DR. JOHN R. DE ALMEIDA (Orcid ID : 0000-0002-1546-5033)

DR. CHRISTOPHER W NOEL (Orcid ID : 0000-0002-6625-0809)

DR. DAVID FORNER (Orcid ID : 0000-0002-1014-3545)

DR. EVAN MICHAEL GRABOYES (Orcid ID : 0000-0003-3766-468X)

DR. ANDREW G SHUMAN (Orcid ID : 0000-0002-9305-7860)

DR. ANTOINE ESKANDER (Orcid ID : 0000-0003-4326-1393)

DR. DOUGLAS B CHEPEHA (Orcid ID : 0000-0001-9062-3719)

Article type : Original Article

Development and Validation of a Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN) in a Scarce Resource Setting: Response to the COVID-19 Pandemic

John R. de Almeida[1,2*], MD, MSc, Christopher W. Noel[1,2*], MD, David Forner[1,2,3*], MD, Han Zhang[4], MD, Anthony C. Nichols[5], MD, Marc A. Cohen[6], MD, Richard J. Wong[6], MD, Caitlin McMullen[7], MD, Evan M. Graboyes[8], MD, MPH, Vasu Divi[9], MD, Andrew G. Shuman[10], MD, Andrew J. Rosko[10], MD, Carol M. Lewis[11], MD, MPH, Ehab Y. Hanna[11], MD, Jeffrey Myers[11], MD, PhD, Vinidh Paleri[12] , MBBS, MS, Brett Miles[13], DDS, MD, Eric Genden[13], MD, Antoine Eskander[14], MD, ScM, Danny .J Enepekides[14], MD, MPH Kevin M. Higgins[14], MD, MSc, Dale Brown[1], MD, Douglas B. Chepeha[1], MD, MScPH, Ian J. Witterick[15], MD, MSc, Patrick J. Gullane[1] , MD, Jonathan C. Irish[1], MD, MSc, Eric Monteiro[15], MD, MSc, David P. Goldstein[1], MD, MSc, Ralph Gilbert[1], MD

*contributed equally

1. Department of Otolaryngology-Head and Neck Surgery/Surgical Oncology, University Health Network/Princess Margaret Cancer Center/University Health Network, University of Toronto, Toronto, Canada

2. Institute of Health Policy, Management & Evaluation Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
3. Division of Otolaryngology – Head & Neck Surgery, Dalhousie University, Halifax, NS, Canada
4. Division of Otolaryngology – Head & Neck Surgery, McMaster University, Hamilton, ON, Canada
5. Department of Otolaryngology – Head & Neck Surgery, Western University, London, ON, Canada
6. Department of Otolaryngology – Head & Neck Surgery, Memorial Sloan Kettering Cancer Center, New York, New York, USA
7. Head & Neck -Endocrine Oncology, Moffitt Cancer Center, Tampa, FL, USA
8. Department of Otolaryngology – Head & Neck Surgery, Medical University of South Carolina, Charleston, SC, USA
9. Department of Otolaryngology – Head & Neck Surgery, Stanford University, Paolo Alto, CA, USA
10. Department of Otolaryngology – Head & Neck Surgery, University of Michigan, Ann Arbor, MI, USA
11. Department of Otolaryngology – Head & Neck Surgery, MD Anderson Cancer Centre, Houston, TX, USA
12. Head and Neck Unit, The Royal Marsden Hospital, London, UK
13. Department of Otolaryngology – Head & Neck Surgery, Icahn School of Medicine at Mount Sinai, New York, New York, USA
14. Department of Otolaryngology-Head and Neck Surgery, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada
15. Department of Otolaryngology-Head and Neck Surgery, Sinai Health System, University of Toronto, Toronto, ON, Canada

Corresponding Author: Dr. John R. de Almeida
8NU-883 Toronto General Hospital/ University Health Network
200 Elizabeth Street
Toronto ON, M5G 2C4
P: 416-340-3138
F: 416-946-2300
John.dealmeida@uhn.ca

Conflicts of Interest: None to disclose

Funding: None

Running Head: SPARTAN-HN

Precis: The Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN) is the first cancer surgery specific prioritization tool for use during the COVID-19 pandemic. The SPARTAN-HN algorithm is reliable and valid for the stratification of patients with head and neck cancer who require urgent cancer care in resource restricted practice environments.

Word Count: 2982

Keywords: COVID-19; head and neck cancer; waiting lists; surgical procedures; patient selection; health priorities; delivery of health care

AUTHORSHIP CONTRIBUTIONS - CNCR-20-1974.R1

John R de Almeida Study idea and design, writing, and editing.

Christopher W Noel Study design, writing, data analysis, and editing.

