

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, MD, MSc ^{1,2}; Christopher W. Noel, MD ^{1,2}; David Forner, MD ^{1,2,3}; Han Zhang, MD⁴; Anthony C. Nichols, MD⁵; Marc A. Cohen, MD⁶; Richard J. Wong, MD⁶; Caitlin McMullen, MD⁷; Evan M. Graboyes, MD, MPH ⁸; Vasu Divi, MD⁹; Andrew G. Shuman, MD ¹⁰; Andrew J. Rosko, MD¹⁰; Carol M. Lewis, MD, MPH¹¹; Ehab Y. Hanna, MD¹¹; Jeffrey Myers, MD, PhD¹¹; Vinidh Paleri, MBBS, MS¹²; Brett Miles, DDS, MD¹³; Eric Genden, MD¹³; Antoine Eskander, MD, ScM ¹⁴; Danny J. Enepekides, MD, MPH¹⁴; Kevin M. Higgins, MD, MSc¹⁴; Dale Brown, MD¹; Douglas B. Chepeha, MD, MScPH ¹; Ian J. Witterick, MD, MSc¹⁵; Patrick J. Gullane, MD¹; Jonathan C. Irish, MD, MSc¹; Eric Monteiro, MD, MSc¹⁵; David P. Goldstein, MD, MSc¹; and Ralph Gilbert, MD¹

BACKGROUND: In the wake of the coronavirus disease 2019 (COVID-19) pandemic, access to surgical care for patients with head and neck cancer (HNC) is limited and unpredictable. Determining which patients should be prioritized is inherently subjective and difficult to assess. The authors have proposed an algorithm to fairly and consistently triage patients and mitigate the risk of adverse outcomes. **METHODS:** Two separate expert panels, a consensus panel (11 participants) and a validation panel (15 participants), were constructed among international HNC surgeons. Using a modified Delphi process and RAND Corporation/University of California at Los Angeles methodology with 4 consensus rounds and 2 meetings, groupings of high-priority, intermediate-priority, and low-priority indications for surgery were established and subdivided. 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 using the Krippendorff alpha. Rankings from the algorithm were compared with expert rankings of 12 case vignettes using the Spearman rank correlation coefficient. **RESULTS:** A total of 62 indications for surgical priority were rated. Weights for each indication ranged from -4 to +4 (scale range; -17 to 20). The response rate for the validation exercise was 100%. The SPARTAN-HN demonstrated excellent agreement and correlation with expert rankings (Krippendorff alpha, .91 [95% CI, 0.88-0.93]; and 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. *Cancer* 2020;126:4895-4904. © 2020 American Cancer Society.

LAY SUMMARY:

- Many countries have enacted strict rules regarding the use of hospital resources during the coronavirus disease 2019 (COVID-19) pandemic. Facing delays in surgery, patients may experience worse functional outcomes, stage migration, and eventual inoperability.
- Treatment prioritization tools have shown benefit in helping to triage patients equitably with minimal provider cognitive burden.
- The current study sought to develop what to the authors' knowledge is the first cancer-specific surgical prioritization tool for use in the COVID-19 era, the Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (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.

KEYWORDS: coronavirus disease 2019 (COVID-19), delivery of health care, head and neck cancer, health priorities, patient selection, surgical procedures, waiting lists.

Corresponding Author: John R. de Almeida, MD, MSc, Division of Surgical Oncology, Department of Otolaryngology–Head and Neck Surgery, 8NU-883 Toronto General Hospital, University Health Network, 200 Elizabeth St, Toronto, ON, M5G 2C4 Canada; (John.dealmeida@uhn.ca).

