A New Personalized Oral Cancer Survival Calculator to Estimate Risk of Death From Both Oral Cancer and Other Causes

Louise Davies, MD, MS; Benjamin F. Hankey, PhD; Zhuoqiao Wang, MS; Zhaohui Zou, MS; Susan Scott, MPH; Minjung Lee, PhD; Hyunsoon Cho, PhD; Eric J. Feuer, PhD

IMPORTANCE Standard cancer prognosis models typically do not include much specificity in characterizing competing illnesses or general health status when providing prognosis estimates, limiting their utility for individuals, who must consider their cancer in the context of their overall health. This is especially true for patients with oral cancer, who frequently have competing illnesses.

OBJECTIVE To describe a statistical framework and accompanying new publicly available calculator that provides personalized estimates of the probability of a patient surviving or dying from cancer or other causes, using oral cancer as the first data set.

DESIGN, SETTING, AND PARTICIPANTS The models used data from the Surveillance, Epidemiology, and End Results (SEER) 18 registry (2000 to 2011), SEER-Medicare linked files, and the National Health Interview Survey (NHIS) (1986 to 2009). Statistical methods developed to calculate natural life expectancy in the absence of the cancer, cancer-specific survival, and other-cause survival were applied to oral cancer data and internally validated with 10-fold cross-validation. Eligible participants were aged between 20 and 94 years with oral squamous cell carcinoma.

EXPOSURES Histologically confirmed oral cancer, general health status, smoking, and selected serious comorbid conditions.

MAIN OUTCOMES AND MEASURES Probabilities of surviving or dying from the cancer or from other causes, and life expectancy in the absence of the cancer.

RESULTS A total of 22,392 patients with oral squamous cell carcinoma (13,544 male [60.5%]; 1,476 Asian and Pacific Islander [6.7%]; 1,792 Black [8.0%], 1,589 Hispanic [7.2%], 17,300 White [78.1%]) and 402,626 NHIS interviewees were included in this calculator designed for public use for patients ages 20 to 86 years with newly diagnosed oral cancer to obtain estimates of health status-adjusted age, life expectancy in the absence of the cancer, and the probability of surviving, dying from the cancer, or dying from other causes within 1 to 10 years after diagnosis. The models in the calculator estimated that patients with oral cancer have a higher risk of death from other causes than their matched US population, and that this risk increases by stage.

CONCLUSIONS AND RELEVANCE The models developed for the calculator demonstrate that survival estimates that exclude the effects of coexisting conditions can lead to underestimates or overestimates of survival. This new calculator approach will be broadly applicable for developing future prognostic models of cancer and noncancer aspects of a person’s health in other cancers; as registries develop more linkages, available covariates will become broader, strengthening future tools.
When a person receives a new cancer diagnosis, a natural tendency is to focus on the cancer as the main threat to survival. But the person may have other conditions that pose an equal or even greater threat than their cancer—a competing risk of death. In practice, if the person has quite severe comorbid or coexisting conditions, even a relatively serious cancer diagnosis may not end up being the cause of their death. Competing risks are particularly important for cancers in which risk factors for the cancer can also cause coexisting conditions. Personalized data on the competing risks to a person’s survival could provide more highly contextualized data for the patient and their clinician as they discuss cancer treatment options.

To develop this first calculator, we chose oral cancer because prolonged exposures to alcohol and tobacco are risk factors for cancer in this location, but also can result in medical conditions with the potential to shorten life expectancy. Competing as a cause of death that may intervene in conjunction with or before the cancer. The resulting product of this work is the Surveillance, Epidemiology, and End Results Oral Cancer Survival Calculator (SEER OCSC). It provides personalized estimates of the likelihood of surviving or dying from oral cancer or other causes, for newly diagnosed patients ages 20 to 86 years. It is not intended for patients with cancers of the tonsil and tongue base, which are epidemiologically distinct and not anatomically part of the oral cavity. Users are guided through 1 of 3 pathways based on their age and the clinical data available about their general health status. The main aim of this article was to inform clinicians of the SEER OCSC’s purpose and provide a brief description of the statistical framework and the underlying data. Readers are encouraged to access the accompanying Special Communication for key points on clinical use of the tool.

