Individualized Outcome Prognostication for Patients with Laryngeal Cancer

Running Title: Prognostication for Laryngeal Cancer

Authors: Connor W Hoban[1], Lauren J Beesley[2], Emily L Bellile[2], Yilun Sun[2], Matthew E Spector[1], Gregory T Wolf[1], Jeremy MG Taylor[2], Andrew G Shuman[1]

[1]Department of Otolaryngology-Head & Neck Surgery: University of Michigan Medical School, Ann Arbor, MI 48109 United States of America

[2]Department of Biostatistics: University of Michigan School of Public Health, Ann Arbor, MI 48109 United States of America

Keywords: Calculator, Nomogram, Larynx cancer, Prognostication, Risk prediction.

Corresponding Author:
Andrew G Shuman, MD
Department of Otolaryngology Head & Neck Surgery[1]
Michigan Medicine
1500 East Medical Center Drive
Ann Arbor, MI 48109, United States of America
shumana@med.umich.edu
734 232 0120

Conflict of interest: None declared

Funding:

This work was supported by the University of Michigan Head and Neck Specialized Program of Research Excellence program sponsored by the National Cancer Institute at the National Institutes of Health (P50 CA097248); the University of Michigan Cancer Center (P30 CA046592); the Cancer Biostatistics Training Program (T32 CA083654); and an internal University of Michigan Research Office MCubed grant.

Role of the Funder/Sponsor:

The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgements:

We thank the many investigators in the University of Michigan Head and Neck Specialized Program of Research Excellence for their contributions to patient recruitment, assistance in data collection, and encouragement including Carol R. Bradford, MD, Thomas E. Carey, PhD, Douglas B. Chepeha, MD, Sonia Duffy, PhD, Avraham Eisbruch, MD, Joseph Helman, DDS, Kelly M. Malloy, MD, Jonathan McHugh, MD, Scott A. McLean, MD, Tamara H. Miller, RN, Jeff Moyer, MD, Jacques E. Nor, PhD., Lisa Peterson MPH, Mark E. Prince, MD, Nancy Rogers, RN, Laura Rozek PhD, Nancy E. Wallace, RN, Heather Walline, PhD, Brent Ward, DDS, and Francis Worden, MD. There was no financial compensation. We also thank our patients and their families who tirelessly participated in our survey and specimen collections.

Author Contributions:

Conceptualization: C Hoban, M Spector, G Wolf, J Taylor, A Shuman

Methodology: C Hoban, L Beesley, E Belille, Y Sun, M Spector, G Wolf, J Taylor, A Shuman

Software: L Beesley, E Belille, Y Sun, J Taylor

Formal analysis: C Hoban, L Beesley, E Belille, Y Sun, J Taylor

Investigation: C Hoban, L Beesley, E Belille, Y Sun, M Spector, G Wolf, J Taylor, A Shuman

Resources: M Spector, G Wolf, J Taylor, A Shuman

Data curation: C Hoban, L Beesley, E Belille, Y Sun, G Wolf, J Taylor, A Shuman

Writing – original draft: C Hoban, G Wolf, J Taylor, A Shuman

Writing – review and editing: C Hoban, L Beesley, E Belille, Y Sun, M Spector, G Wolf, J Taylor, A Shuman

Visualization: C Hoban, L Beesley, E Belille, Y Sun

Supervision: G Wolf, J Taylor, A Shuman

Project administration: G Wolf, J Taylor, A Shuman

Funding acquisition: G Wolf, J Taylor, A Shuman

Precis: Accurate prognostication is essential to the optimal management of laryngeal cancer, but suboptimal reliability and accuracy limit the integration of existing individualized prediction tools into routine clinical decision making. Further development of individualized prognostic calculators may improve risk prediction, treatment planning, and counselling for patients with laryngeal cancer.

Abstract

Background: Accurate prognostication is essential to the optimal management of laryngeal cancer. Predictive models have been developed to calculate the risk of oncologic outcomes, but extensive external validation of accuracy and reliability is necessary prior to implementing them into clinical practice.

Methods: Four published prognostic calculators that predict five-year overall survival for patients with laryngeal cancer were evaluated using patient information from a prospective epidemiology study cohort (n=246; median follow-up 60 months) with previously untreated stage I-IVb laryngeal squamous cell carcinoma.

Results: Different calculators can give substantially different predictions for individual patients. The observed 5-year overall survival was significantly higher than the averaged predicted 5-year overall survival of the four calculators (71.9% [95% CI 65%-78%] versus 47.7%). Statistical analyses demonstrated the calculators' limited capacity to discriminate outcomes for risk-stratified patients. The AUC ranged from 0.68 to 0.72. C-index values were similar for each of the four models (0.66 to 0.68). There was a lower than expected hazard of death for patients who received induction (bioselective) chemotherapy (HR 0.46 [0.24, 0.88] p=0.024) or primary surgical intervention (HR 0.43 [0.21, 0.90] p-value 0.024) compared to concurrent chemoradiation.

Conclusions: Suboptimal reliability and accuracy limit the integration of existing individualized prediction tools into routine clinical decision making. The calculators predicted significantly worse-than-observed survival among patients treated with induction chemotherapy and primary surgery, suggesting a need for updated consideration of modern treatment modalities. Further

development of individualized prognostic calculators may improve risk prediction, treatment planning, and counselling for patients with laryngeal cancer.

Introduction

The multidisciplinary management of head and neck cancer (HNC) is critically dependent upon accurate risk stratification and prediction of clinical outcomes.[1] Managing laryngeal cancer introduces additional challenges secondary to debilitating functional impairments that often accompany the primary disease and/or related therapeutic interventions.[2-5] Accepted standards for tumor staging and oncologic prognostication continue to be helpful in confronting these challenges and guiding decision making.[6, 7] However, more sophisticated methods are mandatory in order to capitalize on discoveries related to tumor biology/genomics and patient factors and their ability to further individualize treatment selections that enhance survival and minimize morbidity.[8, 9] Treatment decision making is particularly complex for laryngeal cancer due to the variety of treatment options available and differing short- and long-term functional consequences that affect quality of life and survival. The management of laryngeal cancer has experienced substantial evolution, primarily driven by the implementation of chemotherapeutic modalities and novel organ preservation strategies.[3, 10-12] This further emphasizes the need for modernized tools when calculating prognostic estimates.

The heterogeneity of laryngeal cancer is one of the many factors that impose formidable challenges to the accurate prediction of individual survival.[13] Survival is influenced by numerous variables, including multiple and diverse tumor specific (size, grade, genomics, biological features, and stage) and patient related (age, race, gender, immune status, smoking status and medical comorbidities) factors.[14] The TNM (tumor–node–metastasis) staging system defined by the American Joint Committee on Cancer (AJCC) is the current prognostic standard for head and neck cancer and predicts survival with reasonable accuracy.[15] Despite its trusted reputation and ubiquitous assimilation in clinical practice, TNM staging fails to incorporate many criteria that

demonstrate prognostic value, thereby limiting its ability to tailor risk predictions to an individual patient.[15] Studies suggest that implementing additional tumor, patient and treatment characteristics into risk calculations can promote superior prognostic accuracy across a diverse range of oncologic subspecialties.[16-19]

The current trend in attitudes toward electronic health information suggests that online versions of these calculators would be readily implemented into medical decision making.[20-24] As a means of regulating newly published risk predictors, the AJCC recently published sixteen inclusion and exclusion criteria that are required for endorsement of any probability or risk model.[25] These benchmarks should help to ensure that performance metrics, compatibility, and clinical relevance are robust amidst the expansion of new prediction tools.[25]

Investigators have worked to address the dearth of individualized clinical decision tools currently available to interdisciplinary teams that manage head and neck squamous cell carcinoma (HNSCC) by developing prognostic calculators specific to head and neck cancers.[26-28] Risk calculators wield potential clinical value but have not yet been subjected to sufficient evaluation and validation to warrant their assimilation into routine practice. Moreover, many of these were generated with older data that may not be directly relevant to current patients. In order to assess clinical prognostic tools, analyses that compare the calculators' predictions to each other and to modern observed outcomes are imperative.[29] These validation studies are best performed on independent patient cohorts that encompass diverse geographic regions and patient demographics.[30, 31] Our goal was to utilize an independent patient cohort to externally evaluate and validate published prognostic calculators designed for patients with laryngeal cancer.

Materials and methods

The University of Michigan IRBMED evaluated and approved this study. All subjects provided written informed consent at enrollment in this prospective epidemiology study (typically at the time of diagnosis).

Prognostic Calculators

Prognostic clinical decision tools were identified via online search engines and expert input. Pubmed and Google Scholar were investigated for peer-reviewed publications using a combination of search terms representing disease (larynx, cancer), prognosis (i.e. survival, risk, prediction and outcome) and methodology (calculator, tool, model and nomogram). Multidisciplinary head and neck cancer specialists were also surveyed in order to probe for existing or emerging prognostic tools not identified in the online search.[16]

Potential calculator candidates were evaluated for eligibility. Inclusion criteria mandated that the calculator utilized clinical data to predict five-year overall survival for squamous cell carcinoma of the larynx. Four prognostic calculators were identified (MAASTRO, Lifemath, Leiden, and MyCancerJourney) and each model was reviewed for content and format.[26-28] The calculators' mathematical formulas were acquired from the original publication, supplementary online materials, or computational derivation. Notably, MyCancerJourney does not have an associated peer-reviewed publication.

Table 1 summarizes each calculator and includes the period, sample size, and other characteristics that describe the original study cohorts. Each calculator functions according to an equation constructed to represent the relationship between tumor characteristics, patient demographics, employed treatment modalities and observed survival outcomes. The calculators considered a distinct set of variables in their prognostic equation and assigned differing quantities of statistical weight to these variables (**Table 2 and Supplementary Tables 1 - 7,**

available online). The calculators were modeled from patient data contained in the Surveillance, Epidemiology, and End Results (SEER) registry, regional study cohorts, or a combination of two patient populations. The four study cohorts included patients treated with curative intent between 1973 and 2009 [26-28].

Patients

The analysis data set was derived from a single-institution prospectively maintained head and neck cancer epidemiologic study.[32-35] A total of 246 patients with biopsy-proven, previously untreated, AJCC stage I-IVb squamous cell carcinoma of the larynx diagnosed and treated with curative intent at the University of Michigan Health System between 2003 and 2014 were included. **Table 2** provides additional summary demographics for the cohort. Patients were evaluated by our multidisciplinary team and discussed at our Tumor Board where treatment recommendations were formulated. Patients with Stage I or II disease were generally treated with single modality surgery (33 [35.1%]), radiotherapy alone (50 [53.2%]), or concurrent chemoradiotherapy for deeply invasive T2 lesions (11 [11.7%]). Patients with Stage III or IV disease received either primary surgery (32 [20.1%]), a single cycle of induction chemotherapy (bioselective) followed by either combined chemoradiation for greater than 50% tumor response or total laryngectomy for less than 50% tumor response (70 [46.1%]), or definitive chemoradiation (50 [32.9%]). Median follow-up was 60 months. Tumor, patient and treatment specific variables were exported from the database and confirmed via chart abstraction.

