

UCSF

UC San Francisco Previously Published Works

Title

Nomogram to Predict the Benefit of Intensive Treatment for Locoregionally Advanced Head and Neck Cancer

Permalink

<https://escholarship.org/uc/item/9b4469zc>

Journal

Clinical Cancer Research, 25(23)

ISSN

1078-0432

Authors

Mell, Loren K
Shen, Hanjie
Nguyen-Tân, Phuc Felix
[et al.](#)

Publication Date

2019-12-01

DOI

10.1158/1078-0432.ccr-19-1832

Peer reviewed



Published in final edited form as:

Clin Cancer Res. 2019 December 01; 25(23): 7078–7088. doi:10.1158/1078-0432.CCR-19-1832.

Nomogram to Predict the Benefit of Intensive Treatment for Locoregionally Advanced Head and Neck Cancer

Loren K. Mell¹, Hanjie Shen¹, Phuc Felix Nguyen-Tân², David I. Rosenthal³, Kaveh Zakeri¹, Lucas K. Vitzthum¹, Steven J. Frank³, Peter B. Schiff⁴, Andrea M. Trotti III⁵, James A. Bonner⁶, Christopher U. Jones⁷, Sue S. Yom⁸, Wade L. Thorstad⁹, Stuart J. Wong¹⁰, George Shenouda¹¹, John A. Ridge¹², Qiang E. Zhang¹³, Quynh-Tu Le¹⁴

¹Department of Radiation Medicine and Applied Sciences, University of California San Diego, La Jolla, California

²Department of Radiation Oncology, Centre Hospitalier de l'Université de Montreal, Montreal, Quebec, Canada

³Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX

⁴Department of Radiation Oncology, New York University School of Medicine, New York, NY

⁵Department of Radiation Oncology, Moffitt Cancer Center, Tampa, Florida

⁶Department of Radiation Oncology, Hazelrigg-Salter Radiation Oncology Center, University of Alabama at Birmingham, Birmingham

⁷Radiological Associates of Sacramento, Sacramento, California

⁸Department of Radiation Oncology, University of California, San Francisco, San Francisco, California

⁹Department of Radiation Oncology, Washington University School of Medicine, St. Louis, Missouri

¹⁰Division of Hematology Oncology, Department of Medicine, Medical College of Wisconsin, Milwaukee, Wisconsin

¹¹Department of Radiation Oncology, McGill University Health Centre, Montréal, Québec, Canada

¹²Department of Surgical Oncology, Fox Chase Cancer Center, Philadelphia, Pennsylvania.

¹³NRG Oncology Statistics and Data Management Center, Philadelphia, Pennsylvania

¹⁴Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA

Abstract

Corresponding Author: Loren K. Mell, M.D., Department of Radiation Medicine and Applied Sciences, 3855 Health Sciences Drive, MC0843, La Jolla, CA 92093, lmell@ucsd.edu, (858) 246-0471 / fax (858) 822-5568.

Notes / Conflicts of Interest: None of the authors has any conflicts of interest. The sponsors had no role in the design of the study, the collection, analysis, and interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication.

Background—Previous studies indicate the benefit of therapy depends on patients' risk for cancer recurrence relative to non-cancer mortality (ω ratio). We sought to test the hypothesis that head and neck cancer (HNC) patients with a higher ω ratio selectively benefit from intensive therapy.

Methods—We analyzed 2688 patients with stage III-IVB HNC undergoing primary radiation therapy (RT) with or without systemic therapy on three phase III trials (RTOG 9003, RTOG 0129, and RTOG 0522). We used generalized competing event regression to stratify patients according to ω ratio, and compared the effectiveness of intensive therapy as a function of predicted ω ratio (i.e., ω score). Intensive therapy was defined as treatment on an experimental arm with altered fractionation (AFX) and/or multiagent concurrent systemic therapy. A nomogram was developed to predict patients' ω score based on tumor, demographic, and health factors. Analysis was by intention-to-treat.

Results—Decreasing age, improved performance status, higher body mass index, node positive status, P16 negative status, and oral cavity primary predicted a higher ω ratio. Patients with ω score ≥ 0.80 were more likely to benefit from intensive treatment (5-year OS, 70.0% vs. 56.6%; HR 0.73, 95% CI 0.57–0.94; $P=0.016$) than those with a ω score < 0.80 (5-year OS, 46.7% vs. 45.3%; HR 1.02, 95% CI 0.92–1.14; $P=0.69$; $P=0.019$ for interaction). In contrast, the effectiveness of intensive therapy did not depend on risk of progression.

Conclusion—HNC patients with a higher ω score selectively benefit from intensive treatment. A nomogram was developed to help select patients for intensive therapy.

Introduction

While the effectiveness of intensive therapy (e.g., concurrent chemotherapy or altered fractionation) for locoregionally advanced head and neck cancer has been established, there is considerable controversy surrounding which subsets of patients are most likely to benefit from this approach. In particular, the effectiveness of intensive therapy in patients who are older, or who have comorbidities, or who have relatively favorable risk disease (e.g., human papilloma virus (HPV)-associated disease, non-smokers) is unclear.

Traditionally, risk-stratification models used in cancer outcomes research have focused on the effects of treatments and risk factors on endpoints such as overall survival (OS) or progression-free survival (PFS). A problem is that these endpoints do not differentiate effects on primary events, such as disease recurrence or cancer mortality, from competing events, such as death from comorbid illness. As a result, such models are suboptimal, because they pool patients at high risk for cancer events with patients at high risk for competing events, even though these groups have different expected benefit from intensive therapy.^{1–9} Thus, staging systems and nomograms that predict for OS and PFS are likely to be suboptimal for selecting head and neck cancer (HNC) patients for intensive therapeutic regimens.

Previous studies indicate that in patients with competing risks, the hazard for cancer relative to competing mortality events (i.e., ω ratio) is a key determinant of treatment benefit.^{7,9–11} In particular, older patients with higher ω ratios may be good candidates for more intensive therapy; conversely, younger patients with lower ω ratios may not be. Further work is

needed to define factors that critically affect ω ratios and correlate them with treatment effects. Correspondingly, newer methods have been developed to quantify effects on the ω ratio, with considerable improvement in risk stratification compared to standard models.^{10–13} However, it is not known whether the benefit of more intensive treatment varies according to ω ratio, and in particular whether this is a more effective method to predict which patients are most likely to benefit from intensive treatment. The goal of this study was to develop a model to identify locally advanced HNC patients with a higher ω ratio and to test the hypothesis that such patients selectively benefit from treatment intensification.

Materials and Methods

Population, Sampling Methods, and Treatment

We studied 2688 patients with locoregionally advanced (stage III-IVB) HNC treated on three clinical trials: RTOG 9003 (), RTOG 0129 (), and RTOG 0522 (). Details of these protocols have been published previously.^{14–17} Written informed consent was obtained for all patients. The study was conducted in accordance with recognized ethical guidelines and was approved by the institutional review boards at all participating institutions.

Briefly, patients on RTOG 9003 were randomized to one of four arms: hyperfractionated radiotherapy (HFX: 81.6 Gy in 68 fractions BID over 7 weeks), delayed concomitant boost radiotherapy (DCB: 72 Gy in 42 fractions over 6 weeks), split course radiotherapy (SC: 67.2 Gy in 42 fractions over 6 weeks), or standard fractionation (SFX: 70 Gy in 35 fractions over 7 weeks). For the purpose of this analysis, HFX and DCB were considered altered fractionation (AFX), whereas SC and SFX were not. Patients on RTOG 9003 did not receive chemotherapy. Patients on RTOG 0129 were randomized to either AFX or SFX and received chemotherapy (two cycles of cisplatin 100 mg/m² weeks 1 and 4 of chemoradiotherapy for patients receiving AFX, and three cycles at the same dose weeks 1, 4 and 7 for patients receiving SFX). Patients on RTOG 0522 were randomized to either cetuximab (400 mg/m² loading dose followed by 250 mg/m² weekly) vs. no cetuximab, and all patients received AFX (six fractions per week) along with concurrent cisplatin (two cycles of cisplatin 100 mg/m² weeks 1 and 4 of chemoradiotherapy). All human investigations were performed after approval by a local Human Investigations Committee and in accordance with an assurance filed with and approved by the Department of Health and Human Services.