Methods

The Veterans Institutional Review Board of Northern New England reviewed and approved this study as exempt from human participants review under category 2 for the portions involving human participants and category 4 for the data. The data used in this study are available through signed data use agreements with the relevant managing agencies. For more information, please visit the websites for SEER, the National Cancer Institute’s Healthcare Delivery Research Program, or the US Centers for Disease Control and Prevention.

Data Sources

Cancer data were from SEER 18 (2000-2011), and US life tables by race, ethnicity, and sex are from 2011 from the National Center for Health Statistics (Figure 1). SEER-Medicare linked data were used to obtain coexisting condition information for ages 66 years and above. That information was broadened by adding data from persons aged 35 to 94 years who were interviewed in the National Health Interview Survey (NHIS) 1986 to 2009 and provided data on their smoking status and general health self-assessment. While NHIS represents the general population, by conditioning on salient covariates it becomes applicable to patients with oral cancer. Details about the SEER, SEER-Medicare, and NHIS–Linked Mortality File (NHIS-LMF) data sources are included in eAppendix 1 in Supplement 1.

Statistical Analysis

Two Time Scale Competing Risks Modeling

The SEER OCSC system creates personalized models to estimate prognosis utilizing competing risks models. In this case there are 2 possible events, death from oral cancer and death from other causes. The calculator uses 2 time scales at the same time in 1 analysis: for death due to other causes, the time scale is age; for cancer, the time scale is time since diagnosis. Our novel methodology is described in detail in eMethods and eFigure 1 in Supplement 1. Briefly, Lee et al developed prediction methods in continuous time for competing risks data using 2 different time scales (time since diagnosis for cancer death and age for other-cause deaths), extending the work of Cheng et al. Lee et al then developed discrete time methods (more commonly used for large cancer registry data sets) for competing risks predictions on both a single time scale and 2 time scales. The work was further extended by clarifying how to conduct the modeling if each cause of death was modeled using independent or using partially overlapping data sources, and using data sources with nonstandard sampling schemes. All of these modifications allowed flexibility to model competing risks in novel ways, taking advantage of available data as well as the complexity of the risk factors and their associations with the risk of the 2 event types.

Each competing risks of death model specified separate sets of covariates to estimate the risk of oral cancer death and the risk of death from a cause other than oral cancer. The SEER-developed, cause-specific death classification (Howlader et al) was used to identify the oral cancer deaths and deaths from other causes to be included in the survival analyses.

Models

First, we describe the model for estimating the probability of death due to oral cancer, and then we describe the 3 models for estimating the probability of death due to other causes (Figure 1), which varied with respect to the data and covariates included. Model details are in eAppendix 3, eTable 5, and

Key Points

**Question**: How does the competing risk of death from noncancer causes affect risk of death from oral cancer?

**Findings**: The models in the calculator using health data of 22,392 US adults with confirmed oral cancer estimates that patients with oral cancer have a higher risk of death from other causes than a matched US population, and that this risk increases by stage.

**Meaning**: These results suggest that survival estimates that exclude the effects of coexisting conditions can lead to under- or overestimates of survival; this tool provides personalized data for discussions about the place of cancer treatment in the patient’s life as a whole.
from the coexisting conditions model, which was fit using
t heir general health care prior to their cancer diagnosis were
patients ages 66 to 86 years who see a clinician regularly for
model, which used only demographic and staging data as
patients who are ages 20 to 39 years were from the basic
death due to other causes depending on the patient's age
dicated whether the person was in poor health just before di-
2011), and a variable for those age 66 years and above that in-
dicated whether the person was in poor health just before di-
metrics Arising From the Models and Model Variations
SEER-Medicare linked data that added to the basic model a
composite measure of the type and severity of 14
specifically recognized coexisting conditions affecting life
expectancy.20,21 Estimates for patients ages 40 to 65 years,
or those aged 66 years and older who do not see a
clinician regularly for their general health care prior to
their cancer diagnosis were from the general health self-
assessment model. This model was fit using data from the
National Health Interview Survey, a general survey of the US
population with mortality follow-up. This model was an
adaptation of an earlier model by Cho et al.22 Internal vali-
dation of the models was assessed using calibration plots23
and time-dependent area under the receiver operating char-
acteristic curve based on 10-fold cross-validation,24 as
described in eAppendix 6, eFigures 9-24, and eTables 19-21
in Supplement 1.