The calculators were designed for utility prior to oncologic treatment. Consequently, pretreatment clinical information was used to populate the relevant variables. Pathological information was only used as a substitute when clinical information was not available. Missing variables were populated using established algorithms described in the supplementary online

materials. Exclusion criteria included carcinoma in-situ, distant metastasis at the time of diagnosis and synchronous primary tumors not including basal or squamous cell carcinoma of the skin.

Statistical Analysis

Each calculator was used to individually predict five-year overall survival for 246 patients in our independent cohort. The arithmetic average of the predictions from the four calculators was tested as a distinct (fifth) calculator, referred to as the "mean" in subsequent analyses. The agreement between these predictions was compared using scatterplots, Spearman's correlation coefficients and the proportion of five-year overall survival predictions that differed by less than 0.10 between separate calculators. The calibration of each calculator was assessed using Kaplan-Meier plots that stratified patients into equal-sized quintiles according to calculator-predicted risk. The average predicted risk for each quintile was compared to the estimated five-year survival for that quintile in a calibration plot. The discriminatory ability of each calculator was assessed using both the area under the ROC curve for the binary outcome of survival at five years[36] and the C-index. Both the C-index and the five year AUC measure the concordance between the predicted risk and the survival outcome and are frequently reported in the literature.

To assess which factors may be responsible for discrepancies between the predicted outcomes and observed survival, a separate Cox model was fit for each possible factor and adjusted for the predicted risk as measured by the mean prediction. All tests for statistical significance utilized 95% confidence intervals and were two-sided. For treatment factors, a calibration plot was used to elucidate which treatment modalities were not well calibrated with

the predicted risk. The method used to construct the calibration plot is described in the supplementary online materials.

Results

The patient cohort represents the typical distribution and epidemiology of patients with laryngeal cancer. Most patients were Caucasian, current/former smokers, and male with no or mild medical comorbidities. Slightly less than half had a glottic subsite whereas the majority originated in the supraglottic larynx.

Observed 5-year overall survival was 71.9% [95% CI of 65% to 78%] whereas each of the calculators predicted significantly worse outcomes and contributed to a mean predicted 5-year overall survival of 47.7%. **Figure 1** describes each calculator's predicted 5-year survival for our patient population. Visual assessment of the estimates suggests that MAASTRO and MyCancerJourney have a tendency to predict worse outcomes; both have more predictions clustered at lower values than the other calculators. MAASTRO showed the greatest discrepancy in its prognostication for patients with high-risk disease. MyCancerJourney has more variation in its predictions. Comparisons of MyCancerJourney and MAASTRO against the remaining two calculators demonstrated less agreement and were characterized by lower correlation coefficients.

The Leiden and Lifemath calculators demonstrated the closest correlation coefficient (rho=0.816) and the weakest association occurred between Lifemath and MyCancerJourney (rho=0.644). **Supplementary Table 8,** available online, reports the percentage of patients for which a selected pair of calculators predicted 5-year overall survival within 0.10. For example, if one calculator predicted a 50% 5-year overall survival for a given patient, the paired calculator was considered to be in consensus if it predicted between 40 and 60% 5-year overall survival for

the same patient. Prognostic consensus between calculators was found to be variable and survival estimates typically agreed within 0.10 for less than 50% of the cohort.

Figure 2a displays Kaplan-Meier survival plots that used calculator predictions to risk-stratify the patients into equal-portioned quintiles. The four calculators were reasonably effective in discriminating between these risk quintiles but Lifemath and Leiden were less adroit in stratifying low and high risk patients whereas MyCancerJourney and MAASTRO were less able to discriminate mid-tiered risk. The mean of the calculators more accurately stratified risk for each of the quintiles.

Figure 2b displays ROC curves and their accompanying AUC and C-index scores for each of the four calculators. These values provide informative measures of prognostic discrimination. The AUC values were similar for each of the four models and fell into a range from 0.68 (Leiden) to 0.72 (MAASTRO and Mean). The C-index values ranged from 0.66 (MAASTRO, Leiden and MyCancerJourney) to 0.68 (Lifemath and Mean).

Calibration studies are summarized in **Figure 3a** and demonstrate lower survival estimates in relation to observed outcomes for each risk-stratified quintile. MAASTRO was especially pessimistic for high risk patients; Lifemath was well calibrated to observed outcomes for patients with mid-tiered risk.

Cox modeling identified male gender (HR 2.03 [1.00, 4.11] p-value 0.031) and initial planned treatment as additional factors that were found to add predictive value even after adjusting for the predicted risk from the calculators. Laryngeal subsite (glottis vs. supraglottic) was not significantly related to survival after adjusting for calculator predictions (eTable 9). **Figure 3b** demonstrates that there was a lower than expected hazard of death for those patients in the cohort who underwent induction (bioselective) chemotherapy (HR 0.46 [0.24, 0.88] p-

value 0.024) or primary surgical intervention (HR 0.43 [0.21, 0.90] p-value 0.024) compared to primary chemoradiation. Adjusted hazard ratios for all other factors are shown in **Supplementary Table 9,** available online.

Discussion:

Available prognostic calculators generated variable predictions with inconsistent accuracy when compared to observed outcomes in an external prospectively maintained cohort of patients with laryngeal cancer. The four calculators were designed with varying degrees of similarity but substantial disparity in performance was evident. We have reported both the discriminatory ability and the calibration properties of the calculators, these can be thought of as measures of the accuracy of the relative and absolute predictions respectively. Relative predictions are whether the patients can be correctly ranked according to risk, and absolute predictions are whether the predicted probabilities of survival are correct. Both are important, and neither is good for any of the four calculators.

Differences in patient cohorts, prognostic variables, statistical modeling and inherent calculator limitations contributed to this variation and highlight many of the challenges associated with oncologic prognostication. LifeMath and MyCancerJourney both utilized the SEER population as their study cohort but exhibited the weakest association in risk prediction. This observation emphasizes the importance of incorporating and weighing prognostic variables accurately.[37] Comorbidity status has proven to be useful in predicting outcomes but was only integrated into the Leiden and MyCancerJourney calculators, providing further explanation for observed differences in calculator performance.[38] The AJCC established guidelines in 2016 that work to synchronize the development of future prognostic calculators, but meaningful discrepancies in the perceived importance of prognostic variables will likely persist.[25]

The MAASTRO calculator was developed in the Netherlands using a study cohort of 994 patients with laryngeal cancer from 1977 to 2008.[27] This calculator was designed for patients who only received radiation therapy as their primary treatment modality and exclusion criteria included carcinoma in-situ, distant metastases, and chemotherapy. MAASTRO's internal validation demonstrated an AUC of 0.73 whereas traditional TNM staging demonstrated an AUC of 0.62 for the study cohort.[27] Previous external validations yielded AUC values of 0.68, 0.74, 0.76, and 0.71. When compared to survival predictions for our entire external cohort, predictions for the 55 subjects who exclusively received radiation demonstrated similar calibration, better correlation to predictions from the other calculators, and substantially improved AUC and C-index values (**Supplementary Table 10 and Figures 1, 2 and 3,** available online**)**. Our study produced the first non-European external validation with an AUC of 0.72 when applied to all 246 patients and an AUC of 0.81 when applied to just those 55 patients who received radiation only. This calculator's pessimistic tendency could be explained by the evolving role of multimodality therapy, although the absolute incremental overall survival benefit of chemotherapy over radiation alone in laryngeal cancer is fairly modest (5%) outside of bioselection[39], The calculator also has a noticeable unfavorable hazard ratio for male gender and low hemoglobin count, however many patients in our cohort were anemic but had better-than-predicted outcomes. In summary, the calculator's performance was fair considering the fundamental differences in the study cohort from which it was developed.

The LifeMath calculator was developed using 50,145 American patients with all sites of head and neck cancer in the SEER database from 1980 to 2009.[26] Internal validation and external validation on 1,362 patients from the Massachusetts General Hospital were performed using an incomparable correlation metric. This study provides an external validation with a C-index of

0.68 and AUC of 0.71. The accuracy of this calculator may have been limited by its neglect of comorbidity status, a variable that is not available in the SEER database.

The Leiden calculator was developed using a study cohort of 1371 Dutch patients with several different sites of head and neck cancer from 1981 to 1998.[28] Internal validation demonstrated a C-index of 0.73 whereas an external validation on 598 patients from the Siteman Cancer Center yielded a C-index of 0.69. This study provides an additional external validation with a C-index of 0.66. The model likely produced pessimistic survival estimates due to evolution and refinement of treatment over the past four decades.

MyCancerJourney used SEER data from 1973 to 1996 and Barnes-Jewish Hospital data from 1995 to 2001 to construct its model but did not have a publication to accompany its online calculator. The calculator performed with a C-index of 0.66 and AUC of 0.70 in this study. MyCancerJourney employed a novel comorbidity metric to characterize patients. The majority of variability in outcome predictions could be explained by the grouping of patients according to treatment modality and other parameters that resulted in substantial prognostic fluctuation.

Each of the calculators was designed using a training dataset that included patients from over 35 years ago. Consequently, many of the patients in the study cohort were not treated according to modern strategies. Reliance on older patient data is likely to have contributed to the accumulative tendency to estimate worse-than-observed survival in the study cohort and underscores the importance of ensuring reliability and accuracy before adopting these into clinical practice.

In contrast to a previous study of currently available oral cavity cancer calculators, the laryngeal cancer patients in our cohort demonstrated consistently better survival when compared to the calculators' predicted outcomes.[30] This observation reinforces the hypothesis that

individualized treatment paradigms for laryngeal cancer need to be considered when predicting survival. Whether this represents differences among the patients themselves or the individualized treatment approaches utilized remains speculative. However, induction chemotherapy (bioselective) for subsequent treatment selection and/or primary surgical intervention were both associated with survival benefits that were greater than expected following adjustments for calculator-estimated survival. Individualized treatment paradigms that integrate neoadjuvant bioselection are associated with a significant survival benefit and may account for the superior outcomes observed.[35] These findings reinforce the need for updated survival calculators and provide further evidence that oncologic interventions and institution-specific care are independent variables that affect the prognosis of patients with laryngeal cancer.

The absence of laryngeal cancer clinical practice guidelines that clearly specify preferred treatment modality may be contributing to disparities in calculator performance. Established practice guidelines help to optimize patient outcomes and standardize the value of care.[40-42] However, emerging evidence suggests that adherence to current guidelines established by the American Head and Neck Society (AHNS) and National Comprehensive Cancer Network (NCCN) does not significantly improve outcomes.[43] Therapeutic regimens are complicated by the range of available treatment modalities and the need to individualize these based upon patient, tumor and institutional factors, making population-level recommendations challenging.[44]

There is a growing impetus for evaluating value of cancer care. Judging value involves the balanced consideration of quality and outcome delivered, and remains especially difficult to measure in oncology due to its multidisciplinary nature, need for prolonged follow-up, and consideration of post-treatment function as well as survival.[45] In order to address these

challenges, prognostic calculators could be modeled from patient cohorts treated in accordance to optimized quality metrics. Once appropriately accurate, precise and calibrated, such calculators could help to establish standardized expected outcomes for individually risk-stratified patients. Comparing an institution's outcomes to calculator predictions on an individualized and risk-stratified basis may serve as an effective method for evaluating and comparing relative quality and value.[45, 46]

There are inherent limitations to this study, chiefly involving the single-institutional data, which may not reflect practice patterns or outcomes in other populations. The patients in our cohort were also treated by an experienced multidisciplinary team and had comprehensive follow up allowing appropriate salvage treatment for recurrence when necessary. Additional limitations may be linked to the improved outcomes generally associated with treatment in academic centers.[35, 40, 42, 47-49] The accuracy of the SEER data, as well as the calculators relying on it, may be confounded if incurable patients treated for palliation were included or other inaccuracies in treatment details were present. Missing variables were an additional source of error in this evaluation. However, derived values (see missing variable methods in the supplement) for hemoglobin, radiation dosage and tumor diameter rarely led to substantial differences in survival predictions, helping to mitigate this concern. The calculators do not consider the role of human papillomavirus (HPV); the incidence and prognostic impact of HPV in laryngeal cancer is considerably lower than oropharyngeal cancer[50].