David Forner Study design, writing, data analysis, and editing

Anthony Nichols Data acquisition and editing

Marc Cohen Data acquisition and editing

Richard Wong Data acquisition and editing

Caitlin McMullen Data acquisition and editing

Evan Graboyes Data acquisition and editing

Vasu Divi Data acquisition and editing

Andrew G Shuman Writing, data acquisition and editing

Andrew Rosko Data acquisition and editing

Carol M. Lewis Data acquisition and editing

Ehab Hanna Data acquisition and editing

Jeffrey Myers Data acquisition and editing

Vinidh Paleri Data acquisition and editing

Brett Miles Data acquisition and editing

Eric Genden Data acquisition and editing

Antoine Eskander Data acquisition and editing

Danny J Enepekides Data acquisition and editing

Kevin M Higgins Data acquisition and editing

Dale Brown Data acquisition and editing

Douglas Chepeha Data acquisition and editing

Ian Witterick Data acquisition and editing

Patrick J. Gullane Data acquisition and editing

Jonathan C Irish Data acquisition and editing

Eric Monteiro Data acquisition and editing

David P Goldstein Data acquisition and editing

Ralph Gilbert Study idea and design, data acquisition and editing.

Abstract

Background

In the wake of the COVID-19 pandemic, access to surgical care for head and neck cancer (HNC) is limited and unpredictable. Determining which patients should be prioritized is inherently subjective and difficult to assess. We propose an algorithm to fairly and consistently triage patients and mitigate risk of adverse outcomes.

Methods

Two separate expert panels, a consensus panel (n=11) and validation panel (n=15), were constructed among international HNC surgeons. Using a modified Delphi process and RAND/UCLA methodology with four consensus rounds and two meetings, groupings of high, intermediate, and low priority indications for surgery were established and sub-divided. A point-based scoring algorithm was developed, the Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN). Agreement was measured during consensus and for algorithm scoring with Krippendorff's alpha (K-alpha). Rankings from the algorithm were compared with expert rankings of 12 case vignettes using Spearman's rank correlation coefficient.

Results

Sixty-two indications for surgical priority were rated. Weights for each indication ranged from -4 to +4 (scale range; min -17, max 20). Response rate for the validation exercise was 100%. The SPARTAN-HN demonstrated excellent agreement and correlation with expert rankings (K-alpha = 0.91, 95% CI 0.88-0.93; rho=0.81, 95% CI 0.45-0.95).

Conclusions

The SPARTAN-HN surgical prioritization algorithm consistently stratifies patients requiring HNC surgical care in the COVID-19 era. Formal evaluation and implementation are required.

Condensed Abstract

1. The Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN) is the first cancer surgery specific prioritization tool for use during the COVID-19 pandemic.

2. The SPARTAN-HN algorithm is reliable and valid for the stratification of patients with head and neck cancer who require urgent cancer care in resource restricted practice environments.

Lay Summary:

Many countries have enacted strict rules about the use of hospital resources during the COVID-19 pandemic. Facing delays in surgery, patients may face worse functional outcomes, stage migration, and eventual inoperability. Treatment prioritization tools have shown benefit in helping triage patients equitably with minimal provider cognitive burden. This study sought to develop the first cancer specific surgical prioritization tool for use in the COVID-19 era, the SPARTAN-HN. This algorithm consistently stratifies patients requiring head and neck cancer surgery in the COVID-19 era and provides evidence for the initial uptake of the SPARTAN-HN.

Introduction

On March 11, 2020, the World Health Organization declared a global pandemic due to the novel coronavirus, SARS-CoV-2, and the resulting coronavirus disease (COVID-19)[1]. As a result, in many jurisdictions, operating room capacity has been limited to only emergent or urgent surgical procedures[2]. Several advisory bodies have issued recommendations to safeguard access to oncologic surgery while still acknowledging that treatment delays may be necessary. The American College of Surgeons has recommended postponing elective surgery, including for low risk cancers, while recommending that other urgent cancer surgeries proceed[3,4]. Cancer Care Ontario has issued similar guidance recommending that hospitals include cancer surgery in their care delivery plan[5].

The time from diagnosis of head and neck cancer (HNC) to surgery is a metric with prognostic importance, with treatment delays portending poorer oncologic outcomes[6-8]. In a recent systematic review evaluating delays in diagnosis to treatment initiation, nine out of thirteen studies demonstrated a decrease in survival associated with treatment delays[6-8]. These data support the urgency of initiation of treatment for patients with HNC, but do not inform a stratification schema when operating room access is not available for all.

As a result of these new imposed constraints, difficult decisions about prioritization for cancer surgery are obligatory, and require consideration of broader scarce resource allocation principles[9]. Key among these is the need for consistency and transparency in order to achieve fairness and to avoid engendering disparities in both access and outcomes[10,11]. Prioritization on a case-by-case basis using expert clinical judgment can be logistically challenging, carry a cognitive burden, and is susceptible to the biases of practitioners.

Surgical prioritization tools or algorithms offer decision-making transparency and provide equitable and time-sensitive access to care to the patients who need it most[12,13]. While tools for surgical prioritization in COVID-19 are emerging, oncology patients have not been explicitly considered[14]. Herein, we present the development and validation of a novel algorithm (SPARTAN-HN) for prioritization of head and neck cancer surgery.