Outcomes

Progression-free survival time was defined as the time from randomization to the first recurrence of disease, or death from any cause, or censoring. Overall survival time was defined as the time from randomization to death from any cause, or censoring. Time to recurrence and time to cancer-specific mortality were defined as the time from randomization to first recurrence (or cancer-related mortality), with competing mortality events treated as censored. Time to competing mortality for recurrence was defined as time from randomization to death from any cause, in the absence of a recurrence event, with recurrence events treated as censored. Correspondingly, time to competing mortality for cancer mortality was defined as time from randomization to death from any cause, in the absence of a cancer mortality event, with cancer mortality events treated as censored.

Statistical Analysis

The study followed TRIPOD guidelines.¹⁸ The statistical approach involved two main steps: (1) development of a model to separate patients by ω ratio and (2) validation of the model as a method to predict treatment effects (i.e., variation in treatment effects as a function of ω). Overall survival was used as the primary outcome assessment for model validation, since this endpoint was not used in model development and represents an outcome of clear clinical benefit to patients.

Kaplan-Meier functions were used to plot PFS and OS and cumulative incidence functions were used to plot competing events with respect to time. The *basehaz* function in R (version 3.4.2) was used to estimate cumulative hazards. Forest plots were used to analyze treatment effects within risk strata, according to intention-to-treat. Proportional hazards assumptions were tested using the Grambsch-Therneau method (*cox.zph* function in R).

We trained risk scores for recurrence, competing mortality, and PFS using data from the control arms from the three studies (Supplementary Figure 1), based on the linear predictor from a multivariable Cox proportional hazards regression.¹⁹ For RTOG 9003, the SC and SFX arms were collectively considered the control group. For the multivariable models, we selected the following candidate variables for inclusion, based on their availability in all three trials and potential association with disease recurrence^{15–17,22–23} and/or competing mortality^{2–4,11,19–21}: age (per 10 years; continuous), female sex, black/African-American race (vs. other), white/Caucasian race (vs. other), body mass index (BMI) (< 20 kg/m² vs. > 20), ECOG performance status (0 vs. 1–2), marital status (married vs. other/unknown), anemia (yes/no), education history (any college vs. other/unknown) – as a proxy for socioeconomic status (SES), primary site (oral cavity vs. oropharynx vs. hypopharynx vs. larynx), T stage (0–2 vs. 3 vs. 4), and N stage (0 vs. 1–2a vs. 2b–2c vs. 3). Anemia was defined for males as a baseline hemoglobin < 13.5 g/dL and for females as a baseline hemoglobin < 12.5 g/dL. For patients with known smoking and tumor P16 status, we included pack-years (< 10 vs. > 10) and P16 (positive vs. negative) as covariates. P16 was analyzed as a prognostic factor for both oropharyngeal and non-oropharyngeal sites, based on several studies that have found differences in outcomes by P16 or HPV status in both oropharyngeal^{22–23} and non-oropharyngeal HNC.^{24–26}

All variables were normalized by subtracting the sample mean and dividing by the sample standard deviation. The mean BMI value was imputed for 194 patients with missing data, using single imputation. Risk scores based on the linear predictor were generated taking the inner product of the coefficient vector with the individual patient's data vector, as described previously.¹¹ Risk strata were defined according to quantiles of the risk score distribution. We compared results with a standard model developed to stratify oropharyngeal cancer patients.^{22–23} Note that this model also stratifies non-oropharyngeal cancer patients (Supplementary Figure 2).

For modeling effects of covariates, we used generalized competing event (GCE) regression based on a proportional relative hazards model.^{11,27–28} A detailed description of the GCE modeling approach is provided below. In brief, the ratio of the cause-specific hazard for recurrence (λ_1) vs. the cause-specific hazard for competing mortality (λ_2) is represented as

ω^+ , whereas the ratio of the cause-specific hazard for recurrence (λ_1) to the hazard for any progression-free event ($\lambda_1 + \lambda_2$) is represented as ω . We use the terms ω and ω^+ *ratio* to refer to observed values, whereas ω and ω^+ *score* refer to values of ω and ω^+ , respectively, predicted by the GCE model.

For GCE regression, the same variables were used as in the Cox proportional hazards models, after normalization. Separate regression models were built for both cause-specific events (i.e., disease recurrence) and competing mortality (i.e., death in the absence of disease recurrence, with the cause-specific event treated as censored). Treatment-related deaths were classified as competing mortality events. To test the sensitivity of our conclusions to model specification, reduce overfitting, and facilitate clinical implementation, we generated a parsimonious GCE model using backward stepwise regression to exclude variables from the regression if $p > 0.20$, and by consolidating N stage (0 vs. 1–3). In this model, only age, performance status, BMI, oral cavity site, N stage, and P16 status had $p < 0.20$ and thus were retained in the final nomogram, which was trained on the subset of controls with known P16 status (N=602).

GCE risk scores were generated by taking the inner product of the (normalized) individual patient's data vector with the difference of the coefficient vector for cause-specific events and competing mortality. For 95% confidence intervals of estimates, we employed the *gcerisk* package in R.²⁷ Risk strata were defined according to quantiles of the GCE risk score distribution. Tests of treatment effects and interactions included random effects for study and age.²⁹ All P values are 2-sided.

Generalized Competing Event (GCE) Model

For mutually exclusive events of type k , we posit the following proportional relative hazards model:

$$\omega^+_{k}(t|X) = \omega^+_{k0}(t) \exp(\beta^+_{k\ GCE} X) \quad [1]$$

where

$$\omega^+_{k0}(t) = \lambda_{k0}(t) / \sum_{j \neq k} \lambda_{j0}(t) \quad [2]$$

Here, $\lambda_{k0}(t)$ is the baseline hazard for an event of type k , $\sum_{j \neq k} \lambda_{j0}(t)$ is the baseline cause-specific hazard for the set of events competing with event type k , X is a vector of covariates, and $\beta^+_{k\ GCE}$ is the vector of effects (coefficients) on the covariates. From this model it can be shown that

$$\beta^+_{k\ GCE} = (\beta_k - \beta_{j \neq k}) \quad [3]$$

where β_k and $\beta_{j \neq k}$ represent effects on the baseline hazard for event type k and competing events, respectively, from the Cox proportional hazard model. We use $\hat{\beta}^+_{k\ GCE} = \hat{\beta}_k - \hat{\beta}_{j \neq k}$ as the estimator for $\beta^+_{k\ GCE}$ and $\hat{\omega}^+_{k0}(t) = \hat{\Lambda}_{k0}(t) / \hat{\Lambda}_{j \neq k0}(t)$, where $\hat{\Lambda}_{k0}(t)$ and $\hat{\Lambda}_{j \neq k0}(t)$ represent

the Nelson-Aalen estimators¹⁸ for the cumulative hazard for event type k and the set of competing events at time t , respectively. We estimate the predicted value of $\hat{\omega}_k^+(t|\mathbf{d})$ for an individual with given data vector \mathbf{d} as:

$$\hat{\omega}_k^+(t|\mathbf{d}) = \hat{\omega}_{k0}^+(t) \exp(\hat{\beta}_{k\text{GCE}}^+ \cdot \mathbf{d}) \quad [4]$$

Note then that $\exp(\hat{\beta}_{k\text{GCE}}^+)$ is the estimate of the ω^+ ratio, which quantifies how the relative hazards for primary and competing events change in response to changes in covariates.

We define the omega value as the ratio of the hazard for an event of type k to the hazard for all events:

$$\omega_k(t) = \lambda_k(t)/\lambda(t) \quad [5]$$

and estimate the predicted omega value as

$$\hat{\omega}_k(t|\mathbf{d}) = \hat{\omega}_k^+(t|\mathbf{d}) / (1 + \hat{\omega}_k^+(t|\mathbf{d})) \quad [6]$$

Note that while $\hat{\omega}_k$ ranges from 0 to 1 inclusive, $\hat{\omega}_k^+$ ranges from 0 to ∞ . For $k=2$, a value of $\hat{\omega}_1^+ = 1$ means the hazard for event type 1 equals the hazard for event type 2, and therefore $\hat{\omega}_1 = \hat{\omega}_2 = 0.5$. For the purpose of this study, we defined ω^+ as the ratio of the hazard for disease recurrence to the hazard for competing mortality in the absence of recurrence, and ω as the ratio of the hazard for disease recurrence to the hazard for recurrence or death from any cause. All values of ω are unscaled unless otherwise specified. Scaled estimates were obtained by factoring out the baseline ω^+ values.

