



Optimizing classical risk scores to predict complications in head and neck surgery: a new approach

Ana Sousa Menezes¹ · Antero Fernandes² · Jéssica Rocha Rodrigues³ · Carla Salomé⁴ · Firmino Machado⁵ · Luís Antunes^{3,6} · Joaquim Castro Silva⁷ · Eurico Monteiro⁷ · Lúcio Lara Santos^{4,8,9}

Received: 31 December 2019 / Accepted: 12 June 2020 / Published online: 18 June 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Purpose To validate tools to identify patients at risk for perioperative complications to implement prehabilitation programmes in head and neck surgery (H&N).

Methods Retrospective cohort including 128 patients submitted to H&N, with postoperative Intermediate Care Unit admittance. The accuracy of the risk calculators ASA, P-POSSUM, ACS-NSQIP and ARISCAT to predict postoperative complications and mortality was assessed. A multivariable analysis was subsequently performed to create a new risk prediction model for serious postoperative complications in our institution.

Results Our 30-day morbidity and mortality were 45.3% and 0.8%, respectively. The ACS-NSQIP failed to predict complications and had an acceptable discrimination ability for predicting death. The discrimination ability of ARISCAT for predicting respiratory complications was acceptable. ASA and P-POSSUM were poor predictors for mortality and morbidity. Our new prediction model included ACS-NSQIP and ARISCAT (area under the curve 0.750, 95% confidence intervals: 0.63–0.87).

Conclusion Despite the insufficient value of these risk calculators when analysed individually, we designed a risk tool combining them which better predicts the risk of serious complications.

Keywords P-POSSUM · ACS-NSQIP · ASA · ARISCAT · Head and neck

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00405-020-06133-1>) contains supplementary material, which is available to authorized users.

✉ Ana Sousa Menezes
ana4644@gmail.com; ana.menezes@hb.min-saude.pt

¹ Department of Otorhinolaryngology-Head and Neck Surgery, Hospital De Braga, Sete Fontes - São Victor, 4710-243 Braga, Portugal

² Polyvalent Intensive Care Unit, Hospital Garcia de Orta, E.P.E, Almada, Portugal

³ Department of Epidemiology, Portuguese Institute of Oncology of Porto FG (IPO-Porto), Porto, Portugal

⁴ Surgical Intermediate Care Unit, Portuguese Institute of Oncology of Porto FG (IPO-Porto), Porto, Portugal

⁵ Department of Public Health Unit, ACES Porto Ocidental, Porto, Portugal

Introduction

Major surgery implies significant homeostatic disturbance to the patient and it is well established that patients who experience postoperative complications within 30 days of surgery have a reduced long term survival rate [1, 2]. Furthermore, even in the absence of complications there is a

⁶ Cancer Epidemiology Group, IPO Porto Research Center (CI-IPOP), Portuguese Institute of Oncology of Porto FG (IPO-Porto), Porto, Portugal

⁷ Department of Otorhinolaryngology-Head and Neck Surgery, Portuguese Institute of Oncology of Porto FG (IPO-Porto), Porto, Portugal

⁸ Experimental Pathology and Therapeutics Group of Portuguese, Institute of Oncology of Porto FG, EPE (IPO-Porto), Porto, Portugal

⁹ Surgical Oncology Department, Portuguese Institute of Oncology of Porto FG (IPO-Porto), Porto, Portugal

20–40% reduction in postoperative physical function and a significant deterioration in the quality of life after major surgery [1, 3].

It has long been accepted that individuals who have limited pre-operative physical fitness have higher rates of morbidity and mortality during their hospital stay [4, 5]. On the other hand, individuals who have better preoperative physical fitness experience less postoperative pain and have better physical functional status postoperatively [6]. For this reason, there has been a growing interest in the concept of prehabilitation which is defined as a multidimensional programme that aims to optimize physical functionality preoperatively to achieve a quicker recovery of functional status in the postoperative period [7]. Prehabilitation has a patient-centered strategy, focused on optimizing patient eligibility for surgery and improving surgical outcomes [7]. Some authors state that prehabilitation is analogous to marathon training [1]. In fact, they both require training. Similar to marathon training, prehabilitation programmes acknowledge the multidimensional aspects of preoperative preparation to include nutritional, psychological, and behavioural interventions in addition to exercise [1].

Various scoring systems have been developed to estimate the risk of perioperative morbidity and mortality and these might be valuable in the selection of patients for prehabilitation before surgery. This is particularly important in the head and neck oncology discipline where complications such as fistula development can result in the significantly extended length of stay and decreased quality of life [8].

The American Society of Anesthesiologists physical status (ASA) is perhaps the best known and widely used grading system for preoperative health of the surgical patients [9]. It relies on a subjective assessment of a patient's overall health that is based on six classes, with ASA I being a normal healthy patient, and ASA VI a brain-dead patient [9] (Appendix A).

Another scoring system is the Physiological and Operative Severity Score for the enumeration of Mortality and morbidity (POSSUM) which was initially developed by Copeland et al. and later adjusted in 1998, into Portsmouth-POSSUM (P-POSSUM) [10, 11]. It aimed to provide both retrospective and prospective analysis of the risk of mortality and morbidity of surgical patients within 30 days after surgery and to facilitate surgical audit and comparison of the performance of individual units. The P-POSSUM score includes 18 parameters divided into two components; 12 physiological and 6 operative factors, to make a minimum score of 18 and a maximum score of 136 and then converted to a percentage with a logistic regression [10, 11] (Appendix B).

More recently, the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) developed an universal surgical risk calculator

based on preoperative risk and postoperative morbidity and mortality [12, 13]. This risk calculator is an open-access online tool that accepts the input of 21 comorbidity and demographic-related, patient-specific variables, in conjunction with a surgery-specific Current Procedure Terminology code, to predict patients' risk of 12 postoperative outcomes within 30 days after surgery (Appendix C).

Likewise, the acknowledgment that postoperative pulmonary complications can contribute significantly to overall perioperative morbidity and mortality, motivated the development of The Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT) risk scale [14]. The ARISCAT score predicts the overall incidence of postoperative pulmonary complications, by assigning a weighted point score to four patient-related factors and three surgical procedure-related factors [14] (Appendix D). It finally categorizes risk as follows: Low risk (0–25 points), Intermediate risk (26–44 points) and High risk (45–123 points).