Methods

The study was granted a waiver (20-0463) from the Research Ethics Board at the University Health Network.

Participants and Setting

For instrument development, a group of 11 expert head and neck cancer surgeons (JD, DG, RG, JI, DC, DB, AE, DE, KH, EM, IW) from three institutions (University Health Network, Sinai Health Systems, and Sunnybrook Health Sciences Centre) at the University of Toronto participated in the consensus process (consensus panel). At the time of the consensus process, all three institutions were operating under significant resource constraints with limited availability of operating room time. For instrument validation, a group of five participants (JD, CN, DF, DG, EM) completed the scoring algorithm designed following the consensus process. Fifteen external head and neck surgeons (HZ, AN, RW, MC, CM, EG, VD, AS, AR, CL, EH, JM, VP, BM, EG) from 10 institutions across Canada (2), the US (7), and the UK (1) participated in a ranking exercise of clinical vignettes (validation panel).

Scope

The scope of variables considered in the prioritization algorithm was established and vetted by the consensus panel (Supplemental 1). All indications for prioritization were presented to the consensus panel using an online survey platform (Google Forms, https://docs.google.com/forms). With two exceptions, survey respondents were asked to consider each of the indications in isolation. For wait-times, panel members were asked to also consider histologic grade. Similarly, for surgical site, the panel was asked to simultaneously consider extent of surgery. Related indications were presented sequentially to facilitate pairwise comparison (e.g. stage I-II vs. stage III-IV were presented in sequence). The list of indications was pilot tested by 4 surgeons (JD, DG, EM, RG) for sensibility (readability, content validity, language, and comprehensibility).

Consensus Process

The consensus panel participated in a Delphi consensus process with four rounds of rating (Supplemental 2). The first two rounds aimed to achieve consensus on the priority grouping (high, intermediate, low). High priority was defined as an indication to proceed to surgery within 2 weeks. The second two rounds of rating involved ranking each indication (less important, neutral, more important) within their respective priority grouping. Two teleconference meetings were conducted between the first and second rounds and between the third and fourth rounds with anonymized results from the prior round presented for discussion and to address inconsistencies and misinterpretations.

A modification of the RAND/UCLA method was used to achieve consensus[15]. This methodology is typically used for determining appropriateness of an intervention but in this setting was used to determine surgical priority. We used a scale from 0-9 in rounds 1 and 2 to indicate the decision to not operate (0), or low priority (scores 1-3), intermediate priority (scores 4-6), or high priority (scores 7-9). For rounds 3 and 4, we used a scale from 1-9 to rate each indication relative to other indications within each of the priority groupings as either less important (1-3), neutral (4-6), or more important (7-9).

Consensus was determined based on RAND/UCLA criteria[15]. For the first two rounds to determine surgical priority, a hierarchical logic was adopted to determine consensus on whether or not surgery should be performed, and to then determine the priority of surgery based on the given indication. Agreement on the decision to not operate was defined as a minimum of 8 the 11 panelists rating a given indication a zero score. If there was no agreement to avoid surgery, agreement for surgical priority was then defined as no more than 3 panelists rating the indication outside the 3-point range containing the median, as per RAND/UCLA guidelines[15]. For rounds 1 and 2, any indication failing to achieve consensus was classified as intermediate priority, and for rounds 3 and 4, any indication failing to achieve consensus was classified as neutral within the priority grouping.

Development of the Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN)

The algorithm utilizes a point-based system to assign a total score based on the sum of the individual indication scores (Supplemental 3), with higher scores corresponding to higher priority. Scoring weights were based on consensus from both sets of rounds such that higher priority indications were assigned scores ranging from + 2 to + 4, intermediate priority indications ranging from -1 to +1, and low priority indications ranging from -2 to - 4. Within each priority grouping 3-point range, scores were assigned based on the consensus ratings from the third and fourth rounds. For any two patients with the same total score, the patient with the longer surgical wait-time is assigned the higher priority rank.

Clinical Vignettes

Twelve clinical vignettes were constructed (Supplemental 4) following the consensus rounds in order to validate the SPARTAN-HN. The vignettes described a variety of clinical scenarios incorporating multiple prioritization indications and additional clinical information. Experts were asked to consider only the patient-level information provided to them and not their own unique clinical and community practice environments. Twelve scenarios were selected for diversity of cases. The number was considered appropriate while avoiding excessive cognitive burden associated with ranking too many scenarios.

Statistical Analysis

<u>Agreement</u>

Agreement between raters during the Delphi process was calculated at each round and within each priority grouping using Krippendorff's alpha (K-alpha). As typical coefficients of reliability are not suitable for coded data, agreement for the rank-orders generated by 5 coders (JD, CN, DF, EM, DG) applying the SPARTAN-HN algorithm to the 12 clinical vignettes was assessed using K-alpha, calculated with 1000 bootstrap samples[16]. The K-alpha allows estimation of reliability for any number of raters and categories, and may be used when missing data is present[17].