Sample Size Estimates

We used the power calculator described by Pintilie³⁸ to estimate sample sizes for hypothetical randomized trials with a primary endpoint of PFS, assuming balanced randomization, accrual time of 3 years, follow-up time of 2 years, two-sided $\alpha=0.05$, and $\beta=0.20$. We considered two events: cancer recurrence ($k=1$) and competing mortality ($k=2$), and assumed a hazard ratio for cancer recurrence (θ_1) of 0.5 and a hazard ratio for competing mortality (θ_2) of 1. Under varying ω_1 values, we allowed the hazard ratio for any event (θ) to vary according to the equation

$$\theta = \omega_1 * \theta_1 + (1 - \omega_1) = (\omega_1^+ * \theta_1 + 1) / (\omega_1^+ + 1) \quad [7]$$

Final GCE Risk Score

R functions to define the GCE risk score and scaled predicted ω are:

```
exp.risk.score = function(AGE,BMI,ECOG12,OC,N0,P16){ exp(-.3693*((.1*AGE-5.72)/.
9126)+.2044*(((BMI>20)-.88538)/.31883)-.2262*((ECOG12-.377076)/.4851)+.
1684*((OC-.03488)/.18364)-.1274*((N0-.14452)/.3519)-.2147*((P16-.488372)/.5))}
```

```
scaled.omega.predicted = function(AGE,BMI,ECOG12,OC,N0,P16)
{exp.risk.score(AGE,BMI,ECOG12,OC,N0,P16)/
(exp.risk.score(AGE,BMI,ECOG12,OC,N0,P16)+1)}
```

```
omega.score = function(AGE,BMI,ECOG12,OC,N0,P16)
{2.6*exp.risk.score(AGE,BMI,ECOG12,OC,N0,P16)/
(2.6*exp.risk.score(AGE,BMI,ECOG12,OC,N0,P16)+1)}
```

For this calculation, AGE is in years, BMI is in kg/cm², ECOG12 is 1 if ECOG performance status is > 0 and 0 otherwise, OC is 1 for oral cavity tumors and 0 otherwise, N0 is 1 if there is no nodal involvement and 0 otherwise, and P16 is 1 for P16 positive tumors and 0 otherwise. The factor 2.6 is the mean baseline ω^+ estimate from the control sample.

Results

Sample characteristics are provided in Supplementary Table 1. Comparisons of model estimates for the entire control group and the subset with known smoking history and P16 status appear in Table 1. Factors predicting a higher ratio were decreasing age, improved performance status, higher body mass index, node positive status, P16 negative status, and oral cavity primary. It is interesting to compare and contrast effect estimates from Cox vs. GCE models. While patients with poorer OS or PFS are typically identified as candidates for more intensive treatment, the GCE model indicates that patients with advanced age, poorer performance status, hypopharynx site, and advanced T category, for example, have a reduced hazard for cancer events *relative to competing mortality*, implying that such patients are relatively less likely to benefit from treatment intensification. Moreover, some factors, such as N3 category, and marital, education, and smoking status, are attenuated in the GCE model due to offsetting effects on recurrence and competing mortality.

Compared to standard models, GCE models improved stratification according to ω ratio within each risk group, with increasing ω from low risk to high risk according to both model predictions and observations (Supplementary Table 2). Agreement between predicted ω (i.e., ω score) and observed ω ratios was high, indicating excellent model fit and validity. The observed 3-year ω and ω^+ ratios for the whole cohort were 0.719 and 2.56, respectively. The observed 3-year ω and ω^+ ratios for the subset with known p16 and smoking status were 0.738 and 2.87, respectively.

As shown in Figure 1, OS differed markedly across risk groups defined by standard models (A,C), whereas GCE models show little correspondence between OS and risk level when risk is defined by ω score (B,D). This suggests that, paradoxically, patients with a better predicted survival (and higher ω score) could be *more* likely to benefit from intensive treatment (by virtue of being much less likely to die from non-cancer causes). This is further shown in Figure 2, which plots the cumulative incidences of cancer recurrence and competing mortality within risk groups. Note that with standard risk-stratification models,

the probability of both cancer and non-cancer mortality are increased in the highest risk strata relative to the GCE model, while the converse is true of the lower risk strata, further supporting GCE model validity. This is because while standard models are designed to separate groups according to PFS and OS, GCE models are designed to optimize the ratio of competing events in order to favor a particular event of interest.

Patients with the highest ω score (≥ 0.80) – representing the highest quintile - were more likely to benefit from intensive treatment (5-year OS, 70.0% vs. 56.6%; HR 0.73, 95% CI 0.57–0.94; Wald P=0.016) than those with ω score < 0.80 (5-year OS, 46.7% vs. 45.3%; HR 1.02, 95% CI 0.92–1.14; Wald P=0.69) (P=0.019 for interaction). For patients with known P16 status, the GCE nomogram similarly identified a statistically significant benefit from treatment intensification in patients with ω score ≥ 0.80 (HR 0.67; 95% CI, 0.47–0.95, Wald P=0.027); in contrast, we did not find statistically significant treatment effects in the high-risk subgroups defined by standard models overall (Figure 3), or in any of the trials separately. Treatment intensification was also associated with statistically significant improvement in OS in patients with ω score ≥ 0.80 in the RTOG 9003 trial separately (Supplementary Figure 3). These results appeared robust over a range of potential cut points near the ω score of 0.80 (Supplementary Figure 4A–B). Calibration plots showed excellent discriminatory ability, with better fitting at higher predicted ω values (Supplementary Figure 4C). A nomogram for calculating an individual's ω score appears in Figure 4.

Model estimates and performance were similar when patients with missing BMI data were omitted from the analysis. We found evidence of efficiency gains with the GCE model relative to standard models under varying definitions of “high-risk” (Table 2), due to the higher ratio of primary to competing events. However, this analysis does not account for efficiency loss that could result from a lower event rate. While the incidence of competing mortality was lower in the high-risk group using the GCE model, the lower incidence of cancer recurrence offset some of the efficiency gains, indicating correlation between primary and competing events. As such, GCE models could be less efficient than models designed to predict recurrence, but this conclusion was sensitive to the lack of P16 status for the majority of the cohort. It is noteworthy, however, that sample size estimates were similar with the various approaches, despite a marked reduction in the overall event rate in the “high-risk” group defined by the GCE model.

Discussion

In this study we found that intensive treatment differentially benefits patients with a higher relative recurrence risk (ω score ≥ 0.80). Previous studies involving HNC and other disease sites have found that ω scores could be used to identify patients with a greater likelihood to benefit from intensive therapy.^{3,10–13} The present study is the first to examine treatment effects within risk groups defined by this factor. We found evidence to support the hypothesis that relative recurrence risk is an important predictor of treatment effectiveness in the HNC population.

Advantages of this study were its large sample size and large number of known predictors of both cancer-related and competing events. Randomization also mitigated the impact of

selection bias, which presents a problem studying treatment effects in other data sources. A limitation of this study, however, is the heterogeneity in the treatment and populations across the three trials. “Intensive treatment” was defined relative to the baseline (control) group, and thus included altered fractionation (with or without concurrent chemotherapy), or chemoradiotherapy with concurrent targeted therapy, depending on the trial. We were also unable to control directly for some predictors, such as comorbidity and SES, that likely would have helped optimize the model relative to standard approaches. For example, income was prognostic for both cancer recurrence and competing mortality, but had to be omitted since it was not collected for all trials. While previous studies have found increased survival for patients undergoing treatment at high-volume centers,^{30–32} radiotherapy quality in RTOG trials is considered to be high. The incidence of non-cancer mortality in this cohort was also lower than has been observed in prior studies,² indicating the exclusion of many patients at risk for competing events.