Over the years, the aforementioned risk tools have been studied in head and neck surgery with conflicting results. Therefore, we have evaluated these surgical risk calculators comparing several predicted outcomes with the observed outcomes in our population, to develop a model to estimate the degree of risk and afterwards implement prehabilitation programmes in our institution.

Material and methods

The medical records of all patients submitted to head and neck major surgery with Intermediate Care Units (IMCU) admittance for postoperative care, in a tertiary care hospital, from January 2016 to December 2017, were retrospectively reviewed. An additional cohort of patients admitted in 2018, was included for validation purposes of our risk model.

The admittance to the IMCU was due to the complexity of the surgical procedure and/or due to comorbidities of patients.

The following surgical risk calculators were used: P-POSSUM, ACS-NSQIP, ASA and ARISCAT, to predict the risk of postoperative outcomes [9–12, 14]. In this study, on-line calculator tools were used to obtain P-POSSUM scores (www.riskprediction.org.uk) and ACS-NSQIP scores (www.riskcalculator.facs.org) [13, 15]. Clinical information was manually entered into the study database. For the ACS-NSQIP risk calculator, the most relevant Current Procedural Terminology (CPT) codes were selected based on the type, extent, and attributes of the procedure. When multiple procedures were performed that could not be captured by a single code, the principal CPT code, after consultation with the surgeon, that represented the most clinically complex procedure among all procedures done during that operation,

was chosen. To maintain consistency, “Surgeon Adjustment of Risk” was not altered.

The occurrence of postoperative complications within 30 days was registered. The severity of complications was evaluated using Clavien-Dindo (none/minor if inferior or equal to grade 2 and major if equal or superior to grade 3) and ACS-NSQIP (“any complications” or “serious complications”) classifications [12, 16, 17]. The ACS-NSQIP defines “serious complications” as the presence of any of the following: cardiac arrest, myocardial infarction, pneumonia, renal insufficiency and failure, pulmonary embolism, deep venous thrombosis, return to the operating room (OR), deep incisional surgical site infections, organ space surgical site infections, systemic sepsis, unplanned intubation, urinary tract infection, and wound disruption. “Any complication” was defined as superficial incisional surgical site infections, stroke, or ventilator support > 48 h or any of the aforementioned serious complications [12]. In cases with multiple complications, the case was assigned a grade corresponding to the highest graded complication according to Clavien-Dindo and ACS-NSQIP classifications.

The hospital length of stay (LOS) and mortality (at 1 and 12-months after surgery) were also evaluated. Despite the fact that the risk instruments included in this study were designed to predict 1-month mortality, only one patient died in the first month after surgery in our sample. Therefore, the analysis of the 1-month mortality was not feasible. Although not the most recommended, we decided to analyse the 1-year mortality using the same risk tools.

The body mass index, a variable included in the ACS-NSQIP score, was also included and analysed individually in this study.

The study was approved by the Institutional Review Board and the Ethics Committee of the hospital. The protocols conformed to the ethical guidelines of the World Medical Association Declaration of Helsinki (version 2002).

Statistical analysis

Continuous variables are presented as mean \pm standard deviation and categorical variables as frequencies and percentages.

Chi-squared or Fisher’s exact tests were used to evaluate the association between two categorical variables, namely, the occurrence of postoperative complications (within 30 days after surgery) or death (within 30 days and 1 year after surgery) with different categories of ASA score and ARISCAT.

Comparisons between groups were performed, using independent samples t-tests (or Mann–Whitney) and ANOVA (or Kruskal Wallis) tests for continuous variables as appropriate. P-POSSUM and ACS-NSQIP scores were compared between patients with the occurrence of

postoperative complications within 30 days. Body mass index, P-POSSUM and ACS-NSQIP scores were compared between patients with the occurrence of death within 1-year after surgery.

Survival analysis after surgery of patients with different ASA categories was performed using the Kaplan–Meier methodology.

The ability of the risk tools P-POSSUM, ACS-NSQIP and ARISCAT for predicting post-operative complications and death was assessed using Receiver Operating Characteristic (ROC) curves and estimating the area under the curve (AUC). The discrimination ability of each score was considered acceptable for $0.7 \leq \text{AUC} < 0.9$ and excellent for $\text{AUC} \geq 0.9$ [18]. The 95% confidence intervals (CI) are reported.

Binary logistic regression analysis was performed to evaluate the association of each risk score with the occurrence of major complications (Clavien-Dindo classification ≥ 3 and Serious complication according to ACS-NSQIP score). An univariable model was first built for each score. Then a multivariable model was built using a stepwise variable selection algorithm which retained in the final model only the significant variables. The prediction ability of the final panel of variables was also evaluated using ROC curves and the corresponding AUC.

Significance was settled for $p < 0.05$.

Statistical analysis was performed using the SPSS software, version 22.

Results

Demographic and clinical characterization

There were 128 patients submitted to surgery in the first period of study. Of these, 106 (82.8%) were male and 22 (17.2%) were female. The mean (standard deviation) age of patients was 62.6 ± 10.1 years (min–max 41–91 years).

The mean LOS, in days, was 2.1 ± 2.2 (min–max 1–18 days) in the IMCU and globally in the Institution was 22.0 ± 19.1 (min–max 2–113 days).

Most patients were admitted for elective surgery (93%, $n = 119$). Six patients needed unplanned Intensive Care Unit Admission (4.7%, $n = 6$).

Most patients were independent (84.4%, $n = 108$) or partially dependent (9.4%, $n = 12$) with only eight patients being totally dependent (6.3%) for the activities of daily living.

Around 23% of our patients presented *Diabetes Mellitus* ($n = 29$). The majority of these ($n = 26$) was under oral medication and only three patients were performing insulin treatment. Almost 60% ($n = 76$) presented hypertension requiring medication and 21% ($n = 27$) had a history of congestive heart failure 30 days prior to surgery. Over 40% ($n = 53$)

were active smokers in the previous year and 28.1% ($n = 36$) had severe chronic obstructive pulmonary disease history. Around 30% ($n = 40$) presented dyspnea within 30 days of surgery. None had a history of dialysis or systemic sepsis within 48 h prior to the procedure.

The most common procedures which required admittance in the IMCU in our series were laryngectomy (16.4%), glossectomy with surgical excision of the floor of mouth (13.3%), COMMANDO operation (7.8%) and pharyngolaryngectomy (6.3%). The surgical procedures of our sample are summarized in Table 1.

