.02 | 0.33 | 0.98 (0.51-1.89) |
| Reconstruction | -0.24 | 0.64 | 0.78 (0.22-2.75) |

| | | | |
|---|---|---|---|
| BMI | 0.015 | 0.009 | 1.01 (1.00-1.03) |
| COPD: Yes | 0.199 | 0.29 | 1.22 (0.69-2.15) |
| Gupta variables | 1.10 | 0.096 | 3.01 (2.49-3.63) |

**Figure 1**: Calibration plot for the Gupta model temporal validation (year 2019)

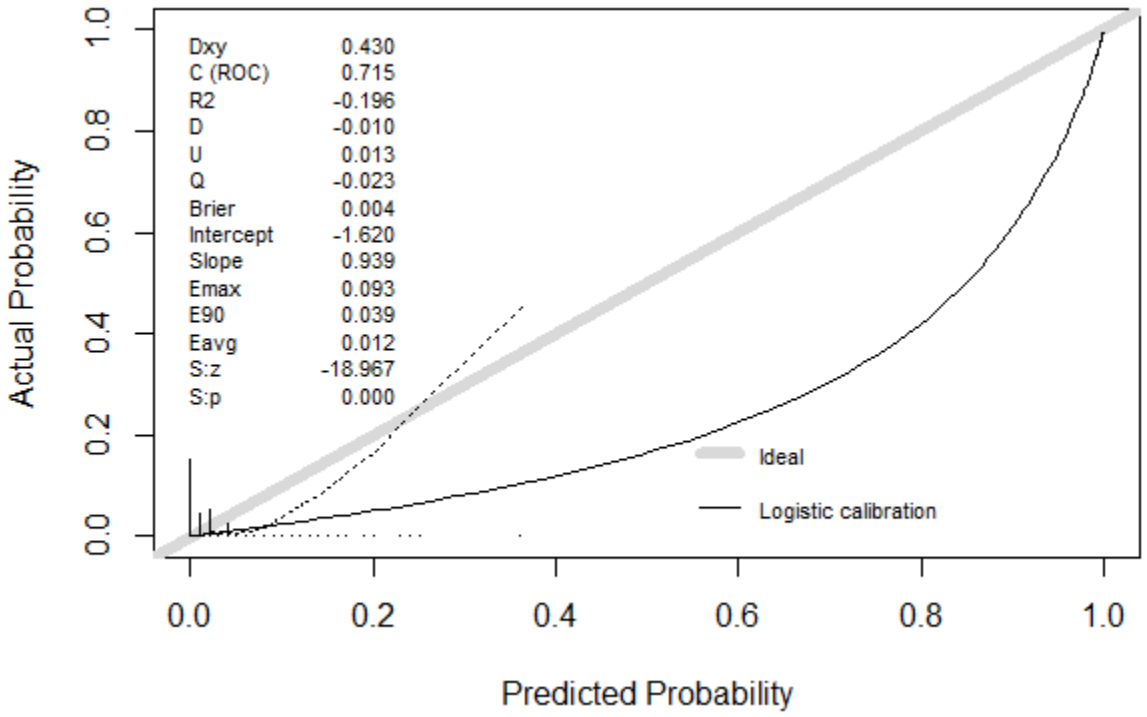