Variation in the definition of “intensive treatment” is a potential limitation of our study; however, the intent was to compare the effectiveness of intensity with respect to the control arm, and our results were unaffected whether we applied fixed or random effects models. Data from large randomized trials or meta-analyses involving homogeneous treatments (in particular, with an established survival benefit) will be important for further model validation. Note that both RTOG 0129 and RTOG 0522 failed to reject the null hypothesis; in the absence of effective therapy, it is not possible to identify subpopulations that would benefit. Future studies involving more trials that met their primary endpoint would be helpful to determine how treatment effects and toxicity vary with the ω ratio. However, we did observe a survival advantage with altered fractionation in the high-risk group from RTOG 9003, lending support to the hypothesis that ω score is a useful predictive marker.

Interpreting the effects of particular covariates in this study should be undertaken with caution, since confidence intervals were fairly wide (leading to some differences in interpretation across samples). Lack of consistency and incomplete collection of key prognostic variables hampers efforts to compare risk models, requiring us to re-train multivariable models in new samples; however, GCE models have previously been validated in population-based studies.^{10,11} It should also be noted that age cutoffs ≤ 50 (and >70) have been previously associated with a selective benefit (or lack thereof) of treatment intensification in HNC, including the RTOG 0522 trial included in this analysis.^{16,33–34} However, age as a sole criterion for treatment selection is generally not favored,³⁵ since other health factors can influence the appropriate intensity of therapy. In this study, age ≤ 50 was not predictive of a treatment benefit in the whole cohort.

GCE regression is a modeling approach with clear differences relative to standard risk-stratification methods. It contrasts with other nomograms^{23,36–37} in that instead of predicting patients’ risk for event-free survival, which is preferable for prognostication, the GCE model seeks to predict the ratio of cancer events to competing events, which is considered preferable as a predictive model. Further studies are required to establish its advantages over standard methods, especially in the postoperative setting and the larger population not participating in trials, who we expect would have differing risk for competing events. An important limitation is that the cutoff of 0.80 for the ω score, although robust, was not

chosen *a priori*; our results should thus be considered hypothesis generating, and should be validated in future studies. Ascertaining optimal cutoffs to define “high-risk” groups remains an area of investigation, especially with models controlling for comorbidity and other geriatric/frailty assessments. Perhaps most interestingly, our findings suggest that a higher absolute risk for recurrence/progression does not necessarily confer a higher likelihood to benefit from intensive therapy (or greater power to detect treatment effects). This is because patients with a low risk for both recurrence and competing mortality may benefit as much from aggressive treatment approaches as patients with high risk for both events.

In summary, here we propose a method to predict an under-reported but meaningful quantity for individual patients (i.e., relative recurrence risk, or ω ratio), with a clinically relevant interpretation (i.e., a value > 50% means the individual’s hazard for cancer recurrence exceeds the hazard for competing mortality). Our findings indicate patients with a higher relative recurrence risk, indicated by a ω score > 0.80, selectively benefit from intensive therapy. This approach is being implemented prospectively in the NRG-HN004 trial, along with a nomogram to inform clinical practice and trial design (comogram.org). Further research, however, is needed to optimize GCE models and to ascertain which patients derive the greatest benefit from intensive therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by grants U10CA180868 (NRG Oncology Operations), U10CA180822 (NRG Oncology SDMC) from the National Cancer Institute (NCI) and Eli Lilly.

Financial Support: This project was supported by grants U10CA180868 (NRG Oncology Operations), U10CA180822 (NRG Oncology SDMC) from the National Cancer Institute (NCI) (NRG Oncology/RTOG 9003, , <https://clinicaltrials.gov/ct2/show/NCT00771641>, NRG Oncology/RTOG 0129, , <https://clinicaltrials.gov/ct2/show/NCT00047008>, and NRG Oncology/RTOG 0522, , <https://clinicaltrials.gov/ct2/show/NCT00265941>) and Eli Lilly.

Disclaimers: None: Dr(s). Ridge, Rosenthal, Shen, Shenouda, Schiff, Trotti, Vitzthum, Wong, and Zakeri. Dr. Bonner discloses honoraria and consulting or advisory roles from Bristol-Myers Squibb Company, Eli Lilly and Company, Merck Serono and Cel-Sci, and research funding from Bristol-Myers Squibb Company. Dr. Frank discloses a leadership role, stock or other ownership at C4 Imaging, honoraria from Varian and Hitachi, consulting or advisory role and travel, accommodations or expenses with Varian. Dr. Jones discloses participation in a speakers bureau with Lilly. Dr. Le discloses research funding from Amgen, Varian and Redhill and travel, accommodations or expenses from BMS (consultant role). Dr. Mell discloses honoraria from BMS, Pfizer, and Varian, consulting or advisory roles from BMS, Pfizer, and Merck, research funding from Merck, and travel, accommodations, or expenses from BMS and Merck. Dr. Nguyen discloses consulting or advisory roles from Bristol Myers Squibb (advisory board). Dr. Thorstad discloses familial employment at Elekta. Dr. Yom discloses honoraria from Lilly, and research funding from Genentech, Merck, Bristol-Myers Squibb. Dr. Zhang discloses stock or other ownership with Pfizer.

References

1. Argiris A, Brockstein BE, Haraf DJ, et al. Competing causes of death and second primary tumors in patients with locoregionally advanced head and neck cancer treated with chemoradiotherapy. *Clin Cancer Res* 2004;10:1956–1962. [PubMed: 15041712]
2. Mell LK, Dignam JJ, Salama JK, et al. Predictors of competing mortality in advanced head and neck cancer. *J Clin Oncol* 2010;28:15–20. [PubMed: 19933920]

3. Rose BS, Jeong JH, Nath SK, Lu SM, Mell LK. Population-based study of competing mortality in head and neck cancer. *J Clin Oncol* 2011;29:3503–3509. [PubMed: 21844503]
4. Kwon M, Roh JL, Song J, et al. Noncancer health events as a leading cause of competing mortality in advanced head and neck cancer. *Ann Oncol* 2014;25:1208–14. [PubMed: 24669018]
5. Mell LK, Weichselbaum RR. More on cetuximab in head and neck cancer. *N Engl J Med* 2007;357:2201–2202. [PubMed: 18032774]
6. Dignam JJ, Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. *J Clin Oncol* 2008;26:4027–4034. [PubMed: 18711194]
7. Mell LK, Jeong JH. Pitfalls of using composite primary end points in the presence of competing risks. *J Clin Oncol* 2010;28:4297–4299. [PubMed: 20713860]
8. Mell LK, Zakeri K, Rose BS. On lumping, splitting, and the nosology of clinical trial populations and end points. *J Clin Oncol* 2014;32:1089–1090. [PubMed: 24550413]
9. Mell LK, Carmona R, Gulaya S, Lu T, Wu J, Saenz CC, Vaida F. Cause-specific effects of radiotherapy and lymphadenectomy in stage I-II endometrial cancer: a population-based study. *J Natl Cancer Inst* 2013;105:1656–66. [PubMed: 24123960]
10. Carmona R, Gulaya S, Murphy JD, et al. Validated competing event model for the stage I-II endometrial cancer population. *Int J Radiat Oncol Biol Phys* 2014;89:888–98. [PubMed: 24969798]
11. Carmona R, Zakeri K, Green G, et al. Improved method to stratify elderly patients with cancer at risk for competing events. *J Clin Oncol* 2016;34:1270–7 [PubMed: 26884579]
12. Zakeri K, Rose BS, D'Amico AV, Jeong JH, Mell LK. Competing events and costs of clinical trials: Analysis of a randomized trial in prostate cancer. *Radiother Oncol* 2015;115:114–9. [PubMed: 25857696]
13. Zakeri K, Rose BS, Gulaya S, D'Amico AV, Mell LK. Competing event risk stratification may improve the design and efficiency of clinical trials: secondary analysis of SWOG 8794. *Contemp Clin Trials* 2013;34:74–9. [PubMed: 23063467]
14. Fu KK, Pajak TF, Trotti A, Jones CU, Spencer SA, Phillips TL, et al. A Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003. *Int J Radiat Oncol Biol Phys* 2000;48:7–16. [PubMed: 10924966]
15. Nguyen-Tân PF, Zhang Q, Ang KK, Weber RS, Rosenthal DI, Soulieres D, et al. Randomized phase III trial to test accelerated versus standard fractionation in combination with concurrent cisplatin for head and neck carcinomas in the Radiation Therapy Oncology Group 0129 trial: long-term report of efficacy and toxicity. *J Clin Oncol*. 2014;32:3858–66 [PubMed: 25366680]
16. Ang KK, Zhang Q, Rosenthal DI, Nguyen-Tân PF, Sherman EJ, Weber RS, et al. Randomized phase III trial of concurrent accelerated radiation plus cisplatin with or without cetuximab for stage III to IV head and neck carcinoma: RTOG 0522. *J Clin Oncol* 2014;32(27):2940–50. [PubMed: 25154822]
17. Beitler JJ, Zhang Q, Fu KK, Trotti A, Spencer SA, Jones CU, et al. Final results of local-regional control and late toxicity of RTOG 9003: a randomized trial of altered fractionation radiation for locally advanced head and neck cancer. *Int J Radiat Oncol Biol Phys* 2014;89:13–20. [PubMed: 24613816]
18. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Br J Cancer*. 2015;112(2):251–9. [PubMed: 25562432]
19. Cox DR. Regression models and life tables. *J R Stat Soc. Series B Stat Methodol* 1972;B34:187–220.
20. Zakeri K, MacEwan I, Vazirnia A, Cohen EE, Spiotto MT, Haraf DJ, et al. Race and competing mortality in advanced head and neck cancer. *Oral Oncol*. 2014;50(1):40–4. [PubMed: 24134948]
21. Park A, Alabaster A, Shen H, Mell LK, Katznel JA. Undertreatment of women with locoregionally advanced head and neck cancer. *Cancer* 2019; in press.