Fifty-eight patients experienced one or more postoperative complication within 30 days after surgery (45.3%) (Fig. 1). The most common postoperative complication was surgical infection (31.0%; $n = 18$), followed by respiratory infection or insufficiency (25.9%, $n = 14$). Respiratory insufficiency was defined as postoperative $\text{PaO}_2 < 60$ mmHg and/or $\text{PaCO}_2 > 50$ mmHg. When we analyse the complications, according to Clavien-Dindo classification, most patients presented grade II (i.e., requires pharmacological treatment, 28.1%), grade IIIb (i.e., requires surgical treatment under general anesthesia, 6.3%) and IVa (i.e., life threatening complication-single organ dysfunction, 3.9%). Globally, the incidence of major postoperative complications according to Clavien-Dindo Classification (grade equal or superior to grade III) was 14.8% ($n = 19$). Twenty-two patients (17.2%) presented a serious complication according to the ACS-NSQIP classification.

Thirty-nine patients died during the study period: one died in the first month after surgery (overall mortality at 1-month of 0.8%) and 29 patients died in the first year (22.7%).

We analysed the effect of the nutritional status in the 1-year mortality. Patients with death in the first year after surgery had a significantly lower mean body mass index (22.9 ± 3.9 Kg/m²) than patients who survived (25.3 ± 4.9 Kg/m²) ($p = 0.006$).

Five patients (3.9%) had prior chemoradiation, six patients (4.7%) had a history of radiotherapy only and seven patients (5.5%) had a history of chemotherapy only. There was no association between history of prior chemo and/or radiotherapy and the occurrence of postoperative complications ($p > 0.05$). On the other hand, prior radiotherapy was associated with 1-year mortality ($\chi^2 = 6.22$; $p = 0.010$).

The cohort of patients subsequently added to the original sample for validation purposes consisted of 45 patients and was similar to the training cohort considering major demographic and clinical characteristics.

ASA physical status

Most of the patients were ASA 2 (50.8%) and 3 (47.7%), with only two patients being ASA 4 (1.6%). Of the 58

patients who presented postoperative complications, 25 were ASA 2, 31 were ASA 3 and 2 were ASA 4. Of the 18 patients presenting major complications according to the Clavien-Dindo classification (grade ≥ 3), the majority were ASA 3 ($n = 13/61\%$). No significant association between ASA score and the occurrence of complications was found ($p = 0.111$).

A higher ASA score was positively associated with 1-year mortality ($p = 0.005$). In fact, in the first year after surgery, 16.2% of the ASA two patients, 34.4% of the ASA three patients and 100% of the ASA four patients, died.

P-POSSUM

The overall P-POSSUM predicted morbidity rate was 47.93 (± 23.93)% and the predicted mortality rate was 6.42 (± 11.36)%. This means that the P-POSSUM scoring system predicted that 61 patients would develop postoperative complications (47.9%), comparing with the 58 (45.3%) patients who did effectively had complications and that eight patients were expected to die (6.42%), but only one patient did effectively died in the first month after surgery (0.8%).

The physiological and operative severity scores were compared between patients with and without death 1-year after surgery, and no significant differences between groups were observed ($p = 0.100$ and $p = 0.253$, respectively) (Table 2). Also, there were no differences between these groups considering the predicted mortality rate ($p = 0.116$).

The predicted morbidity rate was significantly higher in patients who died in the first year after surgery ($p = 0.049$) (Table 2). P-POSSUM mortality discrimination ability was only reasonable according to the analysis of ROC curves (AUC 0.60; 95% CI 0.49–0.71).

The physiological and operative severity scores were compared among patients with and without complications after surgery, and no differences between groups were observed ($p = 0.499$ and $p = 0.698$, respectively) (Table 2). In addition, there were no differences between these groups considering the predicted morbidity rate ($p = 0.675$). P-POSSUM discrimination ability for serious complications according to ACS-NSQIP classification and major complications according to Clavien-Dindo classification was only reasonable (AUC 0.63; 95% CI 0.48–0.77 and AUC 0.69; 95% CI 0.58–0.81, respectively).

ACS-NSQIP

Patients who developed complications in the postoperative period presented a higher predicted ACS-NSQIP risk of complications pre-operatively, specifically the risk of severe complication ($p = 0.001$), any complication ($p < 0.001$), risk of surgical site infection ($p = 0.030$), risk of pneumonia ($p = 0.170$) and risk of cardiac complications ($p = 0.040$) (Table 3).

Table 1 Head and Neck procedures with Intermediate Care Unit admittance

Most common procedures with ICU admittance	<i>n</i>	Relative Frequency (%)	Associated procedures	
Laryngectomy	21	16.4	Neck dissection	16
	Partial 7		Reconstruction (pedicled flap)	1
	Total 14		Thyroidectomy	2
Pharyngolaryngectomy	8	6.3	Neck dissection	7
			Reconstruction (pedicled flap)	2
			Thyroidectomy	1
Glossectomy with surgical excision of floor of mouth	17	13.3	Mandibulectomy	10
			Neck dissection	15
			Reconstruction (pedicled flap)	7
			Reconstruction (free flap)	2
			Tracheostomy	10
Glossectomy	15	11.7	Reconstruction (pedicled flap)	4
			Reconstruction (free flap)	1
			Tracheostomy	1
			Neck dissection	6
			Pharyngectomy	1
COMMANDO operation	10	7.8	Reconstruction (pedicled flap)	7
			Tracheostomy	2
Mandibulectomy	12	9.4	Reconstruction	9
			Tracheostomy	3
			Neck dissection	2
			Surgical excision of floor of mouth	1
			Partial pharyngectomy	11
Maxilectomy	7	5.5	Tracheostomy and glossectomy	1
			Neck dissection	1
			Reconstruction (pedicled flap)	1
			Orbital exenteration	1
			Orbital enucleation	1
Total parotidectomy	1	0.8	Pharyngectomy and glossectomy	1
			Reconstruction (pedicled flap)	1
Endoscopic resection of malignant sinonasal tumours	2	1.6		
Excision of malignancy from the lip with reconstruction	5	3.9	Neck dissection	1
Revision of hemostasis from the surgical wound	3	2.3		
Suspension microlaryngoscopy with biopsy	6	4.7	Neck dissection	1
Total thyroidectomy	1	0.8	Neck dissection	1
Partial rhinectomy	1	0.8	Neck dissection and Reconstruction with local flap	
Neck dissection	4	3.1	Reconstruction (pedicled flap)	1
Tracheostomy	1	0.8		
Deep neck abscess drainage	1	0.8		
Trans-oral epiglottectomy	1	0.8		
Removal of osteosynthesis material from the mandible	1	0.8		

ACS-NSQIP did not show discrimination ability for predicting surgical site infection (AUC 0.47; 95% CI 0.29–0.65) and pneumonia (AUC 0.59; 95% CI 0.40–0.78) and had a reasonable accuracy for cardiac complications in our sample (AUC 0.65; 95% CI 0.48–0.82).