Validity of SPARTAN-HN Algorithm

Convergent validity of the median rankings from the 5 coders of each of the 12 vignettes using the SPARTAN-HN and the expert panel rankings were assessed using Spearman's rank correlation coefficient. Strength of correlation were considered weak if rho <0.3, moderate if 0.3 – 0.7, and strong if >0.7[18].

In addition to SPARTAN-HN, a second algorithm using a decision-making flow-chart was developed (SPARTAN-HN 2). The tool and associated performance characteristics are included in the appendix (Supplemental 5).

Sample Size Considerations

For determination of an adequate sample size for the expert panel, we assumed that for model validity, strong correlation between the model rank order and expert rank order (i.e. $r \geq 0.7$), an alpha of 0.05, power of 0.8, and a non-response rate of 10%. The calculated sample size requirement was therefore 15 participants.

All analyses were two-sided and statistical significance was set at $p \leq 0.05$. Analyses were conducted using SAS University Edition 9.4 (SAS Institute, Cary, NC, USA).

Results

Establishing Consensus Priority Groupings (First Two Consensus Rounds)

After the first two rounds, the panel failed to achieve consensus for any indications that would result in a decision to not operate. More than 3 respondents indicated that they would not operate for the following indications: (1) the availability of alternative non-surgical treatment with similar prognosis (n=6, 54%), (2) poor performance status (i.e. ECOG 3-4) (n=6, 54%), and (3) very severe comorbidity indicated by non-cancer specific survival < 50% at 1 year (n=5, 45%).

In the first round, consensus was achieved for 15 (24%) indications for surgical prioritization, of which 8 (13%) were considered high priority, 4 (6%) were considered intermediate priority, and 3 (5%) were considered low priority. After review of first round results, consensus was achieved for an additional 28 (45%) indications, 25 (40%) were rated as intermediate priority and 3 (5%) were rated low priority (Table 1).

Establishing Ranking within Each Priority Grouping (Second Two Consensus Rounds)

Of 6 low priority indications, consensus for the importance of factors was achieved for 2 (33%), scenarios, both of which were deemed less important (Table 1). Of 48 intermediate priority indications, consensus for the importance of factors was achieved for 1 (2%) and 8 (17%) indications respectively. Of 8 high priority factors, consensus for the importance of factors was achieved for 4 (50%) scenarios, all of which were deemed to be more important.

Agreement during consensus rounds was weak to moderate for all 4 rounds ranging from 0.27-0.40. The agreement was similar when measured per-priority grouping where K-alpha ranged from 0.32-0.35 (Table 2).

SPARTAN-HN: Surgical Prioritization Scoring System

Priority weights for each indication ranged from -4 to +4 spanning a nine-point range and translated from the two rounds of priority groupings into three categories. Four indications were assigned a + 4 weight based on consensus that these factors were both high priority and more important (Supplemental 2, Table 1). All other high priority indications were assigned a + 3 weighted score because there was no consensus that they were either less or more important. For intermediate priority indications, a weighted score of + 1 was assigned for 7 of the 8 indications deemed to be more important by consensus. The other indication deemed to be more important (thyroid cancer with tracheal invasion) was assigned a score of + 4 because of the fact that this indication can be associated with low grade histology, which is assigned a negative weighted score. Three intermediate priority indications that were rated more important were resource use indications which are generally colinear. As such, the decision was made to assign a maximum score of + 1 for the presence of any or all of these indications. One intermediate priority indication was deemed to be less important by consensus and assigned a score of – 1. All other intermediate priority indications were assigned scores of 0. For the low priority indications, those deemed to be less important were assigned a weight of – 4 and all other indications were assigned a weight of – 3. The total scale score ranges from -17 to +20 (Figure 1).

Reliability & Validity Assessment

Agreement between 5 coders for the SPARTAN-HN was excellent (K-alpha = 0.91). Agreement between 15 expert raters was moderate (K-alpha=0.63) Convergent validity was demonstrated by strong correlation between the rank orders generated by the SPARTAN-HN and external experts (rho=0.81, 95% CI 0.45-0.95, p=0.0007). Agreement between expert rankings and SPARTAN-HN rankings for the 12 vignettes can be visualized in Figure 2.

Discussion

In the setting of the COVID-19 pandemic, where availability of operating room time as well as hospital and intensive care unit beds are limited, prioritization of surgical oncology cases is imperative to mitigate downstream adverse outcomes[19,20]. The current methodology was adopted based on expert consensus. Herein, we propose the SPARTAN-HN with an aim to provide transparency and facilitate surgical prioritization for treatment providers.