22. Ang KK, Harris J, Wheeler R, Weber RS, Rosenthal DI, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010;363:24–35. [PubMed: 20530316]
23. Fakhry C, Zhang Q, Nguyen-Tân PF, Rosenthal DI, Weber RS, Lambert L, et al. Development and validation of nomograms predictive of overall and progression-free survival in patients with oropharyngeal cancer. *J Clin Oncol.* 2017;35(36):4057–4065 [PubMed: 28777690]
24. Chung CH, Zhang Q, Kong CS, Harris J, Fertig EJ, Harari PM, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol.* 2014 12 10;32(35):3930–8. [PubMed: 25267748]
25. Bryant AK, Sojourner EJ, Vitzthum LK, Zakeri K, Shen H, Nguyen C, et al. Prognostic Role of p16 in Nonoropharyngeal Head and Neck Cancer. *J Natl Cancer Inst.* 2018;110(12):1393–1399. [PubMed: 29878161]
26. Tian S, Switchenko JM, Jhaveri J, Cassidy RJ, Ferris MJ, Press RH, et al. Survival outcomes by high-risk human papillomavirus status in nonoropharyngeal head and neck squamous cell carcinomas: A propensity-scored analysis of the National Cancer Data Base. *Cancer.* 2019; in press
27. Shen H, Carmona R, Mell LK. Generalized Competing Event Model: gcerisk R package. R (CRAN), <https://cran.r-project.org/>.
28. Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics.* 1995;51:524–32. [PubMed: 7662841]
29. Michiels S, Baujat B, Mahé C, Sargent DJ, Pignon JP. Random effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. *J Clin Epidemiol* 2005;58(3):238–45. [PubMed: 15718112]
30. Wuthrick EJ, Zhang Q, Machtay M, Rosenthal DI, Nguyen-Tan PF, Fortin A, et al. Institutional clinical trial accrual volume and survival of patients with head and neck cancer. *J Clin Oncol.* 2015;33(2):156–64. [PubMed: 25488965]
31. Boero IJ, Paravati AJ, Xu B, Cohen EE, Mell LK, Le QT, Murphy JD. Importance of radiation oncologist experience among patients with head-and-neck cancer treated with intensity-modulated radiation therapy. *J Clin Oncol.* 2016;34(7):684–90. [PubMed: 26729432]
32. David JM, Ho AS, Luu M, Yoshida EJ, Kim S, Mita AC, et al. Treatment at high-volume facilities and academic centers is independently associated with improved survival in patients with locally advanced head and neck cancer. *Cancer.* 2017;123(20):3933–3942. [PubMed: 28640546]
33. Bourhis J, Overgaard J, Audry H, Ang KK, Saunders M, Bernier J, et al.; Meta-Analysis of Radiotherapy in Carcinomas of Head and neck (MARCH) Collaborative Group. Hyperfractionated or accelerated radiotherapy in head and neck cancer: a meta-analysis. *Lancet.* 2006;368(9538): 843–54. [PubMed: 16950362]
34. Pignon JP, le Maître A, Maillard E, Bourhis J; MACH-NC Collaborative Group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol* 2009;92(1):4–14. [PubMed: 19446902]
35. Wildiers H, Mauer M, Pallis A, Hurria A, Mohile SG, Luciani A, et al. End points and trial design in geriatric oncology research: a joint European organisation for research and treatment of cancer--Alliance for Clinical Trials in Oncology--International Society Of Geriatric Oncology position article. *J Clin Oncol* 2013;31(29):3711–8. [PubMed: 24019549]
36. Wang SJ, Patel SG, Shah JP, Goldstein DP, Irish JC, Carvalho AL, et al. An oral cavity carcinoma nomogram to predict benefit of adjuvant radiotherapy. *JAMA Otolaryngol Head Neck Surg.* 2013;139(6):554–9. [PubMed: 23680917]
37. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015;16(4):e173–80. [PubMed: 25846097]
38. Pintilie M *Competing Risks: A Practical Perspective* (John Wiley & Sons, Chichester, United Kingdom). 2006:115–126.

STATEMENT OF TRANSLATIONAL RELEVANCE

The effectiveness of intensive therapy for head and neck cancer (HNC) patients who are older, or who have comorbidities, or who have favorable risk HPV-associated disease is unclear. Traditional risk-stratification models pool patients at high risk for cancer events with patients at high risk for competing events, even though these groups have different expected benefit from intensive therapy. Studies indicate that the hazard for cancer recurrence *relative* to competing mortality (ω ratio) is a key determinant of treatment benefit, with newer regression methods developed to quantify effects on this ratio. This is the first study to examine the effectiveness of intensive therapy for HNC as a function of ω ratio. We found that patients with a predicted $\omega > 0.80$ had improved overall survival with intensive therapy, using pooled data from three randomized controlled trials.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