There is a need for more accurate prognostic calculators that predict individualized outcomes for patients with laryngeal cancer. Currently available prognostic calculators varied in their ability to consistently and accurately predict survival in an external cohort of patients with laryngeal cancer. Suboptimal reliability and accuracy limit the potential integration of existing

individualized prediction tools into routine clinical practice. The calculators estimated significantly worse-than-observed survival among patients treated with induction bioselective chemotherapy and primary surgery, suggesting that modern treatment modalities must be better integrated into revised prediction tools. Deficiencies in calculator performance may be further explained by institutional variation in oncologic outcomes. Potential avenues to improve performance of calculators include utilizing contemporary patient cohorts, integrating biomarkers and harnessing the promise of the genomic frontier as these data emerge. The use of statistical and machine learning approaches when datasets are large is another intriguing possibility to create mechanisms that can more nimbly respond to exponentially complex and evolving data [51].

These data raise questions about the inherent value of oncologic nomograms. We contend that they are useful for patients to estimate individualized prognosis, and perhaps for comparing results across different cohorts. Predictive models that guide treatment selection might be of higher value and could see increasing demand as the arsenal of available therapies continues to proliferate and individualize. Improved individualized calculators may help to assign value of oncologic care and will be critical in refining the ability of multidisciplinary teams to predict risk, plan shared treatment-related decision making, and counsel patients effectively.

References:

1. Siegel R, Miller K, Jemal A. Cancer statistics, 2015. *CA-Cancer J Clin.* 2015;65(1):5-29.
2. Hutcheson KA, Lewin JS. Functional outcomes after chemoradiotherapy of laryngeal and pharyngeal cancers. *Curr Oncol Rep*. 2012;14(2):158–165.
3. Forastiere AA, Zhang Q, Weber RS, et al. Long-term results of RTOG 91-11: A comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer. *J Clin Oncol*. 2013;31(7):845–852.
4. Worden FP, Kumar B, Lee JS, et al. Chemoselection as a strategy for organ preservation in advanced oropharynx cancer: Response and survival positively associated with HPV16 copy number. *J Clin Oncol*. 2008;26(19):3138–3146.
5. Terrell JE, Fisher SG, Wolf GT. Long-term quality of life after treatment of laryngeal cancer. *Arch Otolaryngol*. 1998;124(9):964–971.
6. Amin MB, Greene FL, Edge SB, et al. The eighth edition AJCC cancer staging manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging [published online January 1, 2017]. *CA-Cancer J Clin*. doi: 10.3322/caac.21388
7. Bilimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: A powerful initiative to improve cancer care in the United States. *Ann Surg Oncol*. 2008;15(3):683–690.
8. Bradford CR, Kumar B, Bellile E, et al. Biomarkers in advanced larynx cancer. *Laryngoscope*. 2014;124(1):179–187.
9. Starska K, Forma E, Jóźwiak P, et al. Gene and protein expression of glucose transporter 1 and glucose transporter 3 in human laryngeal cancer—the relationship with regulatory

hypoxia-inducible factor-1α expression, tumor invasiveness, and patient prognosis. *Tumor Biol*. 2015;36(4):2309–2321.

10. Wolf GT, Hong WK, Fisher SG, et al. induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. *N Engl J Med*. 1991;324(24):1685–1690.
11. Lefebvre JL, Andry G, Chevalier D, et al. Laryngeal preservation with induction chemotherapy for hypopharyngeal squamous cell carcinoma: 10-year results of EORTC trial 24891. *Ann Oncol*. 2012;23(10):2708–2714.
12. Calais G, Pointreau Y, Alfonsi M, et al. Randomized phase III trial comparing induction chemotherapy using cisplatin (P) fluorouracil (F) with or without docetaxel (T) for organ preservation in hypopharynx and larynx cancer. Preliminary results of GORTEC 2000–01. *J Clin Oncol*. 2006;24(18S1):Oral presentation n. 5506.
13. Lønning PE. Breast cancer prognostication and prediction: are we making progress? *Ann Oncol.* 2007;18(suppl 8):viii3-viii7.
14. Vuralkan E, Akin I, Kuran G, Saka C, Akçayöz N, Alicura S. Prognostic factors in larynx cancers. *Kulak Burun Bogaz Ihtis Derg*. 2008;18(4):221–226.
15. Sobin LH. TNM: Evolution and relation to other prognostic factors. *Semin Surg Oncol*. 2003;21(1):3–7.
16. Rabin B, Gaglio B, Sanders T, et al. Predicting cancer prognosis using interactive online tools: A systematic review and implications for cancer care providers. *Cancer Epidem Biomar.* 2013;22(10):1645-1656.
17. Ross PL, Gerigk C, Gonen M, et al. Comparisons of nomograms and urologists' predictions in prostate cancer. *Semin Urol Oncol*. 2002;20(2):82-88.

18. Mackillop WJ, Quirt CF. Measuring the accuracy of prognostic judgments in oncology. *J Clin Epidemiol.* 1997;50(1):21–29.

19. Gorlia T, Wu W, Wang M, et al. New validated prognostic models and prognostic calculators in patients with low-grade gliomas diagnosed by central pathology review: a pooled analysis of EORTC/RTOG/NCCTG phase III clinical trials. *Neuro-Oncology*. 2013;15(11):1568–1579.

20. Hesse BW, Moser RP, Rutten LJ. Surveys of physicians and electronic health information. *N Engl J Med*. 2010;362(9):859–860.

21. Kreps GL, Neuhauser L. New directions in eHealth communication: Opportunities and challenges. *Patient Educ Couns*, 2010;78(3):329–336.

22. Higgins O, Sixsmith J, Barry MM, Domegan C. A literature review on health information-seeking behaviour on the web: A health consumer and health professional perspective (published online 2011]. *ECDC*. doi: 10.2900/5788.

23. McClymont K, Lee S, Schonberg M, Widera E, Miao Y, Smith AK. Usefulness and impact of online prognostic calculators. *J Am Geriatr Soc*. 2014;62(12):2444–2445.

24. Nolan T, Dack C, Pal K, et al. Patient reactions to a web-based cardiovascular risk calculator in type 2 diabetes: a qualitative study in primary care. *Brit J Gen Pract*. 2015;65(632):e152-e160.

25. Kattan MW, Hess KR, Amin MB, et al. American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. *CA-Cancer J Clin*. 2016;66(5):370-374.

26. Emerick KS, Leavitt ER, Michaelson JS, Diephuis B, Clark JR, Deschler DG. Initial clinical findings of a mathematical model to predict survival of head and neck cancer. *Otolaryngol Head Neck Surg*. 2013;149(4):572–578.
27. Egelmeer AGTM, Velazquez ER, de Jong JMA, et al. Development and validation of a nomogram for prediction of survival and local control in laryngeal carcinoma patients treated with radiotherapy alone: a cohort study based on 994 patients. *Radiother Oncol*. 2011;100(1):108–115.
28. Datema FR, Ferrier MB, Vergouwe Y, et al. Update and external validation of a head and neck cancer prognostic model. *Head Neck*. 2013;35(9):1232–1237.
29. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14(1):40.
30. Prince V, Bellile EL, Sun Y, et al. Individualized risk prediction of outcomes for oral cavity cancer patients. *Oral Oncol*. 2016;63(1):66–73.
31. Parks C, Heald J, Hall G, Kamaly-Asl I. Can the prognosis of individual patients with glioblastoma be predicted using an online calculator? *Neuro-Oncology*. 2013;15(8):1074-1078.
32. Duffy SA, Taylor JM, Terrell JE, et al. Interleukin-6 predicts recurrence among head and neck cancer patients. *Cancer*. 2008;113(4):750-757.
33. Virani S, Bellile E, Bradford CR, et al. NDN and CD1A are novel prognostic methylation markers and differ by HPV status in patients with head and neck squamous carcinoma. BMC Cancer. 2015;15(1):825-838.

34. Peterson LA, Bellile EL, Wolf GT, et al. University of Michigan Head and Neck Specialized Program of Research Excellence Program: Cigarette use, comorbidities and prognosis in a prospective head and neck squamous cell carcinoma population. *Head Neck Surg.* 2016;38(12)1810-1820.

35. Wolf GT, Bellile E, Eisbruch A, et al. survival rates using individualized bioselection treatment methods in patients with advanced laryngeal cancer [published online February 2, 2017]. *JAMA Otolaryngol Head Neck Surg.* doi:10.1001/jamaoto.2016.3669

36. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56(2):337-344.

37. Wang SJ, Wissel AR, Ord CB, et al. Individualized estimation of conditional survival for patients with head and neck cancer. *Otolaryngol Head Neck Surg*. 2011;145(1):71-73.

38. Piccirillo JF, Lacy PD, Basu A, Spitznagel EL. Development of a new head and neck cancer-specific comorbidity index. *Arch Otolaryngol Head Neck Surg*. 2002;128(10):1172-1179.

39. Blanchard P, Baujat B, Holostenco V, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): A comprehensive analysis by tumour site. *Radiother Oncol*. 2011;100(1):33–40.

40. Lewis CM, Weber RS. Standardizing treatment: a crisis in cancer care. *Otolaryng Clin N Am.* 2013;46(4):671–679.

41. Blayney DW, McNiff K, Eisenberg PD, et al. Development and future of the American Society of Clinical Oncology's Quality Oncology Practice Initiative. *J Clin Oncol*. 2014;32(35):3907–3913.

42. Lewis CM, Weber RS, Hanna EY. Quality of care in head and neck cancer. *Curr Oncol Rep*. 2011;13(2):120–125.

43. Graboyes EM, Townsend ME, Kallogjeri D, Piccirillo JF, Nussenbaum B. Evaluation of quality metrics for surgically treated laryngeal squamous cell carcinoma. *JAMA Otolaryngol Head Neck Surg*. 2016;142(12):1154–1163.

44. Forastiere AA, Weber RS, Trotti A. Organ preservation for advanced larynx cancer: Issues and outcomes. *J Clin Oncol*. 2015;33(29):3262–3268.