The mean predicted ACS-NSQIP risk of death was significantly higher in patients who died (4.58 ± 8.12 versus 1.13 ± 1.93 , $p = 0.020$). The ACS-NSQIP discrimination ability for predicting the risk of death in the postoperative period was acceptable (AUC 0.74; 95% CI 0.65–0.84).

Fig. 1 Postoperative complications in patients submitted to Head and Neck surgery with Intermediate Care Unit admittance (*n* = 58 patients)

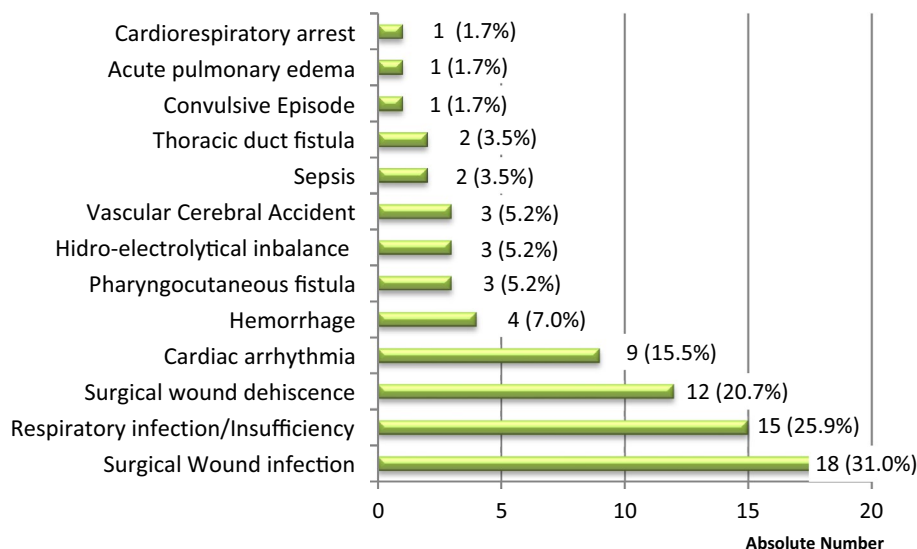


Table 2 Comparison of P-POSSUM scores among patients considering mortality and morbidity parameters

P-POSSUM predicted risk		Physiological Score P-P	Operative severity score	% mortality P-POSSUM	% morbidity P-POSSUM
Death 1-year after surgery	No (<i>n</i> = 94)	20.5 ± 6.0	12.9 ± 3.1	5.2 ± 8.7	45.4 ± 23.1
	Yes (<i>n</i> = 34)	23.2 ± 8.3	13.6 ± 3.5	9.9 ± 16.4	54.8 ± 25.2
	<i>p</i> value	0.100	0.253	0.116	0.049
30-day post-operative complications	No (<i>n</i> = 70)	20.9 ± 6.5	13.0 ± 2.9	5.4 ± 7.5	47.1 ± 23.7
	Yes (<i>n</i> = 58)	21.7 ± 7.1	13.2 ± 3.6	7.6 ± 14.7	48.9 ± 24.3
	<i>p</i> value	0.499	0.698	0.296	0.675

Table 3 Comparison of ACS-NSQIP predicted risks among patients with the occurrence of complications and death

ACS-NSQIP predicted risks	Postoperative Complication		<i>p</i> value	AUC (CI)
	No (<i>n</i> = 70)	Yes (<i>n</i> = 58)		
Severe Complication	17.3 ± 10.6	24.3 ± 11.6	<0.001	0.69 (0.58–0.81)
Any Complication	19.0 ± 12.3	27.6 ± 12.3	<0.001	0.59 (0.48–0.70)
Surgical Site Infection	5.8 ± 5.9	8.2 ± 6.2	0.025	0.47 (0.29–0.65)
Pneumonia	3.7 ± 3.4	5.3 ± 4.0	0.170	0.59 (0.40–0.78)
Cardiac complications	0.7 ± 0.8	1.5 ± 1.9	0.040	0.65 (0.48–0.82)
	Death 1-year after surgery		<i>p</i> value	AUC (CI)
	No (<i>n</i> = 94)	Yes (<i>n</i> = 34)		
Risk of death	1.13 ± 1.93	4.58 ± 8.12	0.020	0.74 (0.65–0.84)

ARISCAT

The mean ARISCAT score in our sample was 16.2 ± 10.3, ranging from 0 to 39. Therefore, there were no patients in the preoperative high-risk category (score ≥ 45).

Patients who developed pulmonary complications had significantly higher (24.1 ± 9.7) preoperative ARISCAT

score than patients without (15.1 ± 9.9) this complication (*p* = 0.001).

In 41 patients with intermediate preoperative ARISCAT score, 12 (29.3%) developed pulmonary complications (Table 4). Patients with low-risk scores had lower rates of pulmonary complications (4.6%) than those in the intermediate-risk group (29.3%), with statistically significant differences (*p* = 0.001) (Table 4).

Table 4 Comparison of ARISCAT predicted risk scores among patients with the occurrence of respiratory complications and death after surgery

Outcome	ARISCAT predicted risk		<i>p</i> value	
	Low risk (<26 points)	Intermediate risk (26–44 points)		
Respiratory Complication	Yes (<i>n</i> = 16)	4	12	0.001
	No (<i>n</i> = 112)	83	29	
Death 1-year after surgery	Yes (<i>n</i> = 34)	22	12	0.905
	No (<i>n</i> = 94)	65	29	

Table 5 Estimated Odds Ratio of Serious Complications (ACS classification) using uni- and multivariable binary logistic regression models

Variable	Univariable OR (95% CI)	Multivariable OR (95% CI)
ARISCAT	1.07 (1.02–1.12)	1.06 (1.01–1.11)
ACS	1.05 (1.02–1.09)	1.05 (1.01–1.09)
P-POSSUM	1.00 (0.88–1.13)	–
ASA		
	2 1	
	3 2.01 (0.87–4.62)	–

The discrimination ability of ARISCAT score for predicting respiratory complications was acceptable (AUC 0.75; 95% CI 0.61–0.88).