¹Division of Surgical Oncology, Department of Otolaryngology–Head and Neck Surgery, Princess Margaret Cancer Center, University Health Network, University of Toronto, Toronto, Ontario, Canada; ²Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; ³Division of Otolaryngology–Head and Neck Surgery, Dalhousie University, Halifax, Nova Scotia, Canada; ⁴Division of Otolaryngology–Head and Neck Surgery, McMaster University, Hamilton, Ontario, Canada; ⁵Department of Otolaryngology–Head and Neck Surgery, Western University, London, Ontario, Canada; ⁶Department of Otolaryngology–Head and Neck Surgery, Memorial Sloan Kettering Cancer Center, New York, New York; ⁷Head and Neck–Endocrine Oncology, Moffitt Cancer Center, Tampa, Florida; ⁸Department of Otolaryngology–Head and Neck Surgery, Medical University of South Carolina, Charleston, South Carolina; ⁹Department of Otolaryngology–Head and Neck Surgery, Stanford University, Palo Alto, California; ¹⁰Department of Otolaryngology–Head and Neck Surgery, University of Michigan, Ann Arbor, Michigan; ¹¹Department of Otolaryngology–Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas; ¹²Head and Neck Unit, The Royal Marsden Hospital, London, United Kingdom; ¹³Department of Otolaryngology–Head and Neck Surgery, Icahn School of Medicine at Mount Sinai, New York, New York; ¹⁴Department of Otolaryngology–Head and Neck Surgery, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada; ¹⁵Department of Otolaryngology–Head and Neck Surgery, Sinai Health System, University of Toronto, Toronto, Ontario, Canada

The first 3 authors contributed equally to this article.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.33114, **Received:** June 8, 2020; **Revised:** June 22, 2020; **Accepted:** June 23, 2020; **Published online** August 11, 2020 in Wiley Online Library (wileyonlinelibrary.com)

INTRODUCTION

On March 11, 2020, the World Health Organization declared a global pandemic due to the novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and the resulting coronavirus disease 2019 (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 patients with 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 the 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 time from diagnosis to treatment initiation, 9 of 13 studies demonstrated a decrease in survival to be associated with treatment delays.⁶⁻⁸ These data support the urgency of initiating treatment for patients with HNC, but to our knowledge do not inform a stratification schema when operating room access is not available for all patients.

As a result of these new imposed constraints, difficult decisions regarding prioritization for cancer surgery are obligatory, and require the consideration of broader principles regarding scarce resource allocation.⁹ Key among these is the need for consistency and transparency 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, carries 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} Although tools for surgical prioritization in the era of COVID-19 continue to emerge, to our knowledge oncology patients have not been explicitly considered.¹⁴ Herein, we have presented the development and validation of a novel algorithm, Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN), for the prioritization of surgery for patients with HNC.

MATERIALS AND METHODS

The current 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 HNC surgeons (J.R.D., D.P.G., R.G., J.C.I., D.B.C., D.B., A.E., D.J.E., K.M.H., E.M., and I.J.W.) from 3 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 3 institutions were operating under significant resource constraints with limited availability of operating room time. For instrument validation, a group of 5 participants (J.R.D., C.W.N., D.E., D.P.G., and E.M.) completed the scoring algorithm designed after the consensus process. Fifteen external head and neck surgeons (H.Z., A.C.N., R.J.W., M.A.C., C.M., E.M.G., V.D., A.G.S., A.J.R., C.M.L., E.Y.H., J.M., V.P., B.M., and E.G.) from 10 institutions across Canada (2 institutions), the United States (7 institutions), and the United Kingdom (1 institution) 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 (see Supporting Information 1). All indications for prioritization were presented to the consensus panel using an online survey platform (Google Forms; <https://docs.google.com/forms>). With 2 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 (eg, stage I and II vs stage III and IV were presented in sequence; AJCC 8th edition). The list of indications was pilot tested by 4 surgeons (J.R.D., D.P.G., E.M., and R.G.) for sensibility (readability, content validity, language, and comprehensibility).