45. Roman BR, Awad MI, Patel SG. Defining value-driven care in head and neck oncology. *Curr Oncol Rep*. 2015;17(1):424.

46. Shellenberger TD, Madero-Visbal R, Weber RS. Quality indicators in head and neck operations: a comparison with published benchmarks. *Arch Otolaryngol Head Neck Surg*. 2011;137(11):1086–1093.

47. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA*. 1998;280(20):1747–1751.

48. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. *JAMA*. 2000;283(9):1159–1166.

49. Birkmeyer JD, Siewers AE, Finlayson EVA, et al. Hospital volume and surgical mortality in the United States. *New Engl J Med*. 2002;346(15):1128–1137.

50. Fakhry C, Westra WH, Wang SJ, et al. The prognostic role of sex, race, and human papillomavirus in oropharyngeal and nonoropharyngeal head and neck squamous cell cancer. *Cancer*. 2017;123(9):1566–1575.

51. Meldolesi E, van Soest J, Damiani A, et al. Standardized data collection to build prediction models in oncology: a prototype for rectal cancer. *Future Oncol*. 2016;12(1):119–136.

Figure Legends:
Figure 1: Scatterplots Comparing Calculator Predictions

Legend: The histograms on the diagonal show the distribution of predicted 5-year survival from each calculator for the 246 University of Michigan patients. The scatterplots below the diagonal show individual predictions of 5-year overall survival from pairs of calculators when applied to the University of Michigan patients. Points close to the 45 degree line are from patients with similar predictions from the two calculators. Correlation coefficients from the scatterplots are shown above the diagonal.

Figure 2: Comparison of Calculator Predictions to Observed Outcomes

Legend: Figure 2a shows the Kaplan-Meier overall survival estimates stratified by quintile of predicted 5-year survival for each calculator. Figure 2b shows the ROC curves of sensitivity and specificity of each calculators predicted 5-year survival compared to the observed 5-year survival.

Figure 3: Calibration of Calculator Predictions to Observed Outcomes

Legend: Figure 3a shows the calibration of each calculator. Each point represents a set of patients with similar predicted probability of 5-year survival (grouped by quintiles). The horizontal axes represent the average predicted probability of the group, the vertical axis show the observed 5-year survival for the group obtained from Kaplan-Meier plots. A well calibrated calculator would have points near the diagonal line. Figure 3b shows the calibration curves stratified by initial treatment plan. Each point represents a small group of patients who have similar predicted 5-year survival and the same initial treatment.

Table 1: Summary of Calculators

Calculator	Cancers in training dataset	Training dataset	Validation dataset	Model Type	Model Details
MAASTRO (Egelmeer et al 2011)	Larynx	994 patients with laryngeal carcinoma treated with RT from 1977-2008 (89.9% N0)	Leuven; 109 patients treated with RT from 2000-2006 (75.2% N0). VU Amsterdam; 178 patients treated with RT from 2001-2007 (92.7% N0). NKI/AML Amsterdam; 205 patients treated with RT from 2000-2008 (89.8% N0). Manchester; 403 patients treated with RT from 1998-2005 (98.8% N0)	Cox Regression	Main effects only.
LifeMath (Emerick et al 2013)	HN	HN patients in SEER up to 2009	Massachusetts General Hospital 1362 patients	Statistical-mechanistic model of cancer metastasis involving separate tumor and node contributions.	Complicated formulas with many parameters and interactions.
Leiden (Datema et al 2013)	HN	1371 patients (638 with laryngeal cancer) at Leiden University Medical Centre 1981-1999	598 pts Barnes-Jewish Hospital between 1995-2000	Cox regression	Main effects only.
MyCancer Journey *	All cancers	SEER 1973-1996 and 11,791 Barnes-Jewish Hospital patients 1995-2001	No validation data.	Cox regression	Main effects and many interactions.

*Calculator available online at https://staging.mycancerjourney.com/myinsights/survival-curves.

Table 2: Patient Characteristics (N=246)

	N (%)* or Mean (SD)	Missing, N (%)	Calculators using Characteristic
Demographics			
Age at Diagnosis	60.0 (10.2)	0 (0)	All four
Gender			
Female	56 (22.7)	0 (0)	All four
Male	190 (77.2)		
Race			
Black	9 (3.6)	0 (0)	LifeMath, MyCJ
Other	9 (3.6)		
White	228 (92.6)		
Smoking Status			
Current (in past 12 mos)	166 (67.4)	1 (0.4)	None
Former (>12 mos)	61 (24.7)		
Never	18 (7.3)		
ACE Comorbidities			
None	49 (19.9)	0 (0)	MyCJ
Mild	112 (45.5)		
Moderate	60 (24.3)		
Severe	25 (10.1)		
ACE Comorbidities (w/o Prior Tumors)			
None	53 (21.5)	0 (0)	Leiden
Mild	118 (47.6)		
Moderate	58 (23.5)		
Severe	17 (6.9)		
Tumor Information			
Primary Site			
Glottic	115 (46.7)	0 (0)	Leiden, MAASTRO
Supraglottic	131 (53.2)		
Subglottic	0 (0)		
AJCC Overall Stage			
I	60 (24.3)	0 (0)	None
II	34 (13.8)		
III	53 (21.5)		
IV	99 (40.2)		
SEER Stage			
Localized	103 (41.7)	0 (0)	MyCJ
Regional	92 (37.2)		
Distant	51 (20.6)		
T Stage			
1	64 (25.9)	0 (0)	Leiden (MAASTRO after transformation)
2	52 (21.1)		
3	71 (28.8)		
4	59 (23.9)		
N Stage			
0	156 (63.4)	0 (0)	LifeMath (transformation used for Leiden, MAASTRO)
1	26 (10.5)		
1b	1 (0.4)		
2	1 (0.4)		
2a	2 (0.8)		
2b	24 (9.7)		
2c	33 (13.4)		
3	3 (1.2)		

Number of Positive Nodes			
0	157 (63.8)	0 (0)	LifeMath
1	33 (13.4)		
2	24 (9.7)		
3	19 (7.7)		
4	10 (4.0)		
5+	3 (1.2)		
Tumor Diameter			
Mean	2.6 (1.49)	103 (41.8)	LifeMath
<1.5cm	31 (12.6)		
1.5-2.5cm	34 (13.8)		
2.5-3.5cm	43 (17.4)		
>3.5cm+	35 (14.2)		
Grade			
1 (well)	38 (15.4)	0 (0)	MyCJ
2 (moderate)	117 (47.5)	(Unknown category)	
3 (poor)	40 (16.2)		
4 (undifferentiated)	1 (0.4)		
Unknown	50 (20.3)		
Extracapsular Spread			
Irrelevant (No Nodes)	157 (63.5)	0 (0)	LifeMath
No	18 (7.3)	(Unknown category)	
Yes	43 (17.4)		
Unknown	28 (11.3)		
Margins			
Negative	78 (31.7)	163 (66.2)	None
Positive	5 (2.0)		
Hemoglobin (g/dl)	13.8 (1.69)	45 (18.2)	MAASTRO
Treatment Information			
Initial Treatment Plan			
Induction Chemo	70 (28.4)	0 (0)	None
No Induction Chemo			
Surgery	65 (26.4)		
Chemoradiation	56 (22.7)		
Radiation Only	55 (22.3)		
Surgery (within 4 months) †			
No	166 (67.4)	0 (0)	MyCJ
Yes	80 (32.5)		
Chemotherapy (within 4 months) †			
No	107 (43.4)	0 (0)	MyCJ
Yes	139 (56.5)		
Radiation (within 4 months) †			
No	42 (17.0)	0 (0)	MyCJ
Yes	204 (82.9)		
Total Radiation Dose	68.1 (4.07)	105 (42.6)	None
Radiation EQD_{2T}	58.6(1.74)	222 (90.2)	MAASTRO

* Percent includes missing values.

† Delivered within four months of original treatment initiation

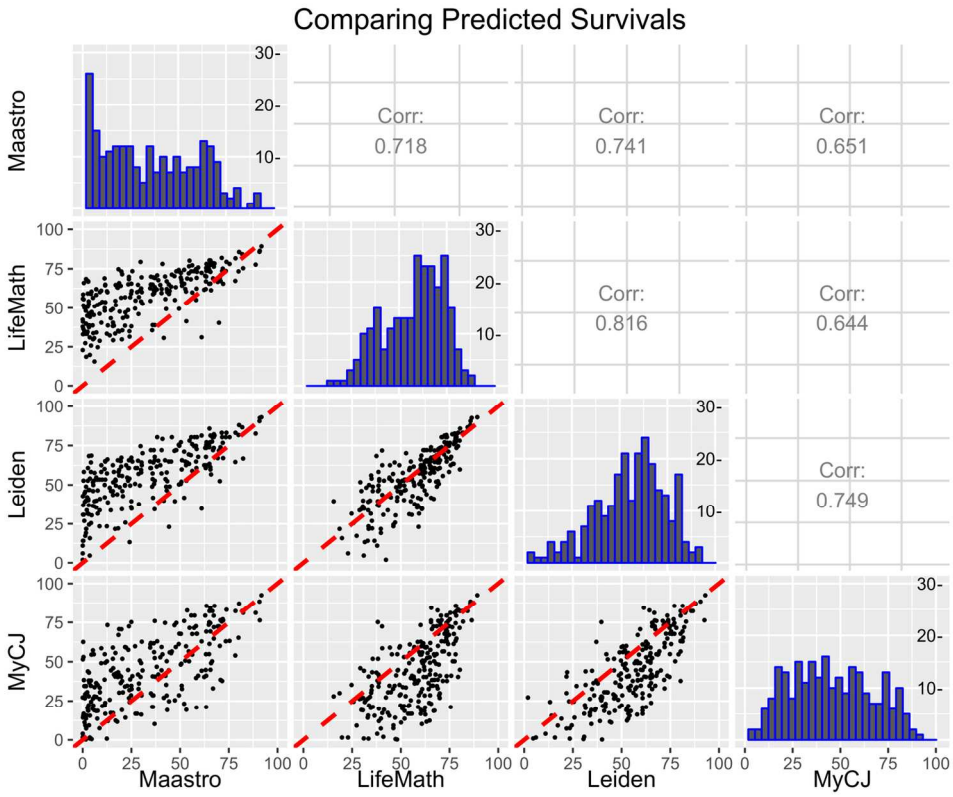


Figure 1

148x119mm (300 x 300 DPI)

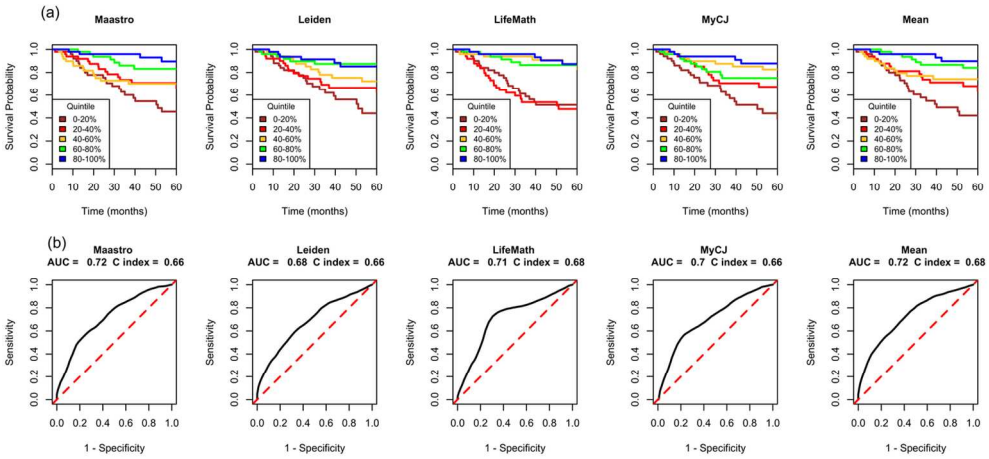


Figure 2

129x61mm (300 x 300 DPI)