The probability of death due to oral cancer was estimated
using SEER data and was stratified by stage (according to the
American Joint Committee on Cancer Staging Manual,
Sixth Edition26 [AJCC 6]). The other covariates were tumor
grade, whether the tumor and lymph node status was diag-
nosed clinically or pathologically, sex, age at diagnosis (fit as
a restricted cubic spline), race and ethnicity, socioeconomic
status of the census tract of residence17-19 (in quintiles based
only on persons with oral cancer), marital status, calendar year
of diagnosis (grouped as 2000-2003, 2004-2007, and 2008-
2011), and a variable for those age 66 years and above that in-
dicated whether the person was in poor health just before di-
agnosis (eTables 1-4 in Supplement 1).

There are 3 models for estimating the probability of
death due to other causes depending on the patient's age and
the availability of health information. Estimates for patients
who are ages 20 to 39 years were from the basic model, which
used only demographic and staging data as described above for SEER oral cancer cases. Estimates for patients ages 66 to 86 years who see a clinician regularly for their general health care prior to their cancer diagnosis were from the coexisting conditions model, which was fit using
relative to an average person of the same sex in the US population. Based on each of these 3 models, the predicted cumulative mortality is estimated for a person with a specific covariate profile. This cumulative mortality curve in the absence of cancer is then used to calculate life expectancy. This modeled life expectancy is compared with life expectancies in a sex-specific US life table, and the age in years, 80% of the Medicare population has no recorded coexisting conditions. In comparison, at age 65 there was a decrease to 11% by age 90 years or greater. In the oral cancer cohort, only 52.8% of patients had none of the major coexisting conditions (chronic obstructive pulmonary disease and diabetes). Chronic obstructive pulmonary disease occurred for 1081 of 6181 patients (17.5%) in the oral cancer cohort age 66 years and older, while in the Medicare population more broadly, as shown by Cho et al,27 the rate starts at only 6% and increases to 11% by age 90 years or greater. In the oral cancer cohort, only 52.8% of patients had none of the major recorded coexisting conditions. In comparison, at age 65 years, 80% of the Medicare population has no recorded coexisting conditions.27 The percentage falls to 60% with no

### Results

Of a total 22,392 included patients, 13,544 (60.5%) were male; 1792 (8.0%) were Asian and Pacific Islander, 1792 (8.0%) Black, 1589 (7.2%) Hispanic, and 17,300 (78.1%) White (Table). The mean age at diagnosis for persons who had stage I cancer was 61.6 years, stage II was 64.7 years, stage III was 62.9 years, and stage IV was 62.3 years. Comorbidity frequencies were estimated from SEER data linked to the Medicare data at the Centers for Medicare & Medicaid Services and are presented in eTable 3 in Supplement 1. The most common coexisting conditions were chronic obstructive pulmonary disease and diabetes. Chronic obstructive pulmonary disease occurred for 1081 of 6181 patients (17.5%) in the oral cancer cohort age 66 years and older, while in the Medicare population more broadly, as shown by Cho et al,27 the rate starts at only 6% and increases to 11% by age 90 years or greater. In the oral cancer cohort, only 52.8% of patients had none of the major recorded coexisting conditions. In comparison, at age 65 years, 80% of the Medicare population has no recorded coexisting conditions.27 The percentage falls to 60% with no
coexisting conditions by age 90 years and older, but this proportion is still lower than in the oral cancer cohort.27

**Estimates of the Risk of Death From Other Causes for Patients With Oral Cancer**

We estimated the risk of death from other causes for patients with oral cancer by creating survival curves in which death from causes other than oral cancer was the event of interest, censoring those who died of oral cancer, and conditional on surviving to 50 years of age using left-truncated models computed on the age scale. The resulting curves represent noncancer cause-of-death life tables for persons diagnosed with oral cancer.28 We then compared them with the gender-specific US life tables for 2011. **Figure 2** shows the estimated likelihood of noncancer survival of patients with oral cancer in the cohort by sex and stage, conditional on being alive at age 50 years. Noncancer survival is the survival estimate given the cohort’s characteristics, but as though they did not have cancer. People with oral cancer have greater competing risks of death (worse noncancer survival) at each stage than the general population, and their noncancer survival worsens by stage. For example, conditional on having survived to age 50 year, a female and male patient diagnosed with stage III cancer would have a 60% and 44% chance, respectively, of being alive at age 70 years, in the absence of their cancer. In the general US population, the corresponding estimates are 86% and 79%, respectively, an absolute difference of 26 and 35 percentage points, respectively.