**Figure 2**: Area Under the Curve (AUC) for the Gupta et al. Model in 2019 HNS patients

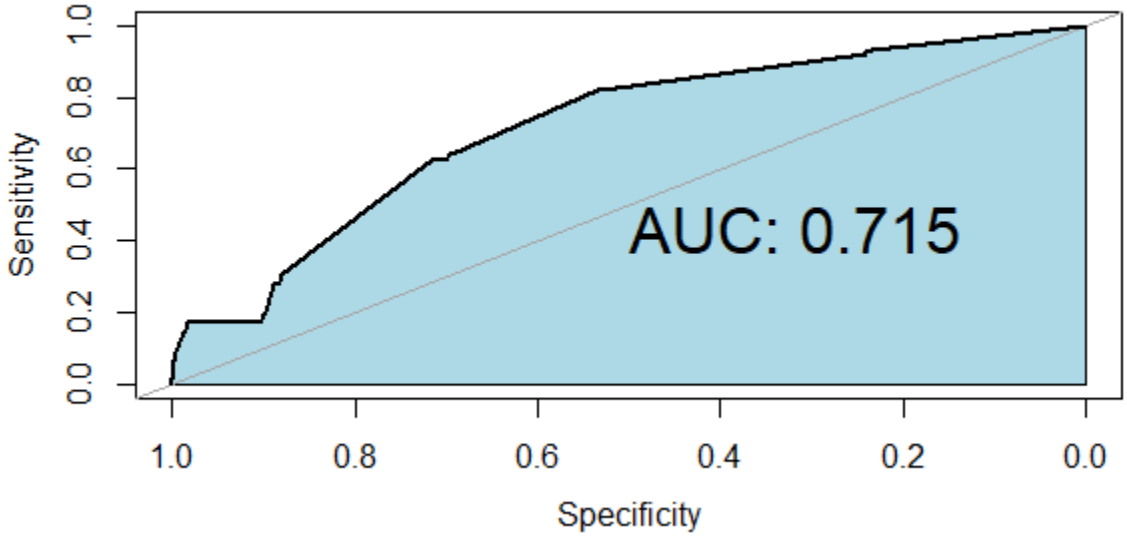



**Figure 3:** AUC for updated model after external validation in 2019 data.

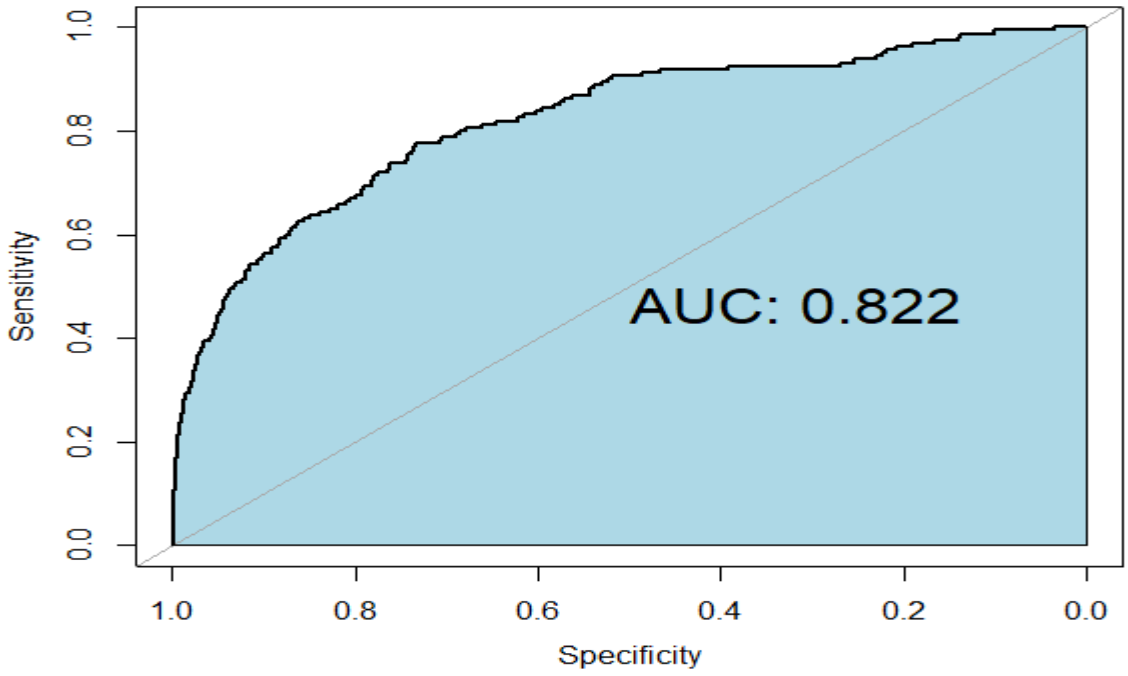
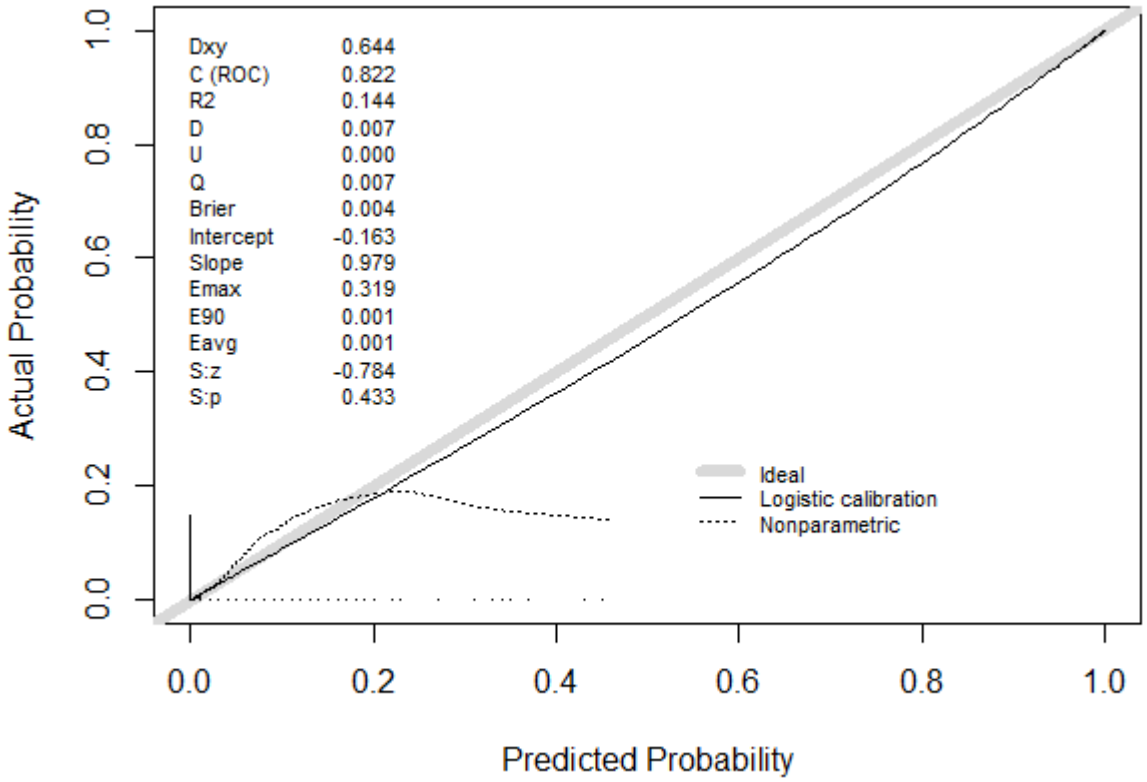