Consensus Process

The consensus panel participated in a Delphi consensus process with 4 rounds of rating (see Supporting Information 2). The first 2 rounds aimed to achieve consensus regarding the priority grouping (high, intermediate, or low). High priority was defined as an indication to

proceed to surgery within 2 weeks. The second 2 rounds of rating involved ranking each indication (less important, neutral, or 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/University of California at Los Angeles (UCLA) method was used to achieve consensus.¹⁵ This methodology typically is used to determine the appropriateness of an intervention but in this setting was used to determine surgical priority. We used a scale ranging from 0 to 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 to 9 to rate each indication compared with 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 2 rounds to determine surgical priority, a hierarchical logic was adopted to determine consensus regarding whether 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 of the 11 panelists rating a given indication with a zero score. If there was no agreement to avoid surgery, agreement for surgical priority then was defined as ≤ 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 that failed to achieve consensus was classified as being of 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 SPARTAN-HN

The algorithm uses a point-based system to assign a total score based on the sum of the individual indication scores (see Supporting Information 3), with higher scores corresponding to higher priority. Scoring weights were based on consensus from both sets of rounds such that high-priority indications were assigned scores ranging from +2 to +4, intermediate-priority indications were assigned scores ranging from -1 to +1, and low-priority indications were assigned scores ranging from -2 to -4. Within each priority grouping 3-point range, the scores were assigned based on the consensus ratings from the third and fourth rounds. For any 2 patients with the same total score, the

patient with the longer surgical wait time was assigned the higher priority rank.

Clinical Vignettes

Twelve clinical vignettes were constructed (see Supporting Information 4) after the consensus rounds 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 the excessive cognitive burden associated with ranking too many scenarios.

Statistical Analysis Agreement

Agreement between raters during the Delphi process was calculated at each round and within each priority grouping using the Krippendorff alpha (K-alpha). Because typical coefficients of reliability are not suitable for coded data, agreement for the rank orders generated by 5 coders (J.R.D., C.W.N., D.F., E.M., and D.P.G.) 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 for estimation of reliability for any number of raters and categories, and may be used when there are missing data.¹⁷

Validity of the 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 the Spearman rank correlation coefficient. The strength of the correlation was considered weak if the rho was < 0.3 , moderate if the rho was between 0.3 and 0.7, and strong if the rho was > 0.7 .¹⁸

In addition to SPARTAN-HN, a second algorithm using a decision-making flowchart was developed (SPARTAN-HN2). The tool and associated performance characteristics are included in Supporting Information 5.

Sample Size Considerations

For determination of an adequate sample size for the expert panel, we assumed that for model validity, there was a strong correlation between the model rank order and expert rank order (ie, $\rho \geq 0.7$), an alpha of .05, power of 0.8, and a nonresponse rate of 10%. Therefore, the calculated sample size requirement was 15 participants.

All analyses were 2-sided and statistical significance was set at $P \leq .05$. Analyses were conducted using SAS University Edition 9.4 statistical software (SAS Institute Inc, Cary, North Carolina).

RESULTS

Establishing Consensus Priority Groupings (First 2 Consensus Rounds)

After the first 2 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 nonsurgical treatment with a similar prognosis (6 respondents; 54%); 2) poor performance status (ie, Eastern Cooperative Oncology Group [ECOG] performance status of 3-4) (6 respondents; 54%); and 3) very severe comorbidity as indicated by a non-cancer-specific survival rate of <50% at 1 year (5 respondents; 45%).

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

Establishing Ranking Within Each Priority Grouping (Second 2 Consensus Rounds)

Of the 6 low-priority indications, consensus for the importance of factors was achieved for 2 scenarios (33%), both of which were deemed less important (Table 1). Of the 48 intermediate-priority indications, consensus for the importance of factors was achieved for 9 (19%) of scenarios. Of 8 high-priority factors, consensus for the importance of factors was achieved for 4 scenarios (50%), all of which were deemed to be more important.

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

SPARTAN-HN: Surgical Prioritization Scoring System

Priority weights for each indication ranged from -4 to $+4$, spanning a 9-point range and translated from the 2 rounds of priority groupings into 3 categories.

Four indications were assigned a weight of $+4$ based on consensus that these factors were both high priority and more important (Supporting Information 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 as more important were resource use indications, which generally are 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 was 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 ranged from -17 to $+20$ (Fig. 1).

Reliability and Validity Assessment

Agreement between the 5 coders for the SPARTAN-HN was excellent (K-alpha, .91). Agreement between the 15 expert raters was moderate (K-alpha, .63). Convergent validity was demonstrated by a strong correlation between the rank orders generated by the SPARTAN-HN and external experts (ρ , 0.81; 95% CI, 0.45-0.95 [$P = .0007$]). Agreement between expert rankings and SPARTAN-HN rankings for the 12 vignettes is shown in Figure 2.