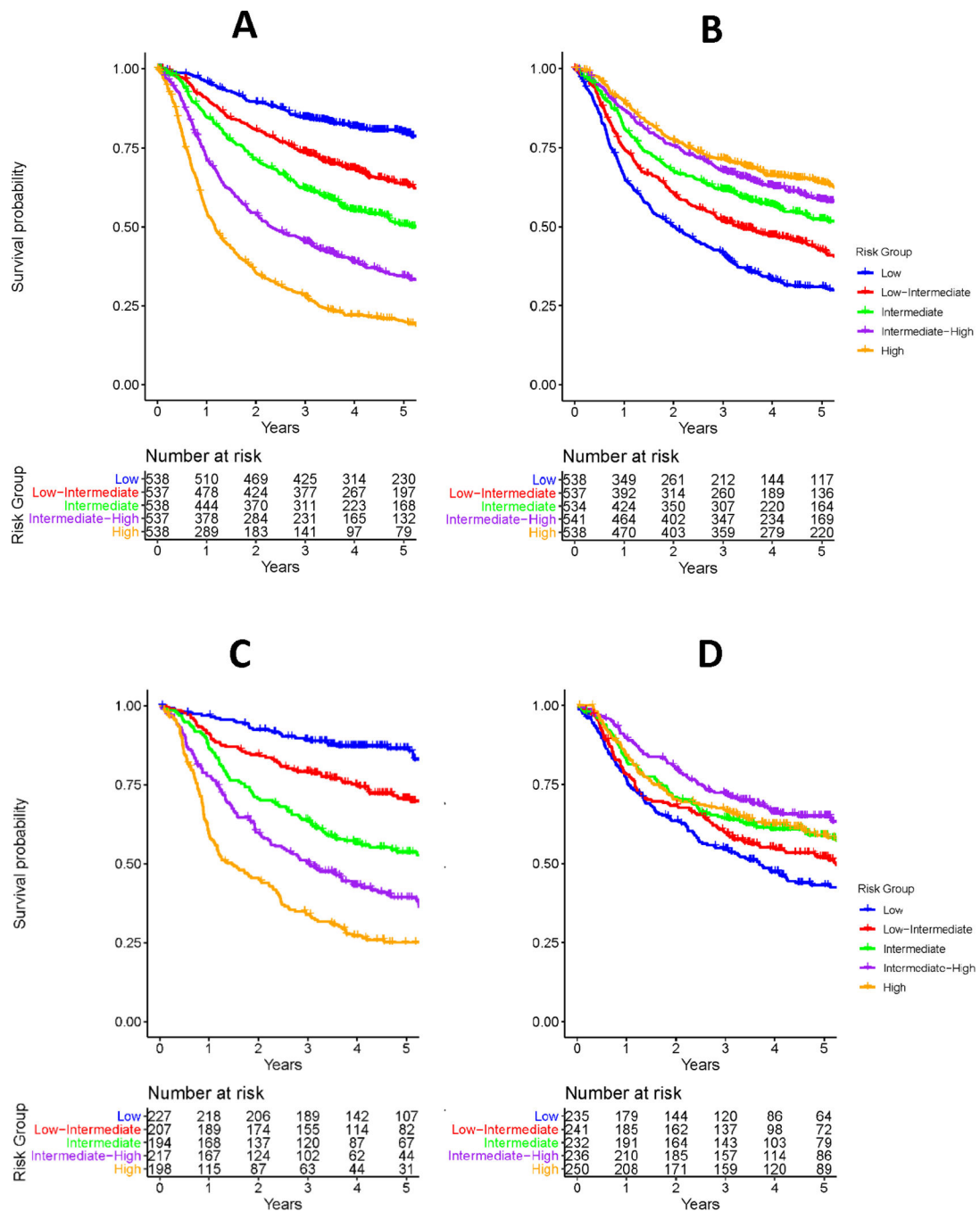


Figure 1. Overall Survival, by Risk Strata. **(A)** Cox Model in the Whole Cohort. **(B)** Generalized Competing Event (GCE) Model in the Whole Cohort. **(C)** Fakhry et al. Nomogram²¹ in Patients with Known Smoking History and P16 Status. **(D)** GCE Nomogram in Patients with Known P16 Status.

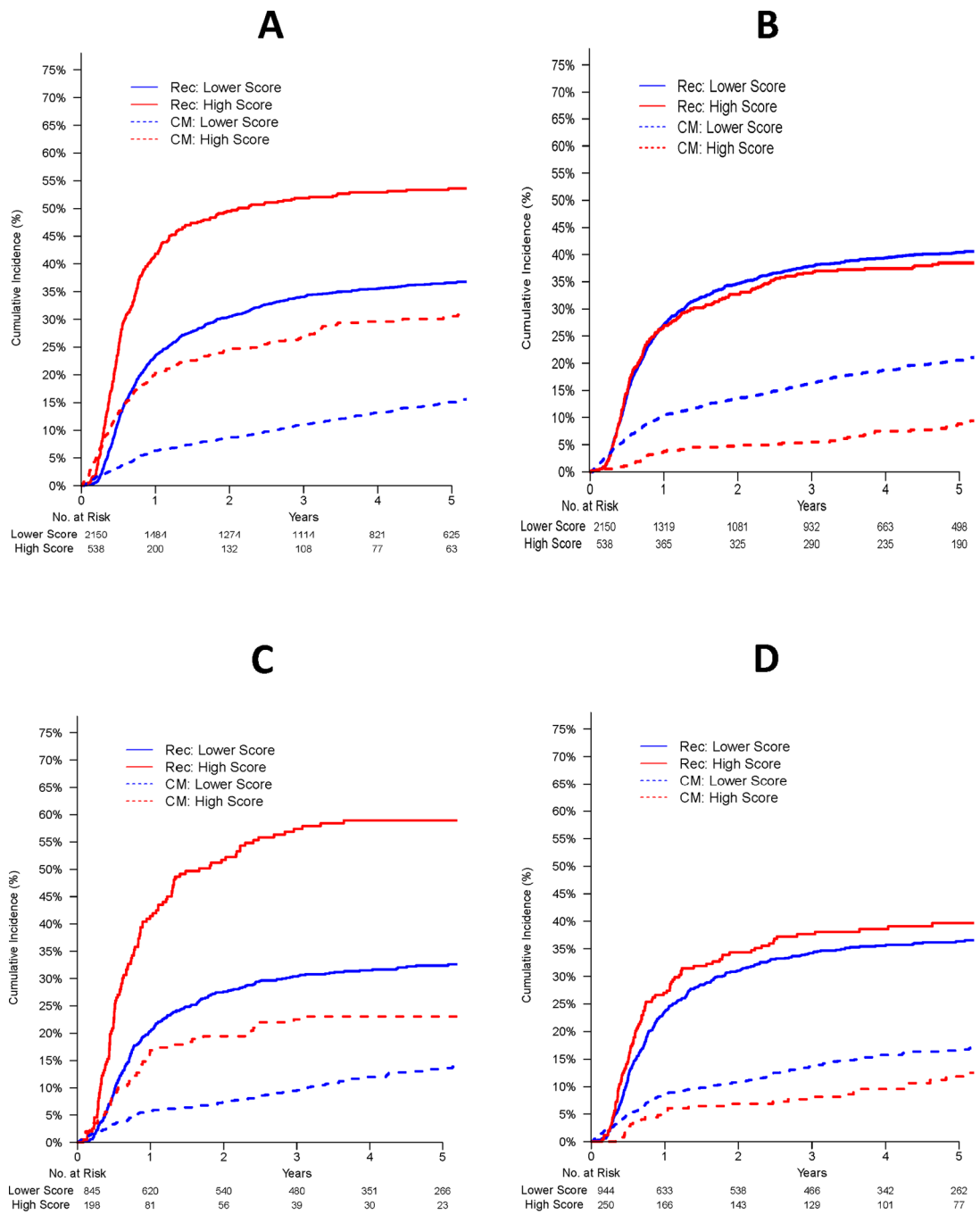
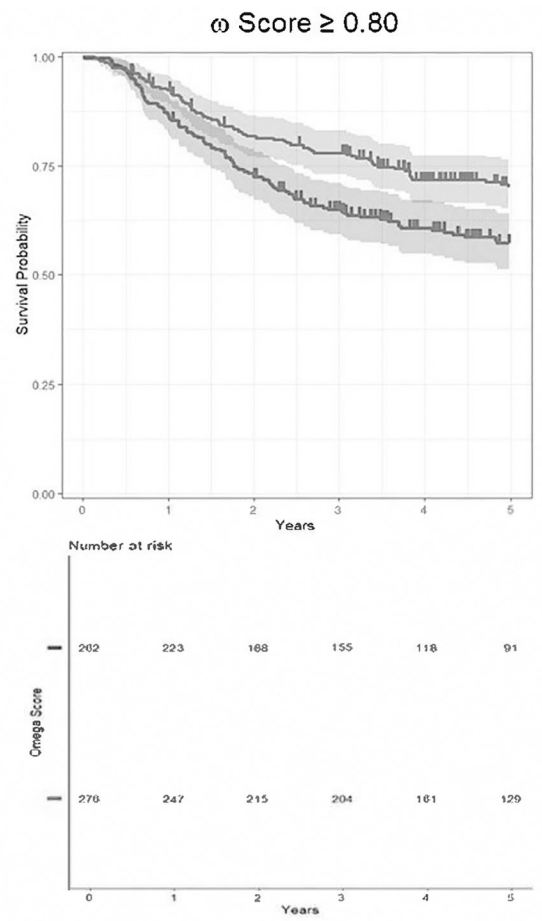
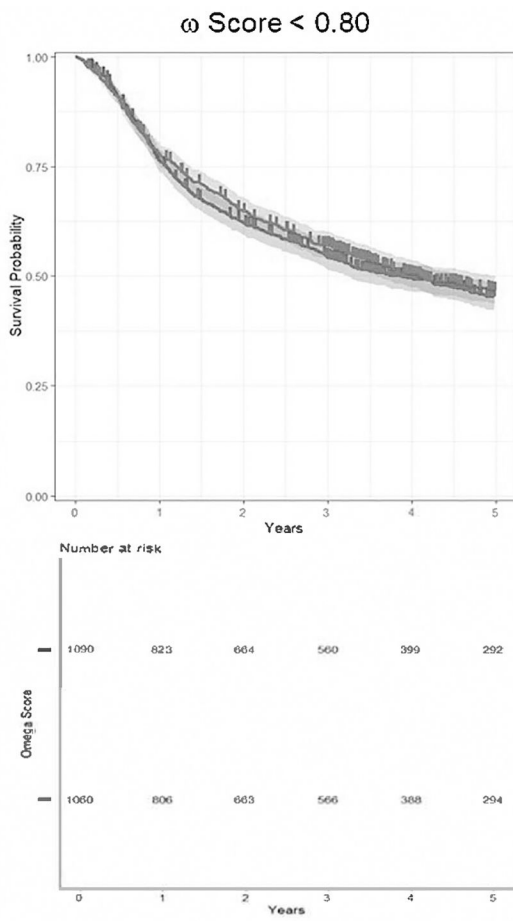