**Model Results and Validation**

At all disease stages, patients with only clinical staging information rather than pathologic staging had a higher risk of death due to oral cancer. Histologic grade was also found to be important in modeling cancer death, with the risk increasing in a stepwise fashion by grade, although the association of grade with death diminished with each increasing stage of disease.29 Among patients aged 20 to 65 years with AJCC 6 stage I or II cancer, the adjusted risk of death was 2.7 times greater (95% CI, 1.72-4.11) if the tumor is poorly differentiated or undifferentiated than if the tumor is well differentiated. Among patients aged 66 to 94 years, the risk of death was 3.0 (95% CI, 2.02-4.54) times greater. Depending on the stage of disease, non-Hispanic Blacks and/or Hispanics patients had an increased risk of cancer death compared with non-Hispanic Whites and non-Hispanic Asian and Pacific Islander patients, with the association of race and ethnicity with death due to oral cancer diminishing slightly for higher stages. For example, for a 70-year old male patient with diabetes, who is not married, has stage I disease, surgically assessed primary tumor and nodes, and a poorly differentiated histologic grade, the cumulative 3-year chance of dying from oral cancer is 11.1% (95% CI 8.5%-14.1%) if he is non-Hispanic White, 12.7% (95% CI, 9.0%-17.2%) for if he is non-Hispanic Asian and Pacific Islander, 14.5% (95% CI, 10.0%-19.7%) if he is non-Hispanic Black, and 15.1% (95% CI, 10.4%-20.6%) if he is Hispanic. For the models for causes of death other than oral cancer, the baseline parameters as a function of age provide a life table that forms the baseline for all other patient covariate profiles. The risk of death from causes other than oral cancer increased as a function of stage in both the basic and coexisting conditions models. For example, in the basic model male patients had a relative risk of death from other causes of 1.26 (95% CI, 1.13-1.41) for stage II, 1.35 (95% CI, 1.19-1.54) for stage III, and 1.52 (95% CI, 1.36-1.69) for stage IV, all as compared with stage I. After adjusting for coexisting conditions, the association of stage was somewhat diminished at higher stages, but socioeconomic status of the census tract of residence was significant in both of these models. For example, in the basic model for men, the risk of death due to other causes was 1.53 (95% CI, 1.33-1.75) in the lowest quintile, 1.37 (95% CI, 1.19-1.57) in the second quintile, 1.26 (95% CI, 1.09-1.45) in the third quintile, and 1.22 (95% CI, 1.06-1.40) in the fourth quintile, all as

---

**Figure 2. Estimated Noncancer Survival Probability Among Patients With Oral Cancer**

![Figure 2](https://example.com/)
compared with the highest quintile. In the general health assessments models, both smoking status and self-assessed overall health status were statistically significant. Considering that risk factors, treatment patterns, and diagnostic precision or care for other conditions may affect survival probability, we analyzed the periods 2008 to 2011 and 2012 to 2017 separately as well as by each year based on time since diagnosis. A comparison of SEER cause-specific survival for oral cancers diagnosed from 2012 to 2017 compared with 2008 to 2011 only showed an improvement of 1.6 percentage points at 5 years (63.4% vs. 65.0%) and minimal changes in mortality over time, reflecting secular trends in tobacco use that affect the population at risk of getting and dying of oral cancer. This system is that it provides numeric estimates that are representative of what a typical person is likely to experience with their cancer, rather than reflecting the outcomes of a particular institution, case series, or randomized trial cohort. The SEER OCSC includes alternative and complementary data sources such as the NHIS, which is a strength be-
cause it provides data on smoking and self-assessed general health status, as well as Medicare data, which include records of claims for health care, allowing estimates of illness type and severity. Socioeconomic status is incorporated into the system by including information about the person’s area of residence (if they choose to provide it). The inclusion of detailed information about coexisting conditions in the prognosis calculations is particularly useful in oral cancer because alcohol intake and tobacco use, which are commonly associated with the development of these cancers, result in other medical problems that can adversely affect survival even in the absence of cancer, such as heart disease, chronic obstructive pulmonary disease, and liver disease.4

The SEER OCSC, like all such tools, has limitations. The coexisting conditions data come from Medicare files, so the information for the models is directly available only for those 66 years of age or older. Some important prognostic variables, such as cancer margin status, data on extracapsular tumor spread, and depth of tumor invasion, and intangible but important factors such as levels of family and social support are not available in registry or other population-based data sources. The cancer staging data are from AJCC 6. Cancer-specific cumulative mortality estimates provided by the SEER OCSC may be slightly higher than those estimated using the eighth edition of the AJCC. There is not yet sufficient follow-up time available to use AJCC 8, which just came into use in 2018. Additionally, there are variables that we have chosen not to include, such as pack-years of smoking and alcohol intake. Lastly, death certificate data are inherently imperfect. Methods have been devised to overcome these problems,6 but misclassification can result in over- or under-estimation of cancer deaths.