DISCUSSION

In the setting of the COVID-19 pandemic, in which the availability of operating room time as well as hospital and intensive care unit beds is limited, the prioritization of surgical oncology cases is imperative to mitigate downstream adverse outcomes.^{19,20} The current methodology was adopted based on expert consensus. In the current study, we have proposed the SPARTAN-HN, with the objective of providing transparency and facilitating surgical prioritization for treatment providers.

Creating COVID-19-era allocation schemas that are ethically sound is both critical and challenging. Emanuel et al have advocated 4 ethical principles with which to

TABLE 1. Prioritization Indications and Scores After 4 Rounds of Ranking

	Intermediate-Priority Factors					High-Priority Factors			
	-4	-3	-2	-1	0	+1	+2	+3	+4
Alternative therapy available		Poor performance status (ie, ECOG PS 3, 4)		Wait time exceeded <2 wk 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 wk for high-grade histology	Wait time exceeded by ≥ 2 wk for high-grade histology	Wait time exceeded by ≥ 2 wk for high-grade histology
Very severe comorbidity (eg noncancer survival <50% at 1 y)		Wait time not exceeded but approached in 1 wk for low-grade histology		Wait time exceeded for low-grade histology (≥ 2 wk)	Wait time exceeded for low-grade histology (≥ 2 wk)	Hypopharyngeal cancer requiring total laryngopharyngectomy	Advanced lymph node disease (eg, N3 or macroscopic ENE)	Advanced lymph node disease (eg, N3 or macroscopic ENE)	Clinical or imaging progression (ie, advancing stage)
		Low-grade parotid malignancy		Oral cavity cancer with soft-tissue resection	Oral cavity cancer with soft-tissue resection	Nasal or paranasal sinus cancer requiring open anterior craniofacial surgical resection	Symptomatic disease progression while on wait list	Symptomatic disease progression while on wait list	Potential significant functional morbidity or inoperability if tumor growth
		Thyroid cancer with lymph node disease		Oral cavity cancer with bone resection	Oral cavity cancer with bone resection	Stage III to IV disease (AJCC 8th edition)	Previous RT	Previous RT	Potential moderate functional or cosmetic impairment if tumor growth
				Oral cavity cancer requiring near-total or total glossectomy	Oral cavity cancer requiring near-total or total glossectomy	Length of surgery <4 h			Thyroid cancer with tracheal invasion
				Oropharyngeal cancer with transoral surgery	Oropharyngeal cancer with transoral surgery	Hospital length of stay 1-3 d			
				Oropharyngeal cancer with mandibulectomy	Oropharyngeal cancer with mandibulectomy	No intensive care unit or step-down unit			
				Laryngeal cancer requiring total laryngectomy	Laryngeal cancer requiring total laryngectomy				
				Hypopharyngeal cancer with total laryngectomy and partial pharyngectomy	Hypopharyngeal cancer with total laryngectomy and partial pharyngectomy				
				Nasopharyngeal cancer requiring endoscopic resection	Nasopharyngeal cancer requiring endoscopic resection				
				Nasopharyngeal cancer requiring maxillectomy	Nasopharyngeal cancer requiring maxillectomy				
				Nasal or paranasal sinus cancer requiring endoscopic resection	Nasal or paranasal sinus cancer requiring endoscopic resection				
				Advanced skin cancer requiring skin resection and regional flap reconstruction	Advanced skin cancer requiring skin resection and regional flap reconstruction				
				Advanced skin cancer requiring free-flap reconstruction	Advanced skin cancer requiring free-flap reconstruction				
				High-grade parotid malignancy	High-grade parotid malignancy				
				Temporal bone malignancy	Temporal bone malignancy				
				Head and neck cancer with no lymph node disease	Head and neck cancer with no lymph node disease				
				Head and neck cancer with limited lymph node disease	Head and neck cancer with limited lymph node disease				
				Stage I to II	Stage I to II				
				Age <50 y	Age <50 y				
				Age 50-64 y	Age 50-64 y				
				Age 65-84 y	Age 65-84 y				
				Age ≥ 85 y	Age ≥ 85 y				
				ECOG PS 0, 1	ECOG PS 0, 1				