Creating COVID-19 era allocation schemas that are ethically sound is both critical and challenging. Emanuel et al. advocate four ethical principles to guide allocation of scarce resources: (1) maximizing the benefits produced by scarce resources, (2) treating people equally, (3) promoting and rewarding instrumental value, and (4) giving priority to the worst off[9]. These have been contextualized for cancer care more broadly, and are manifest in the SPARTAN-HN algorithm[21]. The high priority indications implicitly embrace an underlying premise of saving the most lives and/or preserving the most life-years. Many procedures for head and neck cancer patients are aerosol generating and increase risk to health care workers and other hospitalized patients[22]. Our process accounted for these by giving consideration to these factors in the consensus process, although indications associated with potential exposure to health care workers did not emerge as low priority indications. Indications associated with lower resource use did achieve consensus for higher importance. This may help avoid the opportunity cost of treating fewer patients with longer surgeries.

Anecdotal and institution-specific prioritization schemas for head and neck cancer patients and general otolaryngology have been suggested[2,13]. These parallel similar efforts for general surgery, cardiac surgery, and orthopedic surgery[12,13,23-28]. In many of these, patients are prioritized by scoring several criteria and summing the scores to achieve a total patient score. Many of these systems have been validated against expert rankings of surgical priority[27,28].

We employed a methodology for developing a point-based prioritization system, similar to those previously described[29]. Point-based surgical prioritization systems have been very well studied. Hansen et al. have previously proposed a methodology for developing a point-based prioritization system using the following seven steps: (1) ranking patient case vignettes using clinical judgment, (2) drafting the criteria and categories within each criteria, (3) pretesting the criteria and categories, (4) consulting with patient groups and other clinicians, (5) determining point values for criteria and categories, (6) checking the test-retest reliability and face validity, and (7) revising the points system as new evidence emerges[29]. Our approach for the development of the SPARTAN-HN was similar. However, given the relatively expedited nature of the process, we did not directly involve patients.

One method proposed for establishing priority of all indications in a point-based scoring system is known as Potentially All Pairwise Rankings of all Alternatives (PAPRIKA)[30]. In the current study, we chose to use the RAND/UCLA process instead of pairwise comparison in order to minimize computational burden. We established 62 indications for surgical prioritization which would create an enormous computational burden with pairwise comparison methodology. One problem inherent in the PAPRIKA methodology is the assumption that all indications are not equal and can be ranked. Clinically, however, certain indications may be equivalent in their priority. Furthermore, pairwise comparisons assume mutual exclusivity of each of the indications, which is not always the case. Use of the RAND/UCLA consensus process avoids the need for multiple pairwise comparison and allows for consideration of each factor in isolation. The goal of the consensus rounds was not to establish a rank order for all indications, but mainly to understand which indications result in high, intermediate, or low priority.

The SPARTAN-HN algorithm demonstrates preliminary reliability and validity. We showed good agreement between raters and the SPARTAN-HN algorithm, suggesting minimal interpretive error. Many of the high-priority indications account for some component of interpretation as raters are forced to consider imminent progression that may result in an adverse outcome. Despite the subjective decisions that must be made as part of SPARTAN-HN, agreement remained high. In fact, true inter-rater reliability is higher, as the K-alpha is a conservative measure of reliability. Other measures of reliability, such as Kendall's W coefficient of concordance, tend to overestimate reliability and cannot be applied to missing data[31]. Perhaps most importantly the SPARTAN-HN correlated highly with expert rankings. With established validity, this algorithm may be ready for preliminary clinical use although further testing against real world data in order to validate it with other cancer outcomes, such as survival, is needed.

The results of this study must be interpreted within the context of study design. Although externally validated by other surgeons across North America and United Kingdom, criteria for which consensus was achieved for prioritization was not vetted by patients, advocacy groups or other stakeholders such as medical or radiation oncologists. The latter groups represent essential providers in the multidisciplinary care of patients with HNC and may have important insight into the availability and effectiveness of non-surgical treatments[19,20]. Nonetheless, the actual prioritization of surgical waitlists remains the sole responsibility of surgeons and their practice partners. Additionally, the SPARTAN-HN algorithm is intended for making difficult prioritization decisions and is not intended to make recommendations for the time horizon in which patients should receive treatment. Instead, established guidelines should be adhered to for treatment targets. Patient wait times as they relate to those targets should be considered when using the SPARTAN-HN algorithm. Our validation process used expert opinion as the gold standard of prioritization, which is potentially biased, and reflects the opinions of surgeons practicing in academic medical centers from three resource-rich nations. Subsequently, use of the SPARTAN-HN algorithm in other geographic regions and healthcare systems requires additional investigation, as local treatment paradigms and risk factors may vary substantially.

In conclusion, we present the development and validation of a novel algorithm for prioritization of head and neck cancer surgery. Further evaluation of its implementation in varying practice settings will be obligatory. However, this study provides data to inform real-world use, as the current pandemic obviates our ability to more rigorously study the instrument prior to making necessary and difficult real-time allocation decisions.