**Figure 4:** Calibration plot for updated model after external validation in 2019 data.




)

**References**


1. Meara JG, Leather AJ, Hagander L, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; **386**(9993): 569-624.

2. Longo WE, Virgo KS, Johnson FE, et al. Risk factors for morbidity and mortality after colectomy for colon cancer. *Dis Colon Rectum* 2000; **43**(1): 83-91.

3. Mayo NE, Feldman L, Scott S, et al. Impact of preoperative change in physical function on postoperative recovery: argument supporting prehabilitation for colorectal surgery. *Surgery* 2011; **150**(3): 505-14.

4. Shander A, Fleisher LA, Barie PS, Bigatello LM, Sladen RN, Watson CB. Clinical and economic burden of postoperative pulmonary complications: patient safety summit on definition, risk-reducing interventions, and preventive strategies. *Crit Care Med* 2011; **39**(9): 2163-72.

5. Peroni DG, Boner AL. Atelectasis: mechanisms, diagnosis and management. *Paediatr Respir Rev* 2000; **1**(3): 274-8.

6. Smetana GW, Lawrence VA, Cornell JE. Preoperative pulmonary risk stratification for noncardiothoracic surgery: Systematic review for the American College of Physicians. *Ann Intern Med* 2006; **144**(8): 581-95.

7. Gupta H, Gupta PK, Fang X, et al. Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest* 2011; **140**(5): 1207-15.

8. Khan NA, Quan H, Bugar JM, Lemaire JB, Brant R, Ghali WA. Association of postoperative complications with hospital costs and length of stay in a tertiary care center. *J Gen Intern Med* 2006; **21**(2): 177-80.

9. Rao MK, Reilley TE, Schuller DE, Young DC. Analysis of risk factors for postoperative pulmonary complications in head and neck surgery. *Laryngoscope* 1992; **102**(1): 45-7.

10. McCulloch TM, Jensen NF, Girod DA, Tsue TT, Weymuller EA, Jr. Risk factors for pulmonary complications in the postoperative head and neck surgery patient. *Head Neck* 1997; **19**(5): 372-7.

11. Petrar S, Bartlett C, Hart RD, MacDougall P. Pulmonary complications after major head and neck surgery: A retrospective cohort study. *Laryngoscope* 2012; **122**(5): 1057-61.

12. Ong SK, Morton RP, Kolbe J, Whitlock RM, McIvor NP. Pulmonary complications following major head and neck surgery with tracheostomy: a prospective, randomized, controlled trial of prophylactic antibiotics. *Arch Otolaryngol Head Neck Surg* 2004; **130**(9): 1084-7.

13. De Melo GM, De Cassia Braga Ribeiro K, Kowalshi LP, Deheinzelin D. Risk factors for postoperative complications in oral cancer and their prognostic implications. *Archives of Otolaryngology - Head and Neck Surgery* 2001; **127**(7): 828-33.

14. Xu J, Hu J, Yu P, et al. Perioperative risk factors for postoperative pneumonia after major oral cancer surgery: A retrospective analysis of 331 cases. *PLoS ONE* 2017; **12**(11): e0188167.