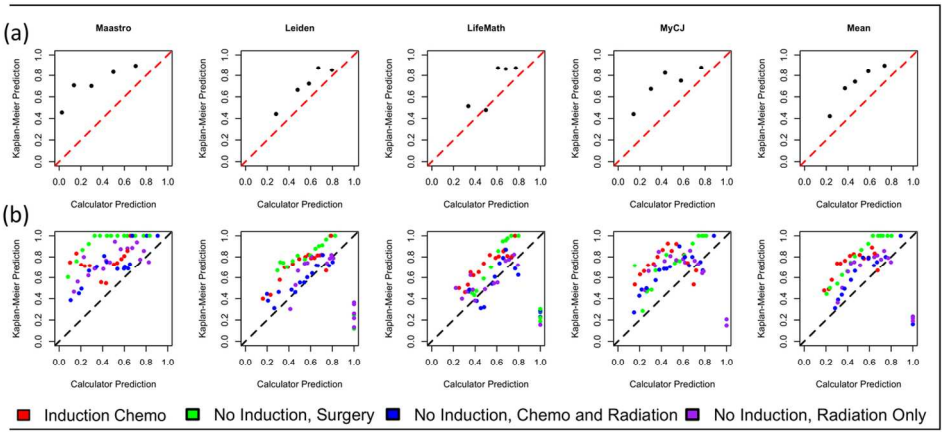


Figure 3

126x62mm (300 x 300 DPI)

Supplementary Materials for "Individualized Outcome Prognostication for Patients with Larynx Cancer"

Section 1: Equations for the Larynx Cancer Outcome Calculators

In the following report, we describe the models used in each of the four evaluated calculators. We present the versions of the four calculators that were available from the respective websites or other sources on 1 November 2016. Previous description of the model structures for the Leiden, LifeMath, and MyCancerJourney calculators can be found in a similar investigation for oral cavity cancer conducted by Prince et al.[1]

Leiden

The Leiden calculator is based on a Cox proportional hazards regression model. Publications by Datema et al. provide the model structure and parameter estimates but not the baseline survival function.[2,3] The equations for the calculator were derived from a study of 1371 Dutch head and neck cancer patients, of which 638 had larynx cancers. The url of the website is https://www.msbi.nl/SV/Chart.aspx?model=Leiden+HNSCC. The baseline 5-year survival was obtained through trial and error, i.e. we obtained 5-year survival predictions for a number of patients using the website, and then applied the published parameters to solve for the 5-year baseline survival probability.

The form of the model is the following for the i^{th} patient: the hazard (rate) of death at time t is

$$\lambda_i(t) = \lambda 0(t)\exp(\beta_{\text{age}} \cdot \text{age}_i + \beta_{\text{gender}} \cdot \text{gender}_i + \beta_{\text{comorbidity}} \cdot \text{comorbidity}_i + \beta_{\text{T}} \cdot \text{T}_i + \beta_{\text{N}} \cdot \text{N}_i + \beta_{\text{M}} \cdot \text{M}_i + \beta_{\text{prior}} \cdot \text{prior}_i + \beta_{\text{location}} \cdot \text{location}_i)$$

where $\lambda 0(t)$ is the baseline hazard rate at time t.

There are 8 covariates in this model, and all of them are categorical except for age. Rather than using subject age directly, this model determines which of several intervals the true age falls into and then assigns a value for the age variable (roughly the midpoint of the age interval) accordingly. In particular, ages 40 or less are assigned a value of 32 years. Ages 41-50, 51-60, 61-70, 71-80, and greater than 80 are assigned values of 47, 56, 66, 75, and 85 respectively. Then, the transformed version of age (with possible values of 32, 47, 56, 66, 75, and 85) is treated as a continuous variable in the model. The Comorbidity (ACE27) variable used in this model was adjusted to exclude the contribution of prior tumors. $\lambda 0(t)$ is the baseline hazard rate at time t, defined for a 32-year-old male patient whose tumor location is lip, with stage T1N0M0 and comorbidity ACE27 score of none, without prior tumors. eTable 1 summarizes the covariates from the model above, how they are coded, and the corresponding β coefficients.

eTable 1: Parameter Values in Leiden Calculator Cox proportional hazards model

Covariate	Code	β
Age†	Continuous	0.04
Gender	Male	Reference, 0
	Female	-0.08
ACE 27*	none	Reference, 0
	mild	0.06
	moderate	0.34
	severe	0.79
T	T1	Reference, 0
	T2	0.25
	T3	0.42
	T4	0.67
N	N0	Reference, 0
	N1	0.37
	N2	0.61
	N3	0.90
M	M0	Reference, 0
	M1	1.85
Prior Tumors	No	Reference, 0
	Yes	0.50
Location	Lip	Reference, 0
	Hypopharynx	0.62
	Oral Cavity	0.41
	Oropharynx	0.47
	Glottic larynx	0.04

	Supraglottic larynx	0.27
	Nasopharynx	0.18

* ACE27 comorbidities excluding contribution for prior tumors
† Age transformed using the "midpoint" rules previously described

Based on the model for the hazard rates, the 5-year survival probabilities are calculated using the formula $S(5) = S_0(5)^{\exp(X\beta)}$ where $\exp(X\beta)$ corresponds to the exponential term in the hazard rate equation above. The baseline 5-year survival is defined (in the online calculator) as the 5-year survival probability for a 32-year-old male patient whose tumor location is lip, with stage T1N0M0 and comorbidity ACE27 score of none, without prior tumors. The calculated value of the 5-year baseline survival is $S_0(5) = 0.9268$.

LifeMath

The LifeMath calculator does not use a Cox proportional hazard model.[4] The predicted 5-year survival is based on a mechanistic model (SNAP) that assesses the influence of primary tumor and nodes separately. The equations for the calculator were developed from a study of 50,145 head and neck cancer patients from the SEER database. The website http://www.lifemath.net/cancer/headneck/outcome/ provides source code in JavaScript, thereby providing all parameters and a table of conversion factors for this calculator. Similar LifeMath calculators are available for breast cancer, melanoma, renal cell cancer, and colon cancer.

The components of the model can be computed using the formulas below: for each patient, let L_{primary} and L_{nodes} denote the contributions to the calculated head and neck cancer specific survival from the primary tumor and the nodes respectively. L_{HNSC10} is the estimated 10-year cancer specific survival, which is adjusted to 5-year overall survival estimate using a table of age and gender specific factors derived from the SEER data. L_{primary}, L_{nodes}, and L_{HNSC10} are defined as follows:

$$L_{\text{primary}} = 1 - \exp(-Q \cdot j_{\text{primary}} \cdot (10 \cdot \text{diameter})^{Z} \cdot g)$$
$$L_{\text{nodes}} = 1 - \exp(-\text{nodes} \cdot R \cdot g)$$
$$L_{\text{HNSC10}} = L_{\text{primary}} + L_{\text{nodes}} - L_{\text{primary}} \cdot L_{\text{nodes}}$$

These equations depend on several constants, which take values (truncated at 4 decimal places) of:

$$Q = 0.1059$$
$$j_{\text{primary}} = 0.6944$$
$$Z = 0.5721$$
$$R = 0.1365$$

In the equation, "diameter" is the primary tumor size at the greatest dimension and "nodes" is the number of nodes. When the number of nodes is unknown, the program sets number of nodes to be 0 and changes jprimary to be 1.The g term in the formulas above represents the effects from all other covariates, including age, race, gender, N stage, extension and ECS (extra capsular spread). The g term for patient i was calculated in the following way:

$$g_i = \exp(\beta_{\text{age}} \cdot \text{age}_i + \beta_{\text{race}} \cdot \text{race}_i + \beta_{\text{site}} \cdot \text{site}_i + \beta_{\text{N}} \cdot \text{N}_i + \beta_{\text{ECS}} \cdot \text{ECS}_i).$$

The covariates, how they are coded and the coefficients (specific to Larynx cancer) are shown in eTable 2.

eTable 2: Lifemath SNAP Model Parameters (parameters for g expression)*

Covariate	Code	β
Age	36 - 45	-0.332
	46 - 55	-0.183
	56 - 65	-0.040
	66 - 75	0.172
	76 - 95	0.393
	else	0
Race	Black	0.275
	White	-0.095
	Other	-0.118
	Unknown	0
Site	Oral Cavity	0.054
	Hypopharynx	0.409
	Larynx	0.051

	Oropharynx	-0.135
	Nasopharynx	-0.002
N Stage	N0/Unknown	0
	N1	0.175
	N2a	0.216
	N2b	0.271
	N2c	0.337
	N3	0.507
ECS	Absent/Unknown	0
	Present	0.378

*Truncated at three decimal places

For other cancer sites, this model contains additional interaction terms, but there are no additional interaction terms for the Larynx cancer subsite. For more details about this model in other subsites, see Prince et al.[1]

MyCancerJourney

The MyCancerJourney (MyCJ) calculator https://staging.mycancerjourney.com/ is based on a Cox proportional hazards regression model, and it is more complicated than the other Cox models considered because it includes many interactions terms. No calculator information is available from the website, and it is not described by a peer reviewed publication, and there is no evidence that the model has been validated in external data. The equations for the calculator were developed from a combination of two datasets. One consisted of 11,791 patients with any cancer treated at the Siteman Cancer Center/Barnes-Jewish hospital and the other included over 2 million cancer patients in the SEER database.

All the parameters and baseline survival estimates that are relevant for larynx cancer patients were obtained by trial and error. The trial and error process is the following: default options from the website as baseline were used (which will be described below), and each variable was separately changed while holding other variables constant. The parameter coefficient for the variable that best corresponds to the results from the website was then determined. All the parameter estimates as well as interaction terms were obtained through repeating the above procedure.