Table 1: Prioritization indications and scores after four rounds of ranking

Low Priority Factors			Intermediate Priority Factors			High Priority Factors		
-4	- 3	- 2	- 1	0	+1	+2	+3	+4
Alternative therapy available	Poor performance status (i.e. ECOG 3, 4)		Wait time exceeded < 2 weeks for low grade histology	Wait-time not exceeded but approaching for high grade histology	Laryngeal cancer requiring partial laryngeal surgery		Wait time exceeded by < 2 weeks for high grade histology	Wait time exceeded by ≥ 2 weeks for high grade histology
Very severe comorbidity (e.g. non-cancer survival < 50% at 1 year)	Wait time not exceed but approach in 1 week for low grade histology			Wait-time exceeded for low grade histology (≥ 2 weeks)	Hypopharyngeal cancer requiring total laryngopharyngectomy		Advanced nodal disease (e.g. N3 or gross ENE)	Clinical or imaging progression (i.e. advancing stage)
	Low grade parotid malignancy			Oral cavity cancer with soft tissue resection	Nasal or paranasal sinus cancer requiring open anterior craniofacial resection		Symptomatic progression while on wait list	Potential significant functional morbidity or inoperability if tumor growth
	Thyroid cancer with nodal disease			Oral cavity cancer with bone resection	Stage III-IV disease		Previous radiotherapy	Potential moderate functional or cosmetic impairment if tumor growth
				Oral cavity cancer requiring near-total or total glossectomy	Length of surgery < 4 hours			Thyroid cancer with tracheal invasion

	Oropharyngeal cancer with transoral surgery	Hospital length of stay 1-3 days	
	Oropharyngeal cancer with mandibulotomy	No intensive care unit or step-down unit	
	Laryngeal cancer requiring total laryngectomy		
	Hypopharyngeal cancer with total laryngectomy and partial pharyngectomy		
	Nasopharyngeal cancer requiring endoscopic resection		
	Nasopharyngeal cancer requiring maxillotomy		
	Nasal or paranasal sinus cancer requiring endoscopic resection		
	Advanced skin cancer requiring skin resection and regional flap reconstruction		
	Advanced skin cancer requiring free flap reconstruction		
	High grade parotid malignancy		
	Temporal bone malignancy		
	Head and neck cancer with no nodal disease		

Head and neck caner with limited nodal disease

Stage I-II

Age < 50

Age 50-64

Age 65-84

Age 85 or older

ECOG 0, 1

ECOG 2

Patient with advanced disease and adjuvant RT is an option

Length of surgery 4-8 hours

Length of surgery > 8 hours

Hospital length of stay 4-7 days

Hospital length of stay > 7 days

Free flap required

Intensive care unit or step-down unit required

No free flap required

No tracheostomy tube required

Table 2. Agreement between experts during the Delphi process

Round	Ordinal Scale	LCL	UCL	Per-Priority Group	LCL	UCL
1	0.38	0.34	0.41	0.34	0.31	0.37
2	0.27	0.24	0.31	0.35	0.32	0.38
3	0.40	0.37	0.43	0.35	0.32	0.38
4	0.34	0.30	0.37	0.32	0.28	0.35

Raters = 11, Agreement measured as Krippendorff's alpha.

LCL: Lower confidence limit (95%), UCL: Upper confidence limit (95% CI), Scale: refers to the 0-9 scale used to rate priority of surgery, Per-Priority Group: refers to the low, medium, and high priority groups related to the scoring scale.

References

1. Rolling updates on coronavirus disease (COVID-19). World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. Accessed April 25, 2020.
2. Topf MC. A Framework for Prioritizing Head and Neck Surgery during the COVID-19 Pandemic. *Head & Neck.* 2020(In press).
3. COVID-19: Elective Case Triage Guidelines for Surgical Care. The American College of Surgeons. https://www.facs.org/covid-19/clinical-guidance/elective-case. Published 2020. Accessed May 6, 2020.
4. COVID-19: Guidance for Triage of Non-Emergent Surgical Procedures. The American College of Surgeons. Published 2020. Accessed May 6, 2020.
5. Pandemic Planning Clinical Guideline for Patients with Cancer. Ontario Health (Cancer Care Ontario). https://www.accc-cancer.org/docs/documents/cancer-program-fundamentals/oh-cco-pandemic-planning-clinical-guideline_final_2020-03-10.pdf?sfvrsn=d2f04347_2. Published 2020. Accessed April 11, 2020.
6. Graboyes EM, Kompelli AR, Neskey DM, et al. Association of treatment delays with survival for patients with head and neck cancer: a systematic review. *JAMA Otolaryngology–Head & Neck Surgery.* 2019;145(2):166-177.
7. Goel AN, Frangos M, Raghavan G, et al. Survival impact of treatment delays in surgically managed oropharyngeal cancer and the role of human papillomavirus status. *Head & neck.* 2019;41(6):1756-1769.
8. Murphy CT, Galloway TJ, Handorf EA, et al. Survival impact of increasing time to treatment initiation for patients with head and neck cancer in the United States. *Journal of Clinical Oncology.* 2016;34(2):169.
9. Emanuel EJ, Persad G, Upshur R, et al. Fair allocation of scarce medical resources in the time of Covid-19. In: Mass Medical Soc; 2020.