Conclusions

The publicly available SEER OCSC uses novel approaches to modeling competing risks of death: the risk of death from other causes is treated on equal footing with the consideration of the probability of death from the cancer. It provides personalized estimates of health-status-adjusted age, life expectancy in the absence of the cancer, and probability of surviving, dying from the cancer, or dying from other causes 1 to 10 years after diagnosis. A novel extension of existing methodology using registry data, the approach will be broadly applicable for developing prognostic models that can capture the cancer and noncancer aspects of a person’s health. As registries develop more linkages, available covariates will become broader, making future tools even more clinically relevant. The models developed for the SEER OCSC demonstrate that survival estimates that exclude the effects of coexisting conditions can lead to under-or overestimates of survival. In the case of oral cancer, the model and this calculator show that many patients with oral cancer have a greater risk of dying of other causes due to the number and type of coexisting conditions they have. Furthermore, their likelihood of dying of other causes increases as their cancer stage increases, even after adjusting for coexisting conditions.

ARTICLE INFORMATION

Accepted for Publication: June 13, 2023.
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2023 Davies L et al. JAMA Otolaryngology–Head & Neck Surgery.

Author Affiliations: VA Outcomes Group, Department of Veterans Affairs Medical Center, White River Junction, Vermont (Davies); Section of Otolaryngology in Geisel School of Medicine at Dartmouth, and The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, New Hampshire (Davies); Statistical Research and Application Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland (Hankey, Feuer); Information Management Services, Calverton, Maryland (Wang, Zou); Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland (Scott); Department of Statistics, Kangwon National University, Chunchu, Gangwon, Korea (Lee); Department of Cancer AI and Digital Health, National Cancer Center Graduate School of Cancer Science and Policy, and the Integrated Biostatistics Branch, Division of Cancer Data Science, National Cancer Center, Goyang, Gyeonggi-do, Korea (Cho).

Author Contributions: Drs Feuer and Wang had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Davies, Wang, Scott, Feuer. Acquisition, analysis, or interpretation of data: Davies, Wang, Zou, Lee, Feuer. Drafting of the manuscript: Davies, Scott, Feuer. Critical review of the manuscript for important intellectual content: Davies, Hankey, Wang, Zou, Lee, Cho, Feuer. Statistical analysis: Wang, Zou, Lee, Feuer. Administrative, technical, or material support: Feuer. Supervision: Feuer.

Conflict of Interest Disclosures: None reported.

Funding/Support: The Department of Veterans Affairs and the National Cancer Institute supported this study as part of an interagency agreement.

Role of the Funder/Sponsor: Dr Feuer, Dr Hankey, Dr Davies, Mr Wang, Mr Zou, and Ms Scott are employees of or contractors for the National Cancer Institute and were involved in the design and conduct of the study and the collection, management, analysis, and interpretation of the data. The National Cancer Institute reviewed and approved the manuscript and the calculator. The National Cancer Institute was not involved in the preparation of the manuscript or the decision to submit the manuscript for publication. The Department of Veterans Affairs 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.

Disclaimers: The views expressed do not necessarily represent the views of the Department of Veterans Affairs, the National Cancer Institute, or the US Government. Also, Dr Davies is Associate Editor of JAMA Otolaryngology–Head & Neck Surgery but was not involved in any of the decisions regarding review of the manuscript or its acceptance.

Meeting Presentation: An early version of this work was presented as part of an instructional conference at the 5th International Conference on Head & Neck Cancer; July 26, 2014; New York, New York. This work was presented simultaneously with publication at the AHNS 11th International Conference on Head & Neck Cancer; July 10, 2023; Montreal, Quebec, Canada.

Data Sharing Statement: See Supplement 2.

REFERENCES


