Multivariate Analyses to Assess Treatment Effectiveness in Advanced Head and Neck Cancer
Urjeet Patel, MD; Edward Spitznagel, PhD; Jay Piccirillo, MD

**Objective:** To assess relative benefit of combined radiotherapy and surgery over single-modality treatment for advanced-stage squamous cell carcinoma of the aerodigestive tract by means of several multivariable analyses to control for patient variables.

**Design:** Medical chart review.

**Setting:** University medical center.

**Patients and Methods:** The study included 532 patients receiving initial therapy between January 1, 1980, and December 31, 1989. Three multivariate techniques (multiple logistic regression, propensity score stratification, and conjunctive consolidation) were used to compare outcomes for treatment groups.

**Main Outcome Measure:** Five-year survival.

**Results:** Survival for radiation, surgery, and combined treatment groups were 24%, 40%, and 46%, respectively. With the use of multiple logistic regression to control patient variables, the radiation group had a significantly lower survival than the combined therapy group (risk ratio, 2.24; 95% confidence interval, 1.32-3.80), while there was no statistical difference for the surgery group compared with the combined therapy group (risk ratio, 1.26; 95% confidence interval, 0.78-2.03). When analyzed by propensity score, 5-year survival was higher in each quintile for the combined therapy group than for the group who received radiation alone ($P=.002$). There was no significant difference in survival between the surgery and combined treatment groups ($P=.25$). Conjunctive consolidation was used to create a clinical staging system to compare outcomes across treatment groups. In each clinical severity stage, radiation alone had a lower survival than combined therapy ($P=.001$), while no statistical difference was noted between surgery and combined therapy ($P=.50$).

**Conclusions:** All 3 statistical techniques showed a significantly lower survival for patients treated with radiation alone vs combined therapy. No significant difference was noted between surgery and combined therapy. Propensity score analysis and conjunctive consolidation are useful techniques to control prognostic variables in cancer database studies and should be used in future outcome studies that address more current treatment dilemmas in head and neck oncology.


Combined surgery and postoperative radiotherapy are often used to treat stages III and IV squamous cell carcinoma of the upper aerodigestive tract. The choice of combined treatment over single-modality treatment often hinges on clinical factors such as primary site, tumor size, and extent of regional metastasis. Some centers report improved locoregional control with combined therapy, while others report improved survival; however, controversy exists over the benefit of combined therapy in various clinical settings.

The criterion standard for assessing the merits of a given treatment is the prospective randomized trial. Accordingly, such trials have long been advocated; however, these trials in head and neck cancer treatment are inherently problematic and rare. The main problem is the heterogeneity of the study population in terms of tumor stage, primary site, histologic grade, age, and small sample size for any given research trial. In addition, it is often difficult to randomize patients to treatments that are so markedly different, as patients are hesitant to leave such grossly dissimilar options to chance alone. Therefore, studies of treatment effectiveness in head and neck cancer are often relegated to the realm of observational studies, where patients are not randomized to particular treatments.

The goal of the observational study is to measure treatment effectiveness. One of the major difficulties in the analysis of results from observational studies is that the
PATIENTS AND METHODS

POPULATION UNDER STUDY

We studied 332 patients with newly diagnosed TNM stage III or IV biopsy-proved squamous cell carcinoma who were first treated at Washington University Medical Center, St Louis, Mo, between January 1, 1980, and December 31, 1989. These patients were initially identified by means of records from the pathology department of Barnes-Jewish Hospital, St Louis. Patients with American Joint Committee on Cancer (AJCC) TNM stage III or IV squamous cell carcinoma of the oral cavity, oropharynx, or larynx who were initially treated with radiotherapy, surgery, or combined radiation and surgery were included in the study population. Patients with metastatic disease at the time of diagnosis were excluded from the study, as were patients who received therapy other than the 3 above-mentioned treatment options. Baseline and follow-up information was obtained from inpatient medical records as well as records from the Departments of Otolaryngology and Radiation Oncology. Full 5-year follow-up information was obtained for all 332 patients. Supplemental date of death and death certificate information was obtained from Equifax National Death Search (Arlington, Va).