ARISCAT score was not associated with death 1-year after surgery ($p = 0.905$), nor were there differences of the ARISCAT score between the individuals with and without death after surgery ($p = 0.905$). (Table 4).

Multivariable analysis

A binary logistic regression model was built to predict the occurrence of major complications within 30 days after surgery (according to ACS-NSQIP classification), considering as potential independent variables the risk tools: P-POSSUM, ACS-NSQIP, ASA and ARISCAT.

Only ACS-NSQIP and ARISCAT were found statistically significant in the multivariable model and have been included in the final model (Table 5). The occurrence of serious complications increases significantly with ACS-NSQIP score and ARISCAT score (OR = 1.05; 95% CI 1.01–1.10 and OR = 1.08; 95% CI 1.02–1.15, respectively). The AUC obtained with this model for the training set (patients admitted in the period 2016–2017) was 0.75 (95% CI 0.63–0.87) (Fig. 2a). A cut-off between low and high risk was chosen to maximize sensitivity with an acceptable specificity. For

the chosen cut-off a sensitivity of 81.8% and a specificity of 54.8% were obtained for the training set.

Given the small sample size of the original training set, we decided to append an additional dataset of 45 patients admitted in 2018, that was subsequently included for validation purposes to have a larger sample and to refit the model. A sensitivity of 82.8% and a specificity of 52.8% were obtained. The results are presented in Fig. 2b and Table 5.

No significant model was obtained considering the classification of Clavien-Dindo and the risk tools P-POSSUM, ACS-NSQIP, ASA and ARISCAT.

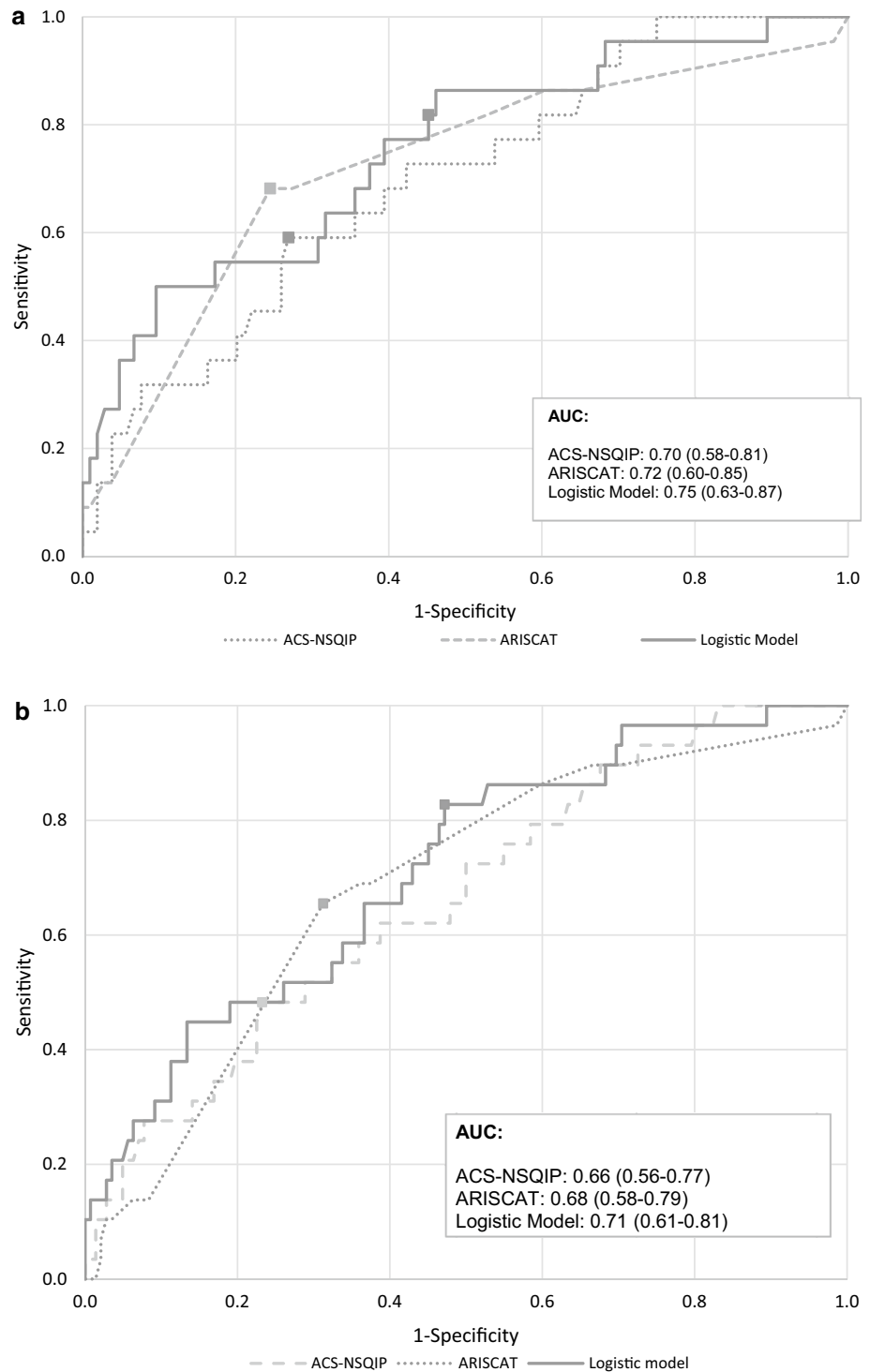
Discussion

Accurate estimates of postoperative complication risks are undoubtedly important to patients, caregivers, and clinicians. To our knowledge, this is the first study performing a comprehensive analysis of the four most important risk scores in the surgical community, ASA, P-POSSUM, ACS-NSQIP and ARISCAT in head and neck procedures. We have studied the theoretically considered high-risk patients, by including patients admitted in the IMCU for postoperative care due to anesthetic or surgical risk. Given the questionable value of the risk scores individually evaluated in this study, we performed a multivariate analysis combining them and designed a new risk tool for our institution which better predicts the risk of serious complications in our patients.

The topic of surgical complications and mortality is a relatively delicate and difficult to report subject. However, our crude 30-day morbidity and mortality were broadly consistent with those in other published reports, at 45.3% and less than 1%, respectively [19–21]. Globally, the incidence of major postoperative complications according to the Clavien-Dindo classification (grade equal or superior to grade 3) was 14.8% ($n = 19$) which is also in line with other reports [19, 20].