TABLE 1. Continued

	Low-Priority Factors	Intermediate-Priority Factors	High-Priority Factors
-4			
-3			
-2			
-1			
0		ECOG PS 2 Patient with advanced disease and adjuvant RT is an option Length of surgery 4-8 h Length of surgery >8 h Hospital length of stay 4-7 d Hospital length of stay >7 d Free flap required Intensive care unit or step-down unit required No free flap required No tracheostomy tube required	
+1			
+2			
+3			
+4			

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group performance status; ENE, extranodal extension; RT, radiotherapy.

TABLE 2. Agreement Between Experts During the Delphi Process

Round	Ordinal Scale ^a	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

Abbreviations: LCL, lower 95% confidence limit; UCL, upper 95% confidence limit.

There were 11 raters and agreement was measured using the Krippendorff alpha.

^a“Ordinal scale” refers to the scale of 0 to 9 used to rate priority of surgery and “Per Priority Group” refers to the low-priority, medium-priority, and high-priority groups related to the scoring scale.

guide the 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 those patients who are 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 patients with HNC are aerosol-generating and increase the risk to health care workers and other hospitalized patients.²² Our process accounted for these by giving consideration to these factors during the consensus process, although indications associated with potential exposure to health care workers did not emerge as low-priority ones. Indications associated with lower resource use did achieve consensus for higher importance. This may help to avoid the opportunity cost of treating fewer patients with longer surgeries.

Anecdotal and institution-specific prioritization schemas for patients with HNC 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 the scoring of several criteria and summing of the scores to achieve a total patient score. Many of these systems have been validated against expert rankings of surgical priority.^{27,28}

We used a methodology for developing a point-based prioritization system, similar to those previously described.²⁹ To the best of our knowledge, point-based surgical prioritization systems have been very well studied to date. Hansen et al previously proposed a methodology for developing a point-based prioritization system using the following 7 steps: 1) ranking patient case vignettes

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

FIGURE 1. The Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer (SPARTAN-HN) scoring system. ECOG indicates Eastern Cooperative Oncology Group; ENE, extranodal extension.

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 to 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 the 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 to minimize computational burden. We established 62

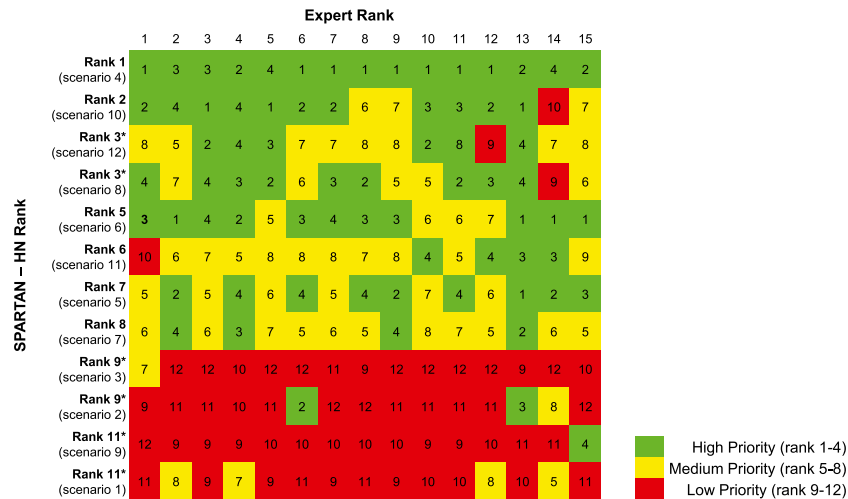


FIGURE 2. External validation rank results. A total of 14 experts were asked to rate the 12 scenarios provided (shown on the x-axis) and the results were compared with the rank generated by models 1 and 2 (shown on the y-axis). Green shading reflects high priority (ranked 1-4), yellow shading indicates medium priority (ranked 5-8), and red shading indicates low priority (ranked 9-12). Asterisk denotes ties from the algorithm. SPARTAN-HN indicates Surgical Prioritization and Ranking Tool and Navigation Aid for Head and Neck Cancer.