10. Xiao R, Ward MC, Yang K, et al. Increased pathologic upstaging with rising time to treatment initiation for head and neck cancer: a mechanism for increased mortality. *Cancer.* 2018;124(7):1400-1414.
11. Shuman AG, Campbell BH, Ethics A, Service P. An Ethical Framework for Head and Neck Cancer Care Impacted by COVID-19. *Head & Neck.* 2020.
12. Hunter RJ, Buckley N, Fitzgerald EL, MacCormick AD, Eglinton TW. General Surgery Prioritization Tool: a pilot study. *ANZ journal of surgery.* 2018;88(12):1279-1283.
13. Silva-Aravena F, Álvarez-Miranda E, Astudillo CA, González-Martínez L, Ledezma JG. On the data to know the prioritization and vulnerability of patients on surgical waiting lists. *Data in brief.* 2020;29:105310.
14. Prachand VN, Milner R, Angelos P, et al. Medically-necessary, time-sensitive procedures: A scoring system to ethically and efficiently manage resource scarcity and provider risk during the COVID-19 pandemic. *Journal of the American College of Surgeons.* 2020.
15. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR. *The RAND/UCLA appropriateness method user's manual.* RAND CORP SANTA MONICA CA;2001.
16. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Communication methods and measures.* 2007;1(1):77-89.
17. Krippendorff K. Agreement and information in the reliability of coding. *Communication Methods and Measures.* 2011;5(2):93-112.
18. McHorney CA, Ware Jr JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical care.* 1993:247-263.
19. Wu V, Noel CW, Forner D, et al. Considerations for head and neck oncology practices during the coronavirus disease 2019 (COVID-19) pandemic: Wuhan and Toronto experience. *Head & Neck.* 2020.
20. Forner D, Noel CW, Wu V, et al. Nonsurgical Management of Resectable Oral Cavity Cancer in the Wake of COVID-19: A Rapid Review and Meta-Analysis. *Oral Oncology.* 2020:104849.

21. Marron JM, Joffe S, Jagsi R, Spence RA, Hlubocky FJ. Ethics and Resource Scarcity: ASCO Recommendations for the Oncology Community During the COVID-19 Pandemic. *Journal of Clinical Oncology.* 2020:JCO. 20.00960.

22. Givi B, Schiff BA, Chinn SB, et al. Safety Recommendations for Evaluation and Surgery of the Head and Neck During the COVID-19 Pandemic. *JAMA Otolaryngology–Head & Neck Surgery.* 2020.

23. Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: overview. *Bmj.* 1997;314(7074):131.

24. MacCormick AD, Collecutt WG, Parry BR. Prioritizing patients for elective surgery: a systematic review. *ANZ Journal of Surgery.* 2003;73(8):633-642.

25. Dennett E, Kipping R, Parry B, Windsor J. Priority access criteria for elective cholecystectomy: a comparison of three scoring methods. *The New Zealand medical journal.* 1998;111(1068):231-233.

26. Noseworthy T, McGurran J, Hadorn D, Project SCotWCWL. Waiting for scheduled services in Canada: development of priority-setting scoring systems. *Journal of evaluation in clinical practice.* 2003;9(1):23-31.

27. Conner-Spady BL, Arnett G, McGurran JJ, Noseworthy TW, Project SCotWCWL. Prioritization of patients on scheduled waiting lists: validation of a scoring system for hip and knee arthroplasty. *Canadian Journal of Surgery.* 2004;47(1):39.

28. Allepuz A, Espallargues M, Moharra M, Comas M, Pons JM, Network RGoSII. Prioritisation of patients on waiting lists for hip and knee arthroplasties and cataract surgery: Instruments validation. *BMC health services research.* 2008;8(1):76.

29. Hansen P, Hendry A, Naden R, Ombler F, Stewart R. A new process for creating points systems for prioritising patients for elective health services. *Clinical governance: An international journal.* 2012.

30. Hansen P, Ombler F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-Criteria Decision Analysis.* 2008;15(3-4):87-107.

31. Sertdemir Y, Burgut H, Alparslan Z, Unal I, Gunasti S. Comparing the methods of measuring multi-rater agreement on an ordinal rating scale: a simulation study with an application to real data. *Journal of Applied Statistics.* 2013;40(7):1506-1519.

Figure Legends

Figure 1. SPARTAN-HN Scoring System.

Figure 2.External validation rank results. 14 experts were asked to rate the 12 scenarios provided (x-axis) and results were compared to rank generated by models 1 and 2 (y-axis). Green reflects high priority (ranked 1-4), yellow medium priority (ranked 5-8), and low priority (ranked 9-12). * denotes ties from algorithm.