COLLECTION OF DATA

Specially designed data extraction forms were used to ensure uniform data collection from the medical records. Data collected before the pretreatment interval and at the time of presentation included basic demographic information, risk factors, medical history, symptom type and duration, complete anatomic description of the tumor including the TNM classification with the 1992 AJCC criteria, pathological description of the biopsy specimen, and details of subsequent therapy. The zero-time for each patient was chosen as the date of first antineoplastic intervention directed at the primary site. Follow-up data, including development of recurrence, new primary, and subsequent treatment, were also collected. Patient and tumor status at last follow-up or death was obtained.

CLASSIFICATION OF DATA

To maintain scientific accuracy and ensure high quality of data, imperfections in data obtained from retrospective studies must be managed in a systematic and consistent manner. The general methods for such management have been previously described.7,8 SYMPTOM SEVERITY

To study the prognostic importance of symptoms for a specific cancer type, the presence of symptoms and their relationship to the primary cancer must be clearly established. To manage possible discrepancies in the medical record, 2 conventions were consistently applied. If a symptom was recorded by at least 1 examiner, the symptom was regarded as present. If different periods of duration were reported, the longer duration was recorded. The details of symptom severity staging as used in our study have been previously described.8 Briefly, the symptoms of dysphagia, otalgia, neck lump, and weight loss were found to be independent predictors of survival. Accordingly, a symptom severity staging system was developed on the basis of the presence of these symptoms. Stage was defined as none if none of the 4 symptoms was recorded, mild if 1 of the 4 symptoms was recorded, moderate if 2 of the 4 symptoms were recorded, and severe if 3 or 4 of the 4 symptoms were recorded.

COMORBIDITY

The presence of concomitant disease unrelated to the disease under study is termed comorbidity. Comorbidity has been shown to clearly impact on survival and treatment selection in several types of cancer.10-12 The Kaplan-Meier index was used to classify comorbidity for this study.13 This scheme was used to classify the patients' comorbidity as none, mild, moderate, or severe (grades 0, 1, 2, or 3, respectively). When a patient's condition was described in the medical record as too sick to tolerate standard antineoplastic therapy, a grade of 3 was assigned regardless of other illnesses. Prognostic comorbidity was defined as grade 3, signifying the presence of concomitant illness that significantly reduces a patient's life expectancy.

CANCER STAGING

The staging criteria for all tumors were reviewed according to the AJCC cancer staging manual.8 All information obtained before the zero-time was used to assess accuracy of the recorded stage as dictated by the AJCC rules. In the case of staging discrepancies between written notes by different physicians where the medical record lacked sufficient anatomic information to accurately restage the tumor, the stage assigned by the most senior otolaryngologist or radiation oncologist was recorded. Information regarding the presence of cervical adenopathy was lacking in 4 members of the final cohort. It was known that they had stage III or IV disease based on T stage alone; subsequently, they were included in the study. These members were omitted from aspects of data analysis requiring exact node status information.

PATHOLOGICAL EXAMINATION

The histologic grade of the primary tumor was recorded from the biopsy or primary specimen for all patients, and grades were grouped into categories of well, moderately, and poorly differentiated. If both biopsy specimen and primary tumor were available, the biopsy specimen was used to define the histologic grade. Specimens graded as moderately to well differentiated were recorded as moderate, and those graded as moderately to poorly differentiated were recorded as poor. Histopathologic grade was absent for 4 patients; these members were omitted from data analyses requiring pathological information.

PRIMARY TREATMENT

Information regarding each patient's initial treatment included type of treatment (radiotherapy, surgery, or combined treatment), type of surgical procedure, timing of radiotherapy (preoperative or postoperative), and therapeutic
complications. Subsequent treatment was defined as treatment initiated secondary to failure of primary therapy and was also recorded.

FOLLOW-UP AND OUTCOME

Each patient was monitored for persistence, recurrence, and development of new primary cancer. Follow-up was considered complete when either a patient's death was documented or a minimum of 5 years survival was obtained. The primary outcome measure presented in this study was 5-year survival.