Figure 2. Competing Event Incidences, by Risk Score. **(A)** Cox Model in the Whole Cohort. **(B)** Generalized Competing Event (GCE) Model in the Whole Cohort. **(C)** Fakhry et al. Nomogram²¹ in Patients with Known Smoking History and P16 Status. **(D)** GCE Model in Patients with Known P16 Status.

A



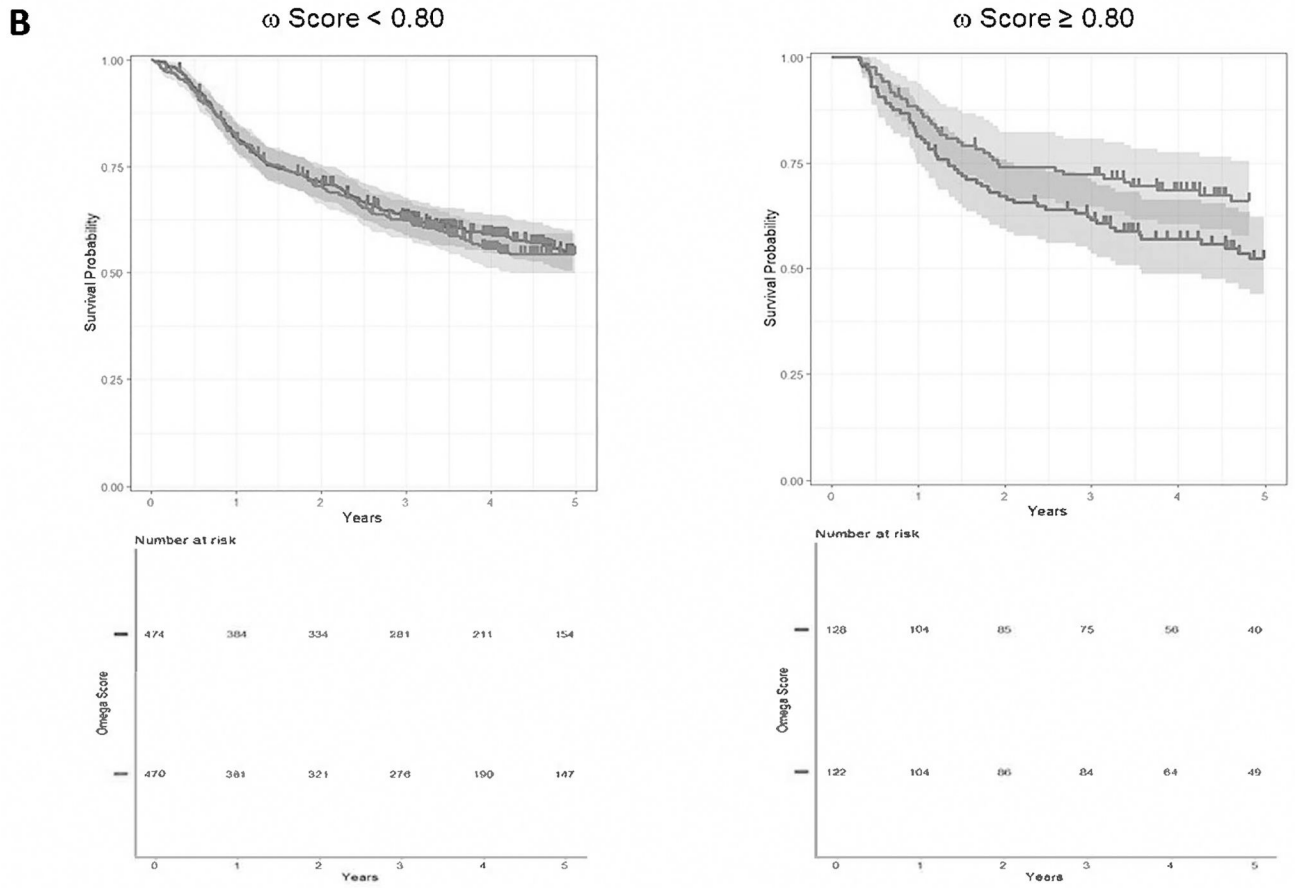


Figure 3. Interaction between experimental therapy and ω score. (A) Whole Cohort. Left: ω score < 0.80; Right: ω score \geq 0.80 (B) Patients with Known P16 Status. Left: ω score < 0.80; Right: ω score \geq 0.80.

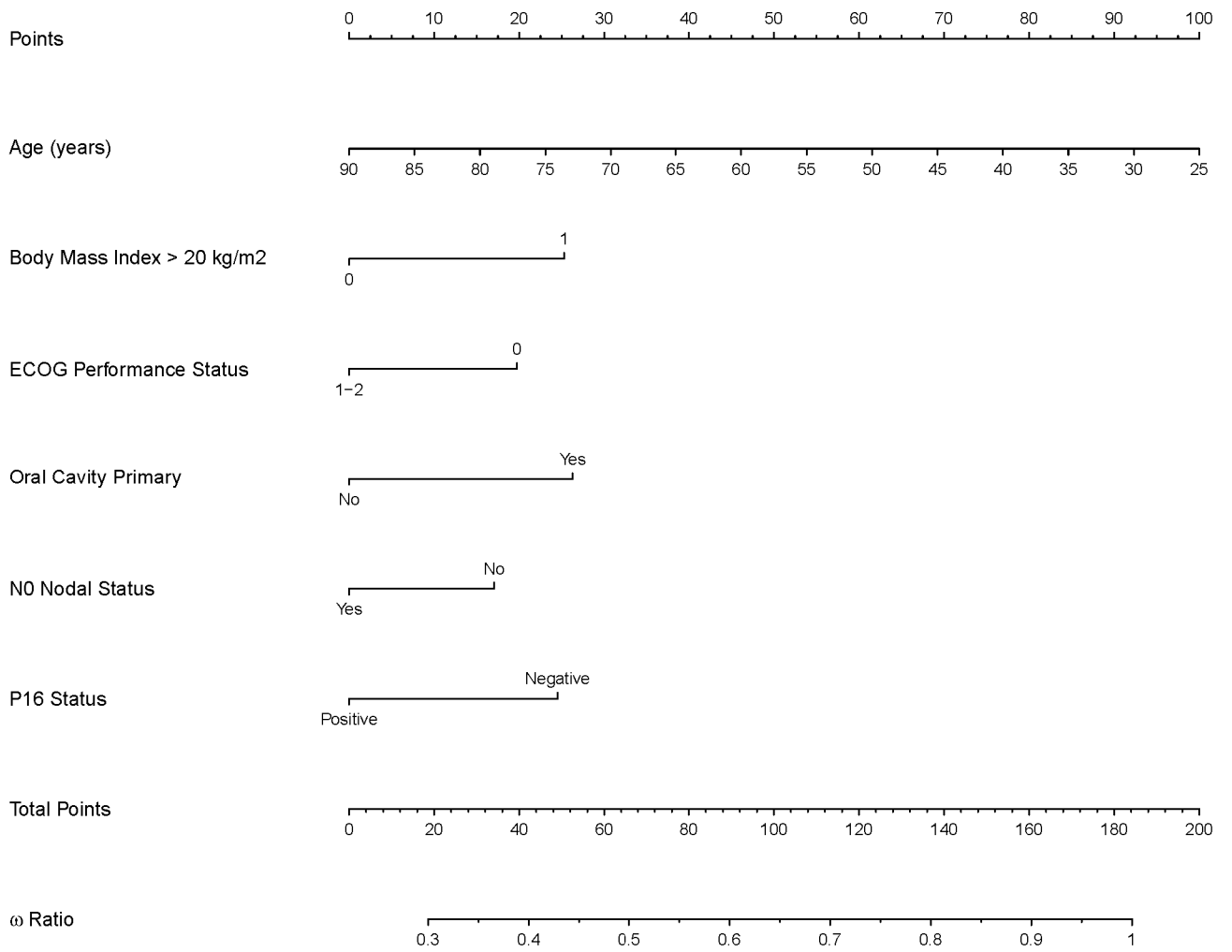


Figure 4. Nomogram to predict patients' relative hazard for recurrence based on generalized competing event (GCE) regression model.