15. Loeffelbein DJ, Julinek A, Wolff KD, Kochs E, Haller B, Haseneder R. Perioperative risk factors for postoperative pulmonary complications after major oral and maxillofacial surgery with microvascular reconstruction: A retrospective analysis of 648 cases. *J Craniomaxillofac Surg* 2016; **44**(8): 952-7.

16. Jones NF, Jarrahy R, Song JI, Kaufman MR, Markowitz B. Postoperative medical complications--not microsurgical complications--negatively influence the morbidity, mortality, and true costs after microsurgical reconstruction for head and neck cancer. *Plast Reconstr Surg* 2007; **119**(7): 2053-60.

17. Fernandez-Bustamante A, Frendl G, Sprung J, et al. Postoperative Pulmonary Complications, Early Mortality, and Hospital Stay Following Noncardiothoracic Surgery: A Multicenter Study by the Perioperative Research Network Investigators. *JAMA Surg* 2017; **152**(2): 157-66.

18. Qaseem A, Snow V, Fitterman N, et al. Risk assessment for and strategies to reduce perioperative pulmonary complications for patients undergoing noncardiothoracic surgery: a guideline from the American College of Physicians. *Ann Intern Med* 2006; **144**(8): 575-80.

19. Canet J, Gallart L, Gomar C, et al. Prediction of postoperative pulmonary complications in a population-based surgical cohort. *Anesthesiology* 2010; **113**(6): 1338-50.

20. Wood CB, Shinn JR, Rees AB, et al. Existing Predictive Models for Postoperative Pulmonary Complications Perform Poorly in a Head and Neck Surgery Population. *J Med Syst* 2019; **43**(10): 312.

21. Sousa Menezes A, Fernandes A, Rocha Rodrigues J, et al. Optimizing classical risk scores to predict complications in head and neck surgery: a new approach. *Eur Arch Otorhinolaryngol* 2021; **278**(1): 191-202.

22. Smith DK, Freundlich RE, Shinn JR, Wood CB, Rohde SL, McEvoy MD. An improved predictive model for postoperative pulmonary complications after free flap reconstructions in the head and neck. *Head Neck* 2021; **30**: 30.

23. Steyerberg EW. Clinical prediction models : a practical approach to development, validation, and updating. Switzerland: Springer

2019.

24. Moons KG, Altman DG, Reitsma JB, Collins GS. New Guideline for the Reporting of Studies Developing, Validating, or Updating a Multivariable Clinical Prediction Model: The TRIPOD Statement. *Adv Anat Pathol* 2015; **22**(5): 303-5.

25. Surgeons ACo. American College of Surgeons National Surgical Quality Improvement Program American College of Surgeons 2020.

26. Coders AAoP. https://www.aapc.com/resources/medical-coding/cpt.aspx.

27. Riley RD, Debray TPA, Collins GS, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Stat Med* 2021; **40**(19): 4230-51.

28. Mukaka M, White SA, Terlouw DJ, Mwapasa V, Kalilani-Phiri L, Faragher EB. Is using multiple imputation better than complete case analysis for estimating a prevalence (risk) difference in randomized controlled trials when binary outcome observations are missing? *Trials* 2016; **17**: 341.

29. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; **98**(9): 691-8.

30. Poldrack RA, Huckins G, Varoquaux G. Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psychiatry* 2020; **77**(5): 534-40.

31. Liu Y, Xue FS, Liu GP, Sun C. The risk factors for postoperative pulmonary complications after head and neck cancer surgery. *J Clin Anesth* 2016; **30**: 33-4.

32. Smith JE, Suh JD, Erman A, Nabili V, Chhetri DK, Blackwell KE. Risk factors predicting aspiration after free flap reconstruction of oral cavity and oropharyngeal defects. *Arch Otolaryngol Head Neck Surg* 2008; **134**(11): 1205-8.

33. Joo YH, Sun DI, Cho JH, Cho KJ, Kim MS. Factors that predict postoperative pulmonary complications after supracricoid partial laryngectomy. *Arch Otolaryngol Head Neck Surg* 2009; **135**(11): 1154-7.

34. Cohen ME, Bilimoria KY, Ko CY, Hall BL. Development of an American College of Surgeons National Surgery Quality Improvement Program: morbidity and mortality risk calculator for colorectal surgery. *J Am Coll Surg* 2009; **208**(6): 1009-16.

35. Mazo V, Sabaté S, Canet J, et al. Prospective external validation of a predictive score for postoperative pulmonary complications. *Anesthesiology* 2014; **121**(2): 219-31.