DATA ANALYSIS

The primary objective of data analysis was to estimate any possible benefit on 5-year survival of combined therapy over either radiation or surgery alone. The possible benefit was estimated by 3 separate multivariable statistical techniques: multivariate logistic regression, propensity score stratification,14 and conjunctive consolidation.12

The information from the data extraction forms was entered into a Paradox database (Borland International, Scotts Valley, Calif). The specially designed database screens were equipped with internal validity checks that facilitated reliable and efficient data entry. Periodic review for internal consistency and comparison with separate databases was performed to ensure accuracy of data entry. Sorting, tabulation, and statistical analyses were performed with the SAS system, release 6.12 (SAS Institute Inc, Cary, NC).

Logistic Regression

The impact of covariates and initial treatment options on 5-year survival was evaluated by multiple logistic regression (PROC LOGIST function). The logistic regression modeled the dependent variable of 5-year survival from the independent patient, tumor, and treatment variables. A regression model was fit with the use of the following covariates: age group, sex, race, prognostic comorbidity, symptom severity, pathological findings, tumor size, presence of adenopathy, primary site, and initial treatment choice. Of the 532 patients, 7 were eliminated from the regression model because of missing information as described in the "Cancer Staging" and "Pathological Examination" subsections. The multivariable regression had an area under the receiver operating characteristic curve of 0.72. This means that the regression model was fairly accurate in discriminating survivors from nonsurvivors on the basis of covariate information. Adjusted risk ratios and corresponding 95% confidence intervals and P values were obtained according to reference groups for each variable.

Propensity Score

The goal of propensity analysis is to reduce the effect of selection bias between 2 treatment options as described by Rosenbaum and Rubin.16-17 Selection bias is clearly problematic in observational studies when clinical covariates (age, comorbidity, tumor stage, etc) impact on both treatment (radiation, surgery, or combined therapy) and outcome (5-year survival). Propensity score stratification seeks to replace the wide host of confounding covariates that may be present in an observational study with a single variable function of these covariates. The covariates are summarized into a single probability function called the propensity score that describes the likelihood of receiving treatment A (surgery plus radiation, for example) vs treatment B (radiation alone). The propensity score can be estimated through logistic regression of the covariates on treatment choice. Accordingly, each individual has a propensity score that represents the probability of being treated with combined therapy rather than radiation alone. The propensity score is then used in further analysis as the single confounding variable.

The study population is then stratified into a discrete number of groups, usually 3, on the basis of the propensity score. Stratification into 5 quintiles has been shown by Rosenbaum18 to eliminate more than 90% of selection bias by covariates. Within each propensity stratum, there will generally be a number of patients who received combined therapy or radiation alone. The rationale behind the propensity score scheme is as follows: If 2 patients have the same propensity score, then it follows that they have the same likelihood of receiving combined treatment as radiation alone on the basis of their given covariates. If the 2 patients receive different treatments, then the choice of treatment can be considered random. The same principle holds for 2 groups with similar propensity scores. Within a given propensity stratum, the group of patients receiving combined therapy will have a distribution of propensity scores similar to that of patients who received surgery alone. Subsequently, the patients composing one treatment group can be considered to be randomly chosen from the entire propensity stratum with regard to their confounding covariate data. Within a propensity stratum, the multivariate distribution of covariates should differ only randomly between the 2 treatment groups as if they had been randomly assigned a treatment option. Thus, use of this technique with stratification into quintiles eliminates selection bias between 2 treatment groups.19