Table 1.

Comparison of Cox vs. GCE models in All Controls (Left Columns) and Complete Cases with Known Smoking and P16 Status (Right Columns).

| Characteristic | All Controls (N=1352) | | Subset with Known Smoking and P16 Status (N=527) | |
|---|-----------------------|----------------------------------|--|----------------------------------|
| | Cox PH Regression | GCE Regression | Cox PH Regression | GCE Regression |
| | HR* (95% CI) | ω^+ Ratio (RHR)** (95%CI) | HR* (95% CI) | ω^+ Ratio (RHR)** (95%CI) |
| Age at diagnosis, per 10 years [¶] | 1.42 (1.31, 1.53) | 0.66 (0.57, 0.77) | 1.37 (1.19, 1.58) | 0.67 (0.51, 0.88) |
| Sex | | | | |
| Female vs. Male | 0.84 (0.70, 1.01) | 1.06 (0.74, 1.51) | 0.94 (0.68, 1.29) | 0.94 (0.51, 1.75) |
| Race | | | | |
| Black | 1.04 (0.74, 1.47) | 0.51 (0.25, 1.06) | 0.65 (0.35, 1.22) | 0.88 (0.25, 3.10) |
| White | 0.80 (0.60, 1.09) | 0.42 (0.22, 0.80) | 0.59 (0.35, 1.00) | 0.83 (0.29, 2.41) |
| Non-Black / Non-White | REF | REF | REF | REF |
| Body Mass Index [¶] | | | | |
| 20 kg/m ² vs. > 20 kg/m ² | 0.55 (0.46, 0.67) | 1.03 (0.71, 1.49) | 0.76 (0.54, 1.09) | 1.66 (0.83, 3.32) |
| ECOG Performance Status [¶] | | | | |
| 1–2 vs. 0 | 1.35 (1.16, 1.58) | 0.49 (0.36, 0.67) | 1.54 (1.19, 2.01) | 0.59 (0.34, 1.00) |
| Anemia | | | | |
| Yes vs. No/Unknown | 1.11 (0.94, 1.30) | 0.79 (0.58, 1.08) | 0.92 (0.69, 1.21) | 0.97 (0.56, 1.67) |
| Married | | | | |
| Yes vs. No/Unknown | 0.75 (0.66, 0.88) | 0.99 (0.74, 1.32) | 0.77 (0.59, 1.00) | 1.16 (0.70, 1.93) |
| Education History | | | | |
| Any College / Vocational / Technical vs. None/Unknown | 0.60 (0.51, 0.71) | 0.98 (0.71, 1.36) | 0.57 (0.43, 0.77) | 0.90 (0.52, 1.55) |
| Anatomic subsite | | | | |
| Oropharynx | REF | REF | REF | REF |
| Larynx | 1.16 (0.96, 1.39) | 1.06 (0.73, 1.53) | 0.93 (0.66, 1.30) | 1.15 (0.59, 2.24) |
| Hypopharynx | 1.65 (1.33, 2.04) | 0.77 (0.50, 1.18) | 1.72 (1.16, 2.55) | 0.87 (0.39, 1.91) |
| Oral Cavity [¶] | 1.47 (1.13, 1.91) | 2.55 (1.40, 4.62) | 2.11 (1.22, 3.64) | 2.46 (0.62, 9.77) |
| T Stage | | | | |
| 0–2 | REF | REF | REF | REF |
| 3 | 1.06 (0.88, 1.28) | 0.85 (0.59, 1.23) | 0.87 (0.63, 1.21) | 0.84 (0.44, 1.61) |
| 4 | 1.53 (1.26, 1.86) | 0.77 (0.52, 1.13) | 1.45 (1.04, 2.01) | 0.84 (0.44, 1.62) |
| N Stage | | | | |
| 0 [¶] | REF | REF | REF | REF |
| 1–2a | 1.26 (1.01, 1.57) | 1.36 (0.88, 2.10) | 1.23 (0.82, 1.84) | 1.53 (0.69, 3.39) |
| 2b–2c | 1.29 (1.06, 1.57) | 1.39 (0.94, 2.06) | 1.54 (1.07, 2.21) | 1.37 (0.67, 2.82) |

| Characteristic | All Controls (N=1352) | | Subset with Known Smoking and P16 Status (N=527) | |
|-----------------------------|-----------------------|----------------------------------|--|----------------------------------|
| | Cox PH Regression | GCE Regression | Cox PH Regression | GCE Regression |
| | HR* (95% CI) | ω^+ Ratio (RHR)** (95%CI) | HR* (95% CI) | ω^+ Ratio (RHR)** (95%CI) |
| 3 | 2.19 (1.65, 2.91) | 1.07 (0.60, 1.91) | 2.93 (1.77, 4.84) | 1.00 (0.36, 2.81) |
| Smoking History, pack-years | | | | |
| 10 vs. > 10 | - | - | 0.50 (0.36, 0.70) | 1.05 (0.56, 1.94) |
| P16 Status [¶] | | | | |
| Positive vs. Negative | - | - | 0.53 (0.39, 0.72) | 0.66 (0.37, 1.18) |

Abbreviations: ECOG, Eastern Cooperative Oncology Group. PH, proportional hazards. GCE, generalized competing event. RHR, relative hazard ratio.

* >1 Indicates increased Hazard Ratio for Progression-Free Survival

** >1 Indicates increased Hazard for Cancer Recurrence Relative to Competing Mortality

[¶] Retained in parsimonious GCE model (nomogram)

Table 2.

Comparison of sample size estimates within variously defined high-risk groups.

| | Cancer Recurrence (3-year cumulative incidence (%)) | Competing Mortality (3-year cumulative incidence (%)) | Hazard Ratio* | N |
|--|---|---|---------------|-----|
| Whole Cohort | | | | |
| Highest Tertile | | | | |
| Cox Model for OS | 48.4 | 25.3 | 0.672 | 442 |
| Cox Model for Recurrence | 50.7 | 21.9 | 0.651 | 364 |
| GCE Model* | 36.6 | 6.4 | 0.574 | 307 |
| Highest Quintile | | | | |
| Cox Model for OS | 51.9 | 26.7 | 0.670 | 409 |
| Cox Model for Recurrence | 54.6 | 24.3 | 0.654 | 346 |
| GCE Model* | 36.6 | 5.5 | 0.565 | 293 |
| Subset with Known P16 and Smoking Status | | | | |
| Highest Tertile | | | | |
| Fakhry Model for OS | 52.3 | 20.7 | .642 | 331 |
| Cox Model for Recurrence | 53.4 | 18.5 | .629 | 298 |
| GCE Model | 39.7 | 8.9 | .592 | 315 |
| Highest Quintile | | | | |
| Fakhry Model for OS | 54.4 | 24.6 | .655 | 351 |
| Cox Model for Recurrence | 58.9 | 19.6 | .625 | 264 |
| GCE Model | 34.6 | 11.0 | .600 | 301 |

Abbreviations: OS, overall survival; GCE, generalized competing event

* Projected hazard ratio for recurrence or death from any cause from equation [7], based on observed ratios.