SPARTAN -HN SCORING SYSTEM

			SCORE
PATIENT FACTORS	A. Performance Status		
	ECOG 0-2	0	
	ECOG 3-4	-3	
	B. Comorbidity		
	No severe comorbidities	0	
	Very severe comorbidities (e.g. < 50% survival at 1 year)	-4	
TUMOR FACTORS	C. Stage		
	I-II	0	
	III-IV	+1	
	D. Histology		
	All intermediate and high-grade histology	0	
	Low grade histology	-3	
	E. Nodal Disease		
	No/limited nodal disease	0	
	Advanced nodal disease (e.g. N3 or ENE)	+3	
	F. Tumor Progression or Potential for Progression		
	Symptomatic progression since decision to treat	+4	
	Clinical progression since decision to treat	+4	
	Potential growth causing inoperability	+4	
	Potential growth causing significant functional morbidity	+4	Max 4
	Potential growth causing moderate functional morbidity or cosmetic impairment	+4	
TREATMENT FACTORS	G. Previous radiotherapy	+3	
	H. Alternative treatment available	-4	
	I. Extent of Surgery		
	Partial laryngeal surgery for organ preservation	+1	
	Total laryngectomy with circumferential pharyngectomy	+1	
	Open anterior craniofacial resection	+1	
	Thyroid cancer requiring tracheal resection	+4	
	All other surgical procedures	0	Max 4
RESOURCE AVAILABILITY	J. Expected Length of Surgery		
	≤ 4 hours	+1	
	> 4 hours	0	
	K. Expected Length of Stay		
	3 days or less	+1	
	> 3 days	0	
	L. Intensive Care or Step-Down Unit Required		
	No	+1	
	Yes	0	Max 1
WAIT TIME FACTORS	M. Wait Time Target Approaching but Not Exceeded		
	Low grade histology	-3	
	High grade histology	0	
	N. Wait Time Target Exceeded by < 2 weeks		
	Low grade histology	-1	
	High grade histology	+3	
	O. Wait Time Target Exceeded ≥ 2 weeks		
	Low grade histology	0	
	High grade histology	+4	Select 1
			TOTAL

cncr_33114_f1.eps

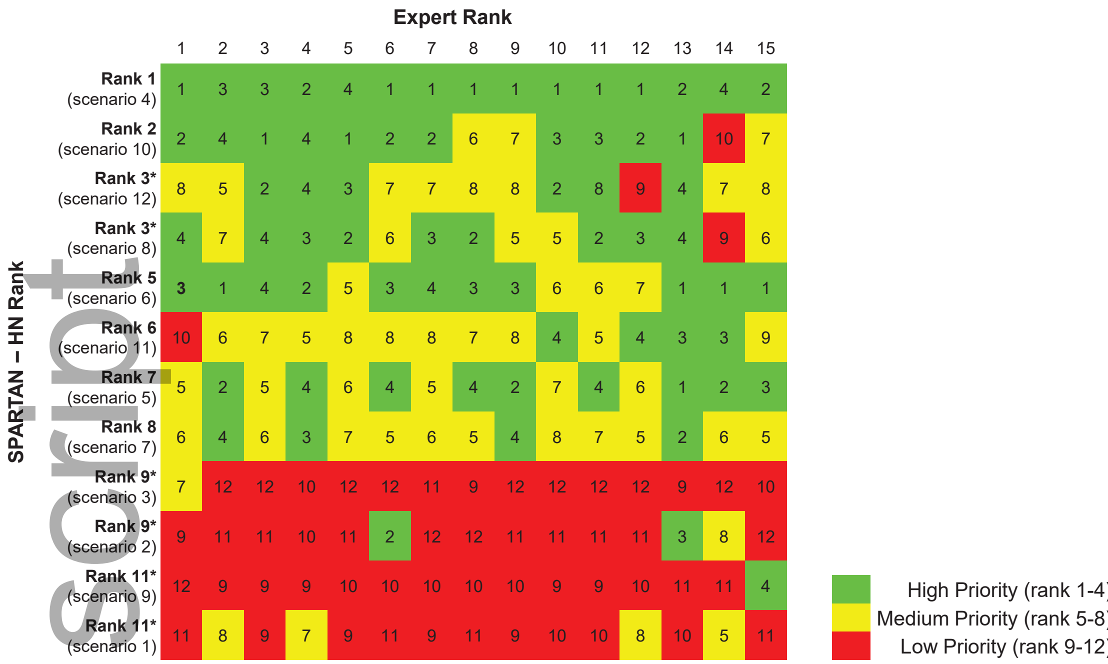

Figure 2: External validation rank results. 14 experts were asked to rank the 12 scenarios provided (x-axis) and results were compared to rank generated by models 1 and 2 (y-axis). Green reflects high priority (ranked 1-4), yellow medium priority (ranked 5-8) and low priority (ranked 9-12). * denotes ties from algorithm

cncr_33114_f2.eps