In our study, propensity score analysis was first used to assess treatment effect between patients receiving combined therapy vs radiation alone. To identify variables that were unbalanced between the 2 treatment groups, bivariate screening was performed for all potential confounding covariates that potentially impact on treatment decision. A multiple logistic regression was performed in a stepwise fashion to determine important predictors of treatment selection. A logistic regression model was then fit with variables found to be significant (P<.15) in the logistic analysis. The area under the receiver operating characteristic curve for this regression model was 0.76, indicating good discrimination between patients receiving combined vs single therapy. With this model, a propensity score was calculated for each patient that predicts the likelihood of being initially treated with combined surgery and radiation therapy vs radiation alone.19 Patients were then sorted by propensity score and clustered into quintiles accordingly. Bivariate screening and logistic regression were then performed within each quintile to identify any remaining bias among covariates after stratification by propensity score. The effect of treatment assignment on 5-year survival was then analyzed within each quintile. The Mantel-Haenszel odds ratio was calculated in addition to the Cochran-Mantel-Haenszel (CMH) x². The Mantel-Haenszel odds ratio represents a composite of the 5 odds ratios derived from

Continued on next page
The characteristics of the 532 study patients are presented in Table 1. The study population was 70% male, and more than 75% were white. Most patients had either absent or mild comorbidity as well as symptom severity. More than 65% had evidence of cervical adenopathy and were equally divided between TNM stage III and stage IV disease. Disease was most prevalent in the oropharynx (40%), followed by the larynx (35%) and then the oral cavity (25%). The most common initial treatment plan was combined surgery plus radiation (52%), followed by radiation alone (25%) and then surgery alone (23%).

The relationship between covariates, including treatment and survival, is also shown in Table 1. Patient characteristics that were associated with decreased survival included increasing age ($P = .01$), increasing comorbidity ($P = .03$), and increasing symptom severity ($P = .002$). Presence of cervical adenopathy was associated with a significant decrease in survival from 47% to 35% ($P = .01$). Laryngeal cancer was associated with the highest 5-year survival at 47%, compared with oral cavity and oropharyngeal disease at 37% and 33%, respectively. Looking at treatment, radiation alone was associated with significantly lower survival of 24% while survival rates for combined therapy and surgery alone were 46% and 40%, respectively.
Multivariate logistic regression was used to assess the impact of covariates and treatment on survival. **Table 2** shows the adjusted risk ratios for covariates and treatment for the study population. Increasing age and male sex were both related to a decreased 5-year survival rate. Similarly, the presence of cervical adenopathy and tumor size greater than stage 1 impact negatively on survival. With regard to primary site, patients with cancer of the oral cavity and oropharynx had significantly higher risk of death than the patients with laryngeal cancer. With regard to treatment, the group of patients treated with radiation alone had a significantly higher risk of death than those receiving combined therapy (risk ratio, 2.24; 95% confidence interval, 1.32-3.80). While treatment with surgery alone did reflect an increased risk of death compared with patients receiving combined treatment, this risk was not found to be statistically significant (risk ratio, 1.26; 95% confidence interval, 0.78-2.03).

Stratification by propensity score and assessment of the treatment effect of each single-modality treatment and combined therapy on survival were performed. The population receiving either radiation alone or combined therapy was examined first (n=410). The population was stratified into propensity quintiles as previously described. **Table 3** shows survival rates for both treatment groups after stratification. The percentage of patients receiving combined therapy decreased from the first propensity quintile to the fifth as predicted by the propensity model. In each of the 5 strata, patients receiving combined therapy had a higher 5-year survival rate than the group receiving...
ing radiation alone. In quintiles 1 and 3, the difference in survival was statistically significant. The \( P \) value for the CMH \( \chi^2 \) comparing survival between the treatment groups while controlling for propensity quintile was .002, suggesting a strong difference in survival between those receiving radiation alone vs combined treatment.

Propensity score analysis was similarly performed for patients initially receiving either surgery alone or combined therapy (n=397). Table 4 shows survival rates for these treatment groups after propensity score stratification. As expected, the percentage of patients receiving combined therapy decreased from the first to the fifth quintile. Within quintile 2, patients receiving surgery alone had a higher survival than those receiving combined treatment, while in the remaining quintiles, patients with combined treatment had more favorable survival. In none of the quintiles was the difference in survival statistically significant. The \( P \) value for the CMH \( \chi^2 \) was .25, suggesting no significant difference between treatment groups across quintiles.