# 8. Discussion

## 8.1 Summary of the results:

The incidence rates of PPCs in HNS vary from 4.5% to 47%[11,14,15,17]. PPCs are associated with increased mortality, morbidity, hospital length of stay, and costs[10,14,18]. Reported risk factors that predicts occurrence of PPCs in HNS patients vary in the literature and no consensus is found for definitive predictors. As identified in manuscript I, several definitions of PPCs are present in the literature, because the studies capture the outcome in different ways. Despite this, we were able to identify a set of risk factors that might be predictors for PPCs development in HNS patients. These factors were categorized into: (i) patient-related risk factors including: obesity, advanced age, history of chronic lung disease, history of heart disease, smoking, male gender, alcohol use, preoperative oxygen saturation, ASA classification, serum albumin level, weight loss, and history of DVT or PE; and (ii) procedure-related risk factors including: Operation time, blood loss and transfusion, anesthesia technique and dose, tongue base resection, tracheostomy, immobilization after surgery, size of the flap with reconstruction, and reconstruction procedure itself.

To mitigate PPCs, risk prediction can be used to identify patients at high risk of developing the outcome. This risk prediction also guides strategies that prompt considerations of nonsurgical alternatives or advisability of postponing surgical procedures[25]. When searching the literature to identify the PPC risk prediction models used in the assessment of patients undergoing HNS, one of the most known models was Gupta Pulmonary index. Taking advantage of having access to the same database that Gupta et al. used to develop and validate their index for general surgical populations, we validate the model for HNS patients of years after the original model's development (2009-2019) (Manuscript II). Subsequently, we followed Steyerberg[22] step-by-step instruction and updated Gupta et al's model with new predictors: age, sex, operation time, BMI, history of smoking, history of COPD, reconstruction procedure, neck dissection, and site of the procedure.

The updated model showed moderate improvement from other attempts to develop new prediction models, Smith and colleagues' model, for example[27]. Smith and colleagues had an AUC

0.65 (95%CI: 0.59, 0.71) when they validated Gupta et al.'s model on their cohort of HNS patients[27]. While our validation for the same model yielded AUC 0.72 (95% CI: 0.68-.075). This moderate discrimination ability enhancement might be due to the fact we used the same database the model was developed with, while other studies converted their cohort's data to be coded as NSQIP predictors.

Interestingly, we were able to come up with a model that has a moderately good performance, discriminatory abilities, and calibration as presented in Manuscript II. This is the first model for PPC prediction in HNS patients (AUC: 0.82) with such metrics near the original Gupta et al.'s index metrics (AUC: 0.897)[9].

## 8.2 Strengths, limitations and challenges:

In manuscript I, as it was a review of the literature, we were able to highlight the risk of bias in analyses of the studies that reported PPC risk factors in HNS. Majority of the literature did not report adjusted odds ratio for the concluded predictors, which suggests that the authors did not account for confounding. Thus, an issue of questionable causality of these predictors arises[22]. Another reason might be the limitation of retrospective study designs that accounted for majority of the studies.

Another challenge when selecting the appropriate risk factors to be included in the updated model in manuscript II, availability of these predictors and how they are registered within the NSQIP. As NSQIP has a rigorous and calibrated process of data entry, the outcome assessment might not be always optimal for such registry databases[22]. This might have limited our access to predictors that might be useful in our model, like exact amount of smoking as it is only recorded (Yes/No) not as (pack per year). Also, for the definition of the outcome, we were restricted by the Gupta definition of respiratory failure[9], although we had data for updating the definition to be more comprehensive for PPCs as recommended to include Pneumonia[23]. This had decreased the number of participants with the outcome included in the analysis, as the outcome (PPCs) now has a clear and broad definition. Hence, we might have a better and more generalizable PPC prediction model.

Defining the HNS population in NSQIP datasets was another challenge. We were only able to do that by selecting participants with CPT® codes in ranges that are defined for HNS procedures. However, we cannot measure the accuracy of this selection process due to the patient confidentiality and anonymous hospitals participated in the database. Thus, increasing the risk of selection bias and information bias, as we might have missed some patients who had HNS procedures but was not coded as the main surgical procedure, as well as we might have included patients who had other than HNS procedures just because they were coded under HNS as the main surgical procedure.

Despite these limitations and challenges, we conduct the first literature review (scoping review) that summarized all PPC risk factors for HNS patients. Another advantage was assessing the risk of bias in the included literature as this is not mandatory for scoping reviews according to the PRISMA guidelines[64], which adds value to the review. As for Manuscript II, we validated the Gupta et al.'s model for the first time in HNS patients using ACS NSQIP database, and showed similar results reported in other studies as mentioned before. Another advantage was the large numbers for validation samples for HNS patients. Also, our updated model is performing better than the original model given all the limitations and challenges mentioned earlier.

### 8.3 Future steps and considerations for future development of risk prediction model:

This thesis project provided an overview of what is there in the literature on PPCs risk factors for HNS and the possibility of generalizing one of the existing models in the HNS

Based on the results of *Manuscript I*, most existing PPC risk factors studies have not accounted for confounding factors, making it challenging to generalize their results. Therefore, this thesis project emphasizes the importance of adjustment for confounding in the analysis of future studies. Also, the use of recommended definition of PPC might help generalizing the results on other cohorts.