The ASA classification was established in the 1940s and has since undergone multiple revisions [22]. Today, the ASA classification is universally recorded for any surgical case performed under anesthesia. While not intended to predict risk, increasing ASA class has been associated with increased perioperative morbidity and mortality [9, 22, 23]. It is also included in other surgical risk calculators, as ACS-NSQIP [12]. In our sample, the ASA score was not associated with the occurrence of complications ($p = 0.111$). However, the association between the ASA score and postoperative complications has been reported in the literature in many surgical specialties, including Otorhinolaryngology. Hackett and colleagues confirmed the potential of ASA score for risk stratification not only for medical complications but also for mortality after surgery. The same study

Fig. 2 a Receiver-operating characteristic curves and performance metrics for our algorithm in predicting the occurrence of serious complications in the training set of patients ($n = 128$). **b** Receiver-operating characteristic curves and performance metrics for our algorithm in predicting the occurrence of serious complications in a larger sample of patients (original training set plus an additional dataset of 45 patients, $n = 173$)



reported that patients with greater ASA classes developed substantially higher rates of postoperative medical complications and mortality when compared to patients in lower ASA classes [22]. In our study, a higher ASA score was positively associated with 1-year mortality ($p = 0.005$) and a lower survival time was observed in patients with higher ASA grades. In spite of the ASA classification being simple and widely

understood, a great variability between assessments has been reported as it relies on a subjective evaluation [24]. Also, it does not describe individual patient risk and cannot, therefore, account for a surgical procedure, preoperative optimization or individual differences in postoperative care setting [24, 25]. Nevertheless, the ASA classification system is a simple, valid metric for determining the risk

of complications and mortality, being extremely useful for clinical communication between colleagues. However, we agree that for more detailed case analysis, for auditing, risk management and funding allocation purposes, ASA classification is insufficient [24].

The P-POSSUM system has been recommended as an accurate method in evaluating surgical outcomes and allowing direct comparisons, despite distinct patterns of referral and populations. It has the advantage of being simple and including variables that are easy to collect. It considers the physiological condition of the patient at admission and the severity of the surgical procedure to predict the rates of morbidity and mortality. It has been already evaluated in head and neck surgery with controversial results. In our study, P-POSSUM discrimination ability for mortality and morbidity was only reasonable according to the analysis of ROC curves. Also, there were no differences between groups with and without complications within 30 days after surgery or death until 1-year after surgery considering the physiological and operative severity scores and also considering the predicted morbidity and mortality rates, respectively. In our sample, P-POSSUM overpredicted 30-day mortality, as a total of eight deaths were predicted but only one occurred. Other authors reported that P-POSSUM overpredicted the occurrence of death and had no relevance in predicting mortality in a population undergoing head and neck surgery [20, 26]. However, the analysis of mortality in our study is limited by the small number of patients who died in a relatively reduced sample. Also, P-POSSUM slightly overpredicted morbidity, as a total of 61 patients were predicted to develop postoperative complications (47.9% morbidity rate) but only 58 effectively did (45.3%). Other colleagues reported divergent results in head and neck surgery for P-POSSUM. Ribeiro and Kowalski used the original POSSUM score to predict complications in 530 patients having orofacial surgery for cancer [21]. The findings in this study mirror those of Griffiths et al. who, in a similar population, reported that POSSUM under-predicted morbidity in the low to moderate risk categories [21, 26]. More recently, Tighe audited 360 operations in 245 patients submitted to orofacial surgery for cancer and concluded that P-POSSUM under-predicted morbidity in the low-risk groups and over-predicted mortality in all risk groups [20]. Unfortunately, in our study, P-POSSUM has revealed itself not suited to predict outcomes in head and neck surgery. Indeed, the variables that comprise the P-POSSUM scoring system were designed for a general surgical population, and variables like “peritoneal soiling” and the “Glasgow Coma Scale” are probably not relevant to head and neck surgery. Remarkably, poor nutritional status is another factor shown to be significantly associated with postoperative mortality in our sample. Considering the similar results of other studies in the head and neck cancer population, we consider that the inclusion of the variable “nutritional

status” in the P-POSSUM score should be equated [26, 27]. Griffiths et al. also suggested that radiotherapy and previous surgery were both significant for the development of postoperative complications and were worthy of inclusion in the original Possum score for head and neck surgery [26]. In our sample, previous radiotherapy was not associated with the occurrence of postoperative complications, possibly due to our small sample size. On the other hand, prior radiotherapy was associated with 1-year mortality.

Considering the ACS-NSQIP calculator in our patients, one may conclude that it had an insufficient accuracy for predicting complications. In spite of the existence of significant differences for the ACS-NSQIP predicted risks between groups with and without specific complications, the discrimination ability for predicting the most common complications (surgical site infection, pneumonia and cardiac complications) is nearly reasonable or worse. Regarding mortality, although ACS-NSQIP was not designed to predict 1-year mortality, it showed an acceptable discrimination ability for predicting the risk of death in the postoperative period in our sample. Also, the predicted ACS-NSQIP risk of death was significantly higher in patients who died ($p = 0.020$). There have been studies previously showing that the ACS-NSQIP database may not adequately predict postoperative complications in complex surgical procedures. Prasad et al. concluded, in a cohort of 98 patients, that ACS-NSQIP risk calculator was a poor predictor of perioperative complications following major head and neck operations [28]. Other two recent studies pertaining to microvascular head and neck reconstruction showed poor prediction performance of ACS-NSQIP [29, 30]. In addition, Schneider et al. have added total laryngectomy to the list of complex procedures for which the NSQIP risk calculator may not be as accurate in predicting postoperative adverse events [31]. More recently, Vosler et al. evaluated 131 patients and reported efficacy of ACS-NSQIP surgical calculator for predicting postoperative complications in head and neck oncology surgeries that do not require microvascular reconstruction [8]. The same authors suggest that this surgical calculator can be improved by the inclusion of several factors important for risk stratification in head and neck oncology, namely, the performance of free flap reconstruction. Furthermore, other study concluded that ACS-NSQIP calculator may be insufficiently calibrated to accurately predict postoperative complication risk for patients previously exposed to chemotherapy undergoing salvage laryngectomy [32]. The same authors, advised caution when estimating postoperative risk among patients undergoing salvage procedures, especially those of older age, poorer functional status, and those requiring neck dissection [32]. There is increasing recognition of the importance of a specialty-specific ACS-NSQIP and were many the studies proving that this scale is currently inadequate for head and neck surgery. Indeed, efforts to develop

disease- and procedure-specific preoperative, intraoperative, and postoperative variables specific to head and neck surgery have been undertaken [8, 32]. We agree with other authors who state that an essential first step in mitigating the inaccuracy of ACS-NSQIP for head and neck procedures, is the combination of CPT codes [28]. In fact, many of the operations performed included multiple high-risk procedures done concurrently, and the final CPT code attributed was not truly representative of the actual complexity of each surgery.