indications for surgical prioritization that would create an enormous computational burden using pairwise comparison methodology. One problem inherent in the PAPRIKA methodology is the assumption that all indications are not equal and can be ranked. However, clinically, certain indications may be equivalent in 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 has demonstrated preliminary reliability and validity. We demonstrated good agreement between raters and the SPARTAN-HN algorithm, suggesting minimal interpretive error. Many of the high-priority indications accounted for some component of interpretation because raters were forced to consider imminent disease 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 interrater reliability was found to be higher because the K-alpha is a conservative measure of reliability. Other measures of reliability, such as the Kendall coefficient of concordance, tend to overestimate reliability and cannot be applied to missing data.³¹ Perhaps most important, 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 to validate it with other cancer outcomes, such as survival, is needed.

The results of the current study must be interpreted within the context of the study design. Although externally validated by other surgeons across North America and the United Kingdom, the criteria for which consensus was achieved for prioritization were 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 nonsurgical treatments.^{19,20} Nonetheless, the actual prioritization of surgical waitlists remains the sole responsibility of surgeons and their practice partners. In addition, the SPARTAN-HN algorithm is intended to assist in making difficult prioritization decisions and is not intended to make recommendations for the time frame 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. The validation process in the current study used expert opinion as the gold standard of prioritization, which is potentially biased, and reflected the opinions of surgeons practicing in academic medical centers from 3 resource-rich nations. Subsequently, use of the SPARTAN-HN algorithm in other geographic regions

and health care systems requires additional investigation because local treatment paradigms and risk factors may vary substantially.

The current study has presented the development and validation of a novel algorithm for the prioritization of surgery for patients with HNC. Further evaluation of its implementation in various practice settings will be obligatory. However, the results of the current study have provided data with which to inform real-world use, as the current pandemic has obviated our ability to more rigorously study the instrument prior to making necessary and difficult real-time allocation decisions.

FUNDING SUPPORT

No specific funding was disclosed.

CONFLICT OF INTEREST DISCLOSURES

Evan M. Graboyes has received grants from the National Cancer Institute and the Doris Duke Charitable Foundation for work performed outside of the current study. Vinidh Paleri offers his services as a proctor for a transoral robotic surgery proctoring program run by Intuitive Surgical and has been remunerated for his time. Antoine Eskander has received a research grant from Merck and acted as a paid consultant for Bristol-Myers Squibb for work performed outside of the current study. Ian J. Witterick has stock in Proteocyte Diagnostics Inc and has received honoraria from Sanofi Genzyme and Medtronic Canada for work performed outside of the current study. The other authors made no disclosures.

AUTHOR CONTRIBUTIONS

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. **Han Zhang:** Data acquisition and editing. **Anthony C. Nichols:** Data acquisition and editing. **Marc A. Cohen:** Data acquisition and editing. **Richard J. Wong:** Data acquisition and editing. **Caitlin McMullen:** Data acquisition and editing. **Evan M. Graboyes:** Data acquisition and editing. **Vasu Divi:** Data acquisition and editing. **Andrew G. Shuman:** Writing, data acquisition, and editing. **Andrew J. Rosko:** Data acquisition and editing. **Carol M. Lewis:** Data acquisition and editing. **Ehab Y. 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 B. Chepeha:** Data acquisition and editing. **Ian J. 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.