The second part results (*Manuscript II*) also presented the possibility of implementing an existing model for PPC in HNS patients. The validated model displayed moderately high overall performance, discrimination ability, and calibration. Future studies are needed to externally

validate this updated model and maybe further updating is needed. Future studies are suggested to benefit from the methods followed in the *Manuscript II* and validate the other existing PPC models in HNS patients.

## 8.4 Implications of the results:

The overview of the literature of PPC risk factors in HNS patients that was provided in this thesis may help researchers in different settings. Future studies' investigators can refer to Manuscript I and use its results for designing studies that evaluate PPC risk factors in HNS to avoid the potential risk of bias. Future studies might build on Manuscript I results and move on to do a systematic review and meta-analysis to provide a comprehensive evidence based adjusted ratios for the PPC risk factors.

The updated model in the *Manuscript II* also can be implemented in different settings. In epidemiological research, this model may help in recruiting the participants for the clinical trials PPC preventive interventions. Also, risk prediction model developers could use this model in determining the minimum sample size needed for prospective or retrospective studies related to PPC development in HNS. In the clinical setting in HNS, this model could be a helpful tool for the clinicians to assess the risk of developing PPC in the upcoming HNS procedures, helping in the individualized preventive interventions such as considerations of nonsurgical alternatives or advisability of postponing surgical procedures[25]. The updated model could also be embedded in a web-based or mobile app, assisting clinicians in making critical decisions. This model's predictions could also be incorporated into electronic patient records, assisting future epidemiological research or assisting clinical decision-making.

## 8.5 Knowledge translation:

We intend to present the results of both manuscripts using different platforms as part of this thesis's knowledge translation role. The scoping review project will be published as a journal article, assisting future PPC risk factors in HNS researchers. Presenting it at various conferences may also aid in the publication of our achievements. A portion of this research has already been presented at the 2022 AADOCR/CADR Annual Meeting & Exhibition, as well as at several local

conferences, including the McGill Dentistry Research Day and the Réseau de Recherche en Santé Buccodentaire et Osseuse (RSBO) conferences. Also, our goal is to publish the results of the second manuscript. The updated model could also be presented on various platforms. The model's regression formula will aid researchers in externally validating the model or using it in various epidemiological studies related to PPC in HNS. The updated model, like the original Gupta et al.'smodel[9], could serve as a foundation for paper-based risk assessment tools such as score charts or nomograms, assisting in identifying individuals at high risk of developing PPC in HNS.

## 9. Conclusion:

1- Risk factors for PPC in HNS patients can be mapped as:

a) patient-related risk factors: obesity, advanced age, history of chronic lung disease, history of heart disease, smoking, male gender, alcohol use, preoperative oxygen saturation, ASA classification, serum albumin level, weight loss, and history of DVT or PE

b) procedure-related risk factors: Operation time, blood loss and transfusion, Anesthesia technique and dose, tongue base resection, tracheostomy, immobilization after surgery, size of the flap with reconstruction, and reconstruction procedure itself.

2- There is a need for establishing causal inferences of PPC risk factors for HNS patients using the proper study designs.

3- Gupta Pulmonary Index is performing moderately good in HNS patients.

4- Updated model of Gupta shows high performance, discrimination ability and calibration.

5- Further studies to externally validate the updated model on other HNS cohorts are needed and maybe further updating of the model.