Conjunctive consolidation was performed as previously described. Table 5 shows survival rates for all patients according to composite stage (A, B, C, or D). A prognostic gradient was noted in survival from stage A through stage D, with a significant \( \chi^2 \) for linear trend \( (P=.001) \). Patients were separated into treatment group, and survival rates were then compared within composite staging groups (Table 6). In all 4 composite staging groups, survival rates were higher for patients receiving combined therapy compared with radiation alone, with a statistically significant difference noted in 3 of the 4 stages. The CMH \( \chi^2 \) was highly significant \( (P=.001) \). Comparing groups receiving combined therapy vs surgery alone, survival was higher for combined therapy in stages A, B, and D, while surgery alone was favored for stage C patients. In only stage D was there a statistically significant difference in survival between patients receiving surgery vs combined treatment. No statistical difference was noted between the 2 groups by the CMH \( \chi^2 \) \( (P=.50) \).

**COMMENT**

Our research demonstrates the usefulness of multivariate analysis with regard to head and neck oncology observational studies. The relative benefit of combined therapy over single-modality therapy was assessed by means of multiple logistic regression, propensity score analysis, and conjunctive consolidation. The 3 forms of analysis concurred in their findings that combined therapy offered significantly higher survival at 5 years than radiotherapy alone. In contrast, no significant difference was seen when combined therapy was compared with surgical treatment alone. By using multivariate analysis to eliminate selection bias,
the difference in survival can be attributed to treatment effect without the influence of confounding variables.

Previous studies have compared combined therapy with single-modality treatment for various tumors in the head and neck region. While many of these studies seek to measure treatment effect, few of them compare different treatment options while controlling for selection bias. Patient variables such as comorbidity, pathological grade, symptom severity, and age are often omitted from analysis despite the fact that these variables may influence treatment choice as well as outcome. Subsequently, conclusions are drawn regarding treatment effectiveness without adequately controlling for potential selection bias. Without such control, the conclusions may not accurately assess true treatment effectiveness.

While multivariate analysis does permit a more controlled estimate of treatment effectiveness, there do exist potential inaccuracies in its formulation. Each multivariate model is able to control only the study variables included in the analysis. In our study, there was no variable to quantify the amount of radiotherapy given to each patient, nor was there any measure in the quality of the surgery performed. Subsequently, it is possible that variables exist that would alter the measured treatment effects had they been included in the multivariate analysis. In addition, a given multivariate analysis makes use of statistical models to approximate the data being analyzed. The degree to which a given model fits the data appropriately can vary and needs to be considered when the results of statistical analysis are interpreted. This is especially true when different statistical tools yield varying results when the same data are analyzed.

Multivariate analysis is a highly useful tool to measure treatment effect in observational studies. Multiple logistic regression is a statistical technique that is frequently used in analysis of observational study results. Propensity score analysis and conjunctive consolidation are also highly effective at controlling selection bias to measure treatment effectiveness. The use of these techniques will improve the ability to accurately measure treatment effectiveness in observational studies. These tools may be applied to more current clinical dilemmas, such as chemoradiation protocols compared with surgical resection for treatment of head and neck cancer.

Accepted for publication October 11, 2001.

This study was supported in part by grant R01CA2072 from the National Cancer Institute, Bethesda, Md (Dr Piccirillo).

Table 6. Bivariate Analysis of 5-Year Survival and Treatment for 528 Patients by Conjunctive Consolidation Staging

<table>
<thead>
<tr>
<th>Radiation</th>
<th>Combined Therapy</th>
<th>Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.001†</td>
<td></td>
</tr>
</tbody>
</table>

*P values correspond to comparison between single modality vs combined therapy.
†Cochran-Mantel-Haenszel χ² P values for single vs combined therapy.

This study was presented at the Fifth International Conference on Head and Neck Cancer, San Francisco, Calif. July 31, 2000.

Corresponding author and reprints: Jay Piccirillo, MD, Department of Otolaryngology, 660 S Euclid St, Box 8115, St Louis, MO 63110 (e-mail: piccirij@msnotes.wustl.edu).

REFERENCES