The hazard rate at time t is computed as follows (for larynx cancer patients only):

$$\begin{aligned}\lambda_i(t) = \lambda_0(t)\exp(&\beta_{\text{age}}\cdot \text{age}_i + \beta_{\text{gender}}\cdot \text{gender}_i + \beta_{\text{race}}\cdot \text{race}_i + \beta_{\text{comorbidity}}\cdot \text{comorbidity}_i && \textit{(Main Effects)}\\ &+ \beta_{\text{histology}}\cdot \text{histology}_i + \beta_{\text{stage}}\cdot \text{stage}_i + \beta_{\text{grade}}\cdot \text{grade}_i && \textit{(Main Effects)}\\ &+ \beta_{\text{surgery}}\cdot \text{surgery}_i + \beta_{\text{chemo}}\cdot \text{chemo}_i + \beta_{\text{radiation}}\cdot \text{radiation}_i && \textit{(Main Effects)}\\ &+ \beta_{\text{age-gender}}\cdot \text{age}_i\cdot \text{gender}_i + \beta_{\text{age-race}}\cdot \text{age}_i\cdot \text{race}_i + \beta_{\text{age-stage}}\cdot \text{age}_i\cdot \text{stage}_i && \textit{(Interactions)}\\ &+ \beta_{\text{age-comorbidity}}\cdot \text{age}_i\cdot \text{comorbidity}_i + \beta_{\text{surgery-chemo}}\cdot \text{surgery}_i\cdot \text{chemo}_i && \textit{(Interactions)}\\ &+ \beta_{\text{gender-comorbidity}}\cdot \text{gender}_i\cdot \text{comorbidity}_i + \beta_{\text{race-comorbidity}}\cdot \text{race}_i\cdot \text{comorbidity}_i && \textit{(Interactions)}\\ &+ \beta_{\text{chemo-comorbidity}}\cdot \text{chemo}_i\cdot \text{comorbidity}_i + \beta_{\text{histology-comorbidity}}\cdot \text{histology}_i\cdot \text{comorbidity}_i && \textit{(Interactions)}\\ &+ \beta_{\text{histology-grade}}\cdot \text{histology}_i\cdot \text{grade}_i + \beta_{\text{radiation-stage}}\cdot \text{radiation}_i\cdot \text{stage}_i) && \textit{(Interactions)}\end{aligned}$$

where $\lambda_0(t)$ is the annual rate of death at time t for a 60- to 64-year-old white male patient with squamous cell histology with in situ stage, grade 1 tumor, without comorbidity or any treatment, and with Larynx cancer subsite. The hazard rate is determined by 11 main effects and several interaction terms. Their coding as well as the estimated βs for the main effect and interaction terms are shown in eTables 3 and 4 respectively. Note that the "in situ" and "local" stages and "G1" and "G9" (Unknown) grades have the same hazard ratio. Therefore, both are marked as reference.

eTable 3: Main Effect Parameter Values in MyCancerJourney Cox proportional hazards model

Covariate	Code	β
Age	<35	-0.86
	35 - 39	-0.69
	40 - 44	-0.69
	45 - 49	-0.32
	50 - 54	-0.39
	55 - 59	-0.23
	60 - 64	Reference, 0
	65 - 69	0.06

	70 - 74	0.33
	75 - 79	0.59
	80 - 84	0.34
	>84	0.51
Gender	Male	Reference, 0
	Female	-0.39
Race	White/Other	Reference, 0
	Black	-0.25
Comorbidity	None	Reference, 0
	Mild	0.23
	Moderate	0.49
	Severe	1.14
Histology	Squamous Cell	Reference, 0
	Other	-0.13
Stage (SEER)	in situ/Local	Reference, 0
	Regional	0.95
	Distant	1.65
Grade	G1/G9 (Unknown)	Reference, 0
	G2	0.17
	G3	0.33
	G4	-0.19
Surgery	No	Reference, 0
	Yes	-1.14
Chemotherapy	No	Reference, 0
	Yes	-0.05
Radiation	No	Reference, 0
	Yes	-0.62

eTable 4: Interaction Parameter Values in MyCancerJourney Cox proportional hazards model

Covariate Interaction	**β**
Age - Gender Interaction	
Age in 40-44 and Gender = Female	0.86
Age >84 and Gender = Female	0.80
Else[a]	0
Age - Race	
Age in 40-44 and Race = Black	1.02
Age in 50-54 and Race = Black	0.66
Age in 55-59 and Race = Black	0.57
Age in 65-69 and Race = Black	0.61
Else	0
Age - Stage	
Age in 75-79 and Stage = Regional	-0.46
Age in 80-84 and Stage = Distant	-1.46
Else	0
Age - Comorbidity	
Age in 75-79 and Comorbidity = Severe	-0.73
Age >84 and Comorbidity = Severe	-1.01
Else	0
Treatment - Treatment	
Surgery = Yes and Chemo = Yes	0.76
Else	0
Gender - Comorbidity	
Gender = Female and Comorbidity = Mild	0.31
Else	0
Race - Comorbidity	
Race = Black and Comorbidity = Severe	-0.46
Else	0
Treatment and Comorbidity	

Chemo = Yes and Comorbidity = Severe	-0.37
Else	0
Histology - Comorbidity	
Histology = Other and Comorbidity = Severe	1.18
Else	0
Histology - Grade	
Histology = Other and Grade = G2	-1.11
Else	0
Treatment - Stage	
Radiation = Yes and Stage = Distant	-0.47
Else	0

a Interaction terms are nonzero only for certain covariate combinations

The baseline 5-year survival is defined as the 5-year survival probability for a 60- to 64-year-old white male patient with squamous cell histology with in situ stage, grade 1 tumor, without comorbidity or any treatment, and with Larynx cancer subsite. The value is $S_0(5) = 0.481$. The 5-year survival is then estimated using the equation $S(5) = S_0(5)^{\exp(X\beta)}$ where $\exp(X\beta)$ is the exponential term in the hazard rate expression above.

We note that the large number of interaction terms in the MyCJ model does lead to some strange features and may result from overfitting. For black patients, for example, there is a decrease in the survival only for very specific age ranges 40-44 and 50-59 and 65-69, but not for other age ranges. However, there is no comparable noticeable change in survival for specific age ranges for white patients.

Maastro

The Maastro calculator (http://www.predictcancer.org/Main.php?page=LarynxFollowUpModel) is based on a Cox proportional hazards model. Egelmeer et al. (2011) provides the structure of the model and a set of parameter estimates.[5] However, the predicted survivals from the online calculator did not match the parameter estimates provided in the paper, so we believe the model used in the online calculator as of November 1st 2016 is an updated version of the model published in 2011.[5] According to information available on the website, the model was developed using a dataset of 994 patients with squamous cell laryngeal carcinoma that were treated at the Maastro clinic in the Netherlands from January 1977 to December 2008. This model was validated using data from 895 patients from four different cohorts from Leuven (Belgium), Amsterdam (VU and NKI/ALV Hospital), and Manchester (UK). This calculator was intended to be used only for larynx cancer patients that were cancer stage I-IV, were treated with high dose radiotherapy alone (so no surgery or chemotherapy), and had no distant metastases at the time of diagnosis. It is noted in the main text that we ignored the eligibility criteria by applying the Maastro model to all patients, including patients who received other treatments in addition to radiation. Therefore, we apply this model to some patients who do not meet the eligibility requirements for the model. When we applied the Maastro model strictly to the subset of our patient population that received radiation only (and no induction chemotherapy) as part of the initial treatment plan (N=55), we obtain the results presented in Section 7 of this document.

Model parameters and the baseline survival probability were estimated by trial and error using the online calculator. Hazard ratios were estimated by separately changing the value of each variable, holding the other variables constant. The baseline survival probability was determined using the hazard ratio estimates we obtained and the online 5-year survival predictions for a large group of subjects. The form of the model is the following for the ith patient: the hazard of death at time t is

$$\lambda_i(t) = \lambda_0(t)\exp(\beta_{\text{age}} \cdot \text{age}_i + \beta_{\text{gender}} \cdot \text{gender}_i + \beta_{\text{T}} \cdot \text{Tstage}_i + \beta_{\text{N}} \cdot \text{Nstage}_i + \beta_{\text{location}} \cdot \text{location}_i + \beta_{\text{hemoglobin}} \cdot \text{hemoglobin}_i + \beta_{\text{EQD2T}} \cdot \text{EQD2T}_i)$$

eTable 5: Parameter Values in Maastro Calculator Cox proportional hazards model

Covariate	Hazard Ratio	beta=ln(HR)
Age	1.05	0.045
Gender		
Female (ref)	1	0
Male	2.38	0.865
T Stage		
T1 (ref)	1	0
T2	1.13	0.121
T3	1.96	0.675

T4	3.60	1.279
N Stage		
N0 (ref)	1	0
N+	1.44	0.364
Location		
Glottic (ref)	1	0
Non-Glottic	1.31	0.266
Hemoglobin Level (in mmol/L)	0.73	-0.321
EQD2*T*	0.99	-0.004

The estimated baseline survival probability, $S_0(5) = 0.6545$ represents the 5-year survival probability for a female subject with T1N0, glottic cancer with value 0 for all continuous variables in the model (including age). This survival probability, therefore, does not have much practical meaning. The 5-year survival is then estimated using the equation $S(5) = S_0(5)^{\exp(X\beta)}$ where $\exp(X\beta)$ is the exponential term in the hazard rate expression above.

Section 2: Treatment of Missing Data

Tumor diameter, hemoglobin, and EQD2T were not available for some of the patients. We developed a set of rules for determining reasonable values to assign each patient with missing values. For patients missing tumor diameter, diameters of 1cm, 2cm, 3cm, and 4cm were assigned for T1, T2, T3, and T4 patients respectively. This was consistent with the pattern of diameters seen for subjects with both diameter and T stage observed. We also had a few (18%) subjects missing values for hemoglobin. We used single imputation to fill in the missing values of hemoglobin based on a regression model of hemoglobin that included gender, comorbidities, age, main tumor site, CA stage, SEER stage, and initial treatment as covariates.

Missing values of EQD2T were handled using as much of the available information as possible. EQD2T is a function of total radiation dose, dose per fraction, and treatment time. EQD2T can only be measured in subjects that received radiation. For our primary analyses, we calculate EQD2T as if all subjects received radiation. We note that analyses restricted to subjects that actually received radiation produced similar results. For subjects missing total radiation dose, dose per fraction, and treatment time, a total dose of 63 was assigned to T1N0Glottic patients and a total dose of 70 was assigned for all others. Dose per fraction was assigned to be 2Gy and then treatment time was assigned a reasonable value based on the relationship between duration of treatment and (total dose/dose per fraction) in the observed data. For subjects with observed total dose and either dose per fraction or duration of treatment missing (but not both), we set the missing variable to take a reasonable value based on the observed data. For subjects with total dose observed and missing values for dose per fraction and duration, we assign dose per fraction to 2Gy and set duration to a reasonable value based on the observed data. It is worth noting that the coefficient in the Maastro model corresponding to EQD2T is near zero, so the value of EQD2T (within a reasonable range) will have little impact on the final 5-year survival prediction. Therefore, we do not expect the missing (and subsequently imputed) EQD2T values to have much impact.

Utilization of chemotherapy after definitive treatment was not available for some subjects, and we therefore developed a set of rules to incorporate adjuvant chemotherapy. A subject receiving no induction chemotherapy and surgery as his/her initial treatment plan was listed as having had adjuvant chemotherapy if the subject had extracapsular spread and/or positive margins.