## 10. References


1. Meara JG, Leather AJ, Hagander L, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* 2015; **386**(9993): 569-624.
2. Longo WE, Virgo KS, Johnson FE, et al. Risk factors for morbidity and mortality after colectomy for colon cancer. *Dis Colon Rectum* 2000; **43**(1): 83-91.
3. Mayo NE, Feldman L, Scott S, et al. Impact of preoperative change in physical function on postoperative recovery: argument supporting prehabilitation for colorectal surgery. *Surgery* 2011; **150**(3): 505-14.
4. Tevis SE, Kennedy GD. Postoperative complications and implications on patient-centered outcomes. *J Surg Res* 2013; **181**(1): 106-13.
5. Shander A, Fleisher LA, Barie PS, Bigatello LM, Sladen RN, Watson CB. Clinical and economic burden of postoperative pulmonary complications: patient safety summit on definition, risk-reducing interventions, and preventive strategies. *Crit Care Med* 2011; **39**(9): 2163-72.
6. Peroni DG, Boner AL. Atelectasis: mechanisms, diagnosis and management. *Paediatr Respir Rev* 2000; **1**(3): 274-8.
7. Smetana GW, Lawrence VA, Cornell JE, Smetana GW, Lawrence VA, Cornell JE. Preoperative pulmonary risk stratification for noncardiothoracic surgery: systematic review for the American College of Physicians. *Ann Intern Med* 2006; **144**(8): 581-40.
8. Fernandez-Bustamante A, Frendl G, Sprung J, et al. Postoperative Pulmonary Complications, Early Mortality, and Hospital Stay Following Noncardiothoracic Surgery: A Multicenter Study by the Perioperative Research Network Investigators. *JAMA Surg* 2017; **152**(2): 157-66.
9. Gupta H, Gupta PK, Fang X, et al. Development and validation of a risk calculator predicting postoperative respiratory failure. *Chest* 2011; **140**(5): 1207-15.
10. Khan NA, Quan H, Bugar JM, Lemaire JB, Brant R, Ghali WA. Association of postoperative complications with hospital costs and length of stay in a tertiary care center. *J Gen Intern Med* 2006; **21**(2): 177-80.
11. Xu J, Hu J, Yu P, et al. Perioperative risk factors for postoperative pneumonia after major oral cancer surgery: A retrospective analysis of 331 cases. *PLoS ONE* 2017; **12**(11): e0188167.
12. Rao MK, Reilley TE, Schuller DE, Young DC. Analysis of risk factors for postoperative pulmonary complications in head and neck surgery. *Laryngoscope* 1992; **102**(1): 45-7.
13. McCulloch TM, Jensen NF, Girod DA, Tsue TT, Weymuller Jr EA. Risk factors for pulmonary complications in the postoperative head and neck surgery patient. *Head and Neck* 1997; **19**(5): 372-7.
14. Petrar S, Bartlett C, Hart RD, MacDougall P. Pulmonary complications after major head and neck surgery: A retrospective cohort study. *Laryngoscope* 2012; **122**(5): 1057-61.
15. Ong SK, Morton RP, Kolbe J, Whitlock RML, McIvor NP. Pulmonary complications following major head and neck surgery with tracheostomy: A prospective, randomized, controlled trial of prophylactic antibiotics. *Archives of Otolaryngology - Head and Neck Surgery* 2004; **130**(9): 1084-7.
16. De Melo GM, De Cassia Braga Ribeiro K, Kowalshi LP, Deheinzelin D. Risk factors for postoperative complications in oral cancer and their prognostic implications. *Archives of Otolaryngology - Head and Neck Surgery* 2001; **127**(7): 828-33.
17. Jones NF, Jarrahy R, Song JI, Kaufman MR, Markowitz B. Postoperative medical complications--not microsurgical complications--negatively influence the morbidity, mortality, and true costs after microsurgical reconstruction for head and neck cancer. *Plast Reconstr Surg* 2007; **119**(7): 2053-60.
18. Loeffelbein DJ, Julinek A, Wolff KD, Kochs E, Haller B, Haseneder R. Perioperative risk factors for postoperative pulmonary complications after major oral and maxillofacial surgery with microvascular reconstruction: A retrospective analysis of 648 cases. *J Craniomaxillofac Surg* 2016; **44**(8): 952-7.

19. Qaseem A, Snow V, Fitterman N, et al. Risk assessment for and strategies to reduce perioperative pulmonary complications for patients undergoing noncardiothoracic surgery: a guideline from the American College of Physicians. *Ann Intern Med* 2006; **144**(8): 575-80.
20. Sousa Menezes A, Fernandes A, Rocha Rodrigues J, et al. Optimizing classical risk scores to predict complications in head and neck surgery: a new approach. *Eur Arch Otorhinolaryngol* 2021; **278**(1): 191-202.
21. Wood CB, Shinn JR, Rees AB, et al. Existing Predictive Models for Postoperative Pulmonary Complications Perform Poorly in a Head and Neck Surgery Population. *J Med Syst* 2019; **43**(10): 312.
22. Steyerberg EW. Clinical prediction models : a practical approach to development, validation, and updating. Switzerland: Springer