Our study demonstrated that the ARISCAT score was a reliable risk calculator for predicting postoperative respiratory complications. Other studies had conflicting results regarding the value of ARISCAT scale in head and neck patients. Wood et al. observed poor predictive performance of ARISCAT in a cohort of 794 patients admitted for major head and neck surgery at their institution [33]. These discrepancies in the literature concerning the accuracy of ARISCAT scale in head and neck patients might be explained by a number of factors. First, the ARISCAT score was validated in a large surgical population in which a very reduced fraction were head and neck surgeries. In addition, the variable “surgical site” is not as important for major head and neck surgery as it is for other surgical specialities where chest wall and diaphragm manipulation might occur and significantly contribute to the risk of postoperative pulmonary complications. Furthermore, the distortion of the upper airway, the frequent use of tracheostomy or surgical resections that might alter the upper aerodigestive anatomy and imply the potential risk for aspiration are unique features in this subgroup of patients [33]. Nevertheless, we consider that ARISCAT score might be useful to stratify risk when advising patients before surgery and, to identify patients most likely to benefit from risk-reduction interventions.

We managed to design a new risk tool for our institution which better predicts the risk of serious complications in our patients. It should be emphasized that apparently modest predictive values for the risk scores and for our regression model that would not be acceptable in diagnostic tests, where accuracy is essential, may still be very helpful in prognostic models, which are used in preoperative visits to predict a complication risk higher than average. Therefore, these results allow us to define our model as a tool with moderate to good clinical utility to estimate the risk of complications. Our next goal is to implement prehabilitation programmes, including its four dimensions, in our high-risk patients undergoing major head and neck surgery.

There are several limitations to this study worthy of discussion. First, it is a retrospective study with a low population number from a single institution. Secondly, we have used general surgical risk calculators which are not yet adapted to head and neck procedures. Other important limitation to mention is that our final model was developed from the compilation of other risk models already

existent which results in a high number of variables to be collected and computed. Nevertheless, we have created a validated risk tool adapted to our population which successfully selects high-risk patients who may require additional care to preempt complications or to resolve them after they occur.

Further research is needed to understand whether additional patient attributes should be supplemented in the calculator to improve its predictive value.

Despite all that has been said, and recognizing the valuable benefits of the risk tools we have analysed, risk prediction models cannot take into account subtleties in patients, their diseases or the technical difficulties of every single operation, the individual performance of the surgeon and the fulfilment of good care standards of every institution. Therefore, whenever necessary, clinical judgment should override any predicted outcome of any risk scales.

Author contributions ASM and LLS designed the study. AF, CS and ASM collected and analysed the data. ASM takes full responsibility for the integrity of the data presented. All authors contributed to the analysis and interpretation of the data for the work. JR, LA and FM made valuable contributions to the statistical analysis. EM and LLS were also responsible for the study supervision. ASM was responsible for the drafting of the manuscript and all the authors for revising it for important intellectual content.

Funding This project has received funding from the FCT (Fundação para a Ciência e a Tecnologia)—Project IPO Score- D-SAIPA/DS/0042-2018. This study was presented as an oral communication at the “69 Congreso Nacional de la Sociedad Española de Otorrinolaringología y Cirugía de Cabeza y Cuello”, in October, in Madrid, Spain.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors. This study was approved by the Ethical Committee of our institution.

Appendices

Appendix A- The American Society of Anesthesiologists physical status

ASA I	Normal healthy patient
ASA II	Patients with mild systemic disease

ASA III	Patients with severe systemic disease
ASA IV	Patients with severe systemic disease that is a constant threat to life
ASA V	Moribund patients who are not expected to survive without the operation
ASA VI	A declared brain-dead patient whose organs are being removed for donor purposes

Preoperative patient and surgical NSQIP risk outcomes inputs

Congestive heart failure in 30 days prior to surgery
Dyspnea
Current smoker within 1 year
History of severe COPD
Dialysis
Acute renal failure
Body mass index calculation (weight and height)

Appendix B- P-POSSUM SCORE

Physiologic score	Operative score
Age (Years)	Operative severity Score
Cardiac signs	Multiple procedures
Respiratory history	Total blood loss
Systolic blood pressure (mm Hg)	Peritoneal soiling
Pulse (beats/min)	Presence of malignancy
Glasgow coma score	Mode of surgery
Haemoglobin g/dL	
White cell count ($\times 10^{12}/L$)	
Urea (mg/dL)	
Sodium (mmol/L)	
Potassium (mmol/L)	
Electrocardiogram	

Appendix C- ACS NSQIP risk calculator

Preoperative patient and surgical inputs	NSQIP risk outcomes
Age group	Serious complication
Sex	Any complication
Functional status	Pneumonia
Emergency case	Cardiac complication
ASA class	Surgical site infection
Wound class contamination	Urinary tract Infection
Steroid use for chronic condition	Venous thromboembolism
Ascites within 30 days prior to surgery	Renal failure
Systemic sepsis within 48 h prior to surgery	Readmission
Ventilator dependent	Return to OR
Disseminated cancer	Death
Diabetes	Discharge to nursing or Rehab facility
Hypertension requiring medication	Predicted length of hospital stay

Appendix D—The seven ariscat risk predictors

1. Age (year) (≤ 50 ; 51–80, > 80 year)
2. Preoperative Spo₂ (≥ 96 , 91–95%, $\leq 90\%$)
3. Respiratory infection in the last month
4. Preoperative anemia (Hb ≤ 10 g/dl)
5. Surgical incision (Peripheral; upper abdominal, intrathoracic)
6. Duration of surgery (h)
7. Emergency procedure