Section 3: Reconstructed Model Performance

It was not necessary to characterize the model (determine the exact model form and all parameter values) for the various calculators, but it was very convenient. To check the accuracy of the reconstructed calculators we compared the results of the online calculator (when available) with our program. Thirty patients with larynx cancer were randomly drawn from the University of Michigan dataset, and their information was used to predict 5-year survival using both the web calculator and our reconstructed models. We did not include the LifeMath model as the exact form of the model and all associated parameter values were available, allowing for exact duplication of the online calculator.

eTable 6 shows the predicted survival probabilities from the website and our reconstructed calculator for the Maastro, LifeMath, and Leiden calculators. The results show that the Leiden calculator has some minor prediction errors for a small number of patients, the maximum absolute value of prediction error (the maximum error in the "Difference" column) between the online prediction and a rounded version of the reconstructed prediction is 1% in 5-year survival probability. The reason for this difference is likely due to rounding error as the publication only provides the coefficient values to 2 significant digits. Similarly, the maximum absolute value of prediction error for the Maastro calculator is 1% in 5-year survival probability. A greater number of patients have an absolute difference equal to 1 for the Maastro calculator than the Leiden calculator, which is likely due to issues of rounding or truncation of the

predictions for the online calculator. MyCancerJourney calculator also has minor prediction errors for larynx cancer patients, the maximum absolute value of prediction error is 0.3%.

eTable 6: Comparison of 5-Year Survival Predictions from the Reconstructed Models and Web Calculators

Maastro[a]			Leiden			MyCJ		
Web	Recon	Difference	Web	Recon	Difference	Web	Recon	Difference
10	9	1	26	25	1	9.1	9.2	-0.1
67	66	1	77	77	0	50	50.3	-0.3
47	47	0	66	66	0	16.7	16.8	-0.1
10	10	0	53	53	0	21.6	21.8	-0.2
29	30	-1	38	38	0	36	36.2	-0.2
61	61	0	62	63	-1	40.2	40.4	-0.2
12	12	0	24	23	1	0.8	0.8	0.0
35	35	0	43	43	0	32.8	32.8	0.0
20	20	0	46	46	0	51.8	51.9	-0.1
34	35	-1	54	54	0	58.8	59.0	-0.2
<5	2	NA	49	49	0	59.2	59.4	-0.2
<5	1	NA	27	27	0	34.4	34.6	-0.2
67	67	0	75	75	0	72	71.9	0.1
56	56	0	76	76	0	67.6	67.6	0.0
<5	3	NA	47	47	0	28	28.1	-0.1
35	36	-1	54	54	0	40.6	40.7	-0.1
35	35	0	42	42	0	21.1	21.2	-0.1
<5	2	NA	21	21	0	15.5	15.6	-0.1
23	23	0	50	50	0	38.8	38.9	-0.1
72	72	0	77	77	0	77	77.0	0.0
16	17	-1	22	22	0	40.6	40.6	0.0
23	24	-1	49	49	0	26.9	27.0	-0.1
9	10	-1	50	50	0	58.4	58.5	-0.1
<5	5	NA	31	30	1	0.5	0.5	0.0
19	19	0	49	49	0	41.3	41.4	-0.1
42	42	0	47	46	1	28.3	28.6	-0.3
50	51	-1	52	52	0	48.7	48.8	-0.1
25	25	0	65	65	0	40.8	40.9	-0.1
91	91	0	91	91	0	83.5	83.5	0.0
<5	1	NA	46	46	0	29.3	29.3	0.0

a The values in the table correspond to the 5-year survival probability predictions. The "Web" column lists the (rounded) predicted 5-year survival probability from the corresponding website and "Recon" column lists the 5-year survival probability predicted using the reconstructed models. The differences between the online value and the rounded version of the reconstructed value are also included. All the values in this table have unit % and represent probabilities.

Section 4: Summary of Models and Pairwise Agreement

eTable 7 summarizes the variables included in each of the four larynx cancer models. We also take a look at the pairwise agreement between the calculators. For each subject in our dataset, we obtain a calculator prediction of 5-year survival using each of the calculators. eTable 8 shows the proportion of predictions that are within 10% of one another.

eTable 7: Summary of Input Variables for Larynx Cancer Calculator

	Calculator	MAASTRO	Lifemath	Leiden	MyCJ
Demographic	Age	Y	Y	Y	Y
	Sex	Y		Y	Y
	Race		Y		Y
	Comorbidity			Y	Y
	Prior Tumors			Y	
Site	Laryngeal Subsites	Y		Y	
	Other HNC		Y	Y	Y
Staging	Histology				Y
	Grade				Y
	Tumor Diameter		Y		
	T	Y		Y	
	N	Y	Y	Y	
	Number of Positive Nodes		Y		
	M			Y	
	SEER Stage				Y
Other	Hemoglobin	Y			
	Extracapsular Spread		Y		
	EQD2T	Y			
	Treatment				Y

eTable 8: Proportion of Survival Predictions within 10% for Pairs of Calculators[a]

	LifeMath	Leiden	MyCJ	Mean[b]
Maastro	18.2%	23.1%	30.4%	32.1%
LifeMath		70.3%	39.4%	55.2%
Leiden			39.8%	58.9%
MyCJ				60.9%

[a] Eg. For a predicted survival of 0.5, this is between 0.4 and 0.6.

b The Mean is the arithmetic average of the LifeMath, Leiden, Maastro, and MyCJ 5-year survival probability predictions

Section 5: Cox Regressions using Calculator Predictions

We are interested in identifying subjects in our cohort that demonstrate improved survival compared to the calculator predictions. We define S to be the mean of the estimated 5-year survivals for the Leiden and LifeMath calculators. e Table 9 shows Cox regressions of Overall Survival (censored at 70 months). Each row corresponds to a separate Cox regression adjusting for $-\log(-\log(S))$ and the corresponding covariate. The results suggest that, even adjusting for the predicted survivals, we have lower rates of death in subjects that are female compared to male and subjects that receive "No Induction and Surgery" or "Induction Chemo" compared to "No Induction and Chemoradiation" as their initial treatment plan.

eTable 9: Cox Model Regression Results: Importance of Each Factor After Adjusting for the Leiden and LifeMath Predictions

Covariate	Covariate HR (95% CI)[a]	Covariate P-value	−log(−log(S)) HR (95% CI)
T Stage		0.676	
T1 (ref)	1		
T2	1.35 (0.61, 2.97)		0.25 (0.14, 0.43)
T3	0.94 (0.44, 2.04)		
T4	0.88 (0.39, 1.95)		
Gender		0.031	
Female (ref)	1		
Male	2.03 (1.00, 4.11)		0.26 (0.16, 0.44)
Comorbidities		0.734	
None (ref)	1		
Mild	1.05 (0.49, 2.26)		0.29 (0.17, 0.48)
Moderate	1.33 (0.59, 2.97)		
Severe	1.48 (0.58, 3.73)		
N Stage		0.469	
N0 (ref)	1		
N1	1.10 (0.52, 2.33)		0.22 (0.12, 0.41)
N2/N3	0.71 (0.38, 1.34)		
Initial Treatment Plan		0.024	
No Induction, Chemoradiation	1		0.24 (0.14, 0.41)
No Induction, Radiation Only	0.97 (0.49, 1.88)		
No Induction, Surgery	0.43 (0.21, 0.90)		
Induction Chemo	0.46 (0.24, 0.88)		
Age		0.448	
10 Years Increase (ref)	1.10 (0.85, 1.44)		0.30 (0.17, 0.53)
Race		0.347	
White (ref)	1		
Black	1.70 (0.61, 4.68)		0.28 (0.17, 0.46)
Other	0.39 (0.05, 2.88)		
SEER Stage		0.982	
Local (ref)	1		
Regional	0.94 (0.49, 1.78)		0.26 (0.15, 0.47)
Distant	0.96 (0.47, 1.97)		
CA Stage		0.547	
I (ref)	1		
II	1.86 (0.74, 4.71)		0.24 (0.13, 0.45)
III	1.22 (0.51, 2.90)		
IV	1.06 (0.47, 2.38)		
Subsite			
Glottic (ref)	1		
Supraglottic	0.89 (0.52, 1.52)	0.671	0.26 (0.15, 0.44)
Grade		0.732	
Well Differentiated (ref)	1		
Moderately Differentiated	1.33 (0.50, 3.54)		0.29 (0.17, 0.49)
Poorly	1.66 (0.58, 4.74)		
Unknown	1.58 (0.56, 4.43)		

a If the covariate is significantly associated with overall survival even after adjusting for the 5-year survival predictions, this suggests that the calculator predictions are not fully capturing the covariate effect on the outcome.

Section 6: Creating the Calibration Plots

In order to assess how well each of the calculator's predictions corresponded to the observed data, we construct calibration plots. We use two different methods.

Method 1: For each calculator, we separate subjects into 5 groups based on quintiles of the calculator's predicted 5-year survival probabilities. Within each group, we estimate the 5-year survival probability from the observed data for subjects in that group. We plot the average calculator predictions for each of the groups against the estimated survival probabilities from the observed data. Points closer to the y=x line correspond to better calibration.

Method 2: Method 1 only breaks the subjects into 5 groups for each calculator, which results in only 5 points from which to evaluate calibration. Together, these points provide a rough idea of the relationship between the calculator predictions and the Kaplan-Meier predictions. Alternatively, we might identify a smoother and more refined version of the relationship between the calculator and Kaplan-Meier 5-year survival predictions by breaking the predictions into many overlapping groups rather than 5 non-overlapping groups. The overlapping groups are determined based on the subjects with calculator-predicted 5-year survival probabilities within a moving window of 0.25. Within each group, we plot the average predicted 5-year survival probability from the calculator against the Kaplan-Meier-predicted 5-year survival probability for the group. This is a form of a kernel smoother and results in a smoothed version of the calibration plot. We use this approach to identify patient characteristics in our cohort that may be associated with improved survival over the calculator predictions.

Section 7: Maastro Predictions Restricted to Cohort Receiving Radiation

The Maastro calculator was created and validated for subjects receiving radiation only (no surgery or chemotherapy), but our primary analysis applies the Maastro calculator to some subjects that did not receive radiation and patients receiving radiation in addition to other treatments. Therefore, we are applying the Maastro calculator to a population it was not designed for, which does not present a "level playing field" for the various calculators. Therefore, we are interested in comparing the performance of the calculators on the subset of patients receiving radiation only (and no induction chemo) as the initial treatment plan (N=55).