REFERENCES

- World Health Organization. Rolling updates on coronavirus disease (COVID-19). Accessed April 25, 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
- Topf MC, Shenson JA, Holsinger FC, et al. Framework for prioritizing head and neck surgery during the COVID-19 pandemic. *Head Neck*. 2020;42:1159-1167.
- American College of Surgeons. COVID-19: elective case triage guidelines for surgical care. Published 2020. Accessed May 6, 2020. <https://www.facs.org/covid-19/clinical-guidance/elective-case>
- American College of Surgeons. COVID-19: guidance for triage of non-emergent surgical procedures. Published 2020. Accessed May 6, 2020. <https://www.facs.org/covid-19/clinical-guidance/triage>
- Ontario Health (Cancer Care Ontario). Pandemic planning clinical guideline for patients with cancer. Published 2020. Accessed April 11, 2020. https://www.accc-cancer.org/docs/documents/cancer-program-fundamentals/oh-cco-pandemic-planning-clinical-guideline_final_2020-03-10.pdf?sfvrsn=d2f04347_2
- 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 Otolaryngol Head Neck Surg*. 2019;145:166-177.
- 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;41:1756-1769.
- 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. *J Clin Oncol*. 2016;34:169-178.
- Emanuel EJ, Persad G, Upshur R, et al. Fair allocation of scarce medical resources in the time of Covid-19. *N Engl J Med*. 2020;382:2049-2055.
- 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;124:1400-1414.
- Shuman AG, Campbell BH; AHNS Ethics & Professionalism Service. Ethical framework for head and neck cancer care impacted by COVID-19. *Head Neck*. 2020;42:1214-1217.
- Hunter RJ, Buckley N, Fitzgerald EL, MacCormick AD, Eglington TW. General surgery prioritization tool: a pilot study. *ANZ J Surg*. 2018;88:1279-1283.
- Silva-Aravena F, Alvarez-Miranda E, Astudillo CA, Gonzalez-Martinez L, Ledezma JG. On the data to know the prioritization and vulnerability of patients on surgical waiting lists. *Data Brief*. 2020;29:105310.
- Prachand VN, Milner R, Angelos P, et al. Medically necessary, time-sensitive procedures: scoring system to ethically and efficiently manage resource scarcity and provider risk during the COVID-19 pandemic. *J Am Coll Surg*. Published online April 9, 2020. doi:92020;S1072-7515(20)30317-3
- Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR. The RAND/UCLA Appropriateness Method User's Manual. RAND Corporation; 2001.
- Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Meas*. 2007;1:77-89.
- Krippendorff K. Agreement and information in the reliability of coding. *Commun Methods Meas*. 2011;5:93-112.
- McHorney CA, Ware JE Jr, 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. *Med Care*. 1993;31:247-263.
- 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;42:1202-1208.
- 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 Oncol*. 2020;109:104849.
- 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. *J Clin Oncol*. 2020;38:2201-2205.
- 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 Otolaryngol Head Neck Surg*. Published online March 31, 2020. doi:10.1001/jamaoto.2020.0780
- Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: Overview. *BMJ*. 1997;314:131-134.
- MacCormick AD, Collett WG, Parry BR. Prioritizing patients for elective surgery: a systematic review. *ANZ J Surg*. 2003;73:633-642.
- Dennett ER, Kipping RR, Parry BR, Windsor J. Priority access criteria for elective cholecystectomy: a comparison of three scoring methods. *N Z Med J*. 1998;111:231-233.
- Noseworthy TW, McGurran JJ, Hadorn DC, Steering Committee of the Western Canada Waiting List Project. Waiting for scheduled

- services in Canada: development of priority-setting scoring systems. *J Eval Clin Pract.* 2003;9:23-31.
27. Conner-Spady BL, Arnett G, McGurran JJ, Noseworthy TW; Steering Committee of the Western Canada Waiting List Project. Prioritization of patients on scheduled waiting lists: validation of a scoring system for hip and knee arthroplasty. *Can J Surg.* 2004;47:39-46.
28. Allepuz A, Espallargues M, Moharra M, Comas M, Pons JM, Research Group on Support Instruments-IRYSS Network. Prioritisation of patients on waiting lists for hip and knee arthroplasties and cataract surgery: instruments validation. *BMC Health Serv Res.* 2008;8:76.
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. *Clin Gov Int J.* 2012;17:200-209.
30. Hansen P, Ombler F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *J Multi-Criteria Decis Analysis.* 2008;15:87-107.
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. *J Appl Stat.* 2013;40:1506-1519.