2019.
23. Abbott TEF, Fowler AJ, Pelosi P, et al. A systematic review and consensus definitions for standardised end-points in perioperative medicine: pulmonary complications. *Br J Anaesth* 2018; **120**(5): 1066-79.
24. Duggan M, Kavanagh BP. Pulmonary atelectasis: a pathogenic perioperative entity. *Anesthesiology* 2005; **102**(4): 838-54.
25. Canet J, Gallart L, Gomar C, et al. Prediction of postoperative pulmonary complications in a population-based surgical cohort. *Anesthesiology* 2010; **113**(6): 1338-50.
26. Mazo V, Sabaté S, Canet J, et al. Prospective external validation of a predictive score for postoperative pulmonary complications. *Anesthesiology* 2014; **121**(2): 219-31.
27. Thompson DA, Makary MA, Dorman T, Pronovost PJ. Clinical and economic outcomes of hospital acquired pneumonia in intra-abdominal surgery patients. *Ann Surg* 2006; **243**(4): 547-52.
28. Linde-Zwirble W, Bloom J, Mecca R, Hansell D. Postoperative pulmonary complications in adult elective surgery patients in the US: severity, outcomes and resources use. *Critical Care* 2010; **14**(1): 1-2.
29. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg* 2005; **242**(3): 326-41; discussion 41-3.
30. Arozullah AM, Khuri SF, Henderson WG, Daley J. Development and validation of a multifactorial risk index for predicting postoperative pneumonia after major noncardiac surgery. *Ann Intern Med* 2001; **135**(10): 847-57.
31. Smith DK, Freundlich RE, Shinn JR, Wood CB, Rohde SL, McEvoy MD. An improved predictive model for postoperative pulmonary complications after free flap reconstructions in the head and neck. *Head Neck* 2021; **30**: 30.
32. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *Jama* 1997; **277**(6): 488-94.
33. Moons KG, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012; **98**(9): 683-90.
34. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *Bmj* 2009; **338**: b375.
35. Christensen E. Prognostic models including the Child-Pugh, MELD and Mayo risk scores--where are we and where should we go? *J Hepatol* 2004; **41**(2): 344-50.
36. Feinstein AR. "Clinical Judgment" revisited: the distraction of quantitative models. *Ann Intern Med* 1994; **120**(9): 799-805.
37. Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993; **118**(3): 201-10.
38. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med* 1996; **125**(5): 406-12.

39. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *Bmj* 2018; **363**: k4245.
40. Jehi L, Ji X, Milinovich A, et al. Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing: Results From 11,672 Patients. *Chest* 2020; **158**(4): 1364-75.
41. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol* 1995; **48**(12): 1503-10.
42. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* 1994; **69**(6): 979-85.
43. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006; **59**(10): 1087-91.
44. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006; **59**(10): 1092-101.
45. Altman DG, Royston P. The cost of dichotomising continuous variables. *Bmj* 2006; **332**(7549): 1080.
46. Royston P, Sauerbrei W, Altman DG. Modeling the effects of continuous risk factors. *J Clin Epidemiol* 2000; **53**(2): 219-21.
47. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006; **25**(1): 127-41.
48. Harrell F. Regression Modeling Strategies. New York: Springer; 2001.
49. Moons KG, Biesheuvel CJ, Grobbee DE. Test research versus diagnostic research. *Clin Chem* 2004; **50**(3): 473-6.
50. Wang J. Calibration slope versus discrimination slope: shoes on the wrong feet. *J Clin Epidemiol* 2020; **125**: 161-2.
51. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984; **3**(2): 143-52.
52. Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Stat Med* 2006; **25**(20): 3474-86.
53. Efron B TR. An Introduction to the Bootstrap. Monographs on Statistics and

Applied Probability. New York: Chapman & Hall; 1993.
54. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; **98**(9): 691-8.
55. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008; **61**(11): 1085-94.
56. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *Bmj* 2009; **338**: b605.
57. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000; **19**(4): 453-73.
58. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006; **144**(3): 201-9.
59. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *Bmj* 2009; **338**: b606.
60. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004; **23**(16): 2567-86.

61. Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or Remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999; **99**(16): 2098-104.
62. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008; **61**(1): 76-86.
63. @MaartenvSmeden. Should a risk prediction model be developed? Comments welcome. Should a risk prediction model be developed? Comments welcome: Twitter; 2018.
64. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International journal of social research methodology* 2005; **8**(1): 19-32.
65. Grimshaw J. A guide to knowledge synthesis: a knowledge synthesis chapter. *Ottawa: Canadian Institutes of Health Research* 2010; **8**: 1-56.
66. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation* 2015; **13**(3): 141-6.
67. Morris M, Boruff JT, Gore GC. Scoping reviews: establishing the role of the librarian. *J Med Libr Assoc* 2016; **104**(4): 346-54.
68. Surgeons ACo. American College of Surgeons National Surgical Quality Improvement Program American College of Surgeons 2020.
69. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly weather review* 1950; **78**(1): 1-3.
70. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; **21**(1): 128-38.
71. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making* 2006; **26**(6): 565-74.
72. Riley RD, Debray TPA, Collins GS, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Stat Med* 2021; **40**(19): 4230-51.
73. Mukaka M, White SA, Terlouw DJ, Mwapasa V, Kalilani-Phiri L, Faragher EB. Is using multiple imputation better than complete case analysis for estimating a prevalence (risk) difference in randomized controlled trials when binary outcome observations are missing? *Trials* 2016; **17**: 341.