References

1. Wynter-Blyth V, Moorthy K (2017) Prehabilitation: preparing patients for surgery. *BMJ* 358:9–10
2. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ (2005) Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg* 242(3):326–343
3. Lawrence VA, Hazuda HP, Cornell JE, Pederson T, Bradshaw PT, Mulrow CD et al (2004) Functional independence after major abdominal surgery in the elderly. *J Am Coll Surg* 199(5):762–772
4. Cabilan CJ, Hines S, Munday J (2015) The effectiveness of prehabilitation or preoperative exercise for surgical patients: a systematic review. *JBI Database Syst Rev Implement Rep* 13(1):146–187
5. Vukomanović A, Popović Z, Durović A, Krstić L (2008) The effects of short-term preoperative physical therapy and education on early functional recovery of patients younger than 70 undergoing total hip arthroplasty. *Vojnosanit Pregl* 65(4):291–297
6. Walther C, Möbius-Winkler S, Linke A, Bruegel M, Thiery J, Schuler G et al (2008) Regular exercise training compared with percutaneous intervention leads to a reduction of inflammatory markers and cardiovascular events in patients with coronary artery disease. *Eur J Prev Cardiol* 15(1):107–112
7. Banugo P, Amoako D (2017) Prehabilitation. *BJA* 17(12):401–405
8. Vosler PS, Orsini M, Enepekides DJ, Higgins KM (2018) Predicting complications of major head and neck oncological surgery: an evaluation of the ACS NSQIP surgical risk calculator. *J Otolaryngol Head Neck Surg* 47(1):1–10
9. Daabiss M (2011) American society of anaesthesiologists physical status classification. *Indian J Anaesth* 55(2):111–115
10. Copeland GP, Jones D, Walters M (1991) POSSUM : a scoring system for surgical audit. *Br J Surg* 78:356–360

11. Prytherc D, Whiteley M, Higgins B, Weaver P, Prout W, Powell S (1998) POSSUM and Portsmouth POSSUM for predicting mortality. *Br J Surg* 85(9):1217–1220
12. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmieciak TE, Ko CY et al (2013) Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aide and informed consent tool for patients and surgeons. *J Am Coll Surg* 217(5):833–842
13. ACS- NSQIP. ACS NSQIP surgical risk calculator. [cited 2018 Jun 1]. Available from: <https://www.riskcalculator.facs.org/RiskCalculator/>
14. Canet J, Paluzie G, Valle J, Castillo J, Ph D, Sabate S (2010) Prediction of postoperative pulmonary complications in a population-based surgical cohort. *Anesthesiology* 113(6):1338–1350
15. Smith J, Tekkis P. Risk modelling in surgery, Healthcare Sport. [cited 2018 Jun 1]. Available from: www.riskprediction.org.uk
16. Dindo D, Demartines N, Clavien PA (2004) Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 240(2):205–213
17. Clavien PA, Barkun J, De Oliveira ML, Vauthey JN, Dindo D, Schulick RD et al (2009) The clavien-dindo classification of surgical complications: five-year experience. *Ann Surg* 250(2):187–196
18. Hosmer D, Lemeshow S (2000) *Applied logistic regression*, 2nd edn. Wiley, Hoboken
19. Ferrier MB, Spuesens EB, Le Cessie S, Baatenburg De Jong RJ (2005) Comorbidity as a major risk factor for mortality and complications in head and neck surgery. *Arch Otolaryngol Head Neck Surg*. 131(1):27–32
20. Tighe DF (2011) Utility of a generic risk prediction score in predicting outcomes after orofacial surgery for cancer. *Br J Oral Maxillofac Surg* 49(4):281–285
21. de Cássia Braga Ribeiro K, Kowalski L (2003) APACHE II, POSSUM and ASA scores and the risk of perioperative complications in patients with oral or oropharyngeal cancer. *Arch Otolaryngol Head Neck Surg* 129:739–745
22. Hackett NJ, De Oliveira GS, Jain UK, Kim JYS (2015) ASA class is a reliable independent predictor of medical complications and mortality following surgery. *Int J Surg* 18:184–190
23. Sidi A, Lobato EB, Cohen JA (2000) The American Society of Anesthesiologists' physical status: category V revisited. *J Clin Anesth* 12(4):328–334
24. Mak PHK, Campbell RCH, Irwin MG (2002) The ASA physical status classification: Inter-observer consistency. *Anaesth Intensive Care* 30(5):633–640
25. Scott S, Lund JN, Gold S, Elliott R, Vater M, Chakrabarty MP et al (2014) An evaluation of POSSUM and P-POSSUM scoring in predicting post-operative mortality in a level 1 critical care setting. *BMC Anesthesiol* 14(1):1–7
26. Griffiths H, Cuddihy P, Davis S, Parikh S, Tomkinson A (2002) Risk-adjusted comparative audit. Is Possum applicable to head and neck surgery? *Clin Otolaryngol Allied Sci* 27(6):517–520
27. Linn BS, Robinson DS, Klimas NG (1988) Effects of age and nutritional status on surgical outcomes in head and neck cancer. *Ann Surg* 207(3):267–273
28. Prasad KG, Nelson BG, Deig CR, Schneider AL, Moore MG (2016) ACS NSQIP risk calculator: an accurate predictor of complications in major head and neck surgery? *Otolaryngol Head Neck Surg (United States)* 155(5):740–742
29. Ma Y, Laitman BM, Patel V, Teng M, Genden E, DeMaria S et al (2019) Assessment of the NSQIP surgical risk calculator in predicting microvascular head and neck reconstruction outcomes. *Otolaryngol Head Neck Surg (United States)* 160(1):100–106
30. Arce K, Moore EJ, Lohse CM, Reiland MD, Yetzer JG, Ettinger KS (2016) The American College of surgeons national surgical quality improvement program surgical risk calculator does not accurately predict risk of 30-day complications among patients undergoing microvascular head and neck reconstruction. *J Oral Maxillofac Surg* 74(9):1850–1858
31. Schneider AL, Deig CR, Prasad KG, Nelson BG, Mantravadi AV, Brigance JS et al (2016) Ability of the national surgical quality improvement program risk calculator to predict complications following total laryngectomy. *JAMA Otolaryngol Head Neck Surg* 142(10):972–979
32. Cao A, Khayat S, Cash E, Nickel C, Gettelfinger J, Tennant P et al (2018) ACS NSQIP risk calculator reliability in head and neck oncology: the effect of prior chemoradiation on NSQIP risk estimates following laryngectomy. *Am J Otolaryngol Head Neck Med Surg* 39(2):192–196
33. Wood CB, Shinn JR, Rees AB, Patel PN, Freundlich RE, Smith DK et al (2019) Existing predictive models for postoperative pulmonary complications perform poorly in a head and neck surgery population. *J Med Syst* 43(10):312

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.