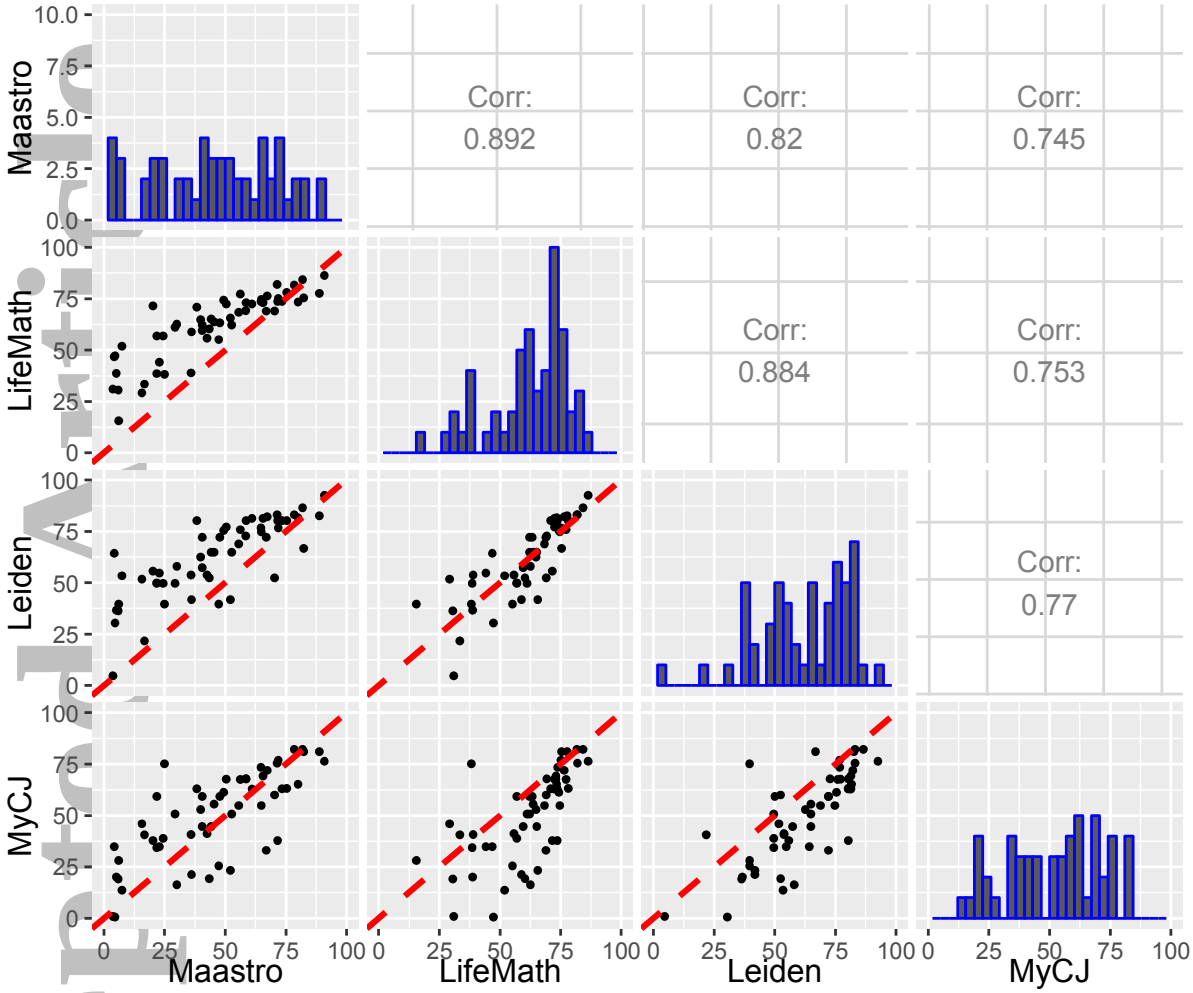
eTable 10 shows the proportion of predictions that are within 10% of one another. The Maastro calculator is more similar to the other calculators when applied to the "radiation only" subset of the population. eFigure 1 shows the comparison of the predicted survivals across calculators. Maastro is more strongly correlated to the other calculators, but it still tends to underestimate the survival probabilities relative to the other calculators. eFigure 2 shows calibration curves for each one of the calculators using only the data from the radiation only subset of the data. Similar to the analysis of the entire sample, the Maastro calculator underestimates the survival probabilities observed in our data. The Leiden and LifeMath calculators appear to have better calibration than the Maastro calculator for the radiation only subset. eFigure 3 displays the resulting ROC curves. Here, the results differ from the results obtained from analysis of the entire study sample. Using the radiation only subset of the data, we see that the Maastro calculator has substantially improved AUC and C Index values.

eTable 10: Proportion of Survival Predictions within 10% for Pairs of Calculators, Restricted to subjects receiving radiation only[a]

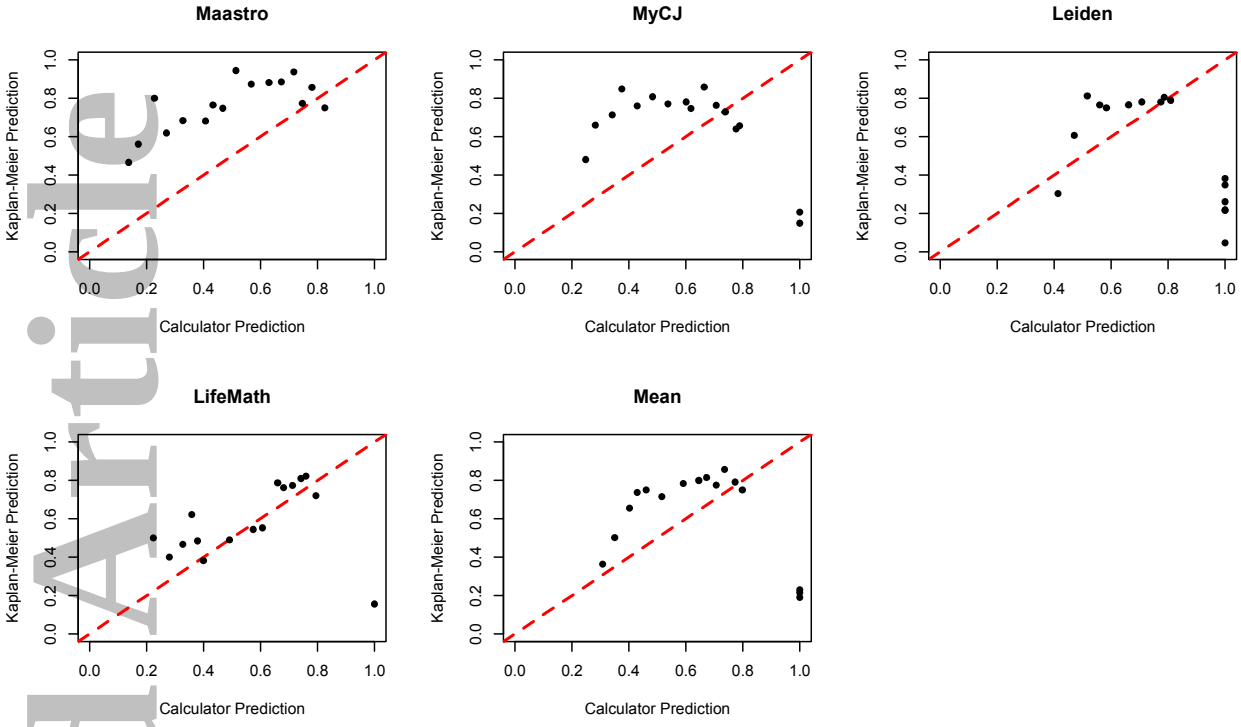
	LifeMath	Leiden	MyCJ	Mean
Maastro	34.5 %	29.0%	38.1%	50.9%
LifeMath		72.7%	45.4%	76.3%
Leiden			29.0%	61.8%
MyCJ				69.0%

[a] Eg. For a predicted survival of 0.5, this is between 0.4 and 0.6.

eFigure 1: Comparing Predicted Survivals Across Calculators using Subjects Receiving Radiation Only

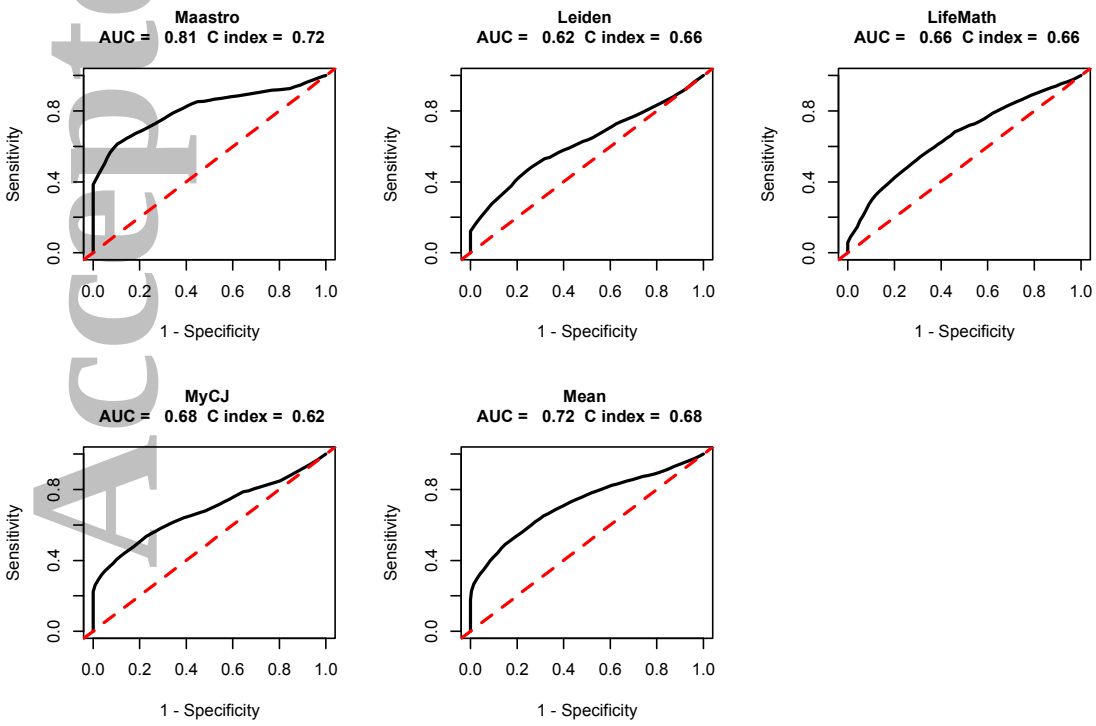


eFigure 2: Calibration Curves using Subjects Receiving Radiation Only[a]



a Calibration curves estimated using smoothing method as described in Section 6 of the Supplementary Materials

eFigure 3: 5-year ROC Curves using Subjects Receiving Radiation Only[a]



a The ROC curve illustrates the ability of each calculator to discriminate subjects with higher and lower 5-year survival. The AUC is the area under the ROC survival curve, with values closer to 1 indicated greater predictive ability. The C-Index (Concordance Index) is a measure of how well the predicted survival probabilities rank subjects with respect to higher and lower observed risk.

References

1. Prince V, Bellile EL, Sun Y, et al. Individualized risk prediction of outcomes for oral cavity cancer patients. *Oral Oncol*. 2016;63(1):66-73.
2. Datema FR, Ferrier MB, van der Schroeff MP, Baatenburg de Jong RJ. Impact of comorbidity on short-term mortality and overall survival of head and neck cancer patients. *Head Neck.* 2010; 32(6):728-736.
3. Datema FR, Ferrier MB, Vergouwe Y, et al. Update and external validation of a head and neck cancer prognostic model. *Head Neck*. 2013;35(9):1232-1237.
4. Emerick KS, Leavitt ER, Michaelson JS, Diephuis B, Clark JR, Deschler DG. Initial clinical findings of a mathematical model to predict survival of head and neck cancer. *Otolaryngol Head Neck Surg*. 2013;149(4):572-578.
5. Egelmeer AGTM, Velazquez ER, Jong JMA, et al. Development and validation of a nomogram for prediction of survival and local control in laryngeal carcinoma patients treated with radiotherapy alone: a cohort study based on 994 patients. *Radiother Oncol*. 2011;100(